Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
Commuter’s Journey to Work Travel Behaviour and the Aggregate Road Passenger Travel Demand in New Zealand

Mingyue Sheng

A thesis submitted for the degree of Doctor of Philosophy in Economics at the University of Auckland

Auckland, New Zealand

2016
Abstract

Economic development has historically been strongly associated with an increase in the demand for transportation and particularly in the number of road vehicles. However, traffic congestion, deterioration in air quality and climate change concerns also arise as a result of this escalation in road transport. Since New Zealand ratified the Kyoto Protocol in 1997, cutting down carbon dioxide equivalent (CO₂-e) emissions has been a priority for contemporary government policy. Thus, reducing CO₂-e emissions from road transport turns out to be critical because this sector alone accounts for 40% of all emissions in the country’s energy emission profile in 2012 (Ministry of Business, Innovation and Employment [MBIE], 2013). To date, local authorities and urban planners have shifted their interest to revitalising public transport as one practical approach of combating the negative externalities generated from road transport. Given that New Zealand, especially Auckland, has a relatively low level of public transport ridership compared to other Australasian cities, the understanding of which variables influence public transport demand at regional level, how travel decisions made by individual commuters, as well as what factors affect the demand for aggregate road passenger travel at national level become key questions to consider.

This thesis contributes to the existing research on the analysis of commuter journey-to-work (JTW) behaviour in a spatial context at both the regional and the individual level. It also fills the research gap in the past literature by examining road passenger’s transport mode choices as a system of equations at the national level. Chapter 2 reviews the literature around traveller's travel behaviour and provides an overview of the methodology used in the following chapters. Using regional level JTW data, chapter 3 examines the relationship between urban form and public transport use in Auckland by applying a spatial Durbin model. Taking network effects into account, chapter 4 investigates individual commuter’s transport mode preferences in Auckland for their JTW travel by estimating a spatially autoregressive logit mode choice model. Chapter 5 develops an aggregate road passenger travel demand model using the seemingly
unrelated regression (SUR) method, and the empirical results from the SUR model deliver some important policy implications in terms of achieving a reduction in the demand for both petrol and diesel cars, and also promoting the use of public transport. Chapter 6 concludes.
Acknowledgements

First, I would like to express my most special thanks to my main supervisor, Professor Basil Sharp, for his constant encouragement, enlightening guidance and prominent enthusiasm throughout the years of my PhD. His generosity and help have been an ongoing inspiration which guided me through the whole process of producing this thesis in the past four years.

Second, I would like to give my sincere appreciation for the constructive comments from my co-supervisor Dr Erwann Sbai, and Professor Mark Greer at the Energy Centre on an early draft of this study.

Last but not least, I would like to take this opportunity to thank my mother, Liping, my husband, Qiping, my sons, Bowen and Damion, my friends and my postgraduate colleagues for their persistent support and patience, never-faded affection and distinguished confidence in me over all these years.

Any errors found in this thesis are my own responsibility.
Table of Contents

Abstract ................................................................................................................................................... ii
Acknowledgements ................................................................................................................................. iv
List of Figures ....................................................................................................................................... viii
List of Tables .......................................................................................................................................... ix
Chapter 1. Introduction .......................................................................................................................... 1
  1.1. The Problems Facing Transport Network in New Zealand........................................................... 1
  1.2. Motivation.................................................................................................................................... 4
  1.3. Objectives and Scope ................................................................................................................... 6
  1.4. Contributions ............................................................................................................................... 9
Chapter 2. Background Information, Literature Review and Overview of Methodology .................... 12
  2.1. Background Information of New Zealand Transport Sector ...................................................... 12
    2.1.1. Physical impacts of climate change .................................................................................... 12
    2.1.2. Climate change challenge on transport sector ................................................................. 13
  2.2. The Studies of the Impact of Urban Form on Travel Behaviour ................................................ 22
  2.3. Applications of Discrete Choice Model to Travel Behaviour ..................................................... 25
  2.4. Research Gap ............................................................................................................................. 29
  2.5. Review of Spatial Effects and Spatial Regression Models .......................................................... 31
    2.5.1. Spatial dependence and spatial heterogeneity ................................................................. 31
    2.5.2. Spatial lag model ................................................................................................................. 34
    2.5.3. Spatial error model ............................................................................................................. 35
    2.5.4. Spatial Durbin model .......................................................................................................... 37
    2.5.5. Spatial autoregressive logit model ...................................................................................... 39
Chapter 3. The Influence of Urban Forms on Transit Behaviour in the Auckland region: A Spatial
  Durbin Analysis ..................................................................................................................................... 41
  Abstract ................................................................................................................................................. 41
  3.1. Introduction ............................................................................................................................... 42
    3.1.1. Background information ..................................................................................................... 42
    3.1.2. Objective, motivation and scope ........................................................................................ 43
  3.2. Data and Empirical Models ........................................................................................................ 46
    3.2.1. Data ..................................................................................................................................... 46
    3.2.2. Variables .............................................................................................................................. 47
    3.2.3. Empirical bus mode share models ...................................................................................... 51
5.4.1 Aggregate road passenger transport demand functions ............................................................ 114
5.4.2 SUR modelling .......................................................................................................................... 115
5.4.3 Demand for aggregate road passenger transport: the SUR model ........................................... 117
5.5. Data Description ....................................................................................................................... 119
5.6 Stationarity Test ....................................................................................................................... 122
5.7 Empirical results and policy implications .................................................................................. 124
5.8 Test for cointegration and structural breaks ............................................................................. 131
5.8.1 Cointegration Test .................................................................................................................. 131
5.8.2 Structural breaks ................................................................................................................... 132
5.9 Conclusion and Recommendations ........................................................................................... 132
Chapter 6. Summary, Conclusions and Directions for Future Research ........................................ 135
6.1. Research objectives ................................................................................................................... 135
6.2. Results summary ...................................................................................................................... 136
6.2.1. The influence of urban forms on transit behaviour .............................................................. 136
6.2.2. Commuter's transport mode preferences and network effects .......................................... 138
6.2.3. Aggregate road passenger travel demand in New Zealand: a seemingly unrelated regression model .......................................................................................................................... 139
6.3. Suggestions for future research ............................................................................................... 142
Appendix A Descriptive Statistics for the Subgroup when Choice = 1 ........................................... 145
Appendix B Geocoded Destination Locations of All Included Trips ................................................. 146
Appendix C log-value of variables vs. Time .................................................................................... 147
Appendix D Residuals vs. Time ........................................................................................................ 149
Appendix E Predicted vs. Actual Demands .................................................................................... 150
List of References .......................................................................................................................... 151
List of Figures

Figure 1.1: Annual Public Transport Boardings per capita in New Zealand, 2005/06-2014/15... 4
Figure 2.1: Global Growth in Energy Sector Emissions from Base Year to 2011.................. 15
Figure 2.2: Energy CO$_2$-e Emissions by Sector, 2012 (in kt CO$_2$-e)................................. 17
Figure 2.3: Energy Emissions by Fuel, 2012 (in kt CO$_2$-e).................................................. 17
Figure 2.4: CO$_2$-e Emissions from the Transport Sector in NZ’s Cities/Regions, 2001........ 19
Figure 2.5: Relationship between SDM, SAR and SEM ......................................................... 38
Figure 3.1: Public Transport Trips and GDP per capita (2007/08, US$).............................. 43
Figure 3.2: Spatial Distribution of Bus Mode Share in the Auckland Region ...................... 48
Figure 4.1: Positive Network Effects for an Individual Commuter ........................................ 80
Figure 4.2: Distribution of All Surveyed Households in the Auckland Region.................. 82
Figure 4.3: Public Transport Boardings and Regular Petrol Prices in Auckland, 2005Q3-2015Q3 ......................................................................................................................... 96
Figure 4.4: The Link between Public Transport Utility and Network Effects ...................... 97
Figure 5.1: Motor Vehicle Ownership among OECD members in 2010............................ 104
Figure 5.2: New Zealand Vehicle Fleet Composition, 2001-2014..................................... 107
Figure 5.3: Fleet Composition Relative to January 2000....................................................... 108
Figure 5.4 (a): Demand for Petrol and Diesel Cars.............................................................. 120
Figure 5.4 (b): Demand for Buses......................................................................................... 120
List of Tables

Table 1.1: Percentage Change in Patronage Levels per capita, 1981-2006 ................................. 2
Table 1.2: Mode Share of JTW Travel for the Morning Peak Period ........................................... 3
Table 2.1: Composition of New Zealand’s GHG Emissions by Gas .......................................... 14
Table 2.2: Transport Emissions by Mode (in kt CO₂-e) .............................................................. 18
Table 2.3: Staged Entry of Sectors into the New Zealand Emission Trading Scheme ............... 21
Table 3.1: Area Unit Level Descriptive Statistics of Variables for Auckland Region ............... 50
Table 3.2: OLS Diagnostics for Spatial Dependence ................................................................. 56
Table 3.3: Non-Spatial OLS and Spatial Durbin Model (Dependent Variable: \(\ln\text{Bus}\)) ...... 58
Table 3.4: Direct, Indirect and Total Effects of the Spatial Durbin Model ................................. 62
Table 3.5: Summary Statistics of Explanatory Variables ............................................................ 85
Table 3.6: Moran's \(I\) Statistics ................................................................................................. 87
Table 4.1: Standard and Spatial Logit Models of Commuter's Travel Mode Choice ................. 91
Table 4.2: Average Marginal Effects of the Spatial Logit Model ............................................... 93
Table 5.1: Summary Statistics of Variables .............................................................................. 121
Table 5.2: ADF Test for Unit Root ........................................................................................... 123
Table 5.3: ADF Test for Unit Root on All Variables in D1 ...................................................... 124
Table 5.4: Test Statistics for Demand Equations in the SUR Model ........................................ 125
Table 5.5: SUR Model Results for the Aggregate Road Transport Demand ............................ 126
Table 5.6: Correlation Matrix of Residuals ............................................................................... 130
Table 5.7: ADF Test for Unit Root on the Residuals ................................................................. 131
Chapter 1. Introduction

1.1. The Problems Facing Transport Network in New Zealand

It is broadly acknowledged that transport has a vital role to play in terms of sustainable development, given the fact that a nation’s economic growth is heavily reliant on the capacity and rationality of its transport system (Banister and Berechman, 2001; Boopen 2006; Munnell, 1992). Nowadays, Greene et al. (1997) claimed that transportation has been involved in almost every good and service produced in any economy. Thus a better designed transport network with a high level of transport accessibility would bring a number of benefits to the society, including travel time savings and a growth in trading opportunities.

However, at the same time, road transport externalities occur whereas someone other than the provider or the user of the transportation service is affected by the act of transportation. Among all the negative externalities that transport sector generate, traffic congestion is probably the most prevalent external cost. It has been evidenced that transport congestion in Auckland, New Zealand’s largest city, becomes a severe issue with the increase of automobiles: during peak hours, Auckland is experiencing lower travel speeds than any Australian city (Grimes, 2007). Although some degree of congestion is desirable, congestion can serve as a signal of success in a metropolitan centre because: roads tend to be congested in places that are attractive to people, but excessive congestion is not. When traffic becomes congested, the speed of flow slows down, thus requiring other drivers to spend more time to complete their trips. The increased number of cars, coupled with the intensifying level of congestion, trap road users endlessly in traffic, cost drivers billions of dollars in fuel, and produce a great amount of air pollution.
Abusah and de Bruyn (2007) revealed that trips made by private vehicles have increased whereas other modes, especially public transport, have decreased since 1960 in Auckland. Several possibilities contributed to this declining trend in public transport patronage, including: 1) the development of automobile-oriented urban forms in 1950s that have had the effect of encouraging car travel; 2) the deregulation of the vehicle industry from 1980s which removed import quotas and reduced import tariffs on vehicle imports from overseas; making imported cars more affordable for domestic consumers; and 3) the elimination of licensing restriction for car imports and privatisation of bus services since 1990.

Furthermore, when compared to other domestic and international cities, Auckland performs poorly in terms of percentage change in public transport patronage. Data from Booz Allen Hamilton (2006) indicates that on a per capita basis, Brisbane, Calgary and Portland have all experienced upturns in their patronage levels, while Christchurch and Perth have had only relatively small drops, as shown in Table 1.1. This contrasts with Auckland, where patronage level per capita has declined by nearly a half. Clearly, in terms of public transport patronage, Auckland is the lowest performer amongst this group of comparison cities. This outcome might be due to the reason that Auckland had been growing rapidly in a sprawling fashion, but other cities not.

### Table 1.1: Percentage Change in Patronage Levels per capita, 1981-2006

<table>
<thead>
<tr>
<th>City</th>
<th>Percentage change in patronage (per capita)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auckland</td>
<td>-49.4</td>
</tr>
<tr>
<td>Christchurch</td>
<td>-6.7</td>
</tr>
<tr>
<td>Perth</td>
<td>-5.7</td>
</tr>
<tr>
<td>Brisbane</td>
<td>15.7</td>
</tr>
<tr>
<td>Calgary</td>
<td>40.9</td>
</tr>
<tr>
<td>Portland</td>
<td></td>
</tr>
</tbody>
</table>

Data source: Booz Allen Hamilton (2006) 1

1 According to the report from Booz Allen Hamilton (2006), the selection of comparison cities to Auckland is primarily based on the availability of data. For Brisbane and Portland, the percentage change in patronage levels was calculated from 1985 and 1987 respectively as these were the earliest datasets available.
Table 1.2 presents some recent figures on commuter’s mode share of JTW travel behaviour in New Zealand. Clearly, commuting pattern on New Zealand’s roads is still dominated by private vehicles as opposed to public transport, with approximately 85% morning peak JTW trips taken by either drivers and/or passengers in private vehicles, compared to only 7% by public transport users, on average.

### Table 1.2: Mode Share of JTW Travel for the Morning Peak Period

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Drive</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
<td>75%</td>
<td>74%</td>
<td>73%</td>
<td>74%</td>
<td>74%</td>
</tr>
<tr>
<td>Drive + walk</td>
<td>3%</td>
<td>3%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>3%</td>
</tr>
<tr>
<td>Passenger</td>
<td>7%</td>
<td>6%</td>
<td>7%</td>
<td>7%</td>
<td>7%</td>
<td>8%</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>Passenger + walk</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Walk only</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Cycle</td>
<td>3%</td>
<td>2%</td>
<td>2%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>PT/ walk or PT</td>
<td>6%</td>
<td>6%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>PT/ car or PT/ car/ walk</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
</tr>
<tr>
<td>Other</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Data source: Ministry of Transport (MoT) (2014).

In recent years, Auckland travellers are making changes to their travel behaviour by shifting from private cars to public transport (Auckland Council, 2012), possibly due to the upgraded infrastructure, improved reliability, greater frequency, and better connected

---

2 The data are based on full-time workers who were aged 16+ and travelled directly to main job from home for journeys commencing 6am - 9.30am. The percentage refers to four year moving average, from 2003-07 to 2010-14.
services for public transport (e.g. the successful story of the Northern Express Lane). Evidence from the New Zealand Transport Agency (NZTA) shows that the number of public transport trips per person per annum has increased from 27 in 2005/06 to 32 in 2014/15, as shown in Figure 1.1. However, it should be noted that the increase is still relatively small compared to the dominance of car travel.

Figure 1.1: Annual Public Transport Boardings per capita in New Zealand, 2005/06-2014/15

Data source: NZTA, 2016

1.2. Motivation

Given that New Zealand ratified the Kyoto Protocol in 1997, cutting down carbon dioxide equivalent (CO$_2$-e) emissions has become a priority in contemporary government policy. Reducing emissions from domestic transport turns out to be a critical element of policy because this sector has a profound influence over the country’s emission profile. Transport alone was responsible for 45.6% of emissions from the energy sector, which was also equivalent to 19.7% of total greenhouse gas (GHG) emissions in 2007. More importantly, although agriculture has been identified as the largest contributing sector to New Zealand's emission profile, carbon dioxide (CO$_2$) emissions from transport have expanded more

---

4 Each non-CO$_2$ Greenhouse Gas has been converted to its CO$_2$ equivalent (CO$_2$-e) using a Global Warming Potential factor.
rapidly than methane emissions from agriculture in recent years (Ministry of Environment [MfE], 2009a). Additionally, available data from Jakob et al. (2006) recorded that the total climate change cost associated with Auckland’s transport sector was $58.4 million in 2001, where private transport accounted for $57.76 million, public transport was only responsible for $0.67 million. The total climate change cost to society of private transport was 86 times that of public transport. Thus, it appears that an efficient public transport system will subsequently reduce the number of cars on the road, lessen the level of carbon dioxide (CO2) emissions and hence lower the climate change cost associated with transport, which mainly came from private vehicles.

In the light of policy concern, Maddison et al. (1996) stated that current transport policies, which rely on increasing road construction in order to accommodate more automobiles, will be ineffective at lessening congestion and thus reducing GHG emissions. Recognising all of the substantial negative aspects and externalities that private transport has brought, local authorities (i.e. the Auckland Regional Council, the MoT), as well as urban planners (i.e. Auckland Transport and NZTA) have shown a propensity to shift their interest to revitalising urban public transport as one practical option to combat the negative externalities generated from road transport such as traffic congestion, deterioration in air quality and climate change issue. Understanding commuter’s transport behaviour and preferences thus becomes the first step of getting closer to this initiative.

Empirically, most studies in the transport literature regarding travel behaviour tend to ignore or underestimate the importance of geographical variation. Mulley and Tanner (2009) argued that the application of non-spatial models such as the ordinary least squares (OLS) method in a spatial context would be biased if there is a spatial dependency in both dependent variable and error terms, due to the fact that they commonly hold a strong assumption that observations in the regressions are independent of one another across all different spatial areas. Moreover, when analysing geographically-coded data, spatial independency in error terms is improbable to occur. This is because observations normally
have a tendency to exhibit positive spatial autocorrelation, that is, a spatial dependency in error terms. In such cases, although the estimated results is still unbiased, they are no longer efficient due to this general ignorance of spatial variation. Therefore, instead of considering the traditional non-spatial or "global" models such as the OLS, an appropriate spatial econometric analysis of travel behaviour is extraordinarily appealing and desirable.

In addition, although a substantial amount of studies has investigated the demand for private vehicles and public transportation, individually and/or jointly worldwide, little evidence was found analysing the demand for different road passenger transport choices as a system of equations. Given the fact that these road passenger transport modes are considered substitutes to one another, there is a strong possibility that an interrelationship exists between the travel demand functions, primarily due to the correlation between their disturbances. This potential correlation between error terms, however, has never been considered when modelling aggregate road passenger transport in New Zealand.

1.3. Objectives and Scope

In order to address the issues around congestion and global climate change on a long-term and sustainable basis, it is expected that government policies should focus *inter alia* on discouraging car travel and promoting the use of public transport on the road. For such policies to be effective it is important that we should properly understand:

- Whether there is a spatial dependency in commuter’s travel behaviours;
- What factors influence aggregate public transport demand for a given region;
- What attributes affect the transport mode choices of individual traveller;
- Whether there is a correlation between the disturbances of the demand for private and public road transportation; and
What factors affect the demand for each road passenger transport mode in New Zealand?

The first objective of this thesis is to conduct a spatial econometric analysis of commuter’s journey-to-work (JTW) behaviour in the Auckland region, from the aspect of aggregate public transport demand at the regional level to disaggregate transport mode choice decision making at the individual level. The JTW travel pattern is a particularly important element to consider in travel behaviour analysis because of three main reasons. First, it is undertaken regularly by a large proportion of the population; therefore, this particular group of travellers could be more effectively targeted by a wide range of policies, whose aim is to reducing the nation's dependence on private vehicles. In such sense, the travel pattern of JTW will provide a measure of how successful such policies are. Second, because commuter’s JTW travel also has strong economic and sociological implications for urban planners and accounts for about a third of all travel in New Zealand. Third, traffic congestion is at its worst during morning and afternoon peak periods, where most trips are undertaken by JTW and journey to school commuting. As a result, it is especially important to ascertain the determinants of JTW (MfE, 2010).

Information on the JTW trips can only be acquired through surveys (or questionnaires) of the population in which participants are questioned about commuter’s travel behaviours. In New Zealand, currently, there are two such surveys undertaken on a regular basis. These are the New Zealand National Census undertaken by Statistics New Zealand, and the continuous New Zealand Household Travel Survey (NZHTS) carried out by the NZTA.

The second objective of this thesis is to develop an aggregate travel demand model of New Zealand’s major road passenger transport modes, by taking into detailed account the potential effect of correlation between their error terms. The empirical results can thus be
used to give various policy suggestions so that we can achieve a reduction in the demand for cars by different fuel types, and also promoting the use of public transport in New Zealand.

The thesis will address the following four main research questions:

1. **What are the research gaps from past literature on commuter’s travel behaviour and the demand for aggregate road passenger travel?**

   For research on commuter’s travel behaviour, the studies of the effect of built environment, or more specifically, the impact of urban form on aggregate travel behaviour and the applications of the discrete choice models to individual travel behaviour are, in general, non-spatial. That is, the majority of the past research does not consider space as a relevant factor when conducting transport-related studies.

2. **At the regional level, how should we examine the impact of urban form variables on aggregate public transport demand?**

   In particular, three sub-questions need to be addressed:

   a. **Should we consider spatial dependence when estimating the relationships between the urban form variables that influence aggregate public transport demand?**

      According to Getis et al. (2004), unless a geographical area is uniform and/or boundless, each location will have some degree of distinctive uniqueness relative to the others. Therefore, the parameters estimated for the model should adequately capture the spatial effect that urban form has on the demand for public transport.

   b. **How can we estimate these geographical components econometrically?**

      This can be addressed by using a spatial econometric method known as spatial Durbin model (SDM), which has the ability to capture the spatial effects in transportation modelling, using a specified weighting scheme.
c. **Does including spatial variability enhance the explanatory power of the model?**

Most studies show a significant improvement in the fit of the data to the model, when spatial effects are incorporated.

3. **At the individual level, what is the probability that a commuter will choose to use public transport to go to work given the transport mode preference of his/her neighbours and the characteristics of the regions where he/she lives?**

   This particular question will be answered by applying a spatial autoregressive logit model to disaggregate travel survey data for the Auckland region.

4. **At the national level, whether there is a correlation between the disturbances of the demand for private and public road transportation? What are the factors that affect the demand for each road passenger transport mode in New Zealand?**

   These empirical problems identified above will be addressed by applying a seemingly unrelated regression (SUR) model, using aggregate quarterly time series data in New Zealand. The advantage of the SUR model is that dependent variables (i.e. petrol cars, diesel cars and buses) which represent aggregate road passenger travel demand are allowed to be considered as a group when they bear a close conceptual relationship to one another, brought forward by the potential correlation between their disturbances.

1.4. **Contributions**

Given that Auckland has relatively low public transport patronage compared to most Australasian cities, an understanding of which variables influence aggregate public transport demand at the regional level, how travel decisions are made by individual commuters when choosing from different transport modes, and the factors that affecting aggregate road passenger travel demand at the national level, become key issues to consider.
The majority of current studies that either examine the determinants of aggregate public transport demand, or investigate how individual commuters choose transport mode share one possible shortcoming in common. They do not capture travel behaviour in a geographical context. This can, in turn, produce biased and inconsistent estimators if spatial effects exist in the model. At the regional level, the literature that takes spatial dependency into account when conducting economic analysis on aggregate public transport demand for JTW travel is quite thin. Moreover, when looking down to an individual level, in New Zealand’s context, there is no up-to-date literature which explicitly investigates how commuters choose their transport modes to go to work, given the transport mode preference of his/her neighbours and also the characteristics of the regions where he/she lives. Additionally, at the national level, little empirical evidence was found analysing the demand for different road passenger transport choices as a system of equations. Given the fact that these road passenger transport modes are considered substitutes to one another, there is a strong possibility that an interrelationship exists between the travel demand functions, primarily due to the correlation between their error terms.

Given the presence of the research gaps, this thesis is thus aimed at conducting a spatial econometric analysis on the commuter's JTW behaviour, at both regional and individual levels. It is also proposed to provide a useful package of policy instruments, which could effectively help us to combat environmental issues, in light of the empirical results mentioned above.

By using the New Zealand Census data and the NZHTS data, this thesis also complements the existing literature in the transportation field by using revealed preference data on how commuters actually travelled to work, rather than stated preference data, which may not accurately reflect what commuters would actually do in a hypothetical situation posed in a
Much of the existing research on this topic uses the latter type of data. This thesis is one of the first such studies to correct for spatial dependency at both aggregate and disaggregate levels, which strengthens the regression model. Additionally, by using aggregate data on a national basis, this thesis also contributes to the analysis of aggregate road passenger transport demand by taking into detailed account the potential existence of correlation between the error terms of different transport modes, so that both private and public transport can be modelled as a group of transport choices available to New Zealand travellers.

The next chapter firstly reviews the literature on traveller's travel behaviour in two categories: the studies of the impact of urban form on travel behaviour and the applications of the discrete choice models on travel behaviour respectively. Secondly, it discusses the concept of spatial effects and presents several associated spatial models used in the following two empirical chapters (i.e. chapter 3 and 4). Chapter 3 examines the relationship between urban form and public transport use in Auckland by applying a spatial Durbin model, using regional level JTW data. Chapter 4 investigates individual commuter’s transport mode preferences in Auckland for their JTW travel by estimating a spatially autoregressive logit mode choice model. Chapter 5 develops an aggregate road passenger travel demand model using the seemingly unrelated regression (SUR) method, and the empirical results from the SUR model deliver various policy implications in terms of achieving a reduction in the demand for both petrol and diesel cars, and also promoting the use of public transport. Chapter 6 provides a summary and future directions.

---

5 As opposed to revealed preference method which is based on actual market behaviour, stated preference experiment, as defined by Kroes and Sheldon (1988), refers to a category of techniques which use individual respondents’ statements about their preferences in a set of transport options to estimate utility functions. In other words, this method is based on hypothetical scenarios.
Chapter 2. Background Information, Literature Review and Overview of Methodology

2.1. Background Information of New Zealand Transport Sector

2.1.1. Physical impacts of climate change

Nowadays, discussion of worldwide climate change has dramatically increased. Research indicates that climate change has undoubtedly developed into a huge threat to the atmosphere and life on earth. For instance, Noyes et al. (2009) noted that climate change has some potentially harmful consequences to the natural environment, including ice melt and organic carbon cycling. Without exception, as the CLIMPACTS (2001) Synthesis report concludes, New Zealand will also likely to experience greenhouse gas (GHG)-induced climate change, although it might not be as fast as the global rate thanks to its maritime location.

Karl and Trenberth (2003) acknowledge that although the Earth has gone through a changing climate by natural forces ever since the beginning of time, anthropogenic, or human activity is certainly responsible for a substantial part of climate warming over a relatively short timeframe, especially during the last century. Evidence from the International Panel for Climate Change [IPCC] (2013) indicates that by the end of the 21st century, temperature of the global surface is projected to exceed 1.5°C relative to the average from year 1850 to 1900.

In New Zealand, the most likely impact resulting from climate change is an increase in temperature of over 1°C by 2050 and of over 2°C, compared with 1990 levels, by the end of the century (New Zealand Climate Change Information [NZCCI], n.d.). In the case of Auckland, temperatures are estimated to increase between 0.6°C and 3.8°C by 2080.

---

6 Climate change usually refers to changes in modern climate. It is commonly known as "global warming" or "anthropogenic global warming".
relative to its 1980 statistics (START, 2006). Consequently, Hannah (2004) warns that based on figures released by a tide gauge analysis using data from the Port of Auckland, this increase in temperature, on average, has led to a linear sea level rise of 1.6 millimetres per year, for the period 1899 to 2000. As Weubles and Jain (2001) claim, greenhouse gases (GHGs), which are mainly accumulated from land use changes and combustion of fossil fuels from transport and other industry sectors over the 20th century, are the major cause of these daunting environmental outcomes. These environmental facts thus lead us to an important question: what should we do when facing the prospect of adverse environmental events associated with global climate change?

2.1.2. Climate change challenge on transport sector

2.1.2.1. An overview of New Zealand Greenhouse Gas Emissions profile

The Ministry of Economic Development (MED) states that New Zealand is a signatory to both the United Nations Framework Convention on Climate Change (UNFCCC) and the Kyoto Protocol. In 1997, New Zealand ratified the Kyoto Protocol and was committed to reducing its GHG emissions to an annual average below or equal to its 1990 emissions level by 2012 for its first commitment period (CP1) or take responsibility for excess emissions.7 According to the MfE, in the base year (i.e. 1990), New Zealand’s total GHG emissions were equal to 59.6 million tons (Mt) of CO2-e. However, in 2011, total GHG emissions were 72.8 Mt of CO2-e, which is equivalent to a 22% rise since 1990. The most recent projection of New Zealand’s net position under the Kyoto Protocol is a surplus of 29.6 Mt of CO2-e, which implies that New Zealand is on track in meeting the Kyoto Protocol target for CP1 (MfE, 2013).

In general, there are six types of direct GHGs included in the Kyoto Protocol to measure and monitor human caused emissions, they are: CO2, hydro fluorocarbons (HFCs),

---

7 The first commitment period of Kyoto Protocol is from 2008 to 2012.
methane (CH₄), nitrous oxide (N₂O), per fluorocarbons (PFCs) and sulphur hexafluoride (SF₆). In addition, the gases are further classified under six sectors, including: agriculture, energy (mainly transport), industrial processes, land use, land-use change and forestry (LULUCF), solvent and other product use, and lastly, waste. Table 2.1 below shows the proportions of GHG emissions, as well as their respective changes between 1990 and 2011.

Table 2.1: Composition of New Zealand’s GHG Emissions by Gas

<table>
<thead>
<tr>
<th>GHG emissions</th>
<th>1990</th>
<th>2011</th>
<th>Change from 1990 in Gg CO₂-e</th>
<th>Change from 1990 in percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂</td>
<td>25047.1</td>
<td>33162.2</td>
<td>8115.2</td>
<td>32.4</td>
</tr>
<tr>
<td>CH₄</td>
<td>25650.3</td>
<td>27050.1</td>
<td>1399.8</td>
<td>5.5</td>
</tr>
<tr>
<td>N₂O</td>
<td>8300.6</td>
<td>10689.7</td>
<td>2389.1</td>
<td>28.8</td>
</tr>
<tr>
<td>HFCs</td>
<td>0.0</td>
<td>1885.1</td>
<td>1885.1</td>
<td>NA</td>
</tr>
<tr>
<td>PFCs</td>
<td>629.9</td>
<td>30.2</td>
<td>-599.7</td>
<td>-95.2</td>
</tr>
<tr>
<td>SF₆</td>
<td>15.2</td>
<td>17.6</td>
<td>2.4</td>
<td>15.9</td>
</tr>
<tr>
<td>Total</td>
<td>59643.1</td>
<td>72834.9</td>
<td>13191.9</td>
<td>22.1</td>
</tr>
</tbody>
</table>


From Table 2.1, it is obvious that CO₂ is one of the major GHGs in NZ’s emissions profile. Moreover, the percentage change from 1990 to 2011 is 32.4%, reflecting the rapid emission growth from the energy sector in comparison to the agriculture sector.

---

8 The indirect GHGs are SO₂, CO, NOₓ and NMVOC. These indirect GHGs do not account towards New Zealand’s GHG emissions total.
9 All GHG emission figures are presented in the unit of Gigagram (Gg).
While New Zealand’s GHG emissions in a global context are small, with only 0.14% of total emissions of global emissions, in terms of emission intensity by population, it was rated as the fifth highest emitter worldwide in 2010 among 40 Annex I countries (MfE, 2005). Take energy sector as an example, Figure 2.1 shows the growth rate in energy sector emissions for each Annex I members. It is clear that the growth rate in New Zealand is high by international standards, as emissions from its energy sector have grown by about 31% from 1990 to 2011. These emissions are not only responsible for contributing to global warming and domestic air pollution, but also cause considerable loss to its economy, such as drops in Gross Domestic Product (GDP).

Figure 2.1: Global Growth in Energy Sector Emissions from Base Year to 2011

Source: MBIE (2013)

---

10 Annex I countries constitutes of OECD countries plus the Eastern European economies in transition and account for 63.7% global GHGs emissions collectively. Annex I countries are those which are full signatories to the UNFCCC.
2.1.2.2. Greenhouse Gas Emissions from transport sector

CO₂ emissions from transport sector have increased swiftly over the past few decades worldwide. In 2004, transport sector alone was responsible for 20% of the world energy-related CO₂ emissions, and this portion is still projected to grow at a rate of 1.7% per annum up to 2030 (International Energy Agency [IEA], 2006). Elkins (1996) claims that mitigating GHG emissions from transport will lead to far-reaching ancillary benefits, as the environmentally harmful gases, particularly NOₓ and SO₂ resulting from combustion of fossil fuel, will be greatly reduced.

In the New Zealand context, the MBIE (2013) states that emissions from the transport continue to dominate in 2012, contributing 43% of total emissions in the energy sector, although it is down from 45% in 2011. This share is much higher than the averaged 30% of CO₂-e emissions from transport in OECD countries (Energy Efficiency and Conservation Authority [EECA], 2004). Figure 2.2 illustrates the energy emissions by sector in 2012. It is clear that emission from transport is the largest: it is even greater than the combined emissions from electricity, fugitive and manufacturing emissions. Figure 2.3 depicts the energy emissions by five different fuel types, including: liquid fuels, gas, coal, geothermal and biomass. It is clear that liquid fuel combustions are responsible for the majority of emissions in 2012, where more than three quarters of liquid fuel emissions come from the transport sector.
Figure 2.2: Energy CO\textsubscript{2}-e Emissions by Sector, 2012 (in kt CO\textsubscript{2}-e)

Source: MBIE (2013).

Figure 2.3: Energy Emissions by Fuel, 2012 (in kt CO\textsubscript{2}-e)

Source: MBIE (2013).
Table 2.2 decomposes emissions from the transport sector by different modes. MBIE (2013) states that road transport emissions constitute the largest portion of domestic transport emission, alone make up nearly 40% of all energy sector emissions in 2012, and have increased by 68% since the base year.

Table 2.2: Transport Emissions by Mode (in kt CO2-e)

<table>
<thead>
<tr>
<th>Calendar Year</th>
<th>Domestic Road</th>
<th>Rail</th>
<th>Aviation</th>
<th>Marine</th>
<th>Total</th>
<th>International Aviation</th>
<th>Marine</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>7,407</td>
<td>79</td>
<td>939</td>
<td>253</td>
<td>8,678</td>
<td>1,320</td>
<td>1,091</td>
<td>2,411</td>
<td>11,088</td>
</tr>
<tr>
<td>2000</td>
<td>10,438</td>
<td>246</td>
<td>1,171</td>
<td>378</td>
<td>12,234</td>
<td>1,798</td>
<td>700</td>
<td>2,498</td>
<td>14,732</td>
</tr>
<tr>
<td>2008</td>
<td>12,454</td>
<td>156</td>
<td>1,078</td>
<td>255</td>
<td>13,942</td>
<td>2,301</td>
<td>972</td>
<td>3,273</td>
<td>17,215</td>
</tr>
<tr>
<td>2009</td>
<td>12,312</td>
<td>164</td>
<td>977</td>
<td>295</td>
<td>13,748</td>
<td>2,193</td>
<td>951</td>
<td>3,144</td>
<td>16,893</td>
</tr>
<tr>
<td>2010</td>
<td>12,443</td>
<td>143</td>
<td>1,004</td>
<td>256</td>
<td>13,847</td>
<td>2,315</td>
<td>1,014</td>
<td>3,329</td>
<td>17,176</td>
</tr>
<tr>
<td>2011</td>
<td>12,569</td>
<td>143</td>
<td>1,040</td>
<td>291</td>
<td>14,043</td>
<td>2,333</td>
<td>1,033</td>
<td>3,366</td>
<td>17,408</td>
</tr>
<tr>
<td>2012</td>
<td>12,437</td>
<td>150</td>
<td>1,016</td>
<td>293</td>
<td>13,897</td>
<td>2,301</td>
<td>1,184</td>
<td>3,484</td>
<td>17,381</td>
</tr>
</tbody>
</table>

\( \Delta \text{1990/2012} \) 67.9% 90.5% 8.2% 15.8% 60.1% 74.3% 8.5% 44.5% 56.7%  
\( \Delta \text{1990/2012 p.a.} \) 2.4% 3.0% 0.4% 0.7% 2.2% 2.6% 0.4% 1.7% 2.1%  
\( \Delta \text{2011/2012} \) -1.0% 5.0% -2.3% 0.8% -1.0% -1.4% 14.6% 3.5% -0.2%  

% of total 2012 energy CO2-e emissions 38.7% 0.5% 3.2% 0.9% 43.3% n.a. n.a. n.a. n.a.

Source: MBIE (2013).

In the case of Auckland, available statistics indicate that total CO2-e emissions from the transport sector were 4.1 Mt in 2001 (National Institute of Water and Atmospheric Research [NIWA], 2002). Figure 2.4 below shows that CO2-e emissions from the transport in the Auckland region ranked first in 2001, accounting for nearly 30% of total emissions, which is consistent with the population size.
Most forms of travel, at least in New Zealand, use liquid fossil fuels (e.g. petrol and diesel), thus emitting GHGs into the atmosphere. In New Zealand, consumption of petrol is largely dominated by household vehicles. Relatively speaking, New Zealanders have a high level of vehicle ownership; the ownership of motor vehicles has increased sharply by nearly 26% from 1990 to almost 2.5 million vehicles. For a country whose population size is just under 4 million, the average number of motor vehicles is around 1.6 per person in 2002 (MoT, 2002). The use of freight transport has also increased due to New Zealand’s unique geographically isolated location. Most exports are moved by ship and, to a lesser extent, planes. As a result, these facts simply imply that if there are no changes in the ways in which we used to travel, or in amount of fuels we used to use, the demand for transport energy and thus the emissions from transport sector, will inevitably grow.
Although agriculture is the largest contributor to New Zealand’s GHG emissions, viz. CH₄ and N₂O emissions, reducing these emissions is unfeasible given present technologies (Huang et al., 2009). According to the New Zealand Herald (2009), New Zealand is already one of the most GHG efficient dairy producers in the world; therefore, mitigating CO₂-e emissions from the transport sector rather than putting effort on reducing CO₂-e emissions from the agriculture sector could be regarded as a feasible, effective and useful measure to achieve the climate change goal for New Zealand. Most importantly, it is critical to have a thorough understanding of the current situation in the transport sector under New Zealand emissions trading scheme (NZETS). Some facts are outlined in the following section.

2.1.2.3. Potential impacts of an Emission Trading Scheme on transport sector

NZCCI states that the NZETS is considered a useful approach of meeting New Zealand’s international obligations around global warming. New Zealand Government has chosen the NZETS as its primary tool to reduce emissions, as it is rated as the least-cost way of reducing emissions. NZETS provides incentives to reduce emissions from main emitters, encourage and support global wide action on global warming, and contribute to New Zealand in efforts to achieve sustainability.

At present the NZETS covers six types of liquid fossil fuels. It comprises aviation gasoline, diesel, jet kerosene, heavy fuel oil, light fuel oil and petrol. However NZCCI notes that emissions from fuel used for international aviation and marine transport are exempted from the scheme; this exemption is consistent with the requirement in the Kyoto Protocol.

Table 2.3 shows that the introduction of the ETS is a staged process. Forestry sector was the first to enter the ETS in New Zealand, followed by transport fuels, electricity
production, industrial processes, synthetic gases, waste and agriculture sectors. The transport sector had been required to report its emissions from 1\textsuperscript{st} January 2010.

Table 2.3: Staged Entry of Sectors into the New Zealand Emission Trading Scheme

<table>
<thead>
<tr>
<th>Sector</th>
<th>Mandatory Reporting of Emission</th>
<th>Full Obligations: Payment for Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forestry</td>
<td>-</td>
<td>1 January 2008</td>
</tr>
<tr>
<td>Transport Fuels</td>
<td>1 January 2010</td>
<td>1 January 2010</td>
</tr>
<tr>
<td>Electricity Production</td>
<td>1 January 2010</td>
<td>1 January 2010</td>
</tr>
<tr>
<td>Industrial Processes</td>
<td>1 January 2010</td>
<td>1 January 2010</td>
</tr>
<tr>
<td>Synthetic Gases</td>
<td>1 January 2012</td>
<td>1 January 2013</td>
</tr>
<tr>
<td>Waste</td>
<td>1 January 2012</td>
<td>1 January 2013</td>
</tr>
<tr>
<td>Agriculture (subject to review)</td>
<td>1 January 2012</td>
<td>1 January 2015</td>
</tr>
</tbody>
</table>

Source: MfE.

For the transport sector, the MfE considers that the most appropriate point of obligation is at the upstream level rather than at the retail level. This implies that individual road users (e.g., motorists or car drivers) are not directly involved in the process of emissions trading. Rather, the scheme applies to liquid fossil fuels further up the supply chain. In other words, according to NZIER (2012), “fuel suppliers who take fuel from the refinery or who import it are required to participate in the ETS by buying emission units to cover the emissions that result from the fuel they purchase and supply to end users”. The petroleum products industry in New Zealand is made up of four multi-national suppliers in petroleum distribution and retailing, they are: BP, Chevron (marketing as Caltex), Mobil (an affiliate
of ExxonMobil) and Z Energy. An independent Australian company, Gull Petroleum, imports refined product into Tauranga and holds retail outlets in the upper North Island (NZIER, 2012). As stated by the MfE, BP, Caltex, Mobil and Z-Energy are the four companies who dominate the entire market and supply a full range of products. Gull on the other hand has a smaller market share and has limited participation in petrol and diesel sales.

Furthermore, for the transport sector, suppliers of liquid fossil fuels were not granted any free allocation of emission units. The reason behind this is that the government believes that suppliers can easily pass on the costs of their ETS obligations to their customers, which in turn implies that the impact of the ETS in terms of net profits of fuel suppliers will be limited. The cost of emission units will be eventually transferred to consumers of liquid fossil fuels. Thus, fuel prices are expected to increase with the NZETS.

2.2. The Studies of the Impact of Urban Form on Travel Behaviour

Density or compactness and distance of residence from the urban centre have proven to be the two most significant urban form determinants of travel behaviour. Among six United States (U.S.) cities, Pushkarev and Zupan (1977; 1982) conduct a series of pioneering studies on travel behaviour after controlling for socioeconomic and demographic characteristics such as vehicle ownership level, household size and income level. The results indicate that residential densities in transit corridors, coupled with the size of the Central Business District (CBD) and the distance of the stations from the CBD, explain a significant amount of the variation in transit demand for local bus, light rail, rapid transit and commuter rail using ordinary least squares (OLS) regression analysis. Their findings also establish the residential density threshold for providing minimum transit service

---

11 New Zealand owned Greenstone Energy bought the former Shell operations and rebranded it as Z Energy in 2011. Z Energy comprises some of the former assets of Shell New Zealand.
levels. For example, they state that four dwelling units per acre is estimated to be the minimum housing density level for supporting regular bus service at 20 buses per day.

In the United Kingdom (UK), by employing cross-sectional 1978/79 UK National Travel Survey data, Hillman and Whalley (1983) report that the total distance travelled by all transport modes decreases with increased population density. Most importantly, the study finds that residents who are resided in very low-density regions (i.e. less than 5 persons per hectare) prefer to travel by car: the distance that they travelled by car is more than twice the distance travelled by the same transport mode by residents of relatively high-density zones (i.e. more than 60 persons per hectare).

According to Concas and DeSalvo (2008), later works show a move from focusing on the research of density threshold levels to the study that examines the effect of urban design and land-use mix on travel behaviour, after controlling for density levels. For instance, using panel data from a sample of 85 urban regions in the U.S. from 1982 to 2003, Su (2009) studies the relationship between travel demand in terms of annual per-capita vehicle-miles-travelled (VMT) and urban spatial characteristics after accounting for gasoline price and socioeconomic variables such as median household income. By applying system dynamic panel data estimation, regression results suggest that the effect of population density, urban congestion, spatial size of urban areas and road density are all statistically significant. The first two variables have a negative impact on per-capita VMT while the latter two are found to be positively related to travel demand.

There is also a relatively small, but growing, body of empirical literature that focuses on the relationship between different characteristics of urban form and public transport demand. Messenger and Ewing (1996) examine the relationship between residential density and bus mode share of work trips based on 698 traffic analysis zones (TAZs) of Metropolitan Dade County, Florida in 1990. The explanatory variables employed are
overall density measures, and some socioeconomic and demographic variables. By applying the full information maximum likelihood estimation (FIML) method, the bus share by place of residence proves primarily dependent on automobile ownership and secondarily on job-housing balance and bus service frequency. Likewise, their study claims that the bus mode share by place of work increases with overall density in terms of persons and employees, cost of parking and transit access to downtown measured in bus travel time. The empirical results are similar to Kitamura et al. (1997), where the authors also find that population density is positively linked to the proportion of public transport trips after controlling for socioeconomic differences.

In contrast with population density, another measure of the intensity between land use and economic activities, employment density, has received less attention. Using multivariate regression analysis, Frank and Pivo (1994) conclude that, based upon a panel study between 1989 and 1994 in the Seattle metropolitan area, employment density, like population density, has a positive impact on the proportion of public transport trips for both shopping and work trips.

Kuby et al. (2004) examine cross-sectional data on average weekday light-rail boardings in 2000 for 268 stations in nine U.S. cities using OLS. Five categories of explanatory variables which describe the characteristics of land use, along with other factors regarding intermodal connection and socioeconomic conditions are used for each city. Their results highlight the importance of land use and accessibility in determining light-rail transit ridership. The most significant finding from this study has to do with population and employment densities. In per day term, it shows that an increase of 100 people residing within walking distance to a light-rail station increases approximately 9.2 boardings, while an increase of 100 people employed within walking distance of a light-rail station raises boardings by about 2.3 commuters. Additionally, a dummy variable for CBD location is not statistically significant at a 5% level, implying that the CBD is no longer a relevant factor in determining light-rail boardings. However, as the authors point out, this
insignificance could be a result of a relatively high correlation between the model’s specification of centrality and the CBD dummy.

### 2.3. Applications of Discrete Choice Model to Travel Behaviour

Individual traveller’s choice of transport mode includes, for instance, private car, public transport, cycling or walking. According to Shen et al. (2009), the discrete choice model (DCM) is the predominant approach. Generally, DCM involves an application of the random utility models (RUMs) which assume that trip makers are perceived to act as rational beings, choosing modes most likely to maximise utility (Abane, 2011). Numerous studies of transport modal choice have used the discrete choice approach (e.g., Asensio, 2002; Bates, 1987; Ben-Akiva and Lerman, 1985; Dissanayake and Morikawa, 2010; Garrod and Millius, 1983; Hensher, 1994; Louviere, 1988; Louviere et al., 2000; Müller et al., 2008; Reilly and Landis, 2002; Cervero, 2002; Schwanen and Moktar, 2005; Shen et al., 2008). Among these applications of discrete choice modelling, research into mode choice for commuter/work trips is possibly the area that has been studied most thoroughly by both transport engineers and transport economists.

One of the earliest applications among travellers’ choice of transport mode was Williams (1978), who uses a cross-section sample of 401 households observed in 1973 for five sections of the Buffalo metropolitan area, the U.S. modal choice is specified as $y_i = f(T_i, U_k)$, where $y_i$ represents the probability of selecting mode $i$, $T_i$ is the $i^{th}$ transportation system characteristics, and $U_k$ signifies the $k^{th}$ user’s characteristics. Both logit and probit estimation techniques were applied to work and non-work trips. Automobile availability (measured as the number of automobiles per household), walking time of public transport relative to waiting time and waiting time of transfer from one mode to another are found to be important determinants of transport mode choice for work trips. However, surprisingly, individual income level has nearly no effect on this choice.
In New Zealand, O’Fallon et al. (2004) apply the discrete choice models (DCMs), namely, multinomial logit and nested logit models, to determine the constraints that affect transport mode choices by morning peak period road users.\(^\text{12}\) They have carried out stated preference experiments in Auckland, Christchurch and Wellington respectively, to examine the impact of a number of policy tools on the commuter’s decision about whether to drive a car to work or study. With a total number of 732 completed and valid questionnaires, which are fairly equally divided among these three cities (i.e., 247 for Auckland, 233 for Wellington and 252 for Christchurch), their study shows that the responsiveness of car drivers to various policy tools is quite different for the three centres. For example, the impact of the toll charge and the surcharge on car parking buildings or lots has the greatest effect on modal shift in Wellington, but in the other two cities, the largest effect on shifting car drivers to another mode is the registration surcharge.

Abane (1993) conducts a multinomial logit (MNL) model analysis to study transport mode choice for the JTW travel behaviour among a group of labours employed by medium and large-scale firms and services located in Accra, the capital city of Ghana, in latter 1990. With 1471 observations, results from this work indicate that the modal choice decisions of employees are largely influenced by perceived service quality of the commercial commuter vehicles and commuters’ personal characteristics, rather than by transport details such as journey distance, access, waiting and in-vehicle times, and fares as revealed in conventional works. Moreover, gender roles, age differences, variations in disposable incomes, occupational status and the reliability of schedules by the specific modes are the most essential factors that workers take into account in choosing transport modes from trotros, taxis, company vehicles, private cars, walking and taking buses for JTW trips.

Palma and Rochat (2000) examine transport mode choice for JTW trips in Geneva, Switzerland, using a dataset which came from a survey undertaken in 1994, under the supervision of the University of Geneva and the National Swiss Funds for Scientific

\(^{12}\) Morning peak period is defined as travels before 10 a.m.
By means of a nested logit approach, the authors derive the same conclusion drawn by those previously arrived by Abane (1993), where the individual’s characteristics are found to be more important determinants of mode of transport to work, compared to the mode-specific factors (i.e. travel time, cost and comfort level). Consisting of 726 commuters working in Geneva, their results also show that for automobile users, elasticities with respect to in-vehicle travel time and cost are -0.27 and -0.29 respectively. While for commuters who travel by public transport, elasticities with respect to on-vehicle travel time and cost are -0.61 and -0.43 respectively. These estimates indicate that the demand for car travel is relatively inelastic and in contrast, the demand for public transport is more sensitive to changes in travel time on average.

Liu (2007) uses data from the largest city in China, Shanghai, to analyse work-trip travellers’ mode choice behaviour by employing combined revealed preference and stated preference surveys conducted in 2001. By estimating several versions of MNL models, their results indicate that in-vehicle time, out-of-vehicle time and one-way work-trip monetary cost are significant in explaining the choice of mode for bus, subway and taxi commuters. In addition, bicycles are found to be an inferior good for all income levels whereas bus and subway appear to be inferior goods for only workers at middle and high income levels; for low income level, however, they are normal goods. Taxis, on the other hand, are a normal good for low income level labour force but for those who are middle or high income workers, they are considered a luxury.

With the aim to encourage more sustainable forms of travel to combat increases in automobile dependence and congestion problems, Commins and Nolan (2011) investigate the determinants of mode of transport to work in the Greater Dublin Area, Ireland. Using data from the 2006 census of population, results of the conditional logit model show that household composition, public transport availability, journey travel time and work location are the four most important attributes when choosing transport mode for JTW.
trips. In terms of gender difference, women are found to be significantly less likely to walk or cycle to work, but are significantly more likely to take public transport. Individuals in households with young children are more likely to drive their own vehicles compared to other modes, mainly because cars can better satisfy the needs and schedules of school-age children. The study also emphasises that with the existence of park-and-ride facilities and a quality bus corridor in an individual’s electoral division of residence, the probability of travelling by public transport increased significantly. 13 Therefore by providing appropriate park-and-ride facilities to public transport stations, the attractiveness of this mode could be enhanced considerably.

One of the most recent research on choice for commuter/work trips came from Habib (2012). Using data from the Greater Toronto Area, Canada, in 2001, this joint trivariate discrete-continuous-continuous study models mode choice of commuting trips jointly with work start time and work duration. Six possible modes are identified in the data set, they are: automobile driver, automobile passenger, local transits, local transit with park-and-ride facilities, GO transit (a regional transit service) with park-and-ride options and walking. By considering workers’ work schedules as a skeletal activity in activity-based travel demand models, their results reveal many behavioural details in regards to their mode choices and work scheduling. For instance, first, for the mode choice model, the variable known as home-to-work distance shows a negative impact on the utility of the walk mode. Second, for the work start time model, automobile travel time coefficient has the highest negative value among all possible transport modes, implying that non-auto users are less sensitive to travel time than are automobile users. Third, for the work duration model, the total number of household vehicles is significant only for the manufacturing occupation group, and this variable had no impact on other employment categories.

13 According to Commins and Nolan (2011), quality bus corridor refers to a route with dedicated road space for buses, like the bus lanes in New Zealand; electoral division is the smallest administrative area for which population statistics are published, which is similar to the meshblock in New Zealand.
Other efforts in estimating travel behaviours include the work from Reilly and Landis (2002), where the authors employ a MNL model using disaggregated travel diary data from the Bay Area Travel Survey 1996 to model transport mode choice given the raster-size measures of land-use and urban design, in the U.S. Their findings reveal that a rise in average density of 10 people per hectare within one mile of an individual’s residence will increase the probability of walking or taking transit by a 7%. In Sweden, Johansson et al. (2005) conduct a survey to collect modal choice and the attitudinal and behavioural indicators from commuters to construct unobservable, or latent, preferences in their DCM. By using survey data of approximately 19,000 commuters between the cities of Stockholm and Uppsala, the authors conclude that a model that incorporated latent variables outperforms traditional DCMs. Furthermore, travel time, cost, and preferences for flexibility, comfort, and environment are all identified as important variables in their study.

2.4. Research Gap

At the regional level, Section 2.1. provides a detailed literature review of the effect of urban form on transport behaviour. Section 2.2., on the other hand, focuses on the discussion of previous empirical applications of DCMs to commuter’s travel behaviour at an individual level. Among all of the empirical works discussed above, one should note that none of them has considered space as a relevant factor. Rather, these past analyses of travel behaviour at either regional or individual levels assume that observations are independent of one another in a given geographical context. To the best of our knowledge, Greer and van Campen (2011) have produced the only published paper which specifically takes spatial effects into account when analysing the determinants of work trip bus ridership in the context of New Zealand, using the spatial error model (SEM).

In practice, however, it seems unlikely that region i’s transport network in terms of vehicles and public transport infrastructure is independent of that of its neighbouring region j, hence this assumption of independence between any nearby geographical units is
likely to be unrealistic. Ignoring spatial characteristics between observations will, in turn, produce biased and inconsistent estimators. In the context of transportation studies, it is unrealistic to assume spatial independence between any given geographical spaces to its neighbours, because the commuting network from one region will definitely have some influence over its neighbouring regions. In particular, the fundamental concept of spatially-defined data analysis is that regions located nearby have a tendency to be more similar than those separated by greater distances. This indicates that positive spatial dependence would seem to be more plausible than spatial independence when analysing geographical data at the regional level (LeSage and Pace, 2010).

When conducting analysis on individual commuter’s transport mode choice, it is also natural to imagine that “people would prefer to use public transit together with other people as a result of social spill-over” (Goetzke, 2008). This particular transport-related behaviour among travellers is known as the “positive network effects”, which has not been brought into consideration in the majority of the past studies. Realistically, positive network effects should exist because a certain commuter’s preference of a specific transportation mode will indeed have some degree of influence in determining transport mode choice to his/her neighbours. This is simple to justify: when making decisions about which transport mode to choose, the individual commuter tends to mimic what their neighbour chooses, as he/she interprets network effects as a signal that a certain transportation mode is safe and reliable. Therefore, this general research gap of ignoring spatial dependence at the regional level, along with the unawareness of positive network effect at the individual level in the literature, lead us to the application of the spatial econometrics models when undertaking travel behaviour analysis. The next section reviews the concept of spatial effects and presents several spatial regression models in detail.

Lastly, at the national level, given the fact that major road passenger transport modes are considered potential substitutes to one another, there is a strong possibility that an
interrelationship exists between the travel demand functions, primarily due to the correlation between their disturbances, a research gap that was thus discussed and addressed in more detail in Chapter 5.

### 2.5. Review of Spatial Effects and Spatial Regression Models

#### 2.5.1. Spatial dependence and spatial heterogeneity

Recently, the economics literature has paid extensive attention to spatial issues when conducting theoretical and applied econometric studies using cross-sectional data of a geographic nature. According to Anselin (1988), spatial data are mainly characterised by two features: spatial dependence and spatial heterogeneity (or spatial non-stationarity). Together, these two particular features have been regarded as spatial effects and nowadays they are perceived as major challenges in spatial analysis (Du and Mulley, 2006). Additional spatial processes also include spatial clustering and feedback forces. The former refers to the process of grouping a set of objects into classes or clusters so that objects within a cluster have high similarity in comparison to one another, but are dissimilar to objects in other clusters (Han et al., 2001), while the latter refers to the fact that individuals and households interact with each other and thereby influence each other for most social processes (Wrigley et al, 1996, Voss et al., 2006, Webber and Pacheco, 2010).

LeSage and Pace (2010) emphasise that data collected from nearby areas are commonly interdependent with each other, thus this spatial dependence requires special consideration when conducting empirical research, because the consequence of ignoring such spatial structure could result in biased estimates. As is well established in the literature, spatial dependence can exist in two forms, substantive and residual (Anselin, 1988). The former is caused by spatial correlation of observed features, which indicates that the explanatory

---

14 This paper will use the synonyms spatial heterogeneity and spatial non-stationarity, interchangeably.
variable in one geographical space is correlated with the variable in adjacent or neighbouring geographical space. The latter, instead, refers to the fact that spatial autocorrelation can also be found among unobserved variables when the error terms are correlated across contiguous geographical space (Case, 1991).

The other spatial effect, spatial heterogeneity, instead, refers to the systematic variation in the behaviour of a given process across geographically-defined space, and typically leads to heteroscedastic error terms, as described in Can (1990). It could exist in a dataset which contains spatial information because a geographical area is uniform and/or boundless (Getis et al., 2004). Spatial heterogeneity thus entails that functional forms, as well as parameters, vary with location; they are not homogeneous across the dataset (Anselin, 1998). As a result, it is arguably inappropriate to estimate the overall parameters for the entire regression, as it is not able to adequately represent and capture the process at any given geographical location due to dissimilarity in spatial units.

In order to accommodate the spatial heterogeneous effect, most empirical applications to transport focus on the application of geographically weighted regression (GWR) model (e.g. Lloyd and Shuttleworth, 2005; Park, 2004; Zhao and Park, 2004; Zhao et al., 2006). Wang et al. (2008) suggest that GWR is a typical technique in exploring the analysis of spatial data, based on its wide-ranging empirical applications. This spatial technique was first introduced by Brunsdon et al. (1996) and further refined by Fotheringham et al. (2002). One benefit of GWR is that it models a local spatial relationship by calibrating a varying coefficient regression model that generates parameters assigned by the spatial units of analysis. Basically, it contains of a group of locally linear regressions that employs distance-weighted intersecting series of the data to examine the potential relationships that vary in geographical space (Farber and Yeates, 2006). This unique feature allows GWR to fully assess the spatial heterogeneity effect between the dependent and independent variables, in the estimated relationship. Another advantage is that one can
visualise the estimated results by transforming the text-stored files into visual maps produced by geographical information system (GIS) packages such as ArcView.

Nevertheless, according to Wheeler et al. (2006), GWR could suffer from substantially stronger multicollinearity and correlation among local regression coefficients. Due to these limitations, the primary use of local regression models is as an exploratory smoothing method only, and it should not be considered an inferential statistical tool (Young et al., 2008). Therefore, even though GWR continues to develop into a standard model for the analysis of spatial data, its adoption among econometricians is fairly limited (Anselin, 2010).

In practice, however, it remains difficult to fully disentangle the effects of spatial non-stationarity from spatial dependency (Bailey and Gatrell, 1995). Florax and Nijkamp (2003) state that in fact, the existence of spatial heterogeneity does not necessarily imply severe implications for the information that can be acquired from a spatial dataset. Spatial dependency, however, does, because a given observation is partly predictable from its neighbouring observations. Therefore this thesis will only focus on discussing the effect of spatial dependency.

LeSage (2004) advocates the application of spatial models when dealing with spatial effects over conventional regressions with augmented binary variables representing geographic dichotomous information. He argues that by simply adding such region dummies, or variables reflecting interaction with locational coordinates that allow variation in the parameters over space to traditional regression models, the estimated results can hardly ever outperform a spatial model.
2.5.2. Spatial lag model

Because of their simplicity, to address the issues of spatial autocorrelation mentioned above, prior spatial studies are mainly concerned with models that contains only one type of spatial interaction effect viz. the spatial lag model and the spatial error model. The former incorporates a spatially lagged dependent variable on the right-hand-side of a regression whereas the latter contains a spatial autoregressive process in the disturbance (Elhorst, 2010). Following Anselin (1992), the point of departure is a simple spatial autoregressive (SAR) model:

\[ y = \rho Wy + X\beta + u \]  

\[ u \sim N(0, I\sigma^2) \]

where \( y \) is an \( n \times 1 \) vector of observations on the dependent variable; \( X \) is an \( n \times k \) matrix of observations on independent variables; and \( \rho \) is a spatial autocorrelation or spatial dependence parameter (Wall, 2004). Additionally, \( \rho \) represents the intensity of the spatial dependence between neighbouring locations, \( W \) is an \( n \times n \) exogenous spatial weights matrix that specifies the assumed spatial structure and also describes the spatial arrangement of the spatial units in the sample. The element \( w_{ij} \) of \( W \) measures the nearness of area units \( i \) and \( j \). \( Wy \) is thus the spatially lagged dependent variable which has the ability to account for various spatially related dependencies. Finally, \( \beta \) represents \( k \times 1 \) vector of estimators to be estimated, and \( u \) is an \( n \times 1 \) vector of independently and identically distributed (i.i.d) random error terms.

\[ 15 \text{ If, however, as indicated in Osland (2010), the assumed spatial structure does not hold for the existing spatial dependence, the estimated value of } \rho \text{ will be low or insignificant. And if } \rho = 1, \text{ then equation (2.1) is equivalent to a standard linear regression model.} \]
Equation (2.1) can be solved for \( y \), and the reduced form is shown in equation (2.2):

\[
y = (I_n - \rho W)^{-1} X \beta + (I_n - \rho W)^{-1} u
\]  

(2.2)

LeSage (2004) points out that in the SAR model, each observation \( y_i \) is a function of the spatial lag term \( W y \) which represents an explanatory variable denoting the weighted average of spatially neighbouring values, e.g., \( y_2 = \rho (w_{12} y_1 + w_{13} y_3) \). Anselin (2003) notes that intuitively, the spatial lag term \( W y \), is correlated with the random error terms \( u \), even when the latter are i.i.d. Consequently, it must be treated as an endogenous variable and the proper estimation technique such as Maximum Likelihood (ML), Spatial Two Stage Least Squares (S2SLS) and/or Generalised Method of Moments (GMM) estimations should apply to account for this endogeneity problem, because the OLS estimates for \( \beta \) are biased and inconsistent.

2.5.3. Spatial error model

If spatial dependence exists in disturbances, a spatial error model (SEM) is usually applied in order to improve the precision of the estimated parameter because this kind of regression involves a non-spherical error term. A general representation of the SEM model can be written as follows (Anselin, 2003):

\[
y = X \beta + \epsilon
\]  

(2.3)

and
\[ \varepsilon = \theta W \varepsilon + u \]  

(2.4)

Substituting Equation (2.4) into (2.3) gives:

\[ y = X\beta + (I_n - \theta W)^{-1} u \]  

(2.5)

where \( \theta \) is known as the spatial autoregressive parameter, which needs to be estimated jointly with the regression coefficients; \( W \) is the spatial weight matrix; and \( \varepsilon \) is a vector of i.i.d random error terms, which is assumed to be uncorrelated to \( u \).

What is notable about this SEM model is that the spatial weights matrix \( W \) now relates to shocks in the unobserved variables (i.e. the error term \( \varepsilon \)) but not to the independent variables of the SEM model (i.e. variable \( X \)). In other words, Equation (2.5) shows that the value of the dependent variable for each geographical location is influenced by the stochastic error terms at all other locations through the spatial multiplier \((I_n - \theta W)^{-1}\).  

Consequently, in the context of public transport, the bus mode share at any spatial area is not only a function of the local characteristics but also of the unobserved variables at adjacent locations. In the case of spatially correlated disturbances, one should notice that even though OLS estimated-results are still unbiased, they are no longer efficient. In addition, according to Anselin (2001), the classical estimators for standard errors are biased.

---

16 Osland (2010) indicates that the smaller the absolute value of \( \lambda \), the smaller the effect of the spatial multiplier.
2.5.4. Spatial Durbin model

In practice, one should realise that spatial dependence can have effects on both dependent and explanatory variables. Hence, according to Osland (2010), a “mixed” spatial Durbin model (SDM) which is introduced by Anselin (1988), offers a more flexible alternative and might be more appropriate to apply by including the “inherent spatial autocorrelation” and the “induced spatial dependence” simultaneously.

The SDM is specified as follows:

\[ y = \rho Wy + X\beta + WX\gamma + u \quad (2.6) \]

This model can be reduced to either Equation (2.1) if \( \gamma = 0 \) or Equation (2.3) if \( \gamma = -\rho \beta \).

The reduced form of Equation (2.6) is:

\[ y = (I_n - \rho W)^{-1}X\beta + (I_n - \rho W)^{-1} WX\gamma + (I_n - \rho W)^{-1}u \quad (2.7) \]

Based on the above equations, an additional term \( WX\gamma \) must be included in the model to capture the \( k \times 1 \) autoregression coefficient vector \( \gamma \) of the spatially lagged explanatory variables \( WX \), which measures the marginal impact of the independent variables from adjacent observations on the dependent variable \( y \) (Kissling and Carl, 2008).

Furthermore, Osland (2010) argues that this SDM could be developed from either an SEM (Anselin, 2006) or from an SAR (Bivand, 1984), and this “mixed” model can be viewed as
an unrestricted model of either SEM or SAR. In other words, the SDM further nests the SAR and the SEM by involving spatial dependence in the error term as well as in the dependent variable. Figure 2.5 illustrates the theoretical relationship between SDM, SAR and SEM in a cross-sectional case.

According to LeSage and Pace (2009), SDM is the only model that will produce unbiased estimates regardless of the true data-generation process (i.e. whether it is a spatial lag or a spatial error model). This is why the SDM is often viewed as the dominant spatial model among others. Although most transport data are geographically linked, past transport studies incorporating spatial effects are relatively scarce compared to their rich applications in other fields, such as agricultural and resource economics (e.g. Benirschka and Binkley, 1994; Hurley et al., 2001; Roe et al., 2002; Weiss, 1996) and housing and real estate analysis (e.g. Basu and Thibodeau, 1998; Berg, 2002; Case et al., 2004; Pace and Gilley, 1997; Smith and Wu, 2009). This empirical gap thus leads us to consider the use of spatial econometric models in the field of transport analysis.
2.5.5. Spatial autoregressive logit model

Spatial lag, spatial error, and spatial Durbin models are the three models of which one can apply when conducting spatial econometrics analysis using cross-sectional data in a spatial context. However, for the spatial analyses of the individual’s transport mode choice, these models will not be applicable because in such cases, as the dependent variable is no longer a continuous variable, rather, it takes the form of a discrete variable. Therefore, another type of spatial model, spatial autoregressive logit model, is introduced.

Following Goetzke (2008), in a binary mode choice case, (i.e. to choose between commuting with public transit or private vehicle), in order to incorporate spatial effects among nearby commuters into transport mode choice model, one should extend the standard logit model, which is derived from the random utility function, by adding a spatially autoregressive mode choice term, $W_{mi}$, who indicates the spatially weighted average mode share of mode $j$ of all the commuters in surrounding locations of individual $i$, in its reduced form. Section 4.3. provides a detailed discussion of how the spatial autoregressive logit model works.

This section highlighted the point that past works that study traveller’s travel behaviour at either aggregate or disaggregate levels, tend to ignore or underestimate the importance of spatial dependency. Given this current research gap from the literature, the following two empirical chapters (i.e. chapter 3 and 4) will conduct some in-depth analyses around commuter’s travel behaviour in a spatial context, with the application of spatial econometric models, at the regional and individual level respectively. Using cross-sectional data mainly extracted from the New Zealand 2006 census, chapter 2 will examine the built environment determinants (i.e. urban forms) of JTW public transport ridership in the context of variant geographical locations, with the application of SDM, based on the hypothesis that there is a spatial dependency in the data. Furthermore, using NZHTS data, chapter 3 will investigate the attributes that affect commuter’s transport
mode preferences by applying a spatially autoregressive model, for the period of 2005/06 – 2008/09, with the assumption that transport mode preference of a certain commuter’s neighbours alongside the characteristics of the regions where he/she lives, do have an impact on how he/she chooses the transport mode to work. These two studies use local data from the Auckland region only.
Chapter 3. The Influence of Urban Forms on Transit Behaviour in the Auckland region: A Spatial Durbin Analysis

Abstract

As well documented in the literature, urban form plays an important role in determining transit ridership. However, among these studies, the majority of empirical work has not considered space as a relevant factor. Instead, most of the findings are based on a strong assumption that there is no spatial effect across the research area. This general negligence of spatial effects will, in turn, produce biased estimators if substantial geographical patterns exist. Given the observational heterogeneous distribution of transit patterns in the Auckland region, it is exceedingly doubtful whether the assumption of no spatial interdependence is valid. Based on cross-sectional data, mainly extracted from the New Zealand 2006 census with additional geographical information compiled by ArcMap for the Auckland region, this chapter contributes to the existing literature by offering insight into the spatial structure of the current public transport sector. The use of a spatial Durbin model provides a better understanding of the urban form factors that influence bus mode share by decomposing the total effect of one explanatory variable into direct and indirect effects. The results show that the total effects are comprised mostly of spatial spill-over impacts. In addition to urban form variables, several other dimensions of potential bus mode share predictors are considered, including transit supply quality, accessibility to other modes of public transport, and variables that describes household characteristics.

Key words: Spatial dependence, spatial Durbin model, spill-over effect, urban form, transit behaviour

A shorter version of this chapter has been presented at the
3.1. Introduction

3.1.1. Background information

Throughout the world, as people’s incomes rise, many shift to faster, more comfortable and more individually flexible means of transportation (Downs, 2003). Not surprisingly, like most modern cities, the recent commuting pattern in Auckland, where one third of New Zealand’s population lives, is dominated by the automobile, with almost 88% of the share for the morning journey to work (JTW) attributed to private motor vehicles, while public transport accounted for only around 8% of the journeys (MoT, 2013a). In comparison to other competitor cities, data from Auckland Regional Transport Agency [ARTA] (n.d.) confirm Auckland’s weak position in terms of public transport use, with only 41 public transport trips made per capita per annum, while Wellington generated almost twice this number at 91, and Sydney had almost threefold (as shown in Figure 3.1).18 Auckland is thus characterised by an elevated level of car-dependence and an extremely low public transport patronage. This lagging in public transport use may imply a relatively less developed public transport system, which could possibly limit Auckland’s potential to become more internationally competitive, in terms of attracting more international investment, events and tourism.

With the aim of reducing automobile dependence and inducing non-automobile commuting, transport planners around the world are attempting to tackle the travel growth problem by implementing transport planning projects that can promote forms of sustainable urban development (e.g. Banister and Marshall, 2000; Barton et al., 1995). In the case of Auckland, transport authorities have also implemented several major projects to facilitate the development of public transport, from both small-scale projects such as expanding bus priority lanes to large-scale development such as bus and rail infrastructure initiatives. Therefore, from the perspective of local government and urban planners, it is

---

18 Auckland Regional Transport Agency, or ARTA, has been replaced by Auckland Transport since the reorganisation at local government on 1st November, 2010. However, this chapter still refers to ARTA since all of the data used here were compiled when ARTA was in existence.
crucial to have a solid understanding of how well the design and layout of urban areas do in terms of contributing to a reduction in automobile use and public transport travel promotion. In other words, what will be the likely impact of urban form on people’s travel behaviour?

![Figure 3.1: Public Transport Trips and GDP per capita (2007/08, US$)](source: ARTA (n.d.))

3.1.2. Objective, motivation and scope

ARTA (2007) reports that in Auckland, the total number of (unlinked) trips travelled by the public transport system in 2007 was 52.4 million, with buses, trains and ferries
contributed to 82%, 10% and 8%, respectively. Given its current dominant position in public transport usage, the analysis undertaken here focuses on the bus only.\footnote{Moreover, New Zealand census data such as ferry usage by commuters in Auckland do not exist.}

The motivation behind this chapter is that in order to properly understand the relationship between urban form and transit ridership, it is necessary to consider the associated spatial structures more specifically. A number of studies have attempted to identify the impact of urban form on different travel behaviours such as mode choices, travel demand and travel patterns, over the past few decades (e.g. Gordon et al., 1989; Headicar and Curtis, 1994; Kitamura et al., 1997; Næss et al., 1995). A key problem with the above literature is the possibility that coefficient estimates of the impact of urban form might be attenuated by spatial dependence. Specifically, these analyses assume that observations are independent of one another in a geographical context. However in reality, it seems unlikely that region $i$’s transport network in terms of vehicles and public transport infrastructure is independent of that of its neighbouring region $j$. Furthermore, from the econometric point of view, ignoring spatial characteristics between observations could, in turn, produce biased and inconsistent estimators (LeSage and Pace, 2010).

This general limitation from past literature gives rise to the need for alternative spatial estimation approaches, such as the spatial Durbin model (SDM), as it has the advantage of separating total effect of a particular variable on the transit ridership into own-region and neighbourhood effects. To the best of our knowledge, Greer and van Campen (2011) have produced the only published paper which specifically takes spatial effects into account when analysing the determinants of work trip bus ridership in the context of New Zealand, using the SEM model. Their paper is also the most relevant study to the current empirical work in this chapter. Greer and van Campen (2011) use cross-sectional data involving 318 area units for the Auckland region in 2006. Once positive spatial autocorrelation is confirmed by a statistically significant Moran’s $I$ value, an SEM model is chosen based on the fact that the robust Lagrange Multiplier (LM) statistic is more significant in the case of...
the spatial error test compared to it alternative, SAR model. Regarding the choice of the spatial weights matrix, a Rook contiguity of order 2, including the lower order of 1, is applied based on the fact that it provides the best fit to the data. The estimation method employed in their study is ML. It is concluded that after adjusting for spatial dependency, the SEM model denotes a substantial enhancement over the OLS model, by providing more accurate parameter estimates and improving the overall predictive power of the empirical model.

However, there remains a potential weakness in interpreting Greer and van Campen’s results. In addition to the spatial lag of the dependent variable included on the right-hand-side of the regression equation, it seems plausible that neighbouring area unit’s characteristics, for instance population density and rush hour frequency, could also play a significant role in explaining variations in a given area unit’s bus ridership. This implies that further investigation of the impact of lagged explanatory variables on transit ridership is required. This study applies the SDM model which has the ability to capture the characteristics of neighbouring regions in order to account for any influence they may exert on their neighbour’s transit ridership patterns.

The remainder of the chapter is organised as follows: section two describes the dataset, outlines the variables used, and specifies the regression models employed. Section three presents some preliminary results from spatial econometric tests. Section four delivers the empirical results of the non-spatial OLS model and the spatial models. The final section provides a conclusion by summarising key findings, outlining limitation and suggesting future works of this study.
3.2. Data and Empirical Models

3.2.1. Data

The major source of data for this study was the New Zealand Census, collected and complied by the Statistics New Zealand on the census day, 6th March, 2006. Additional data, such as distance to Auckland’s CBD, distance to the nearest rail or ferry terminals and census area unit land areas, are calculated using ArcMap. Furthermore, the rush hour frequency, which combines the total number of buses passing through and stopping within each area unit, during both morning and afternoon peak periods, is compiled using the programmes of ArcMap and Microsoft Excel. The data were geocoded at the centroid of each area unit.

The census area unit is the second smallest geographical unit defined by Statistics New Zealand. Area units are aggregations of meshblocks, and they are non–administrative areas that are in between meshblocks and territorial authorities in size (Statistics New Zealand, n.d. a). All data used in this study were compiled at this geographical level. In line with Yu et al. (2010), smaller units such as the meshblocks would render too much variation, and consequently increase analytical instability, while larger units such as territorial authorities would aggregate data too much and are thus incapable of providing useful results. 20 There are 399 census area units within the Auckland region, of which sufficient data could be collected on 317. The final dataset is consistent with the one employed by Greer and van Campen (2011). 21

---

20 There are more than ten thousand mesh blocks and only seven territorial authorities in the Auckland region.
21 Note that in Greer and van Campen (2011), the total number of observation is 318 area units. However, Waiheke Island is dropped from the dataset because although there are some bus services running within this area, they are not connected with any other bus services due to its isolated nature.
3.2.2. Variables

The selection of variables is mainly inspired by previous bus patronage studies. The dependent variable $Bus_i$ is the percentage of workers in area unit $i$ who take bus as their main transport to work, self-reported on the census day. It was obtained by dividing the total number of bus passengers by the total number of JTW commuters in the $i^{th}$ area unit. The percent mode share to bus offers an overall measure of the prominence of bus transport in the Auckland region.

*Figure 3.2* presents the spatial distribution of bus mode share in the Auckland region based on 2006 census data. From this figure, it is evident that the bus mode share is not evenly distributed across area units. More specifically, the observations do not seem to be randomly distributed over space. Area units which have a high level of bus mode share, represented by the darker colour zones, tend to be closely concentrated in the centre, while the area units which have a relatively low bus ridership, shown in the lighter colour parts, are scattered around the boundaries.

Additionally, small clusterings of high values are also detected on the northeast and southeast corners of the map, which further indicates the spatially heterogeneous nature of the distribution of bus mode share. Therefore, spatial autocorrelation is apparently observed, because undoubtedly the probability of a specific value of the bus mode share variable in one specific location (area unit) depends on its value in neighbouring locations.
Figure 3.2: Spatial Distribution of Bus Mode Share in the Auckland Region
Potential bus mode share predictors are divided into three categories: urban form, transit service, and household characteristics. The final dataset includes eight independent variables, where:

1. Urban form variables:
   - *Population Density*$_i$: gross population density in the $i^{th}$ area unit in the Auckland region, measured by the total number of inhabitants per square kilometre;
   - *Employment Density*$_i$: employment density, measured by the total number of full-time and part-time employees per capita in the $i^{th}$ area unit in the Auckland region;
   - *Dwelling*$_i$: total number of private owner occupied dwellings in the $i^{th}$ area unit in the Auckland region; to be used as an indicator of land use patterns;
   - *CBD*$_i$: distance to CBD from the centroid of the $i^{th}$ area unit in the Auckland region, in kilometres;

2. Transit service variables:
   - *Station*$_i$: distance to the nearest public transport terminal/stop other than bus (either train or ferry) from the centroid of the $i^{th}$ area unit in the Auckland region, measured in kilometres;
   - *Frequency*$_i$: frequency of bus service within the $i^{th}$ area unit in the Auckland region;

3. Household characteristic variables:
   - *Income*$_i$: median household income measured in thousands of New Zealand Dollars (NZD) within the $i^{th}$ area unit in the Auckland region;
   - *Car*$_i$: mean number of motor vehicles per household within the $i^{th}$ area unit in the Auckland region;
The Transportation Research Board [TRB] (1996) points out that several urban form variables such as road network type and neighbourhood type, also have some influence on the demand for public transport. In addition, according to Paulley et al. (2006), a few transit service variables which describe the quality of transit service, such as the in-vehicle time and an indicator of the waiting environment, will have some effects on the demand for public transport as well. Unfortunately, these data are not available in this dataset. A summary of key descriptive statistics of the variables used in this analysis are presented in Table 3.1. As can be seen from this table, the bus share for JTW trips in the Auckland region is fairly low; the average figure for all 317 area units is only 5.65%, ranging from a low of 0.13% to a high of 17.43%.

### Table 3.1: Area Unit Level Descriptive Statistics of Variables for Auckland Region

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bus (%)</td>
<td>5.65</td>
<td>3.44</td>
<td>0.13</td>
<td>17.43</td>
</tr>
<tr>
<td><strong>Urban Form variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population Density (per km^2)</td>
<td>833.98</td>
<td>405.64</td>
<td>1.47</td>
<td>1726.74</td>
</tr>
<tr>
<td>Employment Density (per capita)</td>
<td>0.48</td>
<td>0.08</td>
<td>0.27</td>
<td>0.66</td>
</tr>
<tr>
<td>Dwelling</td>
<td>1241.25</td>
<td>503.51</td>
<td>114</td>
<td>3270</td>
</tr>
<tr>
<td>CBD (km^2)</td>
<td>16.68</td>
<td>8.36</td>
<td>2.23</td>
<td>43.29</td>
</tr>
<tr>
<td><strong>Transit Service variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station (km^2)</td>
<td>3.67</td>
<td>4.17</td>
<td>0.14</td>
<td>35.53</td>
</tr>
<tr>
<td>Frequency</td>
<td>130.03</td>
<td>94.48</td>
<td>2</td>
<td>476</td>
</tr>
<tr>
<td><strong>Household characteristic variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income (in thousands of NZD)</td>
<td>27.11</td>
<td>6.32</td>
<td>14.4</td>
<td>48.4</td>
</tr>
<tr>
<td>Car</td>
<td>1.71</td>
<td>0.2</td>
<td>1.18</td>
<td>2.32</td>
</tr>
</tbody>
</table>
3.2.3. Empirical bus mode share models

A logarithmic transformation is applied to both dependent and explanatory variables with the intention of capturing the a priori belief that ceteris paribus, the impact of each explanatory variable on bus mode share is diminishing (Bresson et al., 2004; Gomez-Ibanez, 1996).

Therefore, firstly, the non-spatial bus mode share model in log-log form is specified as below:

\[ \ln(\text{Bus}) = \mathbf{X}\beta_{\text{OLS}} + \epsilon_{\text{OLS}} \]  

(3.1)

The above equation posits that the variation in the natural logarithm of the bus mode share (\(\ln(\text{Bus})\)) in area unit \(i\) is explained by the variables in matrix \(\mathbf{X}\), which comprises a constant term, the natural logarithm of urban form, transit service, and the household characteristic variables. Since the equation of linear regression (3.1) is estimated by ordinary least squares, it is labelled as the OLS model and hence the estimated results from this model serve as a benchmark against the following spatial model estimations. 22

Secondly, the following SAR model is:

\[ \ln(\text{Bus}) = \rho \mathbf{W}\ln(\text{Bus}) + \mathbf{X}\beta_{\text{SAR}} + u_{\text{SAR}} \]  

(3.2)

Similarly, the SEM is:

\[ \ln(\text{Bus}) = \rho \mathbf{W}\ln(\text{Bus}) + \mathbf{X}\beta_{\text{SEM}} + u_{\text{SEM}} \]  

(3.3)

22 All of the empirical models (i.e. OLS, SAR and SDM) are estimated using Stata 11.
\[ \ln(\text{Bus}) = X\beta_{\text{SEM}} + \varepsilon \quad (3.3) \]

where \( \varepsilon = \theta W \varepsilon + u_{\text{SEM}} \)

Lastly, the SDM is given as:

\[ \ln(\text{Bus}) = \rho W \ln(\text{Bus}) + X\beta + WX\gamma + u_{\text{SDM}} \quad (3.4) \]

### 3.3. Preliminary Tests

#### 3.3.1. Spatial weights matrix

In empirical spatial econometric models, the selection of a spatial weights matrix, normally denoted as \( W \), plays an important role. A LeSage (2002) outline that there are many possible means to quantify the structure of spatial dependence between observations. Typical approaches include: distance decay (Anselin, 1980), structure of a social network (Doreian, 1980), economic distance (Case et al., 1993) and \( k \) nearest neighbours (Pinkse and Slade, 1998). However, as Leenders (2002) illustrates, one major challenge facing spatial econometric models is that the spatial weights matrix \( W \) cannot be directly estimated but needs to be explicitly specified \textit{a priori}, and current economic theory provides no formal guidance for this. Although a wide range of literature, echoed by Anselin (2002), has proposed several approaches to create the spatial weights matrix, there barely exists a formal guidance on how to select the “optimal” spatial weights as existing specifications all seem somewhat arbitrary.
Practically, in spite of their lesser theoretical appeal, geographically derived weights are among the most widely applied specification in spatial econometric analysis (Anselin, 2001). In addition, as Manski (1993) argues, this popularity of geographically derived weights is due to the fact that the structure of $W$ is constrained so that the weights are truly exogenous to the model, thus avoiding identification problems. Generally, there are two types of geographically derived weights based on proximities, namely, a binary measure of continuity (when two areas share common borders) and a continuous measure of distance. Following a majority of empirical studies (Fingleton, 1999, 2000; Le Gallo et al., 2003; Rey and Boarnet, 2004), we use a two-dimensional Cartesian coordinate system with the ordered pair $(x, y)$ coordinates to create a spatial weights matrix $W$, given the distance decay specification and its eigenvalues matrix $E$.

By convention, the weights matrix $W$ has been row-standardised such that every row of the matrix sums to one (i.e. $\sum_j w_{ij} = 1$). Each element of $W$ is therefore defined as:

$$w_{ij} = \begin{cases} 
0 & \text{if } i = j \\
\frac{1}{d_{ij}} & \text{if } d_{ij} \leq d^* \text{ and;} \\
0 & \text{if } d_{ij} > d^* \text{ if observation } i \neq j
\end{cases}$$

where $d_{ij}$ is the spherical distance between the centroids of area units $i$ and $j$, and $d^*$ is the critical cut-off distance. This inverse Euclidean distance, $d_{ij}$, contains a maximum threshold band of 24.14 kilometres to guarantee connections between all area units, that is, each spatial unit must have at least one neighbour. This indicates that two area units are considered neighbours when the distance between their centroids is less than 24.14 kilometres, and not neighbours if their centroids lie 24.14 or more kilometres apart.

\[\text{23 The default unit for cut-off length is in miles in Stata 11, by conversion, 15 miles are approximately equal to 24.14 kilometres.}\]
3.3.2. Moran’s I test

A univariate Moran’s I test for residuals is the most commonly employed first-step specification test for spatial autocorrelation (Moran, 1948; Anselin, 1999). The test does not specify an explicit alternative spatial model (i.e. either SAR or SEM models) but has power against both (Anselin and Rey, 1991).

The Moran’s I test for residuals in matrix notation is captured by:

\[ I = (N / S_0)(e'W e / e'e) \]

where \( e \) denotes a vector of OLS residuals, and \( S_0 = \sum_i \sum_j w_{ij} \), a standardisation factor that refers to the sum of the weights for the non-zero cross products (Anselin, 1999).

According to Florax and Nijkamp (2003), the interpretation of Moran’s I should be parallel to a correlation coefficient; however the major distinction is that its value is not bounded by the closed, (-1,+1) interval. A positive value signals positive spatial autocorrelation, measuring the occurrence of similar levels of a variable being found over contiguous or nearby spaces. By contrast, a negative value signals negative spatial autocorrelation, measuring the joint occurrence of high and low attribute values in adjoining locations.

The Moran’s I statistic shows a positive value of 18.733 with a p-value that is lower than 0.0001. As expected, this result indicates that the null hypothesis of no spatial dependence should be rejected. Furthermore, the test statistic indicates that positive spatial
autocorrelation exists, and in order to obtain unbiased and consistent estimators, spatial models should be adapted instead of the non-spatial OLS estimations.

### 3.3.3. The Lagrange Multiplier test

By applying the Lagrange Multiplier ($LM$) test, we select between a spatial lag and a spatial error alternative (Anselin, 2001). Basically there are two major forms of the $LM$ test. The $LM_{lag}$ statistic tests the null hypothesis of no spatial autocorrelation in the dependent variable; the $LM_{error}$ statistic, on the other hand, tests the null hypothesis of no significant spatial autocorrelation in the error terms.

The $LM$ test against a spatial lag alternative ($LM_{lag}$) is demonstrated in Anselin (1988) and takes the following form:

$$LM_{lag} = \left[ e'Wy / (e'e/N) \right]^2 / D$$

where $D = [(WX \beta)'(I_n - X(X'X)^{-1}(WX \beta))/\sigma^2] + tr(W^2 + W'W)$. \(^{24}\)

By contrast, the $LM$ test against a spatial error alternative ($LM_{error}$), which is initially outlined in Burridge (1980), takes the form of:

$$LM_{error} = \left[ e'We / (e'e/N) \right]^2 / [tr(W^2 + W'W)]$$

\(^{24}\) “tr” denotes the trace of the matrix $W$. 

Aside from a scaling factor, this statistic corresponds to the squared value of Moran’s $I$.

As recommended by Florax and Nijkamp (2003), if both hypotheses can be rejected, one should consider constructing the robust forms of these $LM$ tests which have the ability to correct for the presence of local misspecification of the other form (Anselin et al., 1996; Bera and Yoon, 1993). The test procedures of $LM_{lag}^r$ and $LM_{error}^r$ are identical to the one described above$^{25}$. Both the classic and the robust LM tests are based on the residuals of the OLS model and are asymptotically distributed as $\chi^2(1)$.

Table 3.2 presents the diagnostics for spatial dependence. Under the classic $LM$ test, both hypotheses of no spatially lagged dependent variable and of no spatially autocorrelated disturbances can be rejected at a 1% significance level. The robust $LM$ tests consistently show the same results, with rejection of both hypotheses at a 1% significance level. This implies that OLS is rejected in favour of both SAR and SEM models.

Table 3.2: OLS Diagnostics for Spatial Dependence

<table>
<thead>
<tr>
<th>Measure</th>
<th>Statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEM</td>
<td>$LM_{error}$</td>
<td>123.601 ***</td>
</tr>
<tr>
<td></td>
<td>$LM_{lag}^r$</td>
<td>30.395 ***</td>
</tr>
<tr>
<td>SAR</td>
<td>$LM_{lag}$</td>
<td>128.174 ***</td>
</tr>
<tr>
<td></td>
<td>$LM_{lag}^r$</td>
<td>34.968 ***</td>
</tr>
</tbody>
</table>

$^{25}$ The subscript “$r$” denotes “robust”. 
Unlike what holds for the SAR’s counterpart, the Autoregressive (AR) model in time series analysis, the OLS estimation in the presence of spatial dependence will be inconsistent, simply because of the endogeneity issue discussed before. Therefore, in this study, the SAR and SEM models are estimated using ML estimation (Anselin, 1988; Ord, 1975).

3.4. Estimation Results

3.4.1. Non-spatial and spatial models

The results from the non-spatial OLS is reported in Table 3.3, alongside with the estimated results under the SDM. Several distinctions are evident. Firstly, the estimated coefficients of the urban form variables are significant and of the expected signs, in line with earlier findings in the literature. However, against expectations, the variable $\ln(Dwelling)$ is not significant. Additionally, compared with the SDM, most estimated coefficients from the OLS are larger in magnitude, implying that without the inclusion of potential spatial autocorrelation between dependent, independent variables and the error terms, the OLS results simply ignore this spatial variation and produce biased estimates.

The value of R-squared ($R^2$) is 0.730, indicating a reasonable model fit. 26 However, as the results from the Moran’s I statistic and model diagnostic tests in Table 3.2 show, estimates using the OLS method suffers from a major problem: there is evidence of a positive spatial autocorrelation, and the $LM$ test statistic suggests the lag specification as the appropriate alternative. Thus, the above OLS estimates should be interpreted with caution.

---

26 Adjusted R-squared for OLS estimation is 0.725.
Table 3.3: Non-Spatial OLS, Spatial Autoregressive Model and Spatial Durbin Model (Dependent Variable: lnBus)

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>OLS Estimates</th>
<th>SDM Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.749</td>
<td>3.522</td>
</tr>
<tr>
<td>ln(Population Density)</td>
<td>0.138</td>
<td>0.141</td>
</tr>
<tr>
<td>ln(Employment Density)</td>
<td>1.646</td>
<td>-0.285</td>
</tr>
<tr>
<td>ln(Dwelling)</td>
<td>-0.218</td>
<td>-0.166</td>
</tr>
<tr>
<td>ln(CBD)</td>
<td>-0.970</td>
<td>-0.511</td>
</tr>
<tr>
<td>ln(Station)</td>
<td>0.260</td>
<td>0.120</td>
</tr>
<tr>
<td>ln(Frequency)</td>
<td>0.158</td>
<td>0.143</td>
</tr>
<tr>
<td>ln(Income)</td>
<td>-1.053</td>
<td>-0.579</td>
</tr>
<tr>
<td>ln(Car)</td>
<td>-0.865</td>
<td>-0.732</td>
</tr>
<tr>
<td>Lag ln(Population Density)</td>
<td></td>
<td>-0.379</td>
</tr>
<tr>
<td>Lag ln(Employment Density)</td>
<td></td>
<td>3.528</td>
</tr>
<tr>
<td>Lag ln(Dwelling)</td>
<td></td>
<td>-0.066</td>
</tr>
<tr>
<td>Lag ln(CBD)</td>
<td>0.480</td>
<td>**</td>
</tr>
<tr>
<td>Lag ln(Station)</td>
<td>0.098</td>
<td>*</td>
</tr>
<tr>
<td>Lag ln(Frequency)</td>
<td>0.281</td>
<td></td>
</tr>
<tr>
<td>Lag ln(Income)</td>
<td>0.410</td>
<td></td>
</tr>
<tr>
<td>Lag ln(Car)</td>
<td>-3.613</td>
<td>***</td>
</tr>
<tr>
<td>(\rho)</td>
<td>0.823</td>
<td>***</td>
</tr>
<tr>
<td>Squared Correlation</td>
<td>0.730</td>
<td>0.852</td>
</tr>
<tr>
<td>Variance Ratio</td>
<td></td>
<td>0.779</td>
</tr>
<tr>
<td>Log likelihood</td>
<td></td>
<td>-59.929</td>
</tr>
</tbody>
</table>

*** Estimated coefficients significant at 1% level; ** significant at 5%; * significant at 10%
Therefore we should concentrate on the estimated results from the SDM. First of all, upward bias is found in most of the least-squares estimates, suggesting over-estimation of the sensitivity of bus mode share to the urban form, transit supply, and socio-economic and demographic characteristics when spatial dependence is disregarded. Secondly, the spatial lags on $\ln(\text{Population Density})$, $\ln(\text{CBD})$ and $\ln(\text{Station})$ are all significant in the SDM, implying possible omitted variable issue if we do not include them in the non-spatial OLS model. Thirdly, by taking the spatial lag into account, the fit of the model has improved dramatically. The $R^2$ statistic for the SDM model is 0.852, which has higher value compared to the one in OLS. Therefore, after adjusting for spatial dependence, the overall fitness of the model has been improved.

3.4.2. Choosing between alternative spatial dependence models

As Elhorst (2010) describes, if the OLS model is rejected in favour of both SAR and SEM models, then the SDM should be estimated. Therefore, a likelihood ratio (LR) test, also known as the score test, can subsequently be used to test two separate hypotheses that $H_0: \gamma = 0$ and $H_0: \rho \beta + \gamma = 0$.

Recall that the SDM model is reduced to the SAR model if $\gamma = 0$. Osland (2010) proposes that when there is evidence of maintaining the SAR or SEM model, the SDM model specified by Equation (3.4) and the following log likelihood tests may be useful in terms of determining the “true” spatial process. Thus, for the SAR model, one can determine the dominant model by testing the null hypothesis $\gamma = 0$. Rejecting the null hypothesis implies rejecting the SAR. Similarly, a common factor constraint: $\gamma = -\rho \beta$ should be tested in order to determine the best model between the SDM and its SEM. Likewise, if the null is rejected, this indicates statistical evidence in favour of the SDM. With the aid of the LR test, one can decide the better model between the SDM and its restricted versions.
The likelihood ratio ($\lambda$) is defined as:

$$\lambda = 2[\ln (L_U) – \ln (L_R)] \sim \chi^2(m)$$

where $L_U$ is the likelihood function of the unrestricted model (i.e. $L_U = L_{SDM}$) whereas $L_R$ is the likelihood function of the restricted model (i.e. $L_R = L_{SAR}$ or $L_{SEM}$), and $m$ is the number of restrictions imposed. The idea is that if the restrictions are valid, the log likelihood functions should appear to be similar in values and accordingly $\lambda$ should be equal to zero.

The following results are obtained: $L_{SDM} = -59.929$, $L_{SAR} = -86.097$ and $L_{SEM} = -75.867$.

With 8 degrees of freedom, the critical values at 1%, 5% and 10% significance are 1.646, 2.733 and 3.490, respectively. The test statistics exceed the critical values for all cases, therefore we can reject the null hypothesis that the underlying spatial process is SAR or SEM at a 1% significance level. In other words, the restriction on parameter $\gamma$ associated $WX$ and also the common factor constraint are invalid. As a result, the unrestricted SDM should be employed to represent the data-generation process of the spatial dependence. This result further implies that the spatial lags of both the dependent and explanatory variables should be included in the model. In fact, the inclusion of the spatial lags of independent variables makes reasonable sense as area units located near each other should have some degree of similarity in terms of urban form, transit supply and household characteristics variables, because economic activities tend to interact largely across space.

The estimation results for the SDM model are summarised in Table 3.3, alongside the OLS estimates. Overall, the SDM explains over 85% of the variation in the bus mode share.
3.4.3. Decomposing total effect into direct and indirect effects

Interpretation of the SDM model differs from that of its non-spatial regression counterpart, the ordinary least squares, as the $k^{th}$ parameter vector $\beta$ is no longer a partial derivative of $y$ with respect to change in the $k^{th}$ independent variable from the $n \times k$ matrix of $X$ (LeSage and Fischer, 2008). Essentially, the spatial dependence components in the SDM expand the spatial information set to include additional information from neighbouring area units. To see the impact of this additional spatial information, consider the partial derivative of the SDM in equation (3.5) with respect to a particular explanatory variable $x_k$:

$$ M = \frac{\partial y}{\partial x_k} = (1 - \rho W)^{-1} [\beta_k + W\gamma_k] \quad (3.5) $$

The partial derivative results in an $n \times n$ matrix $M$ representing marginal effects, which is shown in Equation (3.5). The impact on the dependent variable from a change in a coefficient can be decomposed into three ways, namely, direct, indirect and total effects. The direct effect is defined as the average of the diagonal elements of matrix $M$ by LeSage and Pace (2009), it provides a summary measure that represents an average of the impacts on bus mode share arising from own-region changes in variable $x_k$. The indirect effect is defined as the average of the off-diagonal elements of matrix $M$; this effect is also known as the spatial spill-over effect as it measures the impact on bus mode share in area unit $i$ arising from changes in variable $x_k$ from all other area units. The total effect is calculated as the average row sums of matrix $M$; it includes both direct plus indirect effects. The total effect measures the average cumulative impact on each observation from changing the $k^{th}$ explanatory variable by one unit across all observations.
Average direct, indirect and total effects estimated are reported in Table 3.4, along with inferential statistics (i.e. the figures in parenthesis are bootstrapped standard errors) calculated using a bootstrap method with 1,000 draws. Because all of the variables are expressed in natural logs, the coefficient estimates can also be interpreted as elasticities.

Table 3.4: Direct, Indirect and Total Effects of the Spatial Durbin Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Direct Effect</th>
<th>Indirect Effect</th>
<th>Total Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\ln(\text{Population Density}))</td>
<td>0.136***</td>
<td>-1.480***</td>
<td>-1.345***</td>
</tr>
<tr>
<td></td>
<td>(4.73E-05)</td>
<td>(4.72E-05)</td>
<td>(5.81E-08)</td>
</tr>
<tr>
<td>(\ln(\text{Dwelling}))</td>
<td>-0.170***</td>
<td>-1.141***</td>
<td>-1.311***</td>
</tr>
<tr>
<td></td>
<td>(3.53E-05)</td>
<td>(3.54E-05)</td>
<td>(3.49E-08)</td>
</tr>
<tr>
<td>(\ln(\text{CBD}))</td>
<td>-0.510***</td>
<td>0.335***</td>
<td>-0.175***</td>
</tr>
<tr>
<td></td>
<td>(1.04E-05)</td>
<td>(1.03E-05)</td>
<td>(1.34E-08)</td>
</tr>
<tr>
<td>(\ln(\text{Station}))</td>
<td>0.124***</td>
<td>1.108***</td>
<td>1.232***</td>
</tr>
<tr>
<td></td>
<td>(3.43E-05)</td>
<td>(3.26E-05)</td>
<td>(5.10E-08)</td>
</tr>
<tr>
<td>(\ln(\text{Frequency}))</td>
<td>0.151***</td>
<td>2.245***</td>
<td>2.395***</td>
</tr>
<tr>
<td></td>
<td>(7.03E-05)</td>
<td>(7.05E-05)</td>
<td>(1.16E-07)</td>
</tr>
<tr>
<td>(\ln(\text{Income}))</td>
<td>-0.580***</td>
<td>-0.374***</td>
<td>-0.955***</td>
</tr>
<tr>
<td></td>
<td>(1.13E-05)</td>
<td>(1.15E-05)</td>
<td>(7.01E-09)</td>
</tr>
<tr>
<td>(\ln(\text{Car}))</td>
<td>-0.816***</td>
<td>-23.732***</td>
<td>-24.548***</td>
</tr>
<tr>
<td></td>
<td>(7.51E-04)</td>
<td>(6.93E-04)</td>
<td>(8.37E-07)</td>
</tr>
</tbody>
</table>

*** Elasticities significant at 1% level
For the total effects, all estimated parameter values have the expected signs, with one exception for $\ln(\text{Population Density})$. The total effects of $\ln(\text{Station})$ and $\ln(\text{Frequency})$ on transit ridership are all positive and significant; while the total effects of $\ln(\text{Dwelling})$, $\ln(\text{CBD})$, $\ln(\text{Income})$ and $\ln(\text{Car})$ and $\ln(\text{Population Density})$ are negative and significant. Separating the total effect of a regressor into direct and indirect effects yields further insights.

For the two transit service variables, first of all, the total effect, which comprise the direct and indirect effects of $\ln(\text{Station})$, is positive and significant, implying that across the Auckland region, as the distance to train station and/or ferry terminal increases, commuters will prefer to choose buses as their transport mode. Next, both the direct and indirect effects of rush hour frequency show a significant positive effect on the bus mode share in a given area. This result provides insights to transport planning viz. that by increasing the number of buses during morning and peak hours, the effect will not only be reflected through a rise in the percentage of commuters who choose to take bus to work in their own district, but also an additional spill-over benefit which can be reflected in nearby areas. The elasticity of total effect of this variable is about 2.4, which indicates that increasing the transit frequency in area unit $i$ by one percent, the average bus mode share across all area units will rise by 2.4%, holding other variables constant.

Secondly, for the group of variables that have negative impact to bus mode shares, the parameter estimate on one of the urban form variables, $\ln(\text{Dwelling})$, indicates that the larger the share of private owner occupied dwellings within an area unit, the lower the share of commuters who take bus to work, which seems intuitively plausible. The estimated coefficient on the total effect of another urban form variable, the distance to the CBD, is negative and significant, suggesting that the propensity to take a bus decreases as the area unit is farther away from the CBD in the Auckland region.
For the two household characteristic variables, both the direct and indirect effects of income level exert a significant negative impact on the bus mode share, reflecting the idea that bus transport is an inferior good: as the commuters become wealthier, they will make fewer bus patronages for their JTW trips. Moreover, the direct effect of $\ln(Car)$ show that there is an inverse relationship between the number of cars owned per household and the bus usage rate, which is in line with the findings on car ownership variable found in Zhao et al. (2006) and Vance and Hedel (2007). The indirect effect of cars exhibits the same tendency, suggesting that with a one percent rise in the number of cars owned in adjacent area unit $j$, the average bus mode share in any given area unit $i$ tends to decline by approximately 23.7%. A possible explanation for this phenomenon is that increaser car use and/or access in neighbouring areas results in feedback forces and a “follow the neighbour” philosophy.

Thirdly, the estimated coefficient on the total effect of $\ln(Population\ Density)$ is negative and significant at the 1% level. This implies that the population density from all observed area units affects negatively the percentage of workers who take bus as their main transport to work, which runs counter to our original hypothesis that high population density leads to high transit ridership.

As discussed earlier, total effect can be unravelled into direct (own-region) and indirect (spatial spill-over) effect. Some notable findings were revealed by our results: two urban form variables: $\ln(Population\ Density)$ and $\ln(CBD)$ have the opposite signs for direct and indirect effect parameters; while the rest stays consistent. The estimated coefficient on the direct effect of $\ln(Population\ Density)$ is positive and significant at the 1% level. The result is consistent with the assumptions made by previous studies without consideration of spatial effects (Maat et al., 2005; Steiner, 1994), where people living in high-density sectors prefer to use more public transport or walk more frequently, but will make fewer and shorter trips by private vehicles. However, the indirect effect is negative and also significant, suggesting that once the population density in nearby regions increases, the
bus mode share in area unit $i$ will decline. This outcome may be due to the fact that commuters in area unit $i$ interpret the rise in population density in their neighbouring regions as a sign of potential congestion issues and/or dissatisfaction of the transit service, since buses might not be running on time, in such cases, taking private vehicles will be a better alternative than using public transport. Because the indirect effect is larger in magnitude, the total effect of $\ln(\text{Population Density})$ is negative.

For the next urban form variable $\ln(\text{CBD})$, the direct effect is negative and significant at the 1% level, suggesting that commuters are less willing to take the bus to work if they live farther away from Auckland city centre. While the spatial spill-over effect is positive and significant, suggesting that the bus mode share in region $i$ will tend to rise if commuters in nearby regions live further away from the CBD. The parameter estimated for the direct effect of distance to CBD is negative and this may reflect the less attractiveness of using public transit for commuters who live far away from the city centre, due to the fact there may be less public transport options available in their region. On the other hand, the positive indirect impact indicates that as a result of the spill-over effect, local commuters may find riding buses is a better option for longer trips, especially when workplace is far from the commuter’s residential address. Moreover, the variable $\ln(\text{CBD})$ may also serve as a proxy for the cost of public transport, therefore, those living further away may find it more cost effective to use their private vehicles to go to work. The sign of the total effect for this variable is negative because the magnitude of the negative direct effect outweighs the positive indirect effect.

Another significant finding from the SDM output is that except for $\ln(\text{CBD})$ and $\ln(\text{Income})$, the total effects are comprised mostly of the spatial spill-over impacts, and only a relatively small portion is attributed to the direct effects on bus mode share that arose from own-region changes in variable $x_k$. For instance, the indirect effect of $\ln(\text{Car})$ constitutes nearly 97% of the total impact of number of cars on bus mode share. Therefore, for the case of spatial dependence considered in the SDM model, least-squares regressions
that ignore this spatial spill-over effect and only produce the coefficients that represent the summary impact measures, result in biased and inconsistent estimates. The result also reveals that spatial spill-overs dominate in transit behaviour analysis and greater attention should be paid from transport and urban planners to both own-region effect and the impact from neighbourhood when evaluating new projects, or making transport investment decisions.

3.5. Conclusion, Limitation and Future Research

This chapter estimated how urban form variables are related to bus mode share and how these effects vary across the Auckland region’s landscapes. Overall, based on area unit level data, the analysis highlighted the complexity and importance of the spatial structure in determining the factors that influence the bus mode share.

The OLS method used in many previous transport-related studies assumes that the observations/regions are independent of one another in a geographical context. OLS thus looks for similarities in different spatial areas and provides an ‘average’ figure to cover the whole space. However, this is not plausible when using spatially-defined data because they are likely to exhibit positive spatial autocorrelation, that is, correlation of a variable with itself through space. Ignoring the spatial characteristics between observations/regions will, in turn, produce biased and inconsistent estimators.

By conducting an in-depth case study using the Auckland region’s data, urban forms, coupled with other factors that affect the bus mode share are explored and all these in turn are related under a spatial context. The Moran’s $I$ test shows that there is statistically significant evidence of the presence of positive spatial autocorrelation. Therefore, by taking spatial dependence into account, spatial regression models are selected over the non-spatial OLS model in order to obtain unbiased and consistent estimators. The
empirical results show that the bus mode share in one area unit exhibits a positive relationship with the share in neighbouring area units.

However, the interpretation of these findings based on SAR and/or SEM models is confounded by the strong spatial autocorrelation of the urban form and other transit characteristics such as transit supply and socioeconomic/demographic differences across area units. By applying the likelihood ratio tests, this chapter confirms the existence of spatial dependency in the lags of both dependent and independent variables. This dominating spatial issue has been addressed by the use of the spatial Durbin model. Estimated results from SDM show that the total effects comprised mostly of spatial spill-over impacts, and only a relatively small percentage is attributed to the direct effects on bus mode share that arose from own-region changes in any given explanatory variable. For planners and developers, the SDM model is not only technically superior, but also preferable for evaluating new projects and making investment decisions (Goetzke and Andrade, 2010). As unlike traditional estimated coefficient interpretations, one can easily unravel total effect into own-region and spatial spill-over effect. The results presented in this chapter indicate that knowledge of a specific spatial lag may provide clues about the importance of future land use patterns on transit ridership.

One limitation is that there is commonly an endogeneity issue with service frequency as an explanatory variable in a regression model with patronage as the dependent variable. Public transport providers often base the frequency they provide on patronage within an area unit; therefore, frequency may depend on patronage, since public transport providers gear service levels according to patronage. The variable $Freq$ in this SDM is thus suspected to be an endogenous variable to the bus mode share. One possible way to investigate the potential endogeneity issue here is to apply the Durbin-Wu-Hausman (DWH) endogeneity test, which is suggested by Davidson and MacKinnon (1993), to test the presence of endogeneity. The DWH endogeneity test requires the use of a valid instrumental variable (IV) to $Freq$. Ideally, this IV is assumed to be correlated with $Freq$.
but uncorrelated with $e_{int}$. However within this dataset, a valid IV variable is impossible to obtain. Another limitation is the lack of enough data to include other public transport modes (i.e. rail and ferry) in this study, unfortunately current dataset (NZ census 2006) is not comprehensive enough to take consideration of alternative transport modes other than buses, and our results might be sensitive to such inclusions.

This empirical analysis suggests several directions for future empirical research. First of all, consistent with Greer and van Campen (2011), this cross-sectional approach will not be sufficient to show the impact of other variables that do not vary across area units. With the presence of panel data, which has the ability to capture time trend, it would be necessary to investigate explicitly the dynamics of these other variables such as the fuel prices. Secondly, current models could be benefited from a more comprehensive dataset which includes more transport supply side variables such as seat capacity and labour/capital cost, when these information becomes available. Lastly, regarding model methodology, another approach, Bayesian estimation, which has the advantage of allowing comparison of various weight matrices based on Bayesian posterior model probability, could be applied in future research and compared with the maximum likelihood estimation.

---

27 Wang (2011) conducts a recent study that attempts to determine the factors that influence the demand for public transport in both short-term and long-run for three cities in New Zealand. Using time series data from the period 1996 to 2008, this paper highlights that considerable fluctuation in fuel price exhibits a positive effect on public transport patronage in all three cities.
Chapter 4. Commuter’s Transport Mode Preferences and Network Effects: A Spatially Autoregressive Discrete Choice Model

Abstract

Recently, public transit has been widely accepted as a socially less expensive alternative to resolve the economic and environmental issues that private transport has caused in many cities around the world. This chapter investigates commuters’ travel demand preferences in the Auckland region by choosing between public transport and private cars, for the period of 2005/06 – 2008/09. It extends the basic logit model by applying a spatially autoregressive logit mode choice model, which allows spatial dependence to vary across geographic regions. Empirical results show that positive network effects exist in transit use, thus the estimated coefficients of a standard logit model which does not reflect spatial dependence would become biased due to omission of relevant variables, leading to biased willingness to pay estimates. This chapter also highlights that all of the variables used in this study were highly significant with expected signs. Personal characteristics including household size, number of household vehicles, variations in household disposable incomes, gender roles, employment status, age differences, together with trip details: total travel distance, trip duration time and destination to Auckland city centre dummy all have significant impacts on influencing commuters’ transport mode preferences.

Key words: Network effects; transport mode preference; spatial weights matrix; spatial autoregressive discrete choice model
4.1. Introduction

4.1.1. Auckland public transport profile

The Auckland region is one of the sixteen regions in New Zealand. It comprises four cities: Auckland, North Shore, Manukau and Waitakere; and also includes three districts: Rodney, Papakura and Franklin. This region is located in the upper east portion of the North Island comprising a total land area of 16,315 square kilometres (km²), with urban areas approximately equal to 1,086 km². Auckland is the most populated region of the country, with more than 1.3 million inhabitants, which is equivalent to approximately 32% of New Zealand's entire population (Statistics New Zealand, 2012). As the most economically thriving and largest metropolitan area, Auckland employs over one-quarter of the workforce throughout New Zealand, based on 2006 Census data.

Auckland Transport is a council-controlled organisation (CCO) of Auckland Council in charge of all of the region’s transport infrastructure (excluding state highways and railways) and most of the region's public transport; namely, public bus, commuter rail and ferry service. It unified the transport functions of the eight former Auckland local authorities and the ARTA from 1st November 2010. According to Lambert (2011), Auckland’s buses carry more than 50 million passenger trips annually. The services are provided by eight main bus operators, namely, Airbus Express, Birkenhead Transport, Howick & Eastern Buses, NZ Bus (trading as Go West, Link, Metrolink, Northstar and Waka Pacific), Ritchies (which also operates Northern Express Services on the Northern Busway), Tranzit, Urban Express and Waiheke bus company (operated by Fullers Group). Around 55% of Auckland buses are built with low-floor, and new buses are being introduced on a regular basis in order to increase the number of buses with wheelchair access in service (Auckland Transport, n.d.).
Auckland’s rail system is operated by Veolia Transport, on behalf of Auckland Transport. The rail network ranges from Pukekohe in the south and Waitakere in the west to the central city in Auckland, with its hub at the Britomart Transport Centre. Currently, there are three rail services operated in Auckland: the Eastern Line (i.e. from Pukekohe to Britomart via Sylvia Park and Glen Innes), the Southern Line (i.e. from Pukekohe to Britomart via Ellerslie and Newmarket) and the Western Line (i.e. from Waitakere to Britomart). Auckland Transport owns the trains and stations and it is in charge of planning and developing train services, stations and vehicles. It coordinates with KiwiRail Infrastructure to provide plans and recommendations for the future developments of the rail corridor (ARTA, 2007).

The location of Downtown Ferry Terminal is in downtown, Auckland. It is located next to the Britomart Transport Centre. This geographically convenient location facilitates most ferry passengers as they can easily transfer to other public transport modes to various destinations across the region. At present, nine commuter ferry facilities are in service. They run between the Downtown Ferry Terminal and the other nine locations, namely, Bayswater, Birkenhead, Devonport, Gulf Harbour, Half Moon Bay, Pine Harbour, Stanley Bay, West Harbour and Waiheke Island (ARTA, 2007). Although ferry services have relatively minor catchments compared to the other two public transport modes, they still play an important role in the public transport system in Auckland: local residents often prefer taking ferries over buses because they cater to coastal communities. The majority of ferry services are operated by the following five firms: 360 Discovery, Belaire Ferries, Fullers Ferries, Pine Harbour Ferries, SeaLink. The responsibility for ferry wharves is shared between Auckland Transport and the Auckland Council. Auckland Transport has also proposed a new public transport network for Auckland, known as the Regional Public Transport Plan (RPTP). The RPTP details the public transport services and initiatives that Auckland Transport intends to offer for the region over the next decade.

---

28 Britomart Transport centre is the transport hub which connects bus, train and ferry services in downtown Auckland. The station is designed to serve up to 10,500 passengers during the peak periods in its current configuration as a terminus.
4.1.2. Econometric issue and objective

During the past few decades, spatial issues have gained an increasing amount of attention in the econometric studies. In contrast to the rapid development of these spatial literatures, limited efforts have been made to uncover how spatial effects influence the transport mode choice. Although the existence of such influences has been acknowledged and some researchers have made initial attempts to examine them both theoretically and empirically from various perspectives (e.g. Dugundi and Walker, 2005; Goetzke, 2008; Páez and Scott, 2007), this growing literature is still largely unresolved and the vision remains somewhat ambiguous. And importantly, to the best of our knowledge, empirical support in the New Zealand is very limited.

Goetzke (2003) argue that it is natural to envision that the knowledge of transport mode share in nearby regions augments information for determining transport mode share in a given study area, because transit patronage depends on both service attributes and accessibility, which are not expected to vary dramatically among neighbouring locations. In the field of transport mode choice studies, according to Goetzke (2008), “positive network effects exist when people prefer to use public transport together with other people as a result of social spill-over”. Therefore, there is a need to closely examine whether spatial autocorrelation or positive network effects exist among work trip commuters when analysing transport mode choice. In other words, an important question that one might ask is: what is the probability that a commuter will choose to use public transport to go to work given the transport mode preference of his/her neighbours and the characteristics of the regions where he/she lives?

The objective of this chapter is thus focused on the application of spatial analysis to transport mode choice, namely, the spatial autoregressive logit model, to understand preference heterogeneity and network effects in public transport among the journey to work (JTW) commuters in the Auckland region. With the availability of pooled cross-
sectional data, the modelling of transit behaviour change over time can also be incorporated into this spatial autoregressive logit model.

The rest of this study is divided into five sections. Section two outlines some existing literature on the concept of network effects. Section three discusses the spatially autoregressive discrete choice model employed in this empirical study. Section four describes the data sources and specifies the variables used. Section five presents some preliminary spatial econometric tests. Section six delivers the empirical results of the transport mode choice analysis. The study is concluded in section seven by highlighting major findings and proposing directions for future research.

4.2. The Concept of Social Network Effects

Generally speaking, social network effects, or simply, network effects, exist if a group of people prefer to do what other people already do. In the context of transportation networks, (Goetzke, 2006) summarises three main reasons why people prefer to use public transit together with others:

- First, there is a utility gain through the network effects coming from some kind of complementarity, since people are not alone. If they travel together, they can meet and communicate with each other, and thus feel safer.

- The second point is based on conformity and can thus be described as avoiding a utility loss by not following others because of the social norm, peer pressure and/or fashion.

- Third, there is a rise in utility level which stems from internalising an information externality, because people using a certain transportation mode (i.e. public transport) send a signal to everyone else that this is a feasible, and/or reliable mode.
The applications of network effects in transportation modelling started from Brock and Durlauf (2001, 2002), who were among the earliest studies to propose both binary and multinomial choice models which incorporate neighbourhood effects. However, as Goetzke and Andrade (2010) mention, the major drawback of their studies is that their approach is inherently non-spatial, while Leenders (2002) and Páez et al. (2008) explicitly point out that the topology of social interactions and neighbourhood impacts should be best captured spatially.

Because of their computational complexity, few spatially autoregressive discrete choice models (DCMs) have been implemented to date. In transportation modelling, Dugundji and Walker (2005), Goetzke and Andrade (2010) and Páez and Scott (2007) are the three most recent applications of spatially autoregressive DCMs. Dugundji and Walker (2005) describe and illustrate spatial interdependencies among transport decision-makers grouped by residential district, using a mixed generalised extreme value model for home to work trips. By adapting an instrumental variable approach, Goetzke and Andrade (2010) develop a spatially binomial mode choice model for home-based commuting trips to investigate the network effects for walking trips.29 Páez and Scott (2007), on the other hand, address the issue of inter-agent interactions to the case of telecommuting, based on spatial relations using a matrix derived from personal relations.

In the field of disaggregated transit use analysis at individual levels, Goetzke (2008) has produced the only published work which explicitly examines the impact of network effects on public versus private transport choice modelling for JTW commuting. Using cross-sectional data based on the 1997/98 regional household travel diary survey in the New York City area, Goetzke (2008) confirms the hypothesis that the network effects play a significant role in mode choice decision-making by urban commuters. By applying a spatial autoregressive logit model with a spatial weights matrix defined as $k$-nearest

---

29 Goetzke and Andrade (2010) limit the trip length to less than 2 miles, which is practical and meaningful for walk trips.
neighbours, the results show that total travel time, accessibility to cars, destination to Manhattan (a proxy to represent high parking cost in Manhattan) and a gender dummy have significant impacts on transit mode choice, while the high-income group variable unexpectedly turn out insignificant even at the 10% significance level. The author also highlights that due to statistical restrictions in the model (i.e. the restricted number of degrees of freedom), the network effects only represent a measure of commuter aggregate network preference (i.e. the network preference by a group of commuters), rather than their individual network effects (i.e. the network effect by an individual commuter).

This chapter builds upon and extends the approach by Goetzke (2008). The main objective is to explore the questions of what factors affect modal choice decisions in urban JTW travel, and to what extent the transport mode choice of one commuter influences the travel choice of his/her neighbours. By explicitly incorporating the spatial dimension using a row-standardised, inverse-distance spatial weights matrix, the empirical method, which is explained in detail in the next section, differs from the previously-used approaches outlined in section 2.1. To the best of our knowledge, this chapter is the first empirical work in New Zealand to explicitly analyse commuter’s transport mode choice behaviour in a spatial context. In addition, transit behaviour over time has never been modelled as most past studies were all purely based on purely cross-sectional data.

4.3. The Empirical Model

4.3.1. Standard logit transport mode choice model

The use of DCMs to analyse individual heterogeneity has a long history. As a part of the family of RUMs, they are used to predict choices an individual will make from a discrete set of alternatives using indirect utility functions which comprise observed (deterministic) and unobserved (stochastic) components (McFadden, 1974). Rational consumers are assumed to choose the alternatives that maximise their utility.
Beville (2009) claims that the variability in the level of attributes across choices and/or alternatives, either in the form of qualities or quantities, is required to estimate parameters who indicate the relative significance of those attributes in relation to the decision outcome. In such sense, the coefficients are interpreted as the individual’s preferences for constituent attributes.

In the context of transport mode choice, the utility $U$ of individual commuter $i = 1, \ldots, N$ for each alternative $j$ is a function of a vector of attributes (both personal characteristics of the commuter and the trip-specific characteristics) $x$ describing the alternative:

$$U_{ij} = \beta x_{ij} + \varepsilon_{ij}$$  \hspace{1cm} (4.1)

where:

- $U_{ij}$ denotes the utility of the $i^{th}$ commuter for alternative $j$;
- $\beta$ is the unknown vector of commuters’ preference to be estimated;
- $x_{ij}$ is the attribute vector of the $i^{th}$ commuter for alternative $j$; and
- $\varepsilon_{ij}$ is a random error component, representing the unobserved portion of utility.

Following Beville (2009), the probability $P$ of alternative $j$ being chosen by a particular individual commuter $i$ can be expressed in Equation (4.2), such that the probability that the utility of alternative $j$ exceeds the utility of all other alternatives $q$:

$$P_{ij} = \text{Prob} \left[ U_{ij} > U_{iq} \right] \hspace{1cm} \forall q \neq j$$  \hspace{1cm} (4.2)
Substituting *Equation (4.1)* into (4.2) gives:

\[
P_{ij} = \text{Prob} \left[ \beta x_{ij} + \epsilon_{ij} > \beta x_{iq} + \epsilon_{iq} \right] \quad \forall \; q \neq j
\]  

(4.3)

Rearranging *equation (4.3)* gives:

\[
P_{ij} = \text{Prob} \left[ \beta x_{ij} - \beta x_{iq} > \epsilon_{iq} - \epsilon_{ij} \right] \quad \forall \; q \neq j
\]  

(4.4)

where *Equation (4.4)* indicates that the probability that a particular individual commuter \( i \) chooses alternative \( j \) is thus equivalent to the greater probability that the difference in observed utility \( \beta x_{ij} - \beta x_{iq} \) than the difference in unobserved utility \( \epsilon_{iq} - \epsilon_{ij} \).

In addition, since data are not based on mode-specific utility levels but on actual transport mode choice decisions by each commuter, *Equation (4.4)* should be expressed in such a way so that the utility \( U_{ij} \) converts to an unobserved, latent variable. By assuming that \( \epsilon_{ij} \) is logistically distributed, a binary logit transport mode choice model can be derived, with \( P_i(j) \) represents the probability of an individual commuter \( i \) choosing mode \( j \) over alternative mode \( q \).

\[
P_i(j) = \frac{\exp(U_{ij})}{\exp(U_{ij})+\exp(U_{iq})}
\]

\[
= \frac{1}{[1+\exp(U_{ij}-U_{iq})]} \quad (4.5)
\]
Equation (4.1) can also be written in its reduced form:

\[ U_j = \mathbf{x}b + \epsilon \quad (4.6) \]

where \( \mathbf{x} \) is an \( n \times b \) matrix of personal/trip characteristic and \( \mathbf{b} \) is a vector of corresponding regression coefficients.

4.3.2. Spatial autoregressive logit transport mode choice model

Following Goetzke (2008), in order to incorporate network effects, Equation (4.6) is extended by adding a spatially autoregressive mode choice term \( \mathbf{Wm}_i \), so that the above random utility model is modified as:

\[ U_j = \mathbf{x}b + \mathbf{Wm}_i \rho + \epsilon_j \quad (4.7) \]

where \( \mathbf{W} \) represents the \( n \times n \) spatial weight matrix; \( \mathbf{m}_i \) is an \( n \times l \) vector of revealed mode choice decisions made by commuters; the spatial lag term, \( \mathbf{Wm}_i \), indicates the spatially weighted average mode share of mode \( j \) of all the commuters in surrounding locations of individual \( i \); and \( \rho \) is the spatial lag term’s regression coefficient to be estimated.

Goetzke (2008) states that the spatially autoregressive structure will become more perceptible when one considers \( \mathbf{m}_i \) as a function of the transport-mode-specific utility vector \( U_j, f(U_j) \). Therefore, by substituting \( f(U_j) \) in equation (4.7) we find:

\[ U_j = \mathbf{x}b + \mathbf{W}f(U_j) \rho + \epsilon_j \quad (4.8) \]
Equation (4.8) is referred as a conditional spatially autoregressive discrete choice model. Utility $U_j$ is conditional upon the observed/revealed neighbourhood’s mode choice decisions, and the assumption for the spatial lag term is exogenous (Anselin, 2002). It is obvious that as the neighbourhood’s mode share in mode $j$ increases, the mode-specific utility of a commuter to also choose this typical transport mode rises as well. Hence, a statistically significant positive value of the spatial autocorrelation parameter $\rho$ can be interpreted as the existence of network effects in this transport mode choice model. 30

In addition, Goetzke (2008) notes that the regression coefficient $\rho$ can also represent the slope of the upward-sloping, single trip willingness to pay (WTP) line with respect to travel mode share as a result of network effects, assuming that the utility derived from using the mode is an approximation of the WTP and that it is also correlated with the WTP for this particular transport mode. In such a way, the positive network effect is thus a positive externality, which has been generated by an additional public transport passenger on the road, as opposed to the negative externality, such as traffic congestion, which has been commonly noted in the transport literature. Figure 4.1 depicts this idea of positive network effects with mode share on the horizontal axis and WTP for a single trip on the vertical axis.

---

30 Note that $\rho$ refers to the aggregate network effects only in the sense that commuters have individual levels of $\rho$ which cannot be captured econometrically, because of the restricted number of degrees of freedom in this model.
4.4. Data Source and Variable Specification

The data employed for this spatially autoregressive DCM were obtained from the New Zealand Household Travel Survey (NZHTS), an official nationwide on-going household travel survey led by Research International on behalf of the MoT in New Zealand every financial year since 2003/04. More than 2,200 households participated in the survey over a two-day period each year around the country. The NZHTS provides a comprehensive source of data on the day-to-day travel behaviour in major cities of New Zealand. \(^{31}\)

Purely due to practical reasons, the full-response JTW data are limited to home-based work trips only (i.e., origin of the trip is the commuter’s residential address and destination of the trip is the commuter’s workplace). Therefore, only data from commuters

\(^{31}\) Between the period of 2003/04 and 2007/08 (inclusive), approximately 2,200 households were invited to participate every year to undertake the NZHTS; however from 2008/09 onwards, this number has increased more than a double - to approximately 4,600 households around the country.
who travel on weekdays, departing from their home to work in the Auckland region, are
used in this study (i.e. 76.8% of the data was removed). 32

Figure 4.2 displays a map with the household locations of all the included trips used in
this study.33 From this map, it can be seen that the household locations are not randomly
distributed. Households who reside in Auckland City, Manukau City, North Shore City
and Waitakere City tend to have greater densities, shown by the clustered red points, while
residential addresses in either the Rodney District, the Papakura District or the Franklin
District appear to be more spread-out. Therefore, by detecting geographical patterns, one
can conclude that the household location density is fairly heterogeneous and the sample is
not spatially random.

The selection of variables for this study is mostly guided by previous research and the
availability of data from NZHTS. This McFadden-type transport mode choice model
consists of the usual constraints affecting transport mode choices by commuters found in
the past literature (McFadden, 1974), and these explanatory variables can be further
divided into two subgroups: personal characteristics and trip details.

---

32 Workers are defined as people who work for full-time, part-time and casual works. The unit of
observation is defined as per trip not as per person.
33 This map is generated by ArcGIS10.0.
Figure 4.2: Distribution of All Surveyed Households in the Auckland Region
Buchanan et al. (2006) argue that the existing literature tends to suggest that people’s personal circumstances including their age, gender, household size, educational attainment and income either directly or indirectly influence travellers’ choice of mode for various trips. Therefore personal characteristics include:

- **HHsize**: household size, measured by the total number of eligible people in household during survey period; 34

- **HHvehicle**: number of household vehicles used by household and usually parked overnight. This variable includes both private and company-owned vehicles, and it serves as an indication of the level of accessibility to private vehicles;

- **Higher-income** = 1 if the commuter’s income is greater than NZD$30,000; 0 otherwise; 35

- **Female**: equals 1 if the commuter is a female; equals 0 if the commuter is a male;

- **Full-time** = 1 if the commuter is a working full-time; 0 otherwise;

- **Young** = 1 if the commuter is younger than 30 years old; 0 otherwise;

- **Old** = 1 if the commuter is older than 55 years old; 0 otherwise

Trip details include:

- **Distance** (in kilometres): trip distance in kilometres; 36

- **Duration** (in minutes): trip duration in minutes;

---

34 Eligible people in household refer to two categories: normal resident in the household and visitors who are living in the household on travel days.

35 This data is calculated based on median weekly income in the Auckland region. Retrieved from http://wdmzpub01.stats.govt.nz/wds/TableViewer/tableView.aspx?ReportName=Incomes/Income by region and sex

36 The trip distance is defined as the best available distance. For instance, trip distances for automobile trips are calculated based on the addresses of origin and destination provided by the participants. This variable is populated for driver, passenger, walk, cycle, bus, train, taxi trips. For train trips, Distance = distance along rail line from station to station.
• $\text{Destination}_\text{AKL} = 1$ if the destination of this trip was located in the city centre of Auckland; 0 otherwise.\(^{37}\)

Moreover, in order to capture the potential year effect from this pooled cross-sectional dataset, four time-effect dummies: \(\text{year} = 1, 2, \ldots, 4\) were used to denote the year effect from the period 2005/06 to 2008/09.

Table 4.1 provides some descriptive statistics of the variables used in this chapter. The total number of observations is 820 trips. The dependent variable \textit{choice} is equal to one if the commuter self-reported that his/her main transport mode to work is local public transport (i.e. bus/ train/ferry) during the survey; and it is equal to zero if this commuter was the car/van driver during JTW. Additionally, local public transport is defined as bus or train trips that are equal to or less than 60 kilometres long, and for ferry trips that are equal to or less than one hour long. As shown in Table 4.1, the usage of public transport is fairly low; only about 7.6% of the people included in this dataset take public transport to work, while the majority of working trips in the Auckland region rely heavily on household vehicles. Additionally, Appendix A presents the descriptive statistics separately for personal and trip characteristics for the subgroup when \textit{choice} = 1.

\(^{37}\) This variable serves as proximity to reflect the parking cost in the Auckland city as there is no information regarding actual parking cost in this area.
Table 4.1: Summary Statistics of Explanatory Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choice</td>
<td>0.076</td>
<td>0.265</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHsize</td>
<td>3.189</td>
<td>1.472</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>HHvehicle</td>
<td>2.362</td>
<td>0.944</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Higher-income</td>
<td>0.455</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Female</td>
<td>0.407</td>
<td>0.492</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Full-time</td>
<td>0.965</td>
<td>0.185</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Young</td>
<td>0.221</td>
<td>0.415</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Old</td>
<td>0.137</td>
<td>0.344</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Trip characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance (kilometres)</td>
<td>12.560</td>
<td>11.337</td>
<td>0</td>
<td>108.716</td>
</tr>
<tr>
<td>Duration (minutes)</td>
<td>23.606</td>
<td>14.674</td>
<td>1</td>
<td>98</td>
</tr>
<tr>
<td>Destination_AKL</td>
<td>0.223</td>
<td>0.417</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

For the explanatory variables that describe the personal characteristics, first of all, the average household size is around three people. However, it is notable to see that the number of private and company-owned vehicles in an average household is slightly above two. This figure implies that accessibility to cars is reasonably high in the Auckland Region, where one person can have access to 1.35 vehicles. The Higher-income variable shows that roughly 46% of the households make more than NZD$30,000 annually. Regarding the workers’ gender, female workers occupy about 41% of the whole sample size. For the workforce’s age distribution, as expected, the majority of the workers are
middle-aged, with only 22% of the workforce are younger than 30 years old, and approximately 14% of the workforce are older than 55 years old.

In addition to personal characteristics variables, three regressors relate to trip characteristics, they are: Distance, Duration and Destination_AKL. The mean value of Distance shows that work travel distances are averaging slightly more than 12.5 kilometres for a one-way trip and with the maximum that is nearly 109 kilometres. The mean value of Duration is around 23.6, which implies that the average trip duration is approximately 23.6 minutes and the maximum value is 98 minutes. The average travel time in Auckland is similar to the one found in the U.S., where in year 2000, the average JTW travel time for Americans is 25 minutes and 30 seconds (McGuckin and Srinivaasan, 2005). In the model with network effects, as explained in Goetzke (2008), total travel time is the main variable which justifies the supply-side of transportation. More importantly, this variable also signifies an aggregation of physical infrastructure, which mainly accounts for transit service frequency for public transport and highway congestion for private vehicles.

According to Commins and Nolan (2011), inner-city workers will generally have more flexible public transport options, coupled with poorer availability of parking choices at workplaces than those who are working in suburban locations. Therefore, a destination dummy, Destination_AKL, is included in this transport mode choice model as it is expected that more people are willing to take public transport instead of cars as their main transport mode to work if their workplace is located in central Auckland. Summary statistics of this variable show that among seven surveyed cities/districts, slightly more than 22% of the workforce is located in Auckland city. A map that illustrates the geocoded destination locations of all trips is visualised in Appendix B.
4.5. Preliminary Tests

4.5.1. Spatial weights matrix and Moran’s I statistics

To account for the heterogeneous household location density observed from figure 4.4, a row-standardised, inverse-distance spatial weights matrix $W$, with a dimension of $n \times n$, is developed using the geographic coordinates of each observation. For all cases, $w_{xy} = 0$ for $x = y$, and non-zero elements indicate neighbours. In addition, this specification of the spatial weights matrix implies that as the distance between household $x$ and $y$ increases (declines), $w_{xy}$ declines (increases), hence placing less (more) spatial weight to the household pair when $x \neq y$ (Coughlin et al., 2004).

Moran’s $I$ serves as an index for dispersion/random/cluster patterns in spatial econometric analysis. A positive Moran’s $I$ statistic indicates a tendency toward clustering, in other words, this positive value signals positive spatial autocorrelation; and vice versa. In line with expectations, Moran’s $I$ statistics under two different formulations for the standard error calculation: normal approximation (or free sampling method) and randomisation (non-free sampling method), both of which are positive and highly significant as shown in Table 4.2 which confirms the fact that positive spatial autocorrelation exists.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Normal Approximation</th>
<th>Randomisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moran’s $I$</td>
<td>0.0837</td>
<td>0.0837</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0012</td>
<td>-0.0012</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0142</td>
<td>0.0141</td>
</tr>
<tr>
<td>Z-score</td>
<td>5.9915</td>
<td>6.0223</td>
</tr>
<tr>
<td>$P$-value*</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

*: Two-tailed test

Note: Under the random permutation procedure: Mean = -0.0011 and Standard deviation = 0.0141
4.5.2. Likelihood ratios test for network effects

The objective of this chapter is to examine whether network effects exist in terms of public transport use. Therefore, the following hypotheses will be tested explicitly:

$$H_0: \rho = 0 \text{ vs. } H_A: \rho \neq 0,$$

which is equivalent to test that:

$$H_0: \text{WTP}_{\text{standard logit}} = \text{WTP}_{\text{spatial autoregressive logit}} \text{ vs. } H_A: \text{WTP}_{\text{standard logit}} \neq \text{WTP}_{\text{spatial autoregressive logit}}$$

If we reject the second null hypothesis, we can conclude that bias in the estimated coefficients has economic significance and that WTP, as estimated by the spatial autoregressive logit model, is statistically different from WTP estimated by the standard logit model. To test the specification of the model, a likelihood ratio (LR) test is applied to test the null hypothesis that the network effects parameter \( \rho \) is equal zero. Let \( L_{\text{standard}} \) be the value of the log-likelihood function of the restricted model (i.e. the standard logit model) and \( L_{\text{spatial}} \) be the value of the log-likelihood function of the unrestricted mode (i.e. the spatial autoregressive logit model). The LR-test statistic is given as:

$$LR = -2(\ln (L_{\text{standard}}) - \ln (L_{\text{spatial}})),$$

which has an asymptotic \( \chi^2 (k-1) \) distribution.

The following results have been obtained after applying likelihood ratios (LR) test, where:

$$L_{\text{standard}} = -342607.33 \text{ and } L_{\text{spatial}} = -337489.19$$
With 13 degrees of freedom, the critical values at 1%, 5% and 10% significance are 4.107, 5.892 and 7.042, respectively.

Obviously, the test statistics exceed the critical values for all cases, therefore we can reject the null hypothesis that the network effects parameter $\rho$ equals zero at a 1% significance level. Or, in other words, the null hypothesis that the WTP from a standard logit model is statistically the same as the WTP in a spatial autoregressive logit model should be rejected. Therefore, one can conclude that the restriction on the network effects variable is invalid. More importantly, the above results show that network effects exist within the data and the estimated coefficients of a standard logit model which do not reflect spatial dependence will become biased, leading to biased estimates of WTP.

4.6. Estimation Results

4.6.1. Comparison between standard logit model and spatial logit model

The estimation results from both standard and spatial logit models of the commuters’ travel mode choice in the Auckland region are shown in table 4.3.\textsuperscript{38} Standard errors are presented in parentheses and the level of statistical significances is marked by asterisks.

The positive and highly significant estimated value of the spatial autoregressive term, $\rho$, gives us the same indication as the result obtained from the LR test. This positive and statistically significant coefficient for the spatial lag variable delivers two significant implications: first, network effects do exist within the data, and, second, the probability that a commuter chooses to take public transport to go to work increases when his/her

\textsuperscript{38} Both standard and spatial logit regression results are adjusted by \textit{pweight} (i.e. person weight), a weighting scheme provided by MoT. As both regressions use multi-year data set, this will give total people across all years, not in per year terms. For example, using a 4 year dataset will give us 4 times national people.
neighbours have a high propensity to do so. In other words, this spatial autoregressive logit model shows that taking public transport to work does not only depend on personal and trip specific characteristics, but also on the spatial interactive variable. The probability of using public transport increases with a higher public transport mode share, as a result of network effects derived from social interactions between neighbours. This empirical result is in line with the transport mode choice analysis from Goetzke (2008), where the author found strong evidence that the transport mode choice decision of a particular individual was associated with the mode choice decisions of spatially connected neighbours.

Given the non-linear nature of this spatial autoregressive logit model, the spatial lag term has a value that is larger than one (i.e. $\rho=2.9998$). However, if we consider a linear form of this spatial autoregressive logit model (i.e. a spatial linear probability regression model), we can confirm that the value of the spatial lag term is right within the range of zero and one. Although the estimated value of 0.3080 from the alternative spatial linear probability regression model seems more intuitive, one should be careful when interpreting this result as some major inherent problems are associated with this kind of regression modelling, frequently known as: (i) heteroscedastic disturbances, (ii) non-linear distribution of the error term (Goldberger, 1964), and the most severe one, (iii) the computed probabilities lying outside the interval (0, 1) of the binary dependent variable (Vogelvang, 2005).

In addition, all of the variables in the standard logit model have the same signs and significance levels as the ones in its spatial version. In terms of magnitude, both models produce quite similar figures, therefore the outcomes from the spatial model that discloses positive aggregate network effects for public transport use remain very robust. As measured by the Pseudo-$R^2$, both the standard and spatial models have a reasonably good fit for a logit model, with 0.2064 for the standard logit model and 0.2182 for the spatial logit model. Nevertheless, the latter performs marginally better.
Table 4.3: Standard and Spatial Logit Models of Commuter’s Travel Mode Choice

<table>
<thead>
<tr>
<th>Variables</th>
<th>Standard Logit</th>
<th>Spatial Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transit mode constant</td>
<td>-2.6983 ***</td>
<td>-3.1985 ***</td>
</tr>
<tr>
<td>(PT=1; drive alone=0)</td>
<td>(0.0232)</td>
<td>(0.0237)</td>
</tr>
<tr>
<td>HHsize</td>
<td>0.1580 ***</td>
<td>0.1881 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0023)</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>HHvehicle</td>
<td>-0.7867 ***</td>
<td>-0.7519 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0039)</td>
<td>(0.0038)</td>
</tr>
<tr>
<td>Higher-income</td>
<td>-0.1486 ***</td>
<td>-0.0947 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0070)</td>
<td>(0.0070)</td>
</tr>
<tr>
<td>Female</td>
<td>0.7586 ***</td>
<td>0.9568 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0071)</td>
</tr>
<tr>
<td>Full-time</td>
<td>0.6934 ***</td>
<td>0.7597 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0189)</td>
<td>(0.0189)</td>
</tr>
<tr>
<td>Young</td>
<td>0.4906 ***</td>
<td>0.4115 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0070)</td>
<td>(0.0071)</td>
</tr>
<tr>
<td>Old</td>
<td>-1.0446 ***</td>
<td>-0.8366 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0173)</td>
<td>(0.0173)</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.1410 ***</td>
<td>-0.1444 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Duration</td>
<td>0.0527 ***</td>
<td>0.0550 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Destination_AKL</td>
<td>0.5771 ***</td>
<td>0.3486 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
<td>(0.0071)</td>
</tr>
<tr>
<td>Year2</td>
<td>1.0131 ***</td>
<td>0.9179 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0101)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>Year3</td>
<td>0.9578 ***</td>
<td>0.7530 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0096)</td>
<td>(0.0098)</td>
</tr>
<tr>
<td>Year4</td>
<td>0.3476 ***</td>
<td>0.2747 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0103)</td>
<td>(0.0104)</td>
</tr>
<tr>
<td>ρ</td>
<td>2.9998 ***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0287)</td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td>820</td>
<td>820</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-342607.33</td>
<td>-337489.19</td>
</tr>
<tr>
<td>Pseudo-R²</td>
<td>0.2064</td>
<td>0.2182</td>
</tr>
</tbody>
</table>

*** Estimated coefficients significant at 1% level; standard errors are in parenthesis
### 4.6.2. Marginal effects of the spatial logit model

As coefficients are difficult to interpret in the logit models, marginal effects of the spatial logit model are reported in Table 4.4.\(^{39}\) One should note that all of the estimated variables are significant at 1% level.

The regression coefficient for \(HHsize\) is positive; this implies that as more people live in a household, the probability that they take public transport to work will increase, on average. Quite logically, if a person prefers to use public transport as opposed to a car for his/her JTW trips, the others in the same household might do the same as they interpret using this particular means of transportation as a “social norm”. Also, there are more demands on the family car, leaving it less available for commuting to work. However in the U.S., empirical results from Mannering and Winston (1985) reveal the opposite: their results show that people who came from larger families tend to have a greater probability of using private vehicles for JTW trips.

The estimated coefficient for \(HHvehicle\) is -0.0555, indicating that if there is an additional household vehicles available to use, people who reside in the household will be about 5.55% less likely to take public transport to work. This result confirms that accessibility or availability to cars is a significant determinant when analysing transport mode choice behaviour in the Auckland region. Researchers have also identified similar patterns in some Asian cities that vehicle ownership is an essential component in the process of making travel-related decisions (Hsu \textit{et al.}, 2007; Tuan and Shimizu, 2005; Senbil \textit{et al.}, 2007).

\(^{39}\)The marginal effects of the spatial logit model show the change in probability when the independent variable increases by one unit, and it is the average marginal effect across the sample. For binary variables, the change is from 0 to 1. The average marginal effects are calculated using Stata command: “margins, dydx(explanatory variables)”.

Table 4.4: Average Marginal Effects of the Spatial Logit Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Spatial Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHsize</td>
<td>0.0139 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
</tr>
<tr>
<td>HHvehicle</td>
<td>-0.0555 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Higher-income</td>
<td>-0.0070 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Female</td>
<td>0.0707 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Full-time</td>
<td>0.0561 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
</tr>
<tr>
<td>Young</td>
<td>0.0304 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Old</td>
<td>-0.0618 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0013)</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.0107 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Duration</td>
<td>0.0041 ***</td>
</tr>
<tr>
<td></td>
<td>(0.00002)</td>
</tr>
<tr>
<td>Destination_AKL</td>
<td>0.0257 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
</tr>
<tr>
<td>Year2</td>
<td>0.0678 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
</tr>
<tr>
<td>Year3</td>
<td>0.0556 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
</tr>
<tr>
<td>Year4</td>
<td>0.0203 ***</td>
</tr>
<tr>
<td></td>
<td>(0.00008)</td>
</tr>
<tr>
<td>( \rho )</td>
<td>0.2216 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0021)</td>
</tr>
<tr>
<td>Sample Size</td>
<td>820</td>
</tr>
</tbody>
</table>

*** Estimated coefficients significant at 1% level; Delta-method standard errors are in parenthesis

Moreover, unlike the insignificance of the high-income dummy revealed in Goetzke (2008), the Higher-income coefficient, -0.0070, is negative and significant. The probability of using public transport for those commuters who earn above the average
wages is 0.7% lower compared to their low-income counterpart; higher-income workers favour automobiles rather than public transport simply because household vehicle is a normal good and public transport is an inferior good.

Furthermore, the positive sign of the estimated coefficient (i.e. 0.0707) on gender indicates that female workers have a greater probability of using public transport for their JTW trips compared to males. They are about 7% more likely to travel by public transport to go to work compared to men. This moderate gender difference towards transport mode choice is supported by Matthies et al. (2002). Based on a survey among 187 inhabitants in Bochum, Germany, this study found that women were more willing than men to increase public transit use and reduce automobile use due to their stronger ecological norms as well as weaker car habits.

Consistent with O’Fallon et al. (2004), we found full-time status to influence transport mode choice, given that the estimated coefficient on this variable is significant and has a positive value of 0.0561. Commuters who work full-time have a higher probability of public transport use: full-time workers are 5.6% more likely to use public transport for work trips compared to part-time casual staff. This might be due to the reason that public transport is likely for part-time workers because they tend to have non-standard hours and therefore public transport may be less likely to cater to these hours.

In regards to the two age variables, younger workers who are 30 years old or under are around 3% more likely to use public transit as their main transport mode to work compared to the reference category of 30 to 55 years-olds; while by contrast, senior workers who are over 55 years old are inclined to make approximately 6% less public transit JTW trips compared to the same reference group. This result differs from Palma and Rochat (2000), who found in the case of Geneva, the opposite outcome: younger commuters whose age was less than 30 years old were more likely to use cars for work
trips, whereas senior people who were older than 50 years old tended to be more committed to public transit. It also differs from Goetzke and Andrade (2010) in their walkability research in the U.S.. Using a spatially autoregressive mode choice model, they find that both age variables are not statistically significant even at the 10% significance level, which implies that age difference does not have an impact on commuter’s choice for selecting “active mode” versus others.

The longer the trip is, the lower the probability for a commuter to take public transport to work, as indicated by the negative sign of the estimated coefficient on Distance, -0.0107. This figure implies that, if the trip length increases by one kilometre, the probability that a person to take public transport to work will reduce by roughly 1.1 %. With the stage-based fare system used in the Auckland region, the longer the trip is, the more expensive is the fare. Commuters are more likely to use their household vehicles for longer trips because they may find automobiles as a more convenient and comfortable transport mode. This result is in line with O’Fallon et al. (2004).

It is also interesting to note that the estimated coefficient of Duration is positive (i.e. 0.0041), which suggests that the indirect utility of commuters derived from using public transit increases by 0.4% with one more minute of in-vehicle time. This outcome, however, contradicts what has been observed in Goetzke (2008), where the regression coefficient of the total travel time variable is negative.

Additionally, since parking costs are much higher in central Auckland area compared to other places, the positive estimate of the last trip characteristic variable Destination_AKL, 0.0257, intuitively makes sense. Inner-city workers are more inclined to take public transit to work in order to avoid excessive parking costs. Commuters whose workplace is situated in the central Auckland area are about 2.6% more inclined to travel by public transport compared to those who work anywhere else in the Auckland region.
Consistent with prior expectations, the estimated coefficients for the individual year-effect dummies, \textit{year2}, \textit{year3} and \textit{year4} are significant and positive. This implies that by comparing to the base year 2005/06, the probability of using public transport for JTW trips improves with time. This outcome can be explained by the boosted fuel prices during this period and the investments on public transit infrastructure made by the Auckland City Council such as the Northern Busway which was completed in February 2008. \textit{Figure 4.3} shows that the public transport boardings in Auckland have increased considerably since mid-late 2005 with a steady increase in the average price of regular petrol.

\textbf{Figure 4.3: Public Transport Boardings and Regular Petrol Prices in Auckland, 2005Q3-2015Q3}

Data source: 1) Public transport boardings: Auckland Transport 2) Average regular petrol price, MBIE
Finally, consistent with Goetzke (2008), not including the network effects variable in transport mode choice analysis also leads to omitted variable bias as is evident from the LR-test discussed previously. By assuming that the relative utility for taking public transport in the suburban areas is lower than average and, in Auckland central city, higher than average, public transport mode share would generally be overestimated in the suburban areas but underestimated in Auckland central city. Figure 4.4 provides a graphical illustration of this situation. In the presence of network effects, public transport utility increases with higher public transport mode share, which is indicated by an upward-sloping line $v_1$. However, if we do not consider the network effects, public transport utility is forced to be constant as represented by a flat line $v_2$ in the graph. Therefore, it is evident to conclude that the non-inclusion of the variable representing network effects will lead to an omitted variable bias in the regression analysis.

**Figure 4.4: The Link between Public Transport Utility and Network Effects**

![Graph depicting the link between public transport utility and network effects](image-url)
4.7. Conclusion, Limitations and Future Research

The main objective of this chapter is to use spatial econometrics approach to analyse individual traveller’s transport mode choice, and to understand preference heterogeneity and network effects in public transport among the journey to work (JTW) commuters in the Auckland region. The spatial autoregressive logit model provided strong empirical evidence that transport mode choice models do suffer from spatial autocorrelation. Ignoring this spatial feature between observations will, in turn, produce biased estimations. Some key findings are summarised below.

First, the spatial sampling issue was examined by developing a row-standardised, inverse-distance spatial weights matrix $W$, with a dimension of $n \times n$, using the geographic coordinates of each observation. As the first step to detect spatial autocorrelation, Moran’s $I$ statistics were computed after the construction of spatial weights matrix under two different formulations for the standard error calculation. Results showed that these statistics were both positive and highly significant, with p-values that are equal to zero, which confirmed the existence of positive spatial autocorrelation.

Second, a likelihood ratios test was applied to test the null hypothesis that the network effects parameter $\rho$ was equal to zero. The LR-test statistic revealed that the spatial autocorrelation parameter $\rho$ was positive and significant, showing that taking public transport exhibited network effects among nearby households. In other words, in the presence of network effects, the process of transport mode choice made by individual commuters does not only involve personal characteristics, transport-mode-specific information, but also the information on transport mode decisions of individual commuter’s neighbours. In addition, this result also indicated that the estimated coefficients of a standard logit model which did not reflect spatial dependence would become biased due to omission of relevant variables, leading to biased WTP estimates.
Third, all of the variables in the spatially autoregressive logit model were found to be highly significant with expected signs. Household size, number of household vehicles, variations in household disposable incomes, gender roles, employment status, age differences, along with total travel distance, trip duration time and destination to Auckland city centre dummy all have significant impacts on influencing commuter’s transport mode choices.

Fourth, given the fact that the process of transport mode choice decision-making is dependent on social network effects, urban and transportation planners should focus not only on infrastructure improvements, but also on strengthening the city’s “green” transport mode culture.

One potential limitation of this chapter is that it is possible that an omitted variable or variables pertaining to service quality that are spatially autocorrelated may account for the high estimated value of $\rho$. For instance, suppose that commuter A lives near a high-frequency bus route. He/she is thus more likely to take public transport than a commuter who does not live near a high-frequency route. Suppose that commuter B lives nearby commuter A. Commuter B therefore also lives near a high-frequency bus route. It might be the case that commuter B’s high probability of taking public transport stems from the fact that he/she lives near the high-frequency route, rather than because of a social network effect. In order to overcome this possible drawback, more information on the supply side of public transport is highly desirable.

For future studies, a more comprehensive survey with additional information on the supply side of transport, such as preferences for flexibility, comfort, and infrastructure quality, would be desirable. With the availability of panel data which describes observations on multiple commuters through multiple years, it will be much easier to investigate explicitly the dynamics of other variables, such as total travel cost, which were
previously identified as important factors in determining transport mode choice in past research. Improvements in data will produce even more robust result.

Moreover, basic transport economics theory states that the demand for movement between particular locations depends both on trip generations in origin and trip attractions in destination. Therefore, as mentioned in Goetzke (2008), another possibility for future research would be to extend the transport mode choice model to embrace network effects on both the origin and destination side of the trip. Intuitively this makes sense, since commuters are more inclined to choose their mode based on both characteristics from origin and destination.

Lastly, this spatial analysis of transport mode use can be applied to examine the travel behaviour of different trip purposes in the Auckland region such as non-work trips; for example, leisure trips or shopping and recreational trips. Or it can be used to analyse the travel behaviour of different travel groups such as school and after-school trips for teenagers, and public transit trips for older and disabled people.
Chapter 5. Aggregate Road Passenger Travel Demand in New Zealand: A Seemingly Unrelated Regression Model

Abstract

Road passenger transportation, which includes both private vehicles and public transport, is regarded a vital link that connects people and economic activities across New Zealand. Although a wealth of past literature examined the demand for private and public transport both individually and jointly worldwide, little evidence was found analysing the demand for different road passenger transport choices as a system of equations. Given the fact that these road passenger transport modes are considered substitutes to one another, there is a strong possibility that an interrelationship exists between the travel demand functions, primarily due to the correlation between their disturbances, a research gap that was thus discussed and addressed in this study. This paper is the first use of seemingly unrelated regression (SUR) method in New Zealand to develop an aggregate road passenger travel demand model. The Breusch-Pagan test of independence confirms the existence of correlated error terms in the three demand equations. The empirical results from SUR model indicated that the values of income elasticity for petrol and diesel cars were 3.99 and 3.45 respectively. However per-capita GDP was found to be an irrelevant variable in determining the demand for public transport. Price elasticities for petrol cars and buses were -0.61 and -0.26 respectively, although for the demand of diesel cars, its own price elasticity was not found to be a significant predictor. The cross price elasticity between diesel cars and petrol cars was 0.44, and between diesel cars and buses was 0.24. Inflation and unemployment were found to be significant only in the demand for diesel cars equation. The age indicators showed a similar trend in the private vehicle demand equations. However for the demand for public transport, the sign on the estimated coefficients for all age groups was positive, suggesting that all New Zealanders are inclined to use public transport as their road transport choice. The above empirical results

---

40In terms of public transport, buses are the most common form of public transport mode in New Zealand as all cities and most towns have bus services available. Rail is excluded from this category as this service is only available in Auckland and Wellington, not across the whole country. Therefore the public transport mode for road passengers only refers to buses in this paper.
from the SUR model also delivers various policy implications in terms of achieving a reduction in the demand for both petrol and diesel cars, and also promoting the use of public transport.

**Key words:** Aggregate road passenger transport demand, car ownership, public transport demand, seemingly unrelated regression, correlated disturbances
5.1. Introduction

Transport is a main component of economic development, both as a sector in itself and as an important factor input to most other economic activities (OECD, 2001). Transport, and its associated infrastructure, therefore, has always been playing a key role in New Zealand's economic prosperity. Based on a report from MfE (2009b), compared to other transport modes, road transport dominates New Zealand’s travel pattern. Within the road transport, according to the MoT (2014), in 2013, light fleet which consists of light passenger vehicles (LPVs)\textsuperscript{41} and light commercial vehicles (LCVs) accounts for most of the travel on New Zealand roads, where LPVs alone contributed over three-quarters of road travel, LCVs a further 16%, and only 8% of road travel was by other types of vehicles. While road transport does provide numerous economic and social benefits, it also generates several negative externalities that have major adverse impacts on our health and environment. For instance, road transport is the primary cause of harmful air pollutants in some urban areas where road traffic and congestion are concentrated. Rivers and streams can be polluted by contaminated run-off from arterial roads and highways, and vehicle wastes such as used batteries and tyres present significant management issues as these require careful disposal.

At the national level, our use of road transport is escalating. In terms of private vehicles, as MfE (2009b) warns, on average, New Zealanders are driving further, owning more cars, choosing an increasing engine size; and the fleet profile is older. Figure 5.1 shows motor vehicle and passenger car ownership among 34 OECD members. New Zealand has the highest motor vehicle ownership (motor vehicle/population ratio) and the fourth highest passenger car ownership (passenger car/population ratio) compared to the other OECD members (OECD, 2013). This high level of automobile dominance in New Zealand is at least in part a result of past government transport policies which makes cars as the “default” form of personal transport for New Zealanders, including the development of automobile-oriented urban forms and highway improvements in 1950s that have had the

\textsuperscript{41} The MoT defines cars as LPVs. LPVs are passenger vehicles that weigh less than 3,500 kg. This group also includes passenger vans, and sport utility vehicles (SUVs).
effect of encouraging car travel, and the deregulation of the vehicle industry from 1980s which removed import quotas and reduced import tariffs on vehicle imports from overseas, making imported cars more affordable for domestic consumers.

![Motor Vehicle Ownership among OECD members in 2010](image)

Data source: OECD (2013)

Although public transport patronage has been growing in recent years and have gained greater social acceptability and aid in the promotion of environmentally friendly lifestyles, the percentage of people travelling by public transport still remains relatively low compared to car trips. Thus, we can conclude that New Zealanders rely primarily on private vehicles for travel, supplemented with public transportation. Unsurprisingly, due to this car-dependence, energy consumption has increased, congestion on local roads and motorways worsens, and CO₂ emission from the road transport sector increases. Globally, the transportation sector is the most rapidly growing sector in terms of energy and particularly fuel consumption, and it is responsible for a substantial share of the global fossil fuel combustion-related CO₂ emission (Dargay and Gately, 1997). Available

---

Note: 2009 data for Canada and Ireland; 2011 data for Australia, Iceland, Japan, Mexico, New Zealand and Switzerland. 2010 data for the other countries. The OECD totals are based on OECD Secretariat's estimates.
statistics indicates that in 2005, CO\textsubscript{2} emissions from transport sector represent 30% of overall CO\textsubscript{2} emissions from fossil fuel combustion in OECD countries. And on a worldwide base, CO\textsubscript{2} emissions from transport sector also experienced an incredibly fast growth rate of 45% from 1990 to 2007 (OECD, 2010). In New Zealand, according to the TomTom traffic index (TomTom, 2015), Auckland and Wellington are considered to be the second and third most congested cities among Australasian metropolitan areas, just behind Sydney. In dollar terms, Wallis and Lupton (2013) estimated that congestion cost the Auckland region 1.25 billion NZD annually when compared to free flow conditions.

Although the total amount of CO\textsubscript{2} emissions that New Zealand emits is relatively small, in terms of per capita measure, it has the OECD’s fourth highest road transport emissions per capita, just behind the United States, Canada and Australia, in 2010. At the same time, not surprisingly, New Zealand is expecting considerable population growth through an increase in natural population and immigration. By 2031, the total population in NZ is forecast to grow to approximately 5.2 million residents (Statistics New Zealand, 2012). This high level of motor vehicle ownership, coupled with the anticipated huge boost in population, will definitely put greater pressure on the New Zealand road transport system and its associated infrastructure.

While significant research has been devoted to the total demand for cars, little research has been done on the demand for cars of different fuel types (i.e. petrol and diesel), a classification that is exceedingly useful when examining road user’s transport mode preference in the light of energy consumption, CO\textsubscript{2} emissions and policy implications. In addition, the demand for public transport has never been modelled as a group along with its major substitute. The purpose of this study, therefore, is to fill the empirical gap in the past literature by developing a multi-equation model of the road passenger transport sector, by proposing and answering the following three key research questions:
1. Whether the disturbances of the demand for the three main road passenger transport choices, namely, petrol cars, diesel cars and buses are correlated?

2. What are the factors affect the demand for road passenger travel in New Zealand?

3. And by how much will these factors impact on the demand for road passenger travel in New Zealand?

To the best of our knowledge, there is no past paper that has explicitly provided satisfactory responses to all of the above questions. The present study is thus the first to address the empirical problems identified above by applying a seemingly unrelated regression (SUR) model, using quarterly time series data in New Zealand from 2001Q1 to 2015Q2. The advantage of the SUR model is that dependent variables which represent aggregate road passenger travel demand (i.e. petrol cars, diesel cars, and buses) are allowed to be considered as a group when they bear a close conceptual relationship to one another, brought forward by the potential correlation between their error terms. The objective of this chapter is thus to develop an aggregate travel demand model of New Zealand’s major road passenger transport modes, by taking into detailed account the effects of correlation between their error terms. The plan of the chapter is as follows. The next section provides background information about overall vehicle fleet composition, light vehicle ownership and some facts about overall car travel in New Zealand. Section three discusses some findings from past literature. Section four specifies the model and section five outlines the data used in this study. Section six presents the results of stationarity test on all of the variables. Section seven shows the estimated results. Section eight presents the test results for cointegration and structural breaks. The last section concludes and suggests some possibilities for future research.
5.2 Background Information

5.2.1. New Zealand Vehicle Fleet Composition

In New Zealand the overall vehicle fleet is made of light passenger vehicles, light commercial vehicles, motorcycles, trucks and buses. Figure 5.2 below shows that the light passenger vehicle is the largest segment among the fleet profile, it alone constitutes nearly 80% of the total vehicle fleet.

Figure 5.2: New Zealand Vehicle Fleet Composition, 2001-2014

![Figure 5.2: New Zealand Vehicle Fleet Composition, 2001-2014](image)

Data Source: MoT (2014)

Figure 5.3 illustrates the annual fleet growth relative to January 2000. Both light passenger and light commercial vehicles experienced a fast growth between 2000 and 2006, but had a relatively steady increase ever since. It should also note that the light fleet
is not the fastest growing part of the overall vehicle fleet as indicated by figure 5.3, bus numbers have risen by 100% since 2000.

**Figure 5.3: Fleet Composition Relative to Jan 2000**

![Fleet Composition Graph]

Data Source: MoT (2014)

### 5.2.2. Light vehicle ownership

Light vehicles ownership per capita, of which the majority were passenger cars in New Zealand increased significantly between 2000 and 2005. As MoT (2014) indicates, this upward trend can be explained by several factors, including the high value of the NZD which made vehicles relatively cheaper to purchase\(^{43}\), high employment levels and the positive economic outlook over that time period. Although between 2007 and 2012, light vehicle ownership per capita declined, it went up again in 2013 and 2014. There were 744 light vehicles for every 1,000 New Zealanders in 2014, compared to 664 in 2001. Notably

\(^{43}\) According to the MoT (2013a), used imported vehicles make up a large proportion of the light vehicle fleet.
there are substantial regional variations between these two islands in New Zealand. Three regions from the South Island: Canterbury, Nelson-Marlborough and Southland are three of the four regions with the highest ownership. The two largest cities from the North Island, Auckland and Wellington, however, have lower ownership rates compared to New Zealand average and most other regions, and contributed to at least in part by the availability and relatively high frequency of public transport services. Additionally, according to data released by the MoT, the motor vehicle access rate has increased over time: 92% of households have access to one motor vehicle in 2013, compared with only 88% in 1991.44

5.2.3. Facts about overall car travel

Data from recent New Zealand Household Travel Survey (NZHTS) indicates that car travel is the main road transport mode, accounting for approximately 70% of the light vehicle45 distance driven by households. On average, New Zealanders aged between 35 and 64 spend about two thirds of their total travel time driving. In terms of gender difference, men drive more than 12,000 km per driver per year, while women drive just over 8,000 km per driver per year. Overall, the amount of driving done per driver increased from 25 km per day in 1989/90 to 29.7 km per day in 2004 – 2007 and then decreased again around 28 km/day in 2009 – 2012 (MoT, 2014). Additionally, one-sixth of household car trips in New Zealand are considered as “short distance” as these car trips are typically less than 2 kilometres (km) and almost half are less than 6km long. It should be noted that short distance car trips are typically more polluting compared to long trips, as cold engines consume around 40% more fuel, therefore producing more emissions (Statistics New Zealand, n.d. b).


45 Excludes motorcycles.
5.3. Literature Review

For the past few decades, extensive effort has been devoted to investigating various factors affecting the demand for private vehicles, with the majority of car ownership, and the demand for public transport around the world. From past research, per capita income, inflation, national unemployment rate, and population indicators are identified as important determinants of both private vehicle ownership and public transport use at the macroeconomic level (Allanson, 1982; Button et al., 1982; Dargay and Gately, 1999; Messenger and Ewing, 2007; Prevedouros and An, 1998; Paravantis and Prevedouros, 2001; Rentziou et al., 2012; Schipper, 1995)

Based on data from a selection of the OECD countries for the time period 1970-1992, Schipper (1995) found that per capita income has a major impact on car ownership. Prevedouros and An (1998) concluded that income, measured by GDP per capita, along with population and unemployment rate, are key determinants of aggregate car ownership in both developing and developed Asian countries. Both studies reveal that the car ownership increases with an increase in the level of income. Paravantis and Georgakellos (2006) estimated an aggregate car ownership model and an aggregate bus fleet model using data from the period 1970 to 2002 in Greece. For the car ownership model, the authors have identified that the percentage of adult population, GDP per capita, bus VKT and car occupancy as important car ownership determinates. For the bus fleet model, percentage of adult population, GDP per capita, and inflation are found to have significant impacts on the bus fleet demand, measured by the number of total buses per 100,000 people.

Furthermore, the rate of unemployment may also be a variable of interest as it is plausible to assume that for those who are unemployed, the ability to possess and afford private transport can be outside their financial budgets, implying there is no other option but to rely on public transport for travel. Therefore we should expect a negative relationship
between the unemployment rate and the demand for car travel, but positive relationship between unemployment rate and the demand for public transport.

Empirically, fuel costs were found to be another important factor that affects road passenger transport demand. By applying an asymmetric error correction model, recent research by Chao et al. (2015) indicates that gasoline prices have significantly positive effects on the two forms of public transport in Taiwan, bus and mass rapid transit (MRT) use. Empirical results from past literature also support the authors’ conclusion that the values of elasticity of public transport demand with respect to gasoline prices are often inelastic. In other words, the absolute value of price elasticity is ranged somewhere between 0 and 1, as found in various previous studies (Wang and Skinner, 1984; Haire and Machemehl, 2007; Currie and Phung, 2007; and Currie and Phung, 2008).

Vehicle costs, as Rive et al. (2015) argue, are inherently hard to measure because the description generally covers not only the cost of purchasing a car, but also the ongoing costs associated with car ownership, such as road costs, repairs, maintenance and insurance. Inflation, therefore, has been used as a proxy to capture such costs in a few studies, because it is plausible to assume that an increase in inflation makes vehicle costs such as maintenance, toll fees, and insurance more expensive, thus reducing the demand for travel. Empirical evidence from Paravantis and Prevedouros (2001) show that inflation had significantly negative effect on the first order autoregressive railway passenger demand models based on the data from 1970 – 1998 in Greece.

Only a few studies have examined the relationship between the usage of private cars and public transportation. For example, by applying a binary probit model, Kitamura (1989) found that changes in car usage could affect public transportation usage; however, on the other hand, changes in public transport usage had only minor impacts on car usage. This
finding implies that the increase in car use, which is a consequence of increasing car ownership, may not be suppressed by improving public transport service.

In the context of New Zealand, Conder (2009) estimated a car ownership model for New Zealand using aggregate national-level time series data. The aggregate car model was an update from previous research by Booz Allen Hamilton (2000) and was split into two parts: car ownership saturation level and path to saturation. GDP per capita, car price index, and a time trend were included as explanatory variables for the analysis of the likely growth path to saturation estimated using OLS.

Wang (2011) conducted a time series analysis to determine the factors that influence the demand for public transport (bus and rail) patronage in both short-term and long-run for New Zealand’s three major cities (Auckland, Wellington and Christchurch), using quarterly data from 1996-2008. Using a partial adjustment (PA) model, the relationship between patronages, measured as the total number of bus or train trips per capita, was modelled as a function of several factors including: service level, real fare, real disposable income per capita, car ownership and real fuel price. The results deliver two important implications. Firstly, fuel price exhibits a positive effect on public transport patronage in all three cities. Secondly, the effect of factors varied across the three cities. For instance, in terms of statistical significance, bus fares were found to have an impact on bus demand in Wellington and Christchurch, but not in Auckland. While in terms of magnitude, income per capita exhibited a negative effect on rail patronage in Wellington, but a positive effect on rail patronage in Auckland.

In addition, past literature indicates that another vehicle fleet variable, vehicle occupancy which represents the vehicle loading factor, should also be considered when estimating car ownership models. However for the New Zealand case, vehicle occupancy is not
considered as a possible factor affecting automobile demand as the vehicle occupancy is mostly characterised by single vehicle occupant (i.e. the driver is the only person in the vehicle), with 63% of the total distance driven by the driver only (MoT, 2013b). Therefore due to this reason, vehicle occupancy was not included in this study. Car price, on the other hand, was not considered as a potential predictor because it is more relevant to influence road users’ decision on whether to own a car or not, but predicting the demand pattern for road passenger travel. Car ownership is also influenced by the fares of alternative transport modes, therefore the bus fare index was included in this study for all of the demand functions. Moreover, percentage of different age groups (i.e. the ratio of total number of people in a certain age group/total population), a related population metric of interest, is considered another potential indicator for road user’s travel demand pattern for both private and public transport demand.

In summary, although a wealth of studies has investigated the demand for cars and public transportation, individually and/or jointly, nationally and internationally, little investigation has been undertaken into account of the demand for cars by different fuel types, a classification that is crucial in examining road user’s transport mode preference in the light of energy consumption, CO₂ emissions, and policy implications. Moreover, given the fact that these private and public transport modes are potential substitutes to one another, to the best of our knowledge, only one study from Jou and Chen (2014) has considered the relationship among the demand of different road transportation modes, including public transportation, car, and motorcycle, in various townships in Taiwan by applying a SUR model. To the best of our knowledge again, no previous study in New Zealand has considered these demands as a system of equations with correlated disturbances, a research gap that was thus discussed and addressed in this paper.
5.4. Model Specification

5.4.1 Aggregate road passenger transport demand functions

The individual demand equation of each aggregate road passenger transport mode in New Zealand is specified as a function of a number of key determinants, including: income of road passenger transport users, price of the road passenger transport mode, price of substitutionary modes, and some socioeconomic and demographic factors. Therefore the individual demand function for petrol cars, diesel cars and buses can be represented as follows:

\[
Q_{it} = f(I_t, P_{it}, P_{it}^S, SD_t) \quad (1)
\]

where

- \(Q_{it}\) = quantity demanded of the \(i^{th}\) road passenger transport mode in the \(t^{th}\) quarter;
- \(I_t\) = income of road passenger transport users in the \(t^{th}\) quarter;
- \(P_{it}\) = price of the \(i^{th}\) road passenger transport mode in the \(t^{th}\) quarter;
- \(P_{it}^S\) = price of substitutes of the \(i^{th}\) road passenger transport mode in the \(t^{th}\) quarter;
- \(SD_t\) = socioeconomic and demographic factors in the \(t^{th}\) quarter

It should be noted that the demand for road passenger travel is normally considered a derived demand as it is not typically demanded just because people prefer travelling (except possibly for a proportion of scenery trips) but because transport supports a range of other activities, which enables passengers to reach a desired destination in order to consume other goods and services. For the case of the demand for buses, in addition to
private vehicles, rail service is also one of the main competitors to buses in the context of public transport.

5.4.2 SUR modelling

Wang and Kockelman (2007) point out that, in many transportation studies, variables of interest are often influenced by similar factors and have correlated disturbances. In such cases, these data are best modelled using a system of interrelated equations because their dependent variables share common characteristics. Some transportation examples include: modelling transportation infrastructure performance (Prozzi and Hong, 2008); analysing the effect of the built environment and residential self-selection on non-work travel (Cao et al., 2006); investigating the impact of anticipated transportation improvement on residential land values (McDonald and Osuji, 1995); and estimating static and dynamic urban travel demands (Gaudry, 1978).

Seemingly unrelated regression (SUR) modelling, firstly discussed by Zellner (1962), is thus appropriate when analysing factors affecting aggregate road passenger transport demand where the dependent variables are considered as a group but do not have a direct interaction. Generally speaking, a SUR system represents a generalisation of a linear regression model which comprises several regression equations, allowing each to have its own dependent variable and same or different sets of exogenous regressors. The key feature of the SUR model is that in referring to responses of the same set of observational units, the errors of these equations are likely to be correlated (Wang and Kockelman, 2007). In this sense, the SUR model can be regarded as either a simplified version of the general linear model where certain coefficients in matrix $\beta$ are restricted to be equal to zero, or as the generalisation of the general linear model where right-hand-side explanatory variables are allowed to be different in each equation. Moreover, as Rentziou et al. (2012) note, although the equations are seemingly unrelated, contemporaneous correlation of error terms exist. Therefore if interlinked equations are estimated using OLS
separately rather than SUR model which amounts to feasible generalised least squares (FGLS) with a specification of the variance-covariance matrix, coefficients are consistent but generally inefficient.

Suppose \( y_{it} \) is a dependent variable, \( x_{it} = (1, x_{it,1}, x_{it,2}, ..., x_{it,K_i})' \) is a \( K_i \)-vector of explanatory variables for observational unit \( i \), and \( u_{it} \) is an unobservable error term, where the double index \( it \) denotes the \( t^{th} \) observation of the \( i^{th} \) equation in the system, \( t \) denotes time and we refer to this as the time dimension. A classical linear SUR model is a system of \( N \) linear regression equations (Zellner, 1962):

\[
y_{it} = X_{it} + \mu_{it}, \quad i = 1, \cdots, N, \text{ and } t = 1, \cdots, T \tag{2}
\]

Denote \( L = K_1 + \cdots + K_N \). Assume that for each \( i = 1, \cdots, N \), \( x_i = [x_{i1}, \ldots, x_{it}]' \) is of full rank \( K_i \), and that conditional on all the regressors \( X' = [X_1, \ldots, X_T] \), the error terms \( U_i \) are i.i.d. over time with mean \( E[u_t|X] = 0 \) and homoscedastic variance \( \Sigma = E(u_t u'_t|X) \). In addition we also assume that \( \Sigma \) is positive definite and denote by \( \sigma_{ij} \) the \((i,j)^{th}\) element of \( \Sigma \), that is, \( \sigma_{ij} = E(u_i u_j|X) \). Under this assumption, the covariance matrix of the entire vector of disturbances \( U' = [U_1, \ldots, U_T] \) is given by \( E[\text{vec}(U)(\text{vec}(U))'] = \Sigma \otimes I_T \).

Moon and Perron (2008) propose further simplification in notation by stacking the observations either in the \( t \) dimension or for each \( i \). For instance, if we stack for each observation \( t \), let \( Y_t = [y_{1t}, \ldots, y_{Nt}]' \), \( \bar{X}_t = diag(x_{1t}, x_{2t}, \ldots x_{Nt}) \), a block-diagonal matrix with \( x_{1t} \ldots x_{Nt} \) on its diagonal, \( U_t = [u_{1t}, \ldots u_{Nt}]' \), and \( \beta = [\beta_1', \ldots \beta_N']' \). Then, \( Y_t = \bar{X}_t \beta + U_t \) \tag{3}
In addition, following most transport economics literature (e.g. see Anwaar, et al., 2012; Kennedy, 2013; Kopits and Cropper, 2005; Rentziou et al., 2012; Zhou and Kockelman, 2009), natural logarithm transformation has been applied for both sides of each aggregate road passenger transport demand equation. The advantage of the log-transformation is that the estimated coefficients can be interpreted as elasticities.

5.4.3 Demand for aggregate road passenger transport: the SUR model

Therefore the demand functions in (1) can be estimated efficiently by the SUR model, by taking into account of the potential correlation among disturbances. The SUR model of demand for aggregate road passenger transport modes can thus be specified as a system of three double-log demand equations as follows, where we use the ownership of petrol/diesel cars (measured by car registrations per 1,000 population) as a measure of demand for private vehicles, and the total vehicle kilometre travelled (VKT) by buses as a measure of demand for buses: 46

\[
\ln(\text{Car}_t) = a_p + \beta_{p1} \ln(\text{GDP}_{pc_t}) + \beta_{p2} \ln(\text{Petrol}_t) + \beta_{p3} \ln(\text{Diesel}_t) + \beta_{p4} \ln(\text{Bus}_t) + \beta_{p5} \ln(\text{Inflation}_t) + \beta_{p6} \ln(\text{Unemployment}_t) + \beta_{p7} \ln(\text{Young people}_t) + \beta_{p8} \ln(\text{Middle} - \text{aged}_t) + \beta_{p9} \ln(\text{Matured}_t) + \beta_{p10} \ln(\text{Senior}_t) + \epsilon_{1t} \\
\ln(\text{Car}_t) = a_d + \beta_{d1} \ln(\text{GDP}_{pc_t}) + \beta_{d2} \ln(\text{Diesel}_t) + \beta_{d3} \ln(\text{Petrol}_t) + \beta_{d4} \ln(\text{Bus}_t) + \beta_{d5} \ln(\text{Inflation}_t) + \beta_{d6} \ln(\text{Unemployment}_t) + \beta_{d7} \ln(\text{Young people}_t) + \beta_{d8} \ln(\text{Middle} - \text{aged}_t) + \beta_{d9} \ln(\text{Matured}_t) + \beta_{d10} \ln(\text{Senior}_t) + \epsilon_{2t}
\]

46 Ideally, VKT is a better measure of demand than counts, but they are not available for petrol and diesel cars (we only have VKT for bus from NZTA). Therefore this study followed several past research who also used counts, instead of VKT, for demand for cars: Bjorner (1999); Paulley et al. (2006); Giuliano and Dargay (2006).
\[ \ln(\text{Bus}_VKT_t) = a_p + \beta_{b1} \ln(\text{GDP}_{pc_t}) + \beta_{b2} \ln(\text{Bus}_t) + \beta_{b3} \ln(\text{Petrol}_t) \\
+ \beta_{b4} \ln(\text{Diesel}_t) \\
+ \beta_{b5} \ln(\text{Rail}_t) + \beta_{b6} \ln(\text{Inflation}_t) + \beta_{b7} \ln(\text{Unemployment}_t) \\
+ \beta_{b8} \ln(\text{Young people}_t) + \beta_{b9} \ln(\text{Middle-aged}_t) \\
+ \beta_{b10} \ln(\text{Matured}_t) + \beta_{b11} \ln(\text{Senior}_t) + u_{3t} \]

where,

\text{Car}_P_t is the total number of registered petrol cars per 1000 people;

\text{Car}_D_t is the total number of registered diesel cars per 1000 people;

\text{Bus}_VKT_t is the VKT by buses per 1000 people;

\text{GDP}_{pc_t} is the seasonally adjusted real Gross Domestic Product (GDP) per capita in 2001Q1 price;

\text{Petrol}_t is the real petrol price adjusted by Consumers Price Index (CPI) with base year 2001Q1;

\text{Diesel}_t is the real diesel price adjusted by CPI with base year 2001Q1;

\text{Bus}_t is the price index for urban bus fares, long distance bus fares (excluding coach tours), taxi fares, shuttle fares, and car hire charges;

\text{Rail}_t is the price index for urban train fares and long distance train fares\(^47\);

\text{Inflation}_t is the seasonally adjusted inflation rate;

\text{Unemployment}_t is the seasonally adjusted unemployment rate;

\(^{47}\) For a detailed description of road transport and rail passenger price indexes, please see “Rail, road, and sea passenger transport service in the CPI” from Statistics New Zealand: http://www.stats.govt.nz/tools_and_services/newsletters/price-index-news/apr-13-article-air-transport.aspx
Young people, is the percentage of 15-24 years old population, referred as “young people”;

Middle-aged, is the percentage of 25-44 years old population, referred as the “middle-aged”;

Matured, is the percentage of 45-64 years old population, referred as the “matured”;

Senior, is the percentage of 65 years old and above population, referred as “senior”.

5.5. Data Description

The analysis was undertaken at a national level using aggregate data. The selection of variables for this study is mostly inspired by previous studies and the availability of data. The following vehicle fleet information, bus VKT and demographic and macroeconomic data required for the estimation of the proposed aggregate model formulations were assembled for the analysis period from the first quarter of 2001 to the second quarter of 2015, a period of time over which data are available for all variables.

Various data sources were collected in this study: vehicle fleet information regarding petrol and diesel car registrations were provided by the MoT. Total bus VKT were provided by the New Zealand Transport Agency (NZTA). Data on petrol and diesel prices came from the MBIE. The two public transport fare indicators, namely, bus fare index and rail fare index, and some socioeconomic and demographic data, including: population, GDP per capita, inflation, unemployment rate, were obtained from Statistics New Zealand through Inforshare.

There are a few highlights on historical patterns for car ownership and bus VKT data. Figure 5.4 (a) shows that the petrol car ownership firstly fluctuated from the beginning of the sample period until around 2007Q4, then dropped sharply and reached its minimum at
2009Q2 and later on increased again. The vehicle ownership for diesel cars however, experienced a rather steady trend across time compared to petrol cars. For the demand for public transport, Figure 5.4 (b) illustrates that there’s an upwards trend for Bus_VKT over the sample period, but the increase is rather slow and steady, with a peak in 2001Q3.

**Figure 5.4 (a): Demand for Petrol & Diesel Cars**

![Graph showing demand for petrol and diesel cars](image)

**Figure 5.4 (b): Demand for Buses**

![Graph showing demand for buses](image)
Table 5.1 presents a summary of descriptive statistics for variables used in this study. The total number of observations is 58 quarters. For the two car ownership variables, the mean of petrol car ownership (10.44) is considerably larger compared to the mean of diesel car ownership (2.18). Turning our attention to the independent variables. Firstly, the income indicator, real GDP per capita showed that the income per person in New Zealand is around $8635NZD in constant 2001Q1 dollar. For the two real fuel prices, petrol price has a higher mean value compared to diesel price. Next, the mean values for bus and rail fare indexes are 1080.24 and 1157.78 respectively.

Table 5.1: Summary Statistics of Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car_P</td>
<td>10.44</td>
<td>2.07</td>
<td>5.95</td>
<td>13.61</td>
</tr>
<tr>
<td>Car_D</td>
<td>2.18</td>
<td>0.54</td>
<td>1.13</td>
<td>3.07</td>
</tr>
<tr>
<td>Bus_VKT</td>
<td>12.98</td>
<td>1.51</td>
<td>9.21</td>
<td>14.88</td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP_pc</td>
<td>8634.55</td>
<td>457.09</td>
<td>7647</td>
<td>9440</td>
</tr>
<tr>
<td>Petrol</td>
<td>126.84</td>
<td>20.79</td>
<td>91.65</td>
<td>159.13</td>
</tr>
<tr>
<td>Diesel</td>
<td>86.81</td>
<td>19.41</td>
<td>53.17</td>
<td>133.75</td>
</tr>
<tr>
<td>Bus</td>
<td>1080.24</td>
<td>137.35</td>
<td>896</td>
<td>1351</td>
</tr>
<tr>
<td>Rail</td>
<td>1157.78</td>
<td>216.95</td>
<td>870</td>
<td>1503</td>
</tr>
<tr>
<td>Inflation</td>
<td>2.36</td>
<td>1.16</td>
<td>0.3</td>
<td>5.3</td>
</tr>
<tr>
<td>Unemployment</td>
<td>5.24</td>
<td>1.14</td>
<td>3.4</td>
<td>7.2</td>
</tr>
<tr>
<td>Young people</td>
<td>0.14</td>
<td>0</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>Middle-aged</td>
<td>0.28</td>
<td>0.01</td>
<td>0.26</td>
<td>0.3</td>
</tr>
<tr>
<td>Matured</td>
<td>0.24</td>
<td>0.01</td>
<td>0.22</td>
<td>0.26</td>
</tr>
<tr>
<td>Senior</td>
<td>0.13</td>
<td>0.01</td>
<td>0.12</td>
<td>0.15</td>
</tr>
</tbody>
</table>
It should be noted that the base period for these two indexes is 2006Q2 not 2001Q1 as these indexes are rebased to reflect changes to the industry/input data etc. Because of the rebase, these index series have to start from the rebase. Moreover, inflation exhibited large variation across the sample time period as the maximum value is 5.3%, while the minimum is only 0.3%. Similar trend was found in the unemployment rate, where the maximum value is 7.2% while the minimum is only 3.4%. Lastly, the mean values for the percentage of young, middle-aged, matured and old people are 14%, 28%, 24% and 13% respectively.

5.6 Stationarity Test

The first step in time series analysis is to determine whether the levels (in this case, log–levels) of the data are stationary. Appendix C shows the plots all the variables against time for visual inspection of stationarity. It can be seen from these curves that none of the series looks stationary in its log-levels, and most of them appear to have an upward-sloping trend. To confirm our hypothesis of non-stationarity among variables, one of the most commonly used tests for stationary in time series, the Augmented Dickey-Fuller (ADF) test, is performed. The null hypothesis of the ADF test is that the series has a unit root. The results from Table 5.2 indicate that all of the variables are non-stationary as shown by their insignificant p-values.
Table 5.2: ADF Test for Unit Root

<table>
<thead>
<tr>
<th>Variables</th>
<th>ADF-t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln_Car_P</td>
<td>-1.327</td>
<td>0.6169</td>
</tr>
<tr>
<td>ln_Car_D</td>
<td>-1.038</td>
<td>0.739</td>
</tr>
<tr>
<td>ln_Bus_VKT</td>
<td>-1.411</td>
<td>0.8576</td>
</tr>
<tr>
<td>ln_GDP_pc</td>
<td>-2.237</td>
<td>0.469</td>
</tr>
<tr>
<td>ln_Petrol</td>
<td>-3.058</td>
<td>0.1166</td>
</tr>
<tr>
<td>ln_Diesel</td>
<td>-1.583</td>
<td>0.4923</td>
</tr>
<tr>
<td>ln_Bus</td>
<td>-2.898</td>
<td>0.1629</td>
</tr>
<tr>
<td>ln_Rail</td>
<td>-2.778</td>
<td>0.2052</td>
</tr>
<tr>
<td>ln_Inflation</td>
<td>-1.721</td>
<td>0.4924</td>
</tr>
<tr>
<td>ln_Unemployment</td>
<td>-1.582</td>
<td>0.7978</td>
</tr>
<tr>
<td>ln_Young people</td>
<td>-2.598</td>
<td>0.2808</td>
</tr>
<tr>
<td>ln_Middle-aged</td>
<td>-0.083</td>
<td>0.9933</td>
</tr>
<tr>
<td>ln_Matured</td>
<td>-0.506</td>
<td>0.9832</td>
</tr>
<tr>
<td>ln_Senior</td>
<td>-2.809</td>
<td>0.1936</td>
</tr>
</tbody>
</table>

Given the fact that all of the variables have unit roots, the next step is to take the first difference of these non-stationary series and apply the ADF test again to see if their first difference is stationary. Generally, if the log-levels of the time series are not stationary, the first differences will be. Results from Table 5.3 confirmed that the all of the first differences become stationary at 1% significant level, with the only exception of \( \text{ln}_{\text{Matured1}} \), which is significant at 10% level. Therefore we conclude that all of the variables are \( I(1) \).

---

48 Whether or not to include constant and/or trend in the ADF test are determined based on visual inspection of the plots of variables against time from Appendix C. In addition, lag lengths are selected using the Schwarz’s Bayesian information criterion (SBIC), as Ivanov and Kilian (2001) suggest, the SBIC works well with any sample size for quarterly data.
Table 5.3: ADF Test for Unit Root on All Variables in D1

<table>
<thead>
<tr>
<th>Variables</th>
<th>ADF-t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln_Car_P1</td>
<td>-5.306</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Car_D1</td>
<td>-6.929</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Bus_VKT1</td>
<td>-6.762</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_GDP_pc1</td>
<td>-7.276</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Petrol1</td>
<td>-7.499</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Diesel1</td>
<td>-5.668</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Bus1</td>
<td>-7.759</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Rail1</td>
<td>-6.803</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Inflation1</td>
<td>-6.812</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Unemployment1</td>
<td>-7.201</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Young people1</td>
<td>-6.526</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Middle-aged1</td>
<td>-5.751</td>
<td>0.0000</td>
</tr>
<tr>
<td>ln_Matured1</td>
<td>-3.305</td>
<td>0.0655</td>
</tr>
<tr>
<td>ln_Senior1</td>
<td>-7.448</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

5.7 Empirical results and policy implications

Table 5.4 summarises some test statistics for demand equations estimated by the SUR model. Regarding the fitness of the model, all *F-statistics* meet the standard statistical test. The values of $R^2$ are ranged from 0.9147 to 0.9814, suggesting that the SUR model explains 91.47% to 98.14% of variability in the data.
The estimation results from the SUR model of the aggregate road passenger transport demand in New Zealand are summarised in Table 5.5. It should be noted that only factors that have a significant impact on the demand for petrol cars, diesel cars, and buses are reported. Standard errors are presented in parentheses and the level of statistical significances is marked by asterisks (** for 1%; ** for 5% level; * for 10% level).

The use of log-log specification enables us to interpret the estimated coefficients as elasticities. For the demand for petrol cars, all of the estimated coefficients have expected signs. Firstly, the income elasticity is positive, indicating that demand for petrol cars go up with an increase in people’s per capita income, holding all else fixed. The estimated value of the income elasticity for petrol cars is 3.99, implying that for every 1% increase in road user’s income, on average, the demand for petrol cars is expected to increase by 3.99%. Secondly, the negative coefficient for ln_Petrol, -0.61, is also in accord with economic theory; there is an inverse relationship between the cost of using petrol cars and the quantity demanded for petrol cars. Thirdly, the cross price elasticity between diesel cars and petrol cars, represented by the estimated coefficient on ln_Diesel, is significant and positive at the 5% level. This implies that when the other variables are unchanged, for every 1% increase in the cost of using diesel cars, the demand for its substitute, petrol cars, is expected to increase by 0.44% on average.
Table 5.5: SUR Model Results for the Aggregate Road Transport Demand

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Petrol Cars</th>
<th>Diesel Cars</th>
<th>Buses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$ln_{Car_P}$</td>
<td>$ln_{Car_D}$</td>
<td>$ln_{Bus_VKT}$</td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td>-52.49***</td>
<td>-87.76***</td>
<td>20.27***</td>
</tr>
<tr>
<td></td>
<td>(15.62)</td>
<td>(23.86)</td>
<td>(5.46)</td>
</tr>
<tr>
<td>$ln_GDP_pc$</td>
<td>3.99***</td>
<td>3.45***</td>
<td>2.89</td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td>(1.19)</td>
<td></td>
</tr>
<tr>
<td>$ln_Petrol$</td>
<td>-0.61*</td>
<td>-0.49***</td>
<td>-4.85</td>
</tr>
<tr>
<td></td>
<td>(0.32)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>$ln_Diesel$</td>
<td>0.44**</td>
<td>0.24***</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>$ln_Bus$</td>
<td>-1.97***</td>
<td>-0.26**</td>
<td>-2.16</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>$ln_Rail$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ln_Inflation$</td>
<td>-0.09**</td>
<td>-2.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ln_Unemployment$</td>
<td>-0.50**</td>
<td>-2.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ln_Young_people$</td>
<td>8.29***</td>
<td>7.54***</td>
<td>2.12***</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
<td>(2.36)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>$ln_Middle_aged$</td>
<td>-9.61**</td>
<td>-30.53***</td>
<td>2.68**</td>
</tr>
<tr>
<td></td>
<td>(3.83)</td>
<td>(5.48)</td>
<td>(1.20)</td>
</tr>
<tr>
<td>$ln_Matured$</td>
<td>-14.87***</td>
<td>-24.94***</td>
<td>4.31***</td>
</tr>
<tr>
<td></td>
<td>(2.19)</td>
<td>(3.14)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>$ln_Senior$</td>
<td>-4.66**</td>
<td>-3.68</td>
<td>1.29**</td>
</tr>
<tr>
<td></td>
<td>(2.19)</td>
<td>(0.51)</td>
<td></td>
</tr>
</tbody>
</table>
Lastly, three out of four age indicators showed an impact on the demand for petrol cars. The relationship between young people and the demand for petrol cars is positive; while the relationship between both middle-aged and mature people, and the demand for petrol cars is negative. This result gives us a possible implication that young people tend to enjoy the convenience and conformability that they derive from car trips, while older generations prefer other modes of transport apart from automobiles, possibly due to a healthier lifestyle choice of travel and considerations on environment.

For the demand of diesel cars, the income elasticity of diesel car demand, represented by the estimated coefficient on $ln\_GDP\_pc$, is significant and positive at the 1% level, indicating that for every 1% increase in road user’s income, on average, the demand for diesel cars is expected to increase by 3.45%. This income effect is only marginally smaller compared to petrol cars. Moreover, inflation and unemployment were both found to have a significant negative impact on the diesel car demand. Firstly, the estimated coefficients on $ln\_Inflation$ is negative, suggesting that an inverse relationship between inflation and the demand for diesel cars. This is in line with our expectation as people will tend to reduce their demand for diesel cars if the associated vehicle costs tend to rise. Secondly, the estimated coefficients on $ln\_Unemployment$ is -0.50, indicating an inverse relationship between unemployment rate and demand for diesel vehicles. As expected, this finding implies that for those who are unemployed, the ability to possess private transport is normally outside of their financial means. The age indicators showed similar trend as they were in the petrol car demand equations, where the estimated coefficient on young population is significant and positive, and the estimated coefficients on middle-aged and mature age groups are significant and negative. Surprisingly, the two fuel cost indicators, $ln\_Petrol$ and $ln\_Diesel$ showed no effect on the demand for diesel cars. In addition, the cross price elasticity between public transport and diesel cars is significant and negative at the 1% level, this unexpected sign might be due to the composition of the bus fare index, where it not only represents the cost of urban bus fares, long distance bus fares, but also taxi fares, shuttle fares, and car hire charges. However this bus fare index was the best available choice to represent the cost of taking bus travel for this study.
For the demand for public transport, first of all, the per-capita income, the price of rail service, inflation and unemployment rate were all found insignificant, suggesting that these variables have no impact on the demand of buses. Secondly, the price elasticity for buses is negative, indicating that if the bus fare decreases, the demand for buses will increase. The cross elasticity between diesel cars and buses is 0.24, indicating that when the other variables are unchanged, for every 1% increase in the cost of using diesel cars, the demand for its substitute, buses, is expected to increase by 0.24% on an average base. Additionally, the sign of \( \ln\) _Petrol_ is significant and negative, this result may be due to the fact that _Bus_VKT_ have not been classified based on their fuel types, therefore the estimated coefficient on \( \ln\) _Petrol_ might be interpreted as the fuel cost for some petrol buses, rather than an indicator for cross elasticity. All of the four age indicators are positive, an indication that New Zealanders are inclined to use public transport as their road transport choice.

The above empirical results from the SUR model also delivers some important policy implications. First of all, in order to achieve a reduction in the demand for automobiles, different policies could be implemented for cars with different fuel types, as the factors that affect the demand for these two major types of private vehicles in New Zealand differ significantly. For instance, policy makers could consider increasing the petrol tax so as to reduce the demand for petrol cars. As the price of petrol increases, with other predictors remaining constant, less petrol cars are demanded since the cost of using this type of road passenger transport mode is higher. The price elasticity for petrol cars is 0.61, suggesting that for every 10% increase in the average real petrol price, we would observe a 6.1% drop in the demand for petrol cars.

Secondly, because fuel prices do not have an impact on the demand for diesel cars, increasing the level of taxes on diesel would not lead to a possible decline in the demand for diesel cars as observed in petrol cars. Rather, policy makers could consider levelling up the taxation on vehicle-related ongoing costs, such as the vehicle registration fees,
annual licensing fees, administration fees, and road user charges for diesel vehicles, in order to achieve an effective reduction of the demand for diesel cars. As suggested by the estimated coefficient on $\ln_{\text{Inflation}}$, if vehicle costs increase by 10%, the quantity of diesel cars demanded is expected to decrease by 0.9%.

Thirdly, for the purpose of promoting the use of public transport, policy makers could consider lowering the fares. This of course, will have to be financed; possibly by recycling the revenue from fuel and car taxation. Price elasticity for buses indicates that on average, when the other variables stay the same, a 10% reduction in bus fares is expected to increase the demand for buses by 2.6%. Given the fact that the current fares on public transport remain relatively high, the government might need to consider granting a subsidy to public transport providers, so that the road transport users can enjoy lower fares and increase their demand for public transport.

Last, as the signs of all of the age indicators are positive from the public transport demand equation, public transport authorities could considering increase the trip frequencies so that young people who mainly ride public transport for educational purpose, middle-aged and matured people who represent the majority of commuters, and senior citizens who mostly rely on public transport because they have less mobility, could all be benefited from an improved service from public transport and hence increase demand.

Moreover, to test whether the estimated correlation between these three equations is statistically significant, we use the Lagrange multiplier (LM) statistic proposed by Breusch and Pagan (1980) and it is shown as the following:

$$
\lambda = N \sum_{m=1}^{M} \sum_{n=1}^{m-1} r_{mn}^2
$$
where $r_{mn}$ is the estimated correlation between the residuals of the $M$ equations (in this case 3) and $N$ is the number of observations (in this case 58). It is distributed as $\chi^2$ with $M(M-1)/2$ degrees of freedom. The results are presented in Table 5.6.

<table>
<thead>
<tr>
<th></th>
<th>$r_{\text{Car}_P}$</th>
<th>$r_{\text{Car}_D}$</th>
<th>$r_{\text{Bus}_\text{VKT}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{\text{Car}_P}$</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{\text{Car}_D}$</td>
<td>-0.1069</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>$r_{\text{Bus}_\text{VKT}}$</td>
<td>-0.2893</td>
<td>-0.3111</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Breusch-Pagan test of independence: $\chi^2 (3) = 11.129, Pr = 0.0110$

Based on the results from Table 5.6, we can reject the null hypothesis that the covariance between the three different equations are equal to zero at 5% significance level. This implies that the residuals from each SUR regression are significantly correlated with each other, representing identical unsystematic influences. Therefore estimating each equation separately using OLS will give us consistent but inefficient coefficients. SUR model, in this case, is superior that OLS as the estimated coefficients are both consistent and efficient. Additionally, because all of the signs are negative, we can conclude that the three road transport modes are substitutes, implying that an increasing effect of the residuals on one mode will decrease the effect of the residuals on the other mode.
5.8 Test for cointegration and structural breaks

5.8.1 Cointegration Test

Cointegration refers to the fact that two or more non-stationary time series possess the same order of integration hence a linear combination of these series is stationary. If a stationary linear combination exists, we can conclude that the non-stationary time series are cointegrated. In other words, even if the variables may wander around in a certain period of time, they cannot drift too far apart from each other in the long-run. The deviations from equilibrium (i.e. residuals) are thus stationary, with finite variance, even though time series variables are not (Chao et al., 2015). Plots of residuals against time for each SUR equation are summarised in Appendix D. From graphical inspection, we conclude that the residuals show little evidence of trend. After estimating the SUR model, follow a two-step cointegration test suggested by Engle and Granger (1987), we firstly obtain the residuals and secondly run an ADF test on them in order to test for unit root.

<table>
<thead>
<tr>
<th>Residuals</th>
<th>ADF-t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>r_Car_P</td>
<td>-4.844</td>
<td>0.0000</td>
</tr>
<tr>
<td>r_Car_D</td>
<td>-4.646</td>
<td>0.0001</td>
</tr>
<tr>
<td>r_Bus_VKT</td>
<td>-3.253</td>
<td>0.0171</td>
</tr>
</tbody>
</table>

Results from Table 5.7 indicate that we can reject the null hypothesis of unit root on the residuals at the 1% level for the equation on petrol and diesel cars, and we can reject the null hypothesis of unit root on the residuals at the 5% level for the equation on buses. Therefore we conclude that the variables are cointegrated, or have a stationary long-run relationship, even though individually they are stochastic. The SUR model is thus valid and can be estimated in its original specification as outlined in section 7.
5.8.2 Structural breaks

Furthermore, based on the graphical illustrations of predicted demand versus actual demand for natural logarithms for petrol cars, diesel cars and buses from Appendix E, we can conclude that none of the equations suffer from visible structural breaks as the predicted and the actual data fit nicely with one another, suggesting that the estimated results from the SUR model is valid and robust for future projections of the demand for aggregate road passenger transport.

5.9 Conclusion and Recommendations

This paper is the first use of SUR in New Zealand to develop an aggregate road passenger travel demand model. It aims at firstly examining whether the error terms of the demand for the three main road passenger transport choices: petrol cars, diesel cars and buses, are correlated and secondly, identifying the factors that have impacts on the demand for each available road passenger transport choice.

The Breusch-Pagan test of independence confirms the existence of correlated error terms of the three demand equations and the empirical results from SUR model indicates that per-capita GDP has an effect on both the demand for petrol and diesel cars, with the values of income elasticity to be 3.99 and 3.45, respectively. However it does not play a role in determining the demand of public transport. Price elasticities for petrol cars and buses are -0.61 and -0.26, respectively, although the own price elasticity for diesel cars was not found to be insignificant. Only one of the cross elasticities for petrol cars is significant. With a positive value of 0.44, the estimated coefficient on $\ln_{Diesel}$ implies that when the other variables are unchanged, for every 1% increase in the cost of using diesel cars, the demand for its substitute, petrol cars, will increase by 0.44% on an average base.
In the case of diesel cars, the cross price elasticity between public transport and diesel cars is negative and significant, this unexpected sign might due to the inclusion of taxi fares, shuttle fares, and car hire charges fares in the bus fare index $ln_{Bus}$. While for the demand of buses, cross price elasticity between diesel cars and buses is positive and significant. However the sign of $ln_{Petrol}$ is significantly negative, indicating the fact that buses have not been classified based on their fuel types, therefore the estimated coefficient on $ln_{Petrol}$ might be interpreted as the fuel cost for some petrol buses included in the dependent variable, rather than an indicator for cross elasticity.

Inflation and unemployment are found to be significant only in the demand for diesel cars equation, implying that the relative importance of these two macroeconomic variables in the demand for the other two road transport modes appears to be relatively minor. The age indicators showed similar trend in the private vehicle demand equations, where the sign on the young population is positive; while for the middle-aged and matured groups, the sign is negative. However for the demand for public transport, the sign on the estimated coefficients for all age groups is positive, suggesting that New Zealanders are inclined to use public transport as their road transport choice. Lastly, in terms of the validity of the SUR model, following a two-step cointegration test suggested by Engle and Granger (1987), we conclude that the non-stationary time series are cointegrated.

The above empirical results from the SUR model also delivers some important policy implications. First of all, in order to achieve a reduction in the demand for automobiles, different policies could be implemented for cars with different fuel types, as the factors that affect the demand for these two major types of private vehicles in New Zealand differ significantly. For achieving a decline in the demand for petrol cars, policy makers could consider increasing the petrol tax; on the other hand, for achieving a reduction in the demand for diesel cars, policy makers could consider levelling up the taxation on vehicle-related ongoing costs, such as the vehicle registration fees, annual licensing fees, administration fees, and road user charges for diesel vehicles. For the purpose of
promoting the use of public transport, policy makers could consider lowering the fares by granting a subsidy to public transport providers. This of course, will have to be financed; possibly by recycling the revenue from fuel and car taxation.

One limitation in this study is that there is possibly an endogeneity issue with petrol and diesel prices as explanatory variables in petrol and diesel car regressions. Future research could consider the use of an instrument such as price of other petroleum based products like kerosene, as it would not be correlated with unobservables affecting vehicle demand, but correlated with fuel prices. Additionally, there are a few recommendations for possible future research. Firstly, in regards to the variables, road density and also road length may be more useful indicators as long as a sufficient quantity of quarterly (not annually) data become available. Secondly, the current study only modelled the total VKT by all types of buses as a measure of road passenger’s demand for public transport. In fact, total VKT by bus represents both the VKT by urban/suburban buses and special use buses (e.g. tourist buses) in New Zealand. These two different types of buses, in nature, should exhibit dissimilar trends and different fluctuation patterns. Thus with the availability of data, separating one of the dependent variables \( \text{Bus}_\text{VKT} \) for passenger bus only may presumably result in better estimations. Thirdly, since public transport can be viewed as a substitute for (at least some types of) private car transport, a sensible course of action when modelling road passenger transport demand would be to investigate regional, rather than national demand patterns. Regional analysis could be explored by extending the existing SUR model to a spatial environment so that potential spatial effects can be incorporated via autocorrelation in spatial error terms. Last but not least, the main purpose of this study is not to predict future road passenger travel demand; the model presented in this research can be used to derive demand implications and construct a forecasting model for private and public transport, given detailed assumptions about energy and economic conditions. In that case, we can thus project and compare energy consumption and CO\(_2\) emissions from both private vehicles and public transport in the future.
Chapter 6. Summary, Conclusions and Directions for Future Research

6.1. Research objectives

It has been commonly recognised that transport congestion in the major cities of New Zealand such as Auckland has become a severe issue with the dramatic increase of motorised transportation since World War II. At present, the Auckland transport network is characterised by a relatively high level of congestion. Given that New Zealand ratified the Kyoto Protocol in 1997, cutting down carbon dioxide equivalent (CO2-e) emissions has become a priority in contemporary government policy. Reducing CO2-e emissions from domestic transport thus plays an important role because this sector has a profound influence over the country’s emission profile. In order to address the issues around congestion and climate change on a long-term and sustainable basis, government should focus *inter alia* on discouraging car drivers and promoting the use of public transport on the road. For such policies to be effective it is important that we understand: whether there is a spatial dependency in commuter’s travel behaviours; what factors influence aggregate bus patronage for a given region; and what determine the transport mode choices for individual traveller.

The overall objective of this thesis is to firstly conduct a spatial econometric analysis of a particular type of travel: commuter’s journey-to-work (JTW) behaviour, from the aspect of aggregate public transport demand at the regional level to disaggregate transport mode choice decision making at the individual level, in the Auckland region. The second objective of this thesis is to analyse the demand for different road passenger transport choices as a system of equations at the national level. The intuition behind this is that given the fact that these road passenger transport modes are considered substitutes to one another, there is a strong possibility that an interrelationship exists between the travel demand functions, primarily due to the correlation between their disturbances.
The thesis outlined four main research questions to be addressed:

1. What is the research gap from past literature on traveller’s travel behaviour?

2. At the regional level, how should we examine the impact of urban form variables on aggregate public transport demand?

3. At the individual level, what is the probability that a commuter will choose to use public transport to go to work given the transport mode preference of his/her neighbours and the characteristics of the regions where he/she lives?

4. At the national level, whether there is a correlation between the disturbances of the demand for private and public road transportation? What are the factors that affect the demand for each road passenger transport mode in New Zealand?

6.2. Results summary

The first research question has been addressed in Chapter 2. Three research gaps from past literature on traveller’s travel behaviour have been identified. The first one is the ignorance of spatial independence at regional level, the second is the unawareness of positive network effects at individual level and the last one is the obliviousness of the correlation between the error terms for the demand for different road passenger transport mode at the national level. These three general limitations from past literature led us to the application of the spatial econometric models and seemingly unrelated regressions when conducting travel behaviour analysis.

6.2.1. The influence of urban forms on transit behaviour

The second research question asked how we should examine the impact of urban form variables on aggregate public transport demand. With the application of a spatial Durbin model, Chapter 3 estimated how urban form variables were related to bus mode share and
how these effects varied across the Auckland region’s landscapes, based on area unit level data.

The Moran’s $I$ statistic showed a positive value of 18.733 with a p-value that is lower than 0.0001. As expected, this result indicated that the null hypothesis of no spatial dependence should be rejected. In other words, the outcome of Moran’s $I$ test confirmed that there is statistically significant evidence of the presence of positive spatial autocorrelation. This empirical outcome implied that the bus mode share in one area unit exhibits a positive relationship with bus mode share in neighbouring area units. Therefore, by taking into account the spatial dependence effect, at this stage, spatial regression models, namely, the SAR and/or the SEM models were selected over the non-spatial OLS model in order to obtain unbiased and consistent estimators.

By applying the likelihood ratio tests, this chapter further confirmed the existence of spatial autocorrelation in the lags of both dependent and independent variables. Thus, the “mixed” spatial Durbin model, which was viewed as an unrestricted model of either SEM or SAR, was estimated as the final spatial model for the aggregate public transport demand analysis.

Empirical outcomes from SDM showed that the total effects comprised mostly of spatial spill-over impacts, and only a relatively small percentage was attributed to the direct effects on bus mode share that arose from own-region changes in most given explanatory variable. For planners and developers, the SDM model was not only technically superior, but also regarded as a preferable way for evaluating policies and making investment decisions. Unlike traditional estimated coefficient interpretations, one could easily unravel total effect into own-region and spatial spill-over effect. The results presented in this chapter highlighted the complexity and importance of the consideration of spatial structure in determining the factors that influence the bus mode share.
6.2.2. Commuter's transport mode preferences and network effects

Chapter 4 answered the third research question through the application of a spatial econometrics approach, namely, the spatial autoregressive logit model. The objective of this chapter was to analyse individual traveller’s transport mode choice, and to understand preference heterogeneity and network effects in public transport among JTW commuters in the Auckland region.

This chapter examined the spatial sampling issue by firstly developing a row-standardised, inverse-distance spatial weights matrix \( W \), with a dimension of \( n \times n \), using the geographic coordinates of each observation. Moran’s \( I \) statistics were computed after the construction of spatial weights matrix, as the first step to detect spatial autocorrelation. Outcomes from this test showed that these statistics were both positive and highly significant, with p-values that are equal to zero, which confirmed that there is statistically significant evidence of the presence of positive spatial autocorrelation.

A likelihood ratios test was applied to test the null hypothesis if the network effects parameter \( \rho \) was equal to zero. The LR-test statistic revealed that the spatial autocorrelation parameter \( \rho \) was positive and significantly different from zero, implying that taking public transport for JTW trips exhibited network effects among nearby households. In other words, with the presence of network effects, the process of individual transport mode choice decision-making not only involved personal characteristics, mode-specific information, but also information on the transport mode decisions of the individual commuter’s neighbours. The spatial autoregressive logit model provided strong empirical evidence that transport mode choice models do suffer from spatial autocorrelation. Additionally, this output also indicated that the estimated coefficients of a standard logit model which did not reflect spatial dependence would become biased due to omission of relevant variables, leading to biased WTP estimates.
All of the variables in the spatially autoregressive logit model were found to be highly significant with expected signs. Household size, number of household vehicles, variations in household disposable incomes, gender roles, employment status, age differences, along with total travel distance, trip duration time and the dummy variable which represents the destination to Auckland city centre all had significant influences on commuter’s transport mode choices.

6.2.3. Aggregate road passenger travel demand in New Zealand: a seemingly unrelated regression model

The last question posed in this thesis is to discover firstly whether there is a correlation between the disturbances of the demand for private and public road transportation. And secondly identify the factors that affect the demand for each road passenger transport mode in New Zealand. This paper is the first use of SUR in New Zealand to develop an aggregate road passenger travel demand model.

First of all, the Breusch-Pagan test of independence confirms the existence of correlated error terms of the three demand equations and the empirical results from SUR model indicates that per-capita GDP has an effect on both the demand for petrol and diesel cars, with the values of income elasticity to be 3.99 and 3.45, respectively. However it does not play a role in determining the demand of public transport. Price elasticities for petrol cars and buses are -0.61 and -0.26, respectively, although the own price elasticity for diesel cars was not found to be insignificant. Only one of the cross elasticities for petrol cars is significant. The cross price elasticity between diesel cars and petrol cars was 0.44, and between diesel cars and buses was 0.24. Inflation and unemployment were found to be significant only in the demand for diesel cars equation. The age indicators showed a similar trend in the private vehicle demand equations. However for the demand for public transport, the sign on the estimated coefficients for all age groups was positive, suggesting that all New Zealanders are inclined to use public transport as their road transport choice.
The above empirical results from the three empirical chapters also deliver some important policy implications. Firstly as Allison et al. (2013) justify, a traditional assumption is that transport passengers take account of the value of the trip to themselves, but not the value of their trip to others on public transport networks. Greater use of a public transport network can convey benefits to others through, for instance, enabling the frequency of services to increase, and thus reduce the service waiting times of existing commuters. Secondly, Goetzke (2008) states that “positive network effects exist when people prefer to use transit together with other people as a result of social spill-over”. In contrast with the negative externalities that road transport generates, the fact that public transport patronage exhibits positive network effects may also have some influential policy implications. It is typical that travel demand forecasting models are the hinge for both transport planners and policy decision-makers when evaluating new transit projects. However, conventionally, these models do not include a variable that explicitly captures the network effects to account for spatial dependency. This non-inclusion of network effects in the stage of model formulation causes an omitted variable bias which becomes visible in the mode-specific constant term, as revealed in Goetzke (2003).

Recall the results of Chapter 3 and 4, the positive and statistically significant coefficient of the spatial lag variable, $\rho$, delivers two significant implications: first, network effects do exist within the data at both the regional and individual levels. Second, the probability that a commuter chooses to take public transport to go to work increases when his/her neighbours have a high propensity to do so. In other words, the spatial autoregressive logit model shows that taking public transport to work does not only depend on personal and trip specific features, but also on the spatial interactive variable. The probability of using public transport increases with a higher public transport mode share, because of network effects which derive from social interactions between neighbours. If the network effect is present however, this systematic forecasting error simply implies that the public ridership in suburbs with low public transport mode share will be overestimated; simultaneously, public transport ridership in the central Auckland with high public transport mode share
will be underestimated. Transport planners and authorities should, therefore, take network effects into consideration when conducting mode choice analysis.

Although building more roads can reduce congestion by increasing traffic capacity in short terms, the benefits are generally temporary: motor vehicles will soon fill up newly built transit lanes, and the vicious cycle starts all over again (Duranton and Turner, 2011). As Maddison et al. (1996) argue, current transport policies which solely rely on increasing road construction will be ineffective in resolving severe congestion issues. In order to reach a stabilisation of GHG emissions from transport and tackle the climate change problem subsequently, behavioural change brought about by appropriate policies will be a compulsory part (Chapman, 2007). This means for transport planners and policymakers, a new tendency should be given to “soft” behavioural change related policies, such as more promotion and encouragement of using public transport in communities, given the fact that a commuter’s travel mode choice is largely affected by what his/her neighbours’ travel decision, due to the presence of positive network effects.

Lastly, in order to achieve a reduction in the demand for automobiles nationwide, different policies could be implemented for cars with different fuel types, as the factors that affect the demand for these two major types of private vehicles in New Zealand differ significantly. For achieving a decline in the demand for petrol cars, policy makers could consider increasing the petrol tax; on the other hand, for achieving a reduction in the demand for diesel cars, policy makers could consider levelling up the taxation on vehicle-related ongoing costs, such as the vehicle registration fees, annual licensing fees, administration fees, and road user charges for diesel vehicles. For the purpose of promoting the use of public transport, policy makers could consider lowering the fares by granting a subsidy to public transport providers. This of course, will have to be financed; possibly by recycling the revenue from fuel and car taxation.
In conclusion, several key implications for policy makers, such as the NZTA are summarised as follows:

Firstly in order to promote the use of public transport:

- Increasing the number of buses running during morning and afternoon peak hours to achieve greater patronage;
- Introducing more “soft” behavioural change related policies such as more promotion and encouragement of using public transport in communities, given the fact that a commuter’s travel mode choice is largely affected by what his/her neighbours’ travel decision, empirically confirmed by a positive value of $\rho$ at both aggregate and individual level;
- Reducing the fares that public transport users have to pay though a subsidy to public transport providers. The subsidy could possibly be financed through recycling the revenue from fuel & car taxation.

Secondly in order to reduce the demand for cars nationwide:

- For petrol cars, increasing the petrol tax;
- For diesel cars, increasing taxation on vehicle-related ongoing costs (i.e. vehicle registration fees, annual licensing fees, administration fees, and road user charges for diesel vehicles).

**6.3. Suggestions for future research**

Because there are only 820 observations for commuter JTW trips between 2005/06 to 2008/09 in chapter 4, estimated results derived for commuter JTW trips could be data specific as they are relatively sensitive to the presence of outliers. Using more data in future analysis could produce more robust results.
Regarding model methodology, another approach other than the ML approach used in the empirical chapters 3 and 4, Bayesian estimation method, could be applied in future research and compare with the performance of the ML estimation. As Leenders (2002) illustrates, one major challenge facing spatial econometric models is that the spatial weights matrix W cannot be directly estimated but needs to be explicitly specified in prior. Moreover, current economic theory which underlies the empirical applications of spatial econometric generally has no formal specification routines of W. As a result, it has become common in practice to explore if the estimated results are robust to the specification of the spatial weights matrix (Elhorst, 2010). More importantly, as demonstrated by Stetzer (1982), a misspecified weights matrix may lead to inconsistent estimates and thus misleading inference, for that reason, a comparable study with regard to spatial weights choice is necessary.

Nevertheless, recall a recent work in Pijnenburg and Kholodilin (2011), spatial models that are estimated with ML lack the ability of performing such comparisons because the likelihood ratio test, which uses the criterion of log likelihood function values in model selection, can only be applied for nested models. While in fact, the two models with different weight matrices are not considered as being nested. For that reason, by applying Bayesian estimation methods, Bayesian posterior model probability has the advantage of allowing such comparison of these non-nested models using various weight matrices. Therefore, it will be rather appealing to examine several alternatives to the inverse distance W used in chapter 3 & 4 (i.e. Rook and Queen continuity, kth nearest neighbours, etc.), with the aim of testing whether the estimation results are sensitive to the choice of the spatial weights matrix based on the Bayesian estimation method.

Lastly, for Chapter 5, in regards to the variables, road density and also road length may be more useful indicators as long as a sufficient quantity of quarterly (not annually) data become available. Secondly, the current study only modelled the total VKT by all types of buses as a measure of road passenger’s demand for public transport. In fact, total VKT by
bus represents both the VKT by urban/suburban buses and special use buses (e.g. tourist buses) in New Zealand. These two different types of buses, in nature, should exhibit dissimilar trends and different fluctuation patterns. Thus with the availability of data, separating one of the dependent variables Bus_VKT for passenger bus only may presumably result in better estimations. Thirdly, since public transport can be viewed as a substitute for (at least some types of) private car transport, a sensible course of action when modelling road passenger transport demand would be to investigate regional, rather than national demand patterns. Regional analysis could be explored by extending the existing SUR model to a spatial environment so that potential spatial effects can be incorporated via autocorrelation in spatial error terms. Last but not least, the main purpose of this study is not to predict future road passenger travel demand; the model presented in this research can be used to derive demand implications and construct a forecasting model for private and public transport, given detailed assumptions about energy and economic conditions. In that case, we can thus project and compare energy consumption and CO₂ emissions from both private vehicles and public transport in the future.
### Appendix A Descriptive Statistics for the Subgroup when *Choice* = 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Choice</em> = 1, obs = 62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>HHsize</em></td>
<td>3.226</td>
<td>1.442</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td><em>HHvehicle</em></td>
<td>1.742</td>
<td>1.100</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><em>Higher-income</em></td>
<td>0.323</td>
<td>0.471</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><em>Female</em></td>
<td>0.645</td>
<td>0.482</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><em>Full-time</em></td>
<td>0.962</td>
<td>0.216</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><em>Young</em></td>
<td>0.306</td>
<td>0.464</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><em>Old</em></td>
<td>0.048</td>
<td>0.216</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Trip characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Distance (kilometres)</em></td>
<td>8.630</td>
<td>11.251</td>
<td>0</td>
<td>30.037</td>
</tr>
<tr>
<td><em>Duration (minutes)</em></td>
<td>23.210</td>
<td>11.251</td>
<td>1</td>
<td>55</td>
</tr>
<tr>
<td><em>Destination_AKL</em></td>
<td>0.484</td>
<td>0.504</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix B Geocoded Destination Locations of All Included Trips
Appendix C log-value of variables vs. Time
Appendix D Residuals vs. Time

Residuals: ln Car_P

Residuals: ln Car_D

Residuals: ln Bus_VKT
Appendix E Predicted vs. Actual Demands

Predicted versus Actual Demand for Petrol Cars

Predicted versus Actual Demand for Diesel Cars

Predicted versus Actual Demand for Buses
List of References


Dissanayake, D., & Morikawa, T. (2010). Investigating household vehicle ownership, mode choice and trip sharing decisions using a combined revealed


http://www.nzta.govt.nz/about/media/releases/2153/news.html


