Appraisal of factors influencing public transport patronage

January 2011

Judith Wang
Booz & Company
Auckland, New Zealand

NZ Transport Agency research report 434

ISBN 978-0-478-35152-9 (print)
ISBN 978-0-478-37151-2 (electronic)
ISSN 1173-3756 (print)
ISSN 1173-3764 (electronic)

NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
research@nzta.govt.nz
www.nzta.govt.nz

Wang, J (2011) Appraisal of factors influencing public transport patronage. *NZ Transport Agency research report no.434.* 176pp.

Keywords: elasticities, public transport demand forecasting, time series data analysis

An important note for the reader

Acknowledgements

The author would like to thank NZ Transport Agency, NZ Statistics, Auckland Regional Transport Authority, Greater Wellington Regional Council and Environment Canterbury for providing the data and their support throughout this project; Richard Scott of Booz & Company, Dr Erwann Sbai of the Department of Economics at The University of Auckland, John Davies of Auckland Regional Council, Dr Andrew Coleman of Motu Economics and Public Policy Research, and Professor Mark Greer of Dowling College for their invaluable advice.

Abbreviations and acronyms

ARTA	Auckland Regional Transport Authority
CPI	Consumer Price Index
ECan	Environment Canterbury
FY	financial year
GWRC	Greater Wellington Regional Council
MED	Ministry of Economic Development
NZ	New Zealand
NZD	New Zealand dollars
NZSTAT	New Zealand Statistics
NZTA	New Zealand Transport Agency
PAM	partial adjustment model
PT	public transport (or passenger transport)
Q1	the first quarter of a year
Q2	the second quarter of a year
Q3	the third quarter of a year
Q4	the fourth quarter of a year
VECM	vector error correlation model

Contents

Executive summary

This report summarises the results of a research project commissioned by the NZ Transport Agency in 2007 and carried out by Booz & Company in 2008/09. The overall objective of the study was to undertake an in-depth analysis of factors influencing public transport patronage and to develop a model for future forecasting of patronage demands. Econometric analyses were applied to annual and quarterly national and regional aggregate data for three major regions in the country: Auckland, Wellington and Canterbury. The following objectives were achieved:

1 Identification of the key factors affecting public transport patronage.

2 Estimation of the elasticities with respect to each of the key factors identified.

3 Development of forecasting models for use by transport operators and transport funding agencies.

The historical trends of public transport patronage for the last decade were explored. The stories behind the changes in trends were revealed and the influencing factors that might have contributed to the changes were identified. A short-term and long-term forecast model was determined for each major transport mode in each region. Six variables were considered in the model as follows:

Dependent variable

1 Patronage (in trips per capita)

Economic determinants

2 Service level (in bus/train kilometre per capita)

3 Real fare (in real revenue per passenger)

4 Real income (in real disposable income per capita)

Structural determinants

5 Car ownership (in cars per capita by region)

6 Real fuel price

Table ES.1 Summary of factors influencing public transport patronage in New Zealand

	Auckland		Wellington		Christchurch bus
	Bus	Rail	Bus	Rail	
Service	positive	positive	n/a	positive	positive
Fare	n/a	negative	negative	n/a	negative
Car ownership	negative	n/a	n/a	negative	n/a
Income	n/a	positive	n/a	negative	n/a
Fuel price	positive	n/a	n/a	positive	positive

Table ES.2 Bus demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland 2003Q3–2008Q2		Wellington		Christchurch	
	short-run	long-run	short-run	long-run	short-run	long-run
Bus service	0.46	0.73	n/a	n/a	0.07	0.62
Bus fare	n/a	n/a	-0.23	-0.46	-0.26	-0.34
Car ownership	-1.96	-3.10	n/a	n/a	n/a	n/a
Income	n/a	n/a	n/a	n/a	n/a	n/a
Fuel price	0.20	0.32	n/a	n/a	0.28	0.37

Table ES.3 Rail demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Rail service	0.99	0.88	1.41	1.63	0.74	0.78	2.39	2.53
Rail fare	-0.97	-0.68	-1.37	-1.25	n/a	n/a	n/a	n/a
Car ownership	n/a	n/a	n/a	n/a	-0.32	n/a	-1.04	n/a
Income	1.61	n/a	2.28	n/a	n/a	-0.22	n/a	-0.70
Fuel price	n/a	n/a	n/a	n/a	0.13	0.12	0.42	0.39

The factors with significant effect on public transport patronage in the three major cities in the last decade were identified as shown in table ES.1. The corresponding best estimate(s) of demand elasticity with respect to the factors identified are summarised in tables ES.2 and ES.3.

The key findings in this study are summarised as follows:

1 Service was identified as the key driving factor among all five factors considered. It had significant influence in all cities and in almost all modes except Wellington bus, and the elasticity estimates were all positive. This was a very encouraging result as it meant the investment in public transport infrastructure, new services and service improvement in the last decade did have a positive influence on public transport patronage in all cities.

2 Fare also had significant influence in all three cities, although not on all modes, and the influence was negative. Bus fare was a significant influencing factor in both Wellington and Christchurch but not in Auckland. As Auckland has always been the biggest city in New Zealand, it has a higher proportion of public transport dependent population. As a result, it appeared that fare did not have a significant influence on bus patronage in Auckland.

3 Rail service and fare elasticities were higher than the corresponding estimates for bus demand in Auckland. Auckland's public transport system, especially the rail system, had improved tremendously over the last decade. As a result, the service and fare elasticities were higher (more elastic) than in other cities.

4 Car ownership had significant negative influences on the demand for Auckland bus and Wellington rail but fare was not an influencing factor in these cities. The car ownership elasticity estimates were also far higher than the estimates from international experience.

5 Income was found to have a positive effect on Auckland rail patronage and on the contrary a negative effect on Wellington rail patronage, while international estimates were negative. The observed unusual positive income effect on Auckland rail could be explained by the change in rail market as a result of significant investment in infrastructure and service improvement. The opening of Britomart in 2003 induced this effect as the central business district (CBD) employment area became within walking distance of the new station. A higher proportion of commuters with higher income were attracted to use rail service. This group also appeared to be more sensitive to quality of service and hence had higher service elasticities.

6 The fluctuation in fuel price in recent years had a positive impact on public transport patronage in all three cities, although not on all modes. In Auckland and Christchurch, the influence on bus patronage was significant. Christchurch had the highest car ownership per capita among the three cities but relatively cheaper bus fares and a more convenient ticketing system. As a result, the estimated fuel price elasticity in Christchurch was higher than in Auckland and Wellington. This implied a higher substitution effect between bus and car in Christchurch compared with other cities.

7 Car ownership was the most elastic among all the factors identified and fuel price was found to have significant influence on Auckland and Christchurch bus patronage but only in the last four or five years. That means in the long run, the most effective policy to encourage use of public transport could be by controlling car ownership or its use. Fuel price, as part of motoring cost, could have an influence in the short run as well.

8 Despite the increase in fares as a result of increases in diesel price, the increase in fuel price was more significant. In other words, public transport was still relatively ‘cheap’ compared with driving. As a result, the increase in patronage was influenced by the increase in fuel price but not influenced by fares for Auckland bus and Wellington rail.

9 The market in Wellington was quite different from Auckland and Christchurch, as Wellington had a more mature commuting market. Wellington had the highest public transport use among the three cities. Wellington also had the highest walk modal share as the CBD employment area is more compact and walkable. On top of that, the council also had a committed parking restraint policy in the CBD. Hence the elasticity estimates for Wellington were generally lower (less elastic) than those for Auckland and Christchurch.

10 Our methodology and analysis in the study was limited by data availability. The missing or inaccurate information could have implications on the quality of the models estimated but we made the most of the information collected. Three technical issues were identified during the peer review process, namely the ‘transport services offered variable’, spurious regression and series length. The referee reports where these technical issues were raised and the responses to them are attached in appendix E.

Abstract

This project examined the demand for local bus and rail services during the period 1996–2008 in the three major cities in New Zealand: Auckland, Wellington and Christchurch. In order to determine the drivers behind changes in public transport ridership over time, econometric analysis techniques were applied to analyse the time series data of patronage of major public transport mode(s) in the three cities, collected for the last decade. A dynamic model was identified for each city by mode relating per capita patronage to fares, service level, car ownership, income and fuel price. The results indicated the three cities all had different characteristics and the drivers behind the long-run and short-run trends were also different. It also appeared that the significant fluctuation in fuel price in recent years had a positive effect on public transport patronage in all three cities.

1 Introduction

1.1 Background

Booz & Company (formerly Booz Allen Hamilton) was commissioned in April 2006 to carry out econometric research into the impact of transport fuel price changes (Kennedy and Wallace 2007). Econometric analysis of the impact of petrol prices in New Zealand was conducted on the following three demand variables:

1. Petrol consumption
2. Highway traffic volumes
3. Public transport (PT) patronage.

From this previous study, statistically significant estimates of the first two direct elasticities were obtained successfully relating petrol price changes to petrol consumption and highway traffic volumes. However, the cross-price elasticity of PT patronage with respect to petrol prices could not be estimated with a satisfactory level of significance. Although, by intuition, one would expect that petrol prices might have a certain level of influence on PT patronage, it was clear there were other factors with more significant influences which had not been considered in the previous study.

Booz & Company was commissioned in 2007 to extend this research to include in-depth analysis of PT patronage. The objectives of this study were as follows:

1. To identify the key factors affecting PT patronage.
2. To estimate the elasticities with respect to each of the key factors identified.
3. To develop forecasting models for use by transport operators and transport funding agencies.

1.2 Scope of the study

The overall objective of this project was to build a sound basis for future forecasting of patronage demand. We explored trends in PT patronage and identified key factors that might have contributed to those trends. Econometric models were applied to estimate the relationships between PT patronage and key factors and hence estimate elasticities for each of the key factors identified. Based on these findings, we recommend econometric forecasting models that could be used by transport operators to explore the impact of changes in key factors on future patronage and revenue.

The analysis was based on annual and quarterly national and regional aggregate data for three major regions in the country:

1. Auckland region[1]
2. Wellington region[2]
3. Canterbury region[3]

[1] The Auckland region will also be referred to as Auckland city in this report.

[2] The Wellington region will also be referred to as Wellington city in this report.

[3] As the major city in the Canterbury region is Christchurch, it will be referred to as Christchurch at a city level in this report.

As the feasibility of econometric analysis is solely dependent on data availability, meaningful forecasting models can only be built if the relevant data is available. We looked into the development of the following models as appropriate:

1 Regional short-term and long-term forecast models

2 National short-term and long-term forecast models (if complete information on more than one region is available).

The above approach has the advantage of being able to capture the different characteristics of different regions. One would expect that the elasticities of different regions with respect to the same factor could vary due to differences in infrastructure provision, urban form, users' perception of PT and other relevant factors.

As elaborated later in this report, due to the differences in the nature of data collected for the three regions, we could only use the first approach in modelling, ie a regional approach. We applied dynamic econometric models so that the changes over time could be captured in the model. With this approach, the short-term and long-term elasticities could be determined simultaneously in one model.

1.3 Report structure

This report is organised as follows. Chapter 2 describes an overview of the study methodology. Chapter 3 summarises the literature review on methodology and the findings from the econometric modelling approach applied in the study. Chapter 4 describes the data acquisition and imputation processes and provides a comprehensive list of data acquired for the study. Chapter 5 provides a detailed preliminary analysis of the data collected. The relationships of the observed trends to the history of changes in the three regions were revealed. Chapter 6 describes the econometric forecast models developed. Chapter 7 compares the results of this study with international experience. Finally, the findings from this study and areas for further research are discussed in chapter 8.

2 Study methodology

2.1 Methodology

Figure 2.1 Overview of study methodology

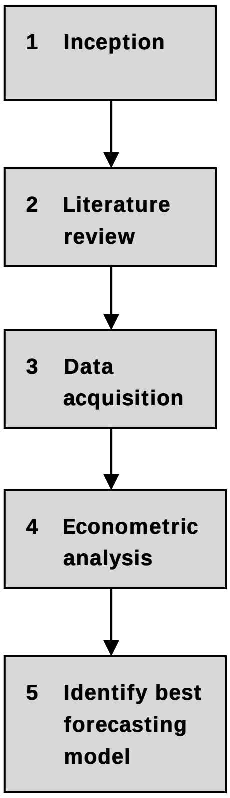


Figure 2.1 illustrates an overview of the study methodology. Details of the five stages are described as follows:

Stage 1 – Inception. This stage of the project involved developing a scoping paper for the proposed approach and then discussing modelling and statistical issues with the steering group and peer reviewers.

Stage 2 – Literature review. The primary purpose of the literature review was to identify key factors that had been incorporated into econometric models in the literature:

- Which key factors have been incorporated?
- How significant are those key factors?

The secondary purpose of the literature review was to identify elasticity estimates from international experience, for the purpose of comparison with our research.

- What are the elasticity estimates for those key factors?

Stage 3 – Data acquisition. The purpose of this stage was to obtain the information required for stages 4 and 5. We expected to obtain time series data on the following:

- PT patronage
- real PT fare and details of historical fare changes

- details of route changes and other disruptions/improvements to PT services
- service level (eg bus vehicle-km)
- real income
- car ownership
- real motoring costs (eg fuel price, parking charges)
- local population
- demographic variables (eg domestic students, overseas students, senior citizens, school age children)
- urban form and accessibility measures.

Stage 4 – Develop econometric models. The purpose of this stage was to apply appropriate econometric modelling techniques in order to identify the most relevant explanatory variables.

Stage 5 – Identify best forecasting model. The purpose of this stage was to identify the best forecasting model among the models developed in stage 4. The difference between the forecasts and actual observations of patronage was analysed in order to evaluate forecasting accuracies of different models.

2.2 Statistical packages

Most of the data and statistical analyses in this study were implemented in R version 2.8.1. R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical and graphical techniques. It was found to be the most suitable software for both preliminary data analysis as described in chapter 5 and econometric modelling as described in chapter 6. A time series analysis software, JMulTi version 4.23, was used instead of R for stationarity tests as described in chapter 6.

3 Econometric analysis techniques

3.1 Literature review

The literature review for this study had two main areas of focus:

1 To ascertain different methodologies

2 To enable a comparison of results from this study with international experience.

In the literature review, we focused on the different methodologies applied in patronage demand analysis. In addition, recent studies on the determinants of the demand for PT, including a study on public bus ridership in Auckland conducted by Greer (2008) of the Energy Centre of the University of Auckland based on the Census 2006 data, were also reviewed. Results from international experience are discussed in chapter 7 of this report.

For more than 50 years, extensive effort has been devoted in the world of research and practice to analysing the impact of changes in fares, service supply, income and other factors on the demand for PT. The key techniques applied in estimates of PT elasticities can be categorised as follows:

1 Empirical analysis (eg Paulley et al 2006; Litman 2004; Goodwin 1992; Oum et al 1992) provides some classic reviews of empirical evidence from studies undertaken prior to 1990.

2 Stated preference or combined revealed preference/stated preference (eg Espino et al 2007; Hensher and King 1998).

3 Econometric analysis (eg Bresson et al 2003; Garcia-Ferrer et al 2006).

As we applied econometric analysis to build our forecasting models in this project, in this part of our literature review, we focused on case studies in other countries where econometric techniques had been applied.

The most relevant to this research project is a study commissioned by the UK Department for Transport in 1998 and conducted by Dargay and Hanly (2002a and 2002b). The main aim of the project was to obtain fare elasticity estimates for use in policy calculations of the projected change in bus patronage from a given average change in fare. The analysis was based on annual national and regional data. The data for the UK as a whole covered the financial years (FY) 1974/75 to 1996/97, whereas data for the regions was limited to the periods 1986/87 to 1996/97. Time series analyses were conducted to take into account the intertemporal nature of the data being used. The basic model developed related per capita bus patronage (all journeys) to real per capita income, real bus fares and service level (bus vehicle-kilometres). In addition, a structural model was estimated to test the interaction of bus patronage, motoring costs and car ownership.

Around the same time, a French research project (Madre and Boulahbal 1999; Bresson and Pirotte 1999), financed by VIA-GTI and the French Department of Transport, was conducted to investigate the effect on PT use of changes in population structure, urban sprawl and increasing car ownership. Unlike the British study, which was only concerned with bus travel, the French study covered all PT modes for urban travel. Bresson et al (2003) combined the data collected for the two projects, ie on the basis of panels[4] of English counties and French urban areas, to analyse the impact of changes in fares, service supply, income and

[4] Panel data, also called longitudinal data, is data for multiple entities in which each entity is observed at two or more time periods.

other factors on the demand for PT. The analysis was based on dynamic econometric models so that both short- and long-term elasticities could be estimated.

More recently, a Spanish research project, financed by the Comunidad Autónoma de Madrid, was carried out by García-Ferrer et al (2006) to analyse data on the choice of alternative PT modes in the Madrid metropolitan area. Monthly data was obtained and econometric analysis was applied to estimate user response to changes in prices and characteristics of the services so a more reliable prediction of demand could be obtained. As in other studies in the literature, dynamic econometric modelling techniques were applied in order to capture the non-stationary characteristics of the data.

Bresson et al (2004) explored the economic and structural determinants of the demand for PT based on a panel data analysis of annual time series from 1975 to 1995 for 62 urban areas in France. Three economic determinants were considered – vehicle-kilometres, income and price. Public transport was found to be an 'inferior good' as the estimated income elasticity was negative. By synthesising the structural determinants (including population ageing, urban sprawl and growing car ownership) in a single indicator, there was an interesting discovery that the 'income effect' was in fact mainly a 'motorisation effect'. It was concluded that the downward trend in PT patronage was mainly due to increasing car ownership. It was also observed that the use of PT was quite sensitive to the service level and its price.

Hensher (2008) analysed the direct elasticities associated with PT demand with respect to changes in three factors: fares, in-vehicle time and headway, based on information from 319 studies. The major influences identified were: time of day (peak, all day versus off-peak), data paradigm (especially combined stated preference/revealed preference versus revealed preference), whether an average fare or class of tickets was included, the unit of analysis (trips versus vehicle-kilometres), specific trip purposes, country and specific-mode (ie bus and train).

Greer (2008) carried out a spatial analysis of cross-sectional data on bus ridership by commuters to work and other related statistics for the Auckland region collected during the 2006 Census. The results from a regression analysis indicated that the following four factors had a positive effect on commuter bus ridership:

1 The total number of commuters from the area unit
2 The distance from the centre of the area unit to the nearest rail or ferry terminal
3 The population density of the area unit
4 The combined morning and evening peak-hour bus service frequency within the area unit.

On the other hand, the following three factors were found to have a negative influence on commuter bus ridership:

1 The average number of cars available to a household within the area unit
2 The distance from the centre of the area unit to the city centre
3 Median household income within the area unit.

Further discussion on the outcomes of international studies compared with New Zealand studies is included in chapter 7 of this report.

3.2 Econometric modelling approach

The data we collected for this study was time series data – data collected for a single entity, eg patronage, at multiple points in time. We wanted to know what the causal effect was on our variable of interest, ie PT patronage (Y), and the effect of a change in another variable, eg fare (X), over time. In other words, what was the dynamic causal effect on Y of a change in X? Another question was, what would be our best forecast of Y at a future date, say in the next quarter? Both these questions could be answered using time series data.

As stated in Stock and Watson (2007), one important assumption in time series regression is the assumption of whether the future will be like the past or not. 'Stationarity' occurs when the future is like the past. That is, a time series is stationary if its probability distribution does not change over time. Based on our findings in the literature review on methodologies, as detailed in section 3.1, we expected stationarity to be very unlikely. The two most common types of non-stationarity are trends and breaks. A trend is a persistent long-term movement of a variable over time and it can be either deterministic or stochastic. A deterministic trend is a non-random function of time. For example, a deterministic trend might be linear in time. In contrast, a stochastic trend is random and varies over time. In general it is more appropriate to model economic time series as having stochastic rather than deterministic trends. The second type of non-stationarity, breaks, can occur for a variety of reasons, such as changes in economic policy. Statistical analysis on the existence of trends and breaks was performed to justify whether the application of dynamic techniques would be necessary.

Dynamic econometric models estimate relationships between explanatory and dependent variables over periods of time, where the concept of time plays a more central role. For example, a dynamic model might have the following form:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots + \beta_p X_{t-p} + u_t \qquad \text{(Equation 3.1)}$$

where Y_t is the dependent variable; X_t is the current value of the independent variable; $X_{t-k}, k \in [1, p]$ is the past value of X_t from k periods before; and u_t is the error term, which represents measurement error and/or omitted factors.

As shown in equation 3.1, in this model, the dependent variable Y_t does not only depend on the current value of X_t, but also on past (lagged) values of X_t.

The two models most commonly used in the literature and applied successfully in Europe (eg Dargay and Hanly 2002a and 2002b in England; Bresson et al 2003 in France and England; García-Ferrer et al 2006 in Spain) to model the demand for PT were: 1) the partial adjustment model (PAM); and 2) the vector error correlation model (VECM). Employing a dynamic methodology enabled us to distinguish between the short- and long-term impact of changes such as fare changes on patronage. It also provided an indication of the time required for the total response to be complete.

3.2.1 Partial adjustment model (PAM)

PAM assumes a geometrically declining adjustment process. The idea behind PAM is that an individual's travel behaviour to a certain extent is based on habit. One's choices today have an effect on one's future decisions. This is modelled by introducing the lagged independent variable on the right-hand side of the equation and the adjustment coefficient. For example, a PAM might be in the following form[5]:

$$Y_t = \beta_1 \lambda + (1 - \lambda) Y_{t-1} + \beta_2 \lambda X_t + u_t \qquad \text{(Equation 3.2)}$$

where

a the short-run reaction of Y to a unit change in X is $\beta_2 \lambda$

[5] Extracted from Asteriou (2006), equation 11.20 on p223.

b the long-run reaction is given by β_1

c an estimate of β_1 can be obtained by dividing the estimate of $\beta_2\lambda$ by one minus the estimate of $(1-\lambda)$.

Further details of the derivation of equation 3.2 are contained in Asteriou (2006), p223.

3.2.2 Vector error correlation model (VECM)

Sometimes two or more series have the same stochastic trend in common. In this special case, referred to as cointegration, regression analysis can reveal long-run relationships among time series variables. VECM is a model we can apply in this case.

For example, if X_t and Y_t are cointegrated, a VECM[6] might be in the following form:

$$\Delta Y_t = \beta_{10} + \beta_{11}\Delta Y_{t-1} + ... + \beta_{1p}\Delta Y_{t-p} + \gamma_{11}\Delta X_{t-1} + ... + \gamma_{1p}\Delta X_{t-p} + \alpha_1\left(Y_{t-1} - \theta X_{t-1}\right) + u_{1t}$$ Equation 3.3

$$\Delta X_t = \beta_{20} + \beta_{21}\Delta Y_{t-1} + ... + \beta_{2p}\Delta Y_{t-p} + \gamma_{21}\Delta X_{2-1} + ... + \gamma_{2p}\Delta X_{t-p} + \alpha_2\left(Y_{t-1} - \theta X_{t-1}\right) + u_{2t}$$ Equation 3.4

The term $\left(Y_t - \theta X_t\right)$ is called the error correction term. The combined model in equations 3.3 and 3.4 is called a VECM. In a VECM, past values of $\left(Y_t - \theta X_t\right)$ help to predict future values of ΔY_t and/or ΔX_t.

As illustrated by equations 3.3 and 3.4, a VECM models the interaction between variables over time with a set of simultaneous equations. In this example, we have two equations because we are only considering two variables, Y_t and X_t. The number of parameters is dependent on the number of lags and the number of variables being considered.

Both PAM and VECM have the capability to utilise time series information to measure long-run and short-run elasticities. Nevertheless, VECM is much more data intensive compared with PAM. As discussed in chapter 6, PAM was, therefore, the only ‘feasible’ model for all three cities due to insufficient information for Wellington and Christchurch.

[6] Extracted from Stock and Watson (2007) pp656–658.

4 Data acquisition

4.1 Data collection

One of our study objectives was to identify the factors influencing PT patronage over time and ultimately to develop forecasting models. Ideally we would need time series data over a time period as long as possible collected at a reasonable frequency. Drawing upon the international experience and previous studies in New Zealand (as summarised in the literature review), we identified the following factors that would potentially have a significant influence on per capita PT patronage in New Zealand.

Economic determinants

1 Service level (in bus-kilometres per capita)

2 Major service improvement (eg provision of Northern Busway in Auckland)

3 Real fare (in real revenue per passenger)

4 Real Income (in real disposable income per capita)

Structural determinants

5 Car ownership (in cars per capita by region)

6 Real motoring costs (eg fuel price, parking charges)

7 Resident population size (national and by region)

8 Demographic variables (eg domestic students, overseas students, senior citizen, school age children)

9 Urban form and accessibility measures (eg census data on journey to work)

We tried to collect information as detailed as possible and for as long as the information was available. From a modelling point of view, this would allow more flexibility and would enable the model to capture the influences of as many factors as possible. For instance, quarterly information enabled the analysis of seasonal variation as well as the influence of changes from one season to the next. Such detailed analysis would not be possible if only annual information was available.

The sources of information investigated included:

- the New Zealand Transport Agency (NZTA)
- regional authorities including: Auckland Regional Transport Authority (ARTA), Greater Wellington Regional Council (GWRC) and Environment Canterbury (ECan)
- New Zealand Statistics (NZSTAT)
- the Ministry of Economic Development (MED)
- the Ministry of Education

4.1.1 Major service improvement

In addition to the regional data set, information on the North Shore bus services in Auckland, including patronage, service and fare, was collected to analyse whether the opening of the Northern Busway had a significant impact on bus patronage in Auckland.

4.1.2 Real fare/real income/real motoring costs

All prices were converted to real prices, ie

$$\text{Real price} = \frac{\text{Price at nominal value *}}{\text{Consumer Price Index (CPI)}}$$

Note: * Nominal value refers to any price or value expressed in money of the day

4.1.3 Resident population size

For consistency, patronage, service level, income and car ownership were all measured on a per capita basis either at a national level or regional level.

4.1.4 Demographic variables/urban form and accessibility measures

Unfortunately, we were not able to obtain full information on all the determinants identified. Information collected for structural determinants (8) and (9) was either not available or not suitable for model development. For (8) demographic variables, international and domestic student statistics were available only nationally on an annual basis from 1994; and quarterly population statistics by age group were not available at a regional level. For (9) urban form and accessibility measures, we considered information on journey-to-work to be relevant to our study but the information was available only in census years at five-year intervals.

4.1.5 Consolidated set of variables

After consolidation, the regional data set included time series of only six variables as follows:

Dependent variable

1 Patronage (in trips per capita)

Economic determinants

2 Service level (in bus/train-kilometres per capita)

3 Real fare (in real revenue per passenger)

4 Real income (in real disposable income per capita)

Structural determinants

5 Car ownership (in cars per capita by region)

6 Real fuel price

4.2 Data availability

The details of information that we collected are listed in table 4.1.

Table 4.1 List and description of data collected

Information	Related variable(s)	Data	Description	Period and frequency of observations	Source
Patronage Auckland	(1) Patronage	Auckland bus patronage for contract services[7]	Total passenger trips per time period	Quarterly from 1996 Q1 to 2008 Q2	ARTA
		Auckland North Shore bus patronage for contract services	Total passenger trips per time period	Monthly from November 2005 to July 2008	ARTA
		Auckland rail patronage	Total passenger trips per time period	Quarterly from 1996 Q1 to 2008 Q2	ARTA
Patronage Wellington	(1) Patronage	Wellington bus patronage (estimated total)	Total passenger trips per time period	Annually from FY 1999/2000 to 2007/08 Quarterly from 2002 Q1	GWRC
		Wellington rail patronage	Total passenger trips per time period	Annually from FY 1999/2000 to 2007/08 Quarterly from 2002 Q1	GWRC
Patronage Christchurch	(1) Patronage	Christchurch PT patronage[8] (bus + ferry)	Total passenger trips per time period	Quarterly from 1997 Q1 to 2008 Q2	ECan
Service Auckland	(2) Service level	Auckland bus service kilometres for contract services	Total bus-kilometres per time period	Quarterly from 1996 Q1 to 2008 Q2	ARTA
		Auckland North Shore bus-kilometres for contract services	Total bus-kilometres per time period	Monthly from November 2005 to July 2008	ARTA
		Auckland rail service kilometres	Total train-kilometres[9] per time period	Quarterly from 1996 Q1 to 2008 Q2	ARTA

[7] For Auckland, only bus contract services data was available. Detailed commercial services information was not available due to commercial confidentiality issues.

[8] For Christchurch, ferry and commercial bus patronage could not be separated from the bus patronage information due to commercial confidentiality issues.

[9] Although there were different train compositions in Auckland (eg 2-car, 4-car and 6-car), the data available for rail services was available only in train-kilometres.

Information	Related variable(s)	Data	Description	Period and frequency of observations	Source
Service Wellington	(2) Service level	Wellington bus-kilometres	Total bus-kilometres per time period	Annually from FY 1999/2000 to 2007/08 Quarterly from 2002 Q1	GWRC
		Wellington rail patronage	Total car-kilometres per time period	Annually from FY 1999/ 2000 to 2007/08 Quarterly from 2002 Q1	GWRC
Service Christchurch	(2) Service level	Christchurch bus-kilometres[10]	Estimated weekly total bus-kilometres	Estimation from bus timetables available from 1997 Q1 to 2008 Q2	ECan
Fare Auckland	(3) Real fare	Auckland bus fare for contract services	Average revenue per passenger[11]	Quarterly from 1996 Q1 to 2008 Q2	ARTA
		Auckland North Shore bus fare for contract services	Average revenue per passenger	Monthly from November 2005 to July 2008	ARTA
		Auckland rail fare	Average revenue per passenger	Quarterly from 1996 Q1 to 2008 Q2	ARTA
Fare Wellington	(3) Real fare	Wellington bus fare	Average revenue per passenger	Annually from FY 1999/ 2000 to 2007/08	GWRC
		Wellington rail fare	Average revenue per passenger	Annually from FY 1999 20/00 to 2007/08	GWRC
Fare Christchurch	(3) Real fare	Christchurch bus fare[12]	Average revenue per passenger	Monthly from November 03 to June 08 Quarterly from 2004 Q1 to 2008 Q2 Annually from 2004	ECan

10 For Christchurch, information on actual bus kilometres was not available.

11 Average revenue per passenger was derived by total revenue divided by total patronage. Revenue from monthly ticket sales and the associated trips made were included in the analysis.

12 The bus fare information for Christchurch was available only for a relatively very short period from November 03 to June 08 whereas the patronage and service information was available from 1997 Q1 to 2008 Q2. The average revenue per passenger information was available only for contract bus services.

Information	Related variable(s)	Data	Description	Period and frequency of observations	Source
CPI	(3) Real fare (4) Real income (6) Real fuel price	Consumer price index (CPI)	CPI all groups	Quarterly from 1978 Q3 to 2008 Q3	NZSTAT
Income	(4) Real income	Real gross national disposable income	Gross disposable income for all institutional sectors in NZD	Quarterly from 1987 Q2 to 2008 Q3	NZSTAT
Car ownership	(5) Car ownership	National licensed vehicles	Currently licensed by vehicle type	Annually as at June of each year from 1962 to 2008	NZTA
		Regional licensed vehicles	Currently licensed by vehicle type by territorial authority	Quarterly from 2002 Q4 to 2008 Q2	NZTA
Fuel price	(6) Real fuel price	Petrol and diesel prices	Cents per litre (real 2007 prices)	Quarterly from 1974 Q1 to 2008 Q3	MED
Population/ demographic variables	(1) Patronage (2) Service level (4) Real income (5) Car ownership	NZ resident population by age	Estimated New Zealand resident population: (1) Under 15 years (2) 15–64 years (3) 65 years and over	Quarterly from 1991 Q1	NZSTAT
		Resident population by region	Estimated New Zealand resident population: (1) North Island: Auckland (2) North Island: Wellington (3) South Island: Canterbury	Annually from FY 1995/96 to 2006/07	NZSTAT

4.3 Data imputation

There were two sets of data that required data imputation, namely, car ownership and regional resident population. An overview of the imputation process is described in this section and the details are documented in appendix A.

4.3.1 Car ownership

As shown in table 4.1, national licensed car statistics were available for the full period under study (1996 Q1 to 2008 Q2) but they were only available on an annual basis. Quarterly regional information was required for modelling purposes. Regional licensed car statistics were available but only from 2002 Q4. Therefore, we needed to impute the quarterly regional licensed cars from 1996 Q1 to 2002 Q3 based on the relationships between the regional and national statistics available for the same period, ie 2002 Q4 to 2008 Q2. This process was carried out by linear regression with the regional licensed cars as the dependent variable and the de-seasonalised national licensed cars as the independent variable. The imputation procedure is documented in appendix A.

4.3.2 Regional resident population

The New Zealand national resident population estimates were available on a quarterly basis while the regional resident population was available on an annual basis. To produce regional resident population estimates, we regressed regional resident population as the dependent variable with the national resident population estimates as the independent variable. The imputation procedure is documented in appendix A.

4.4 Data availability summary

After data imputation, we were able to deduce time series of the six model variables as described in section 4.1. Table 4.2 summarises the resulting data availability for the three regions.

Table 4.2 Summary of data availability

Variables	Auckland	Wellington	Christchurch
1 Patronage (in trips per capita)	Quarterly from 1996 Q1 to 2008 Q2	Annually from FY 1999/2000 to 2007/08 Quarterly from 2002 Q1	Quarterly from 1997 Q1 to 2008 Q2
2 Service level (in bus/train kilometre per capita)	Quarterly from 1996 Q1 to 2008 Q2	Annually from FY 1999/2000 to 2007/08 Quarterly from 2002 Q1	Quarterly from 1997 Q1 to 2008 Q2 (estimated from timetable)
3 Real fare (in real revenue per passenger)	Quarterly from 1996 Q1 to 2008 Q2	Annually from FY 1999/2000 to 07/08	Quarterly from 2004 Q1 to 2008 Q2
4 Real income (in real disposable income per capita)	Quarterly from 1987 Q2 to 2008 Q3		
5 Car ownership (in cars per capita by region)	Quarterly from 1996 Q1 to 2008 Q2 (from data imputation by region)		
6 Real fuel price	Quarterly from 1974 Q1 to 2008 Q3		

4.5 Strengths/weaknesses of the data sets

As discussed earlier, the guiding principle in the data collection was to collect as much detailed information as possible and for as long as data was available. This was because the power of a forecast model could only be as good as the data it was built on. The more data points we had, the more information would be captured. As shown in table 4.2, the Auckland data set was the most comprehensive among the three sets of data. Bus and rail patronage, service and fare information were all available on a quarterly basis from 1996. However, it is important to note that Auckland's bus data set was related to contract services only. As discussed later in section 5.1.1, there were two kinds of bus services in Auckland, namely, contract and commercial services. Thus, the analysis in this report represents only half of the picture. Commercial routes, which were not included in our analysis, were often shorter routes with much higher frequency during peak hours. The characteristics of travellers on these routes could be different. Although Christchurch bus patronage was also available on a quarterly basis from one year later (1997), the data set was not as good as Auckland's as the bus service information was only an estimation from bus timetables and fare information was available only from 2004 based on contract bus services data. The Wellington data set was the least comprehensive of all. As fare information was not available on a quarterly basis, the model had to be developed based on the annual data set available from FY 1999/2000 to FY 2007/08 and only nine data points were available in total.

4.6 Study period identification

The study period was determined based on the principle that detailed information on preferably all six variables was available during this period.

4.6.1 Auckland

As shown in table 4.2, information on patronage, service and fares for Auckland was available on a quarterly basis from 1996 Q1 to 2008 Q2. This became the model period for Auckland as this was the period that information on all six variables was available.

4.6.2 Wellington

As shown in table 4.2, information on patronage, service and fares for Wellington was available only on an annual basis from FY 1999/2000 to FY 2007/08. Thus, the models for Wellington were developed based on annual information for this period.

4.6.3 Christchurch

As shown in table 4.2, information on patronage and estimates of service for Christchurch was available on a quarterly basis from 1997 Q1 to 2008 Q2. Information on fares, however, was available only from 2004 Q1 to 2008 Q2. Therefore, both model periods were tested for Christchurch.

5 Preliminary data analysis

In this section, we discuss the preliminary analysis of the six variables in the data set. As set out in section 4.1, these are:

Dependent variable

1 Patronage (in trips per capita)

Economic determinants

2 Service level (in bus/train kilometre per capita)

3 Real fare (in real revenue per passenger)

4 Real income (in real disposable income per capita)

Structural determinants

5 Car ownership (in cars per capita by region)

6 Real fuel price

Preliminary data analysis was performed by region in two parts: trend analysis and correlation analysis.

Part 1: Trend analysis

The trends of PT patronage, fares and service level in the three cities were studied for the following periods, as identified in section 4.6:

1 Auckland – 1996 Q1 to 2008 Q2.

2 Wellington – FY 99/00 to FY 07/08.

3 Christchurch – 1997 Q1 to 2008 Q2.

Time series plots of the six variables were investigated in order to identify breaks and changes in the trends. The plots were presented in sets of three. The first set consisted of data related to PT, ie patronage, service and fares. The second set consisted of the remaining three variables: car ownership, real disposable income and petrol price. For each variable, the data values were plotted against time together with a trend line showing the rolling average of a one-year period, which was a four-quarter moving average for quarterly data and a 12-month moving average for monthly data. The use of a trend line was to smooth out the seasonal variation so that the trend could be observed easily.

Information on the history of changes or events was analysed with a view to identify their induced effects. The purpose of the trend analysis was to reveal the story behind the observed trends, ie to investigate how the changes or events might have caused the observed trends.

Part 2: Correlation analysis

Scatter plots of each of the five determinants versus the dependent variable, ie patronage, were examined. Correlation analyses were performed by regressing the dependent variable with the determinant variable being examined. A high R-square value here would suggest a possible strong linear relationship. It is important to note that linear regression was applied only to obtain the R-square value as a measure of the strength of any possible linear relationship. It is not part of the forecast model, as described in chapter 6 of this report.

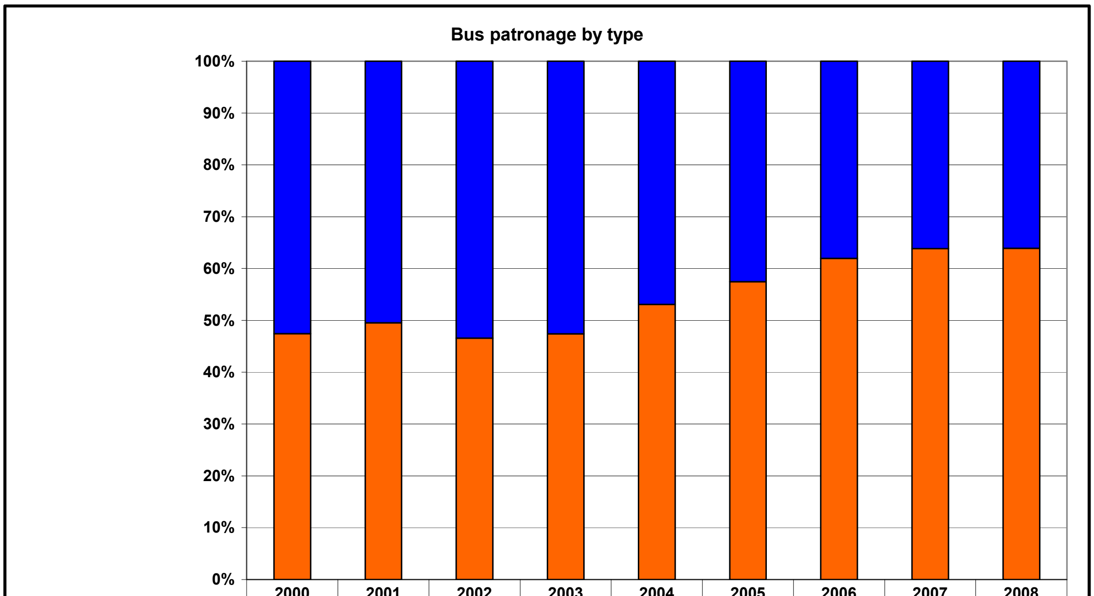
5.1 Auckland

5.1.1 Trend analysis

5.1.1.1 Auckland bus services

In Auckland, there were two kinds of bus services, namely, contract and commercial bus services. Contract bus services were subsidised services monitored by ARTA while commercial services were purely 'commercial' in the sense that no subsidies were provided. As a result, due to commercial confidentiality, information on commercial services was very limited. Only patronage figures were available as a total on an annual basis. Service and fare information was not available at all. As highlighted in section 4.1, detailed information was available only for contract services. The split of bus patronage carried by commercial and contract services is depicted in figure 5.1. The share of patronage on contract services had been increasing over the years. In 2008, contract services carried 64% of total bus passenger trips.

Figure 5.1 Bus contract services versus commercial services



	2000	2001	2002	2003	2004	2005	2006	2007	2008
Commercial Bus patronage	53%	51%	53%	53%	47%	43%	38%	36%	36%
Contracted Bus Patronage	47%	49%	47%	47%	53%	57%	62%	64%	64%

Source: ARTA

The observed patronage, fares and service level and their trend lines of Auckland contract bus services are depicted in figure 5.2(a) to (c). The trends in car ownership in Auckland, national real disposable income per capita and petrol price are depicted in figure 5.3(a) to (c).

Figure 5.2 Trends in Auckland contract bus patronage, service level and revenue per passenger (1996–2008)

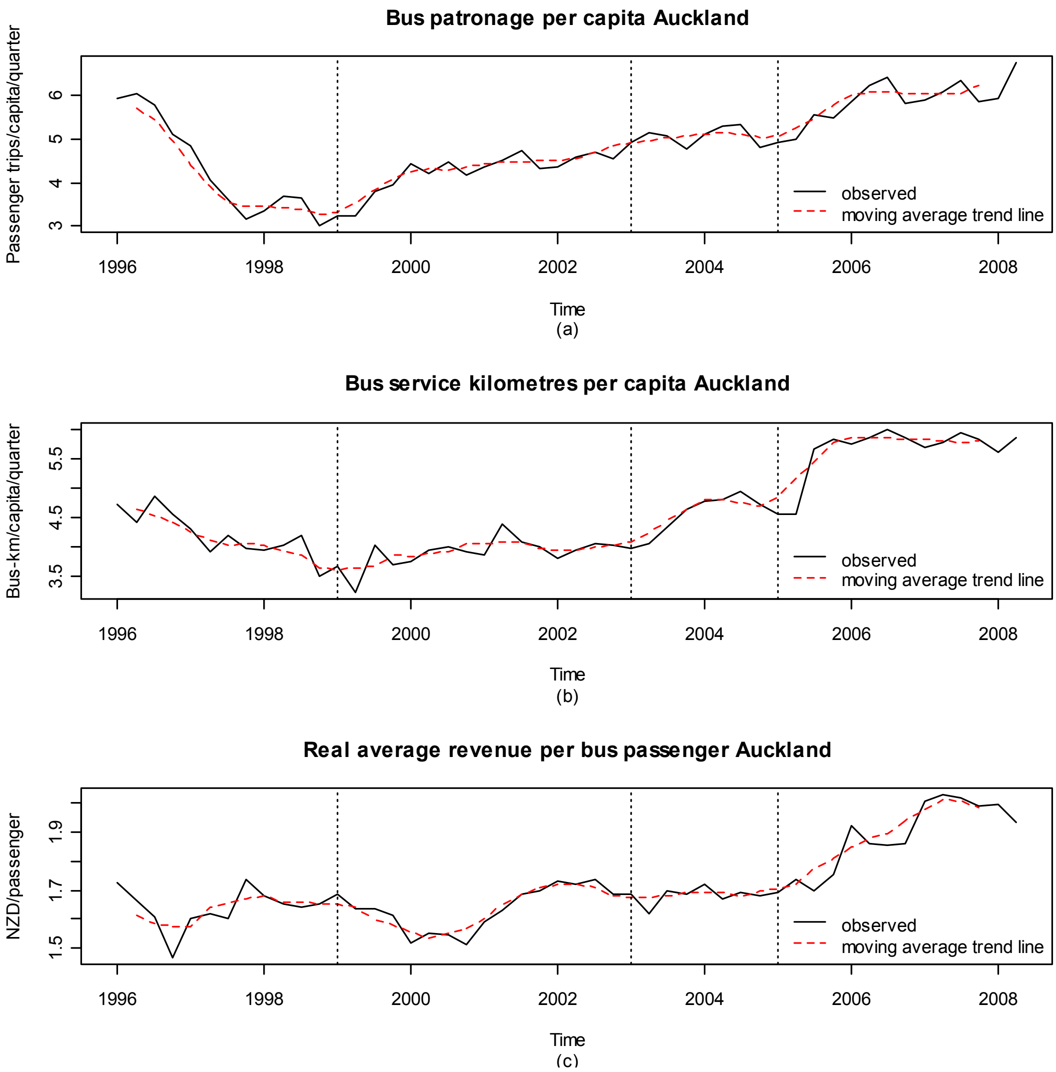
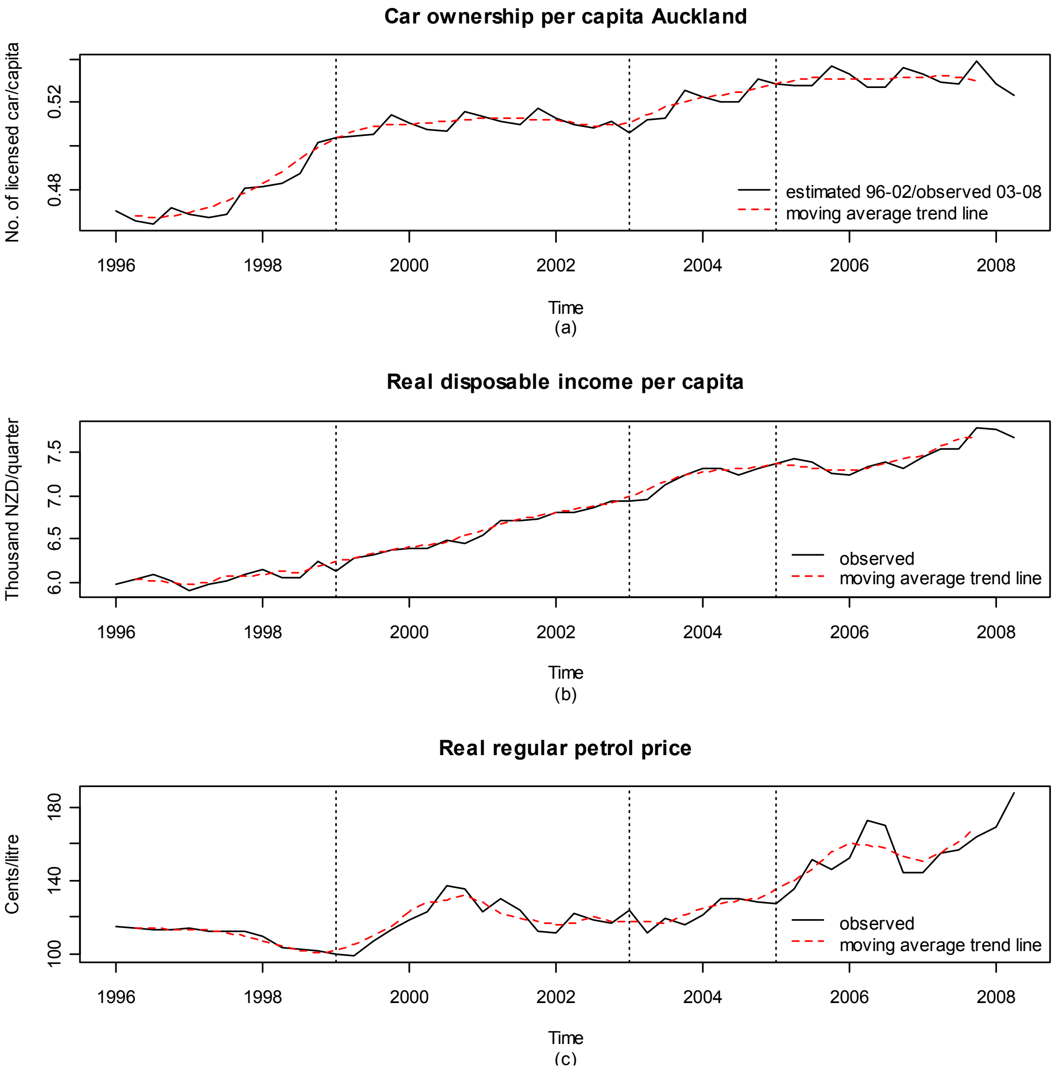


Figure 5.3 Car ownership in Auckland, national disposable income and petrol price



From the bus patronage trend line, as shown in figure 5.2(a), three breaks were identified, as marked by the vertical dotted lines, in 1999, 2003 and 2005. The vertical dotted lines are repeated in all plots as related to Auckland to facilitate easier visualisation of relationships between variables. The first break indicated a change of trend in 1999, where bus patronage started to increase after a decreasing trend since 1996. The decreasing trend in bus patronage might be associated with the increase in car ownership in the 1990s due to the opening of the second-hand imported car market from Japan. As shown in figure 5.3(a), Auckland experienced a steep climb in car ownership from 1996 to 1999.

Bus patronage in Auckland climbed at an almost constant rate from 2000 until 2005. As shown in figure 5.2(a), it appears the trend of patronage changed again in 2005. We believe this is due to two main reasons as follows:

1 The Stagecoach drivers' five-day bus strike in 2005 reduced patronage by around 1 million boardings as it effectively removed one week of patronage, and this had longer-term effects on bus patronage. Many passengers tried travelling by rail during this period and many continued to use the rail service. Also in

2005, commercial service operators reduced the number of bus trips and not all of these were replaced with contract services. As shown in figure 5.1, commercial bus patronage reduced after 2005.

2 The Northern Busway opened in July 2005.

5.1.1.2 The Northern Busway[13]

The Northern Busway is New Zealand's first purpose-built road dedicated to bus PT, and forms a key part of Auckland's rapid transit network. It improved journey speeds, reliability and comfort and provided an attractive alternative to private vehicle use, increasing PT modal share, and reducing the demand for more roading.

Figure 5.4 The Northern Busway in Auckland

Source: Downloaded from www.busway.co.nz/ on 22 May 2009

Opened in February 2008, as shown in figure 5.4, the Northern Busway runs alongside SH1 from Constellation Drive in the north, to the Auckland Harbour Bridge. The Northern Busway is part of a PT

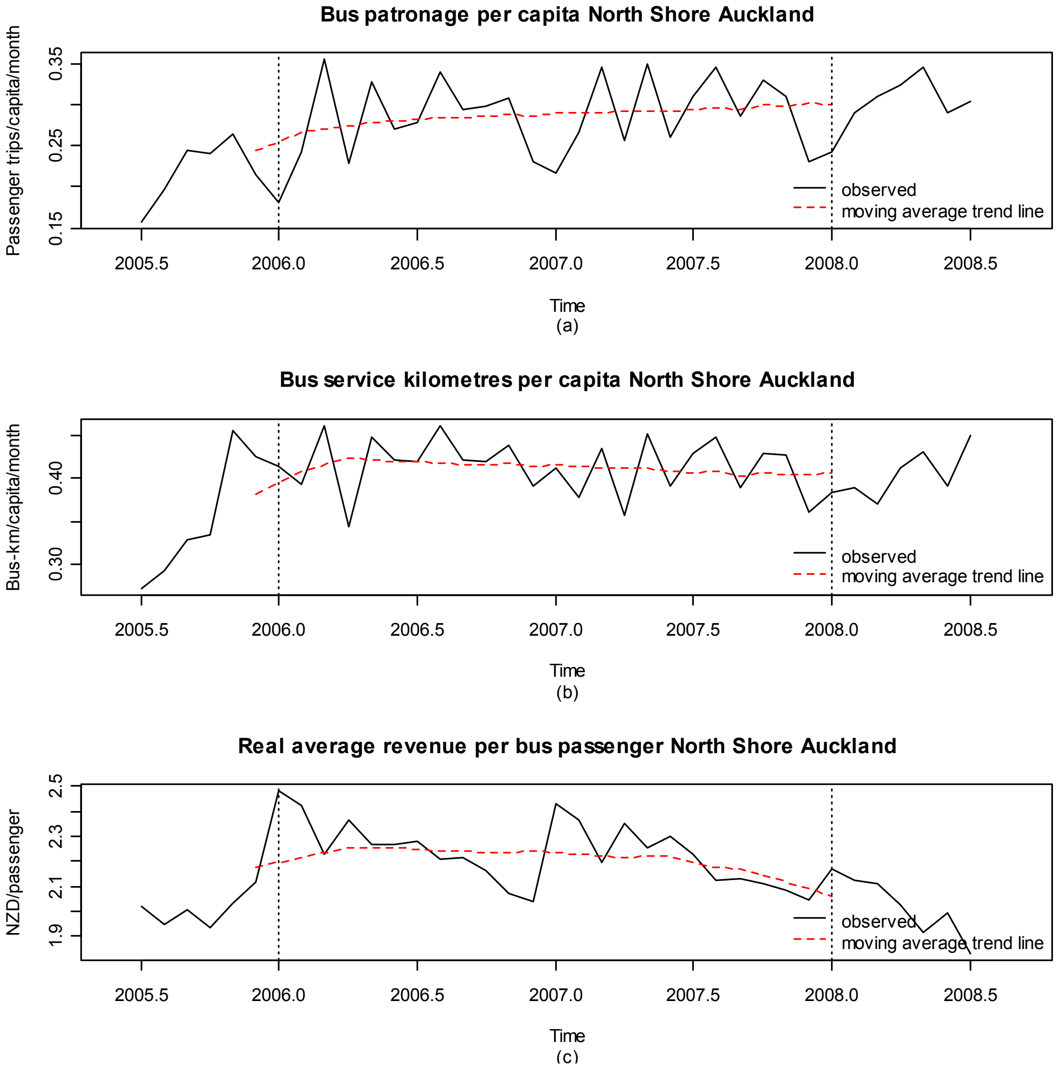
[13] The description of the Northern Busway in this section was extracted from the busway website: www.busway.co.nz/ on 22 May 2009.

network linking North Shore City and the Hibiscus Coast with the CBD. Express services and local bus services link into the busway through five new stations at Albany, Constellation, Sunnynook, Smales Farm and Akoranga.

5.1.1.3 North Shore bus services

As discussed in section 4.2, we collected monthly data on bus services originating or terminating in the catchment area of the Northern Busway. We analysed this data set separately, with a view to identifying the relationships of bus patronage to the other five variables and to see whether the relationships were different from those identified on a regional basis for Auckland.

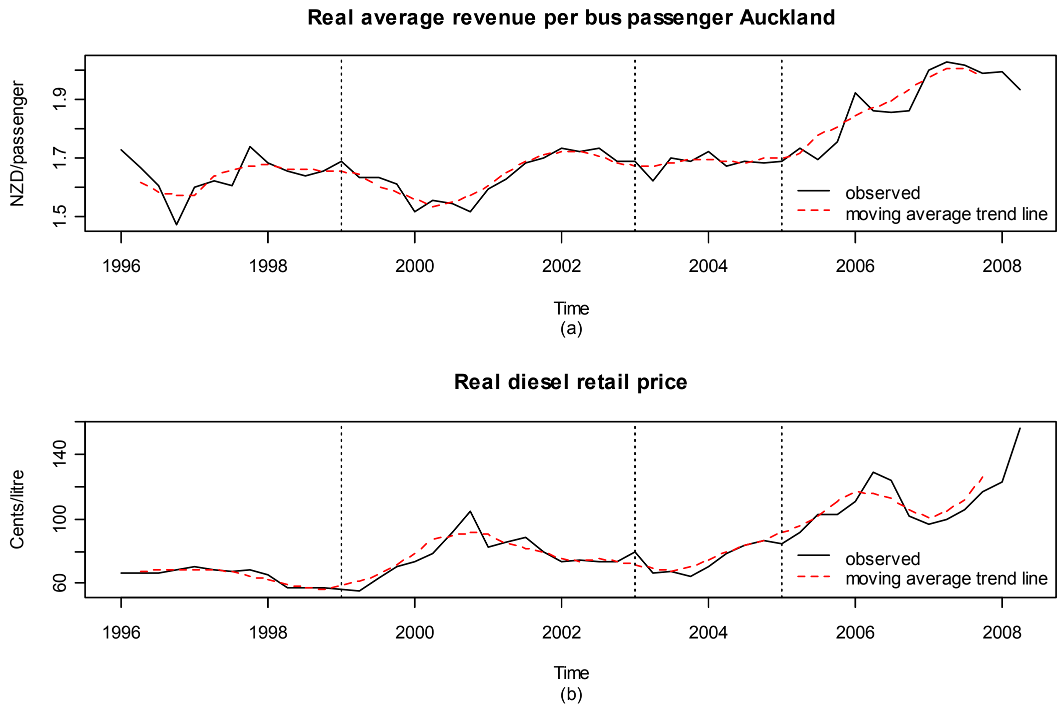
Figure 5.5 Trends in Auckland North Shore contract bus patronage, revenue per passenger and service level (July 05–June 08)



The impact on bus patronage of the opening of the Northern Busway is illustrated in more detail in figure 5.5(a). The breaks in this case were identified to be in 2006 and 2008, corresponding to the effects of opening of the Northern Busway in two stages. Stage 1 of North Shore contract bus services started in November 2005 but did not attract much patronage until March 2006 when school and tertiary students returned. As a result, as shown in figure 5.5(a), a steep climb in patronage was observed in March 2006.

This pattern repeated two years later in March 2008 after the full opening of the Northern Busway in February 2008. As shown in figure 5.5(b), the bus service kilometres followed a similar pattern. These increased steeply in November 2006, remained steady during the first phase and started to increase again in February 2008.

Figure 5.6 Real revenue per bus passenger in Auckland versus diesel retail price



5.1.1.4 Bus fare versus fuel price

Figures 5.2(c) and 5.4(c) are repeated in figure 5.6(a) and (b) to illustrate the relationship between fare and fuel price. As shown in figure 5.6(a), the real average revenue per bus passenger in Auckland had fluctuated over the years. The change in revenue per passenger is dependent on two factors: 1) the distribution of trip length; and 2) the fare. From the operator's point of view, one would expect fares to be directly linked to diesel prices, ie if diesel prices go up, fares will go up. As shown in figure 5.6(b), the increase in diesel prices in recent years has caused a rapid increase in fares.

ARTA was the exception where fares did not increase after fuel prices rose and fell again in 2008. Operators pressured ARTA to raise fares but ARTA avoided doing so (pers.comm). Diesel prices fell quickly during the period of the annual fares review which made it a difficult time to justify increasing fares based on high diesel prices which were no longer relevant as the period of high prices had ended. ARTA and its operators would have been in the difficult position of announcing an increase in fares when diesel prices were lower than before the previous increase in fares.

5.1.1.5 Auckland rail services

The observed patronage, fares and service level and their trend lines of Auckland rail services are depicted in figure 5.7(a) to (c).

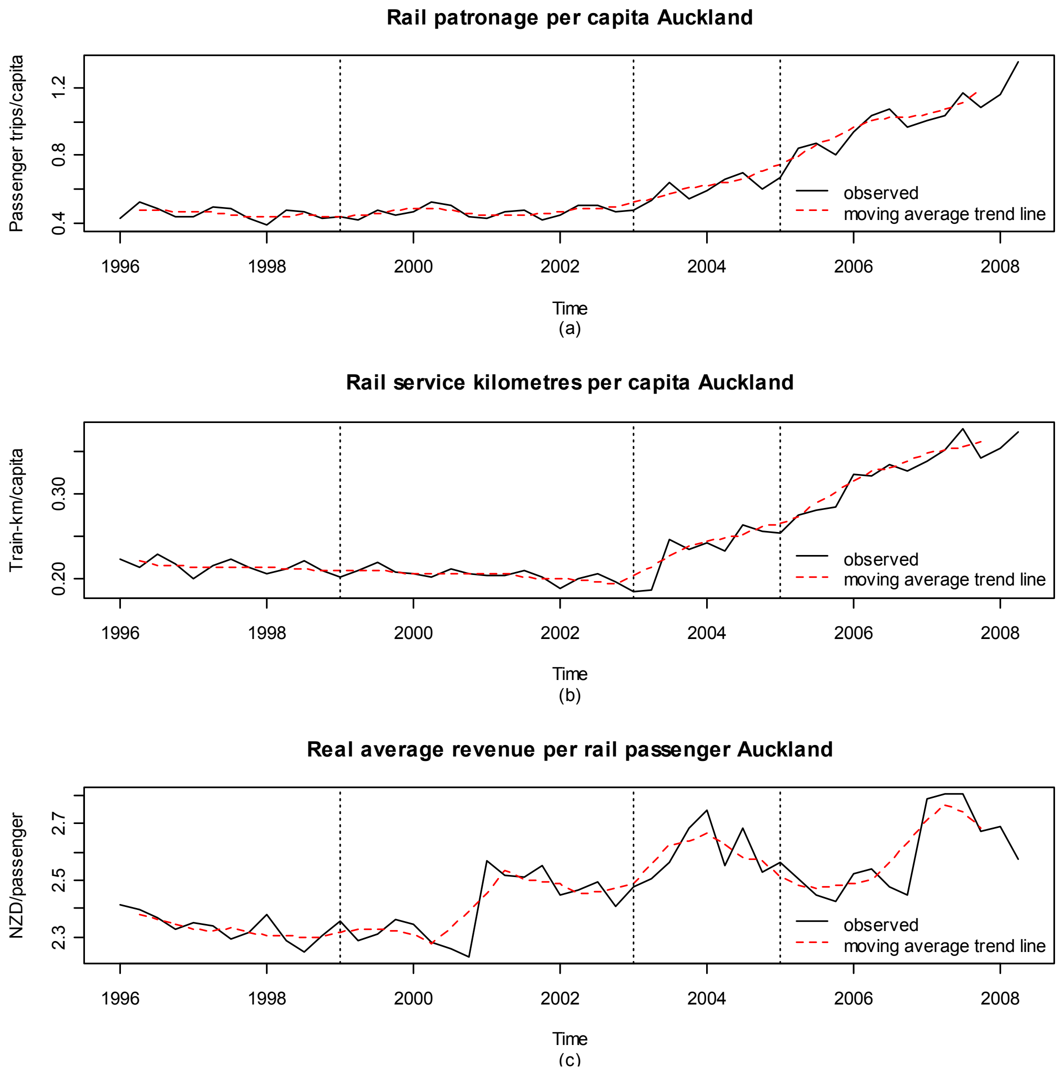
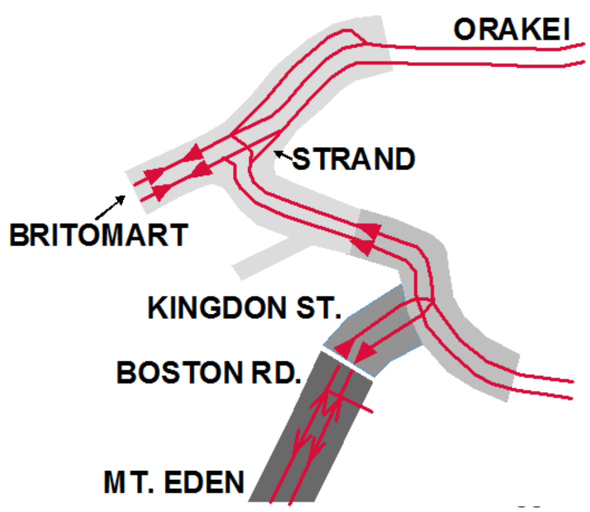


Figure 5.7 Trends in Auckland rail patronage, fares and service level (1996–2008)

Figure 5.8 Railway map of Auckland



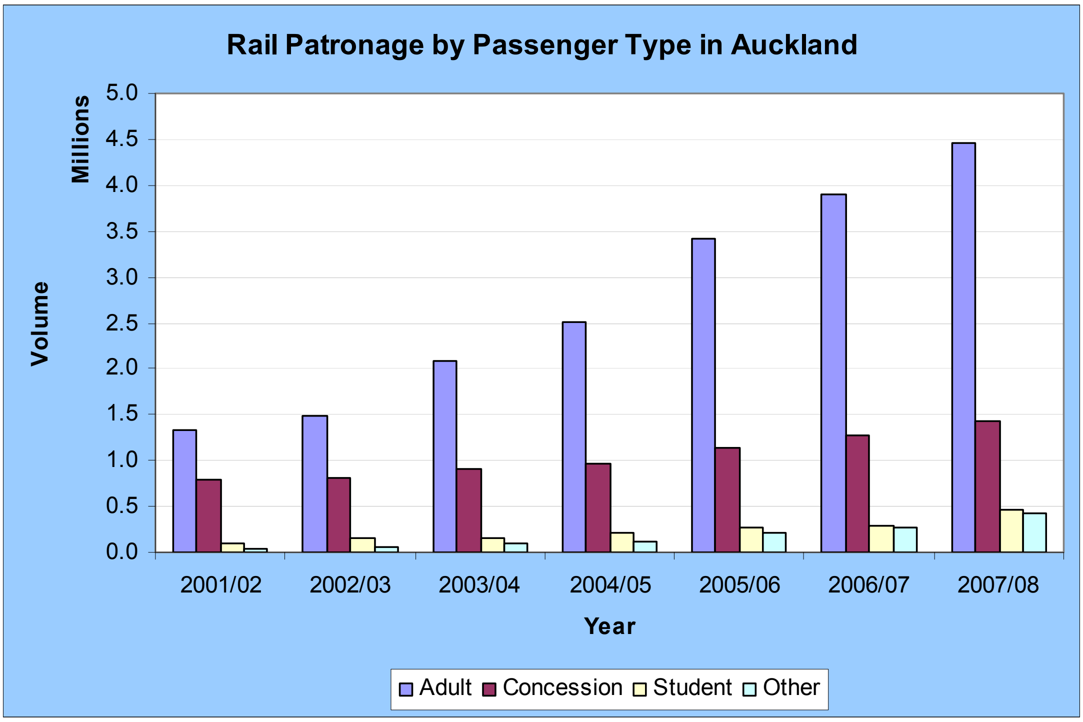
Rail services in Auckland have been a different story from bus services. As shown in figure 5.7(a), the use of rail services in Auckland was constantly low (compared with buses) from 1996 to 2003. Rail patronage has significantly increased since the Britomart Transport Centre was opened in July 2003. The shift of the rail station from The Strand to Britomart made rail a more attractive option as Britomart is located within walking distance of many office buildings in CBD. The Strand is located 1.2km from Britomart. Prior to 2003, most rail passengers had to be transferred from The Strand by bus to get to the central CBD. The locations of Britomart and The Strand are shown in figure 5.8.

Another limitation was the lack of a double track on the western rail line which created limitations on headway. Once double tracking was completed in February 2005, the frequency of service was improved. The service level improvement over time is depicted in figure 5.7(b).

It is important to note the effect on patronage as the result of a 10% increase in rail fare in February 2006 and an additional 15% in January 2007 as shown in figure 5.7(a) and (c). While patronage followed an increasing trend, the rate of increase slowed in 2006 possibly as a result of the fare increase.

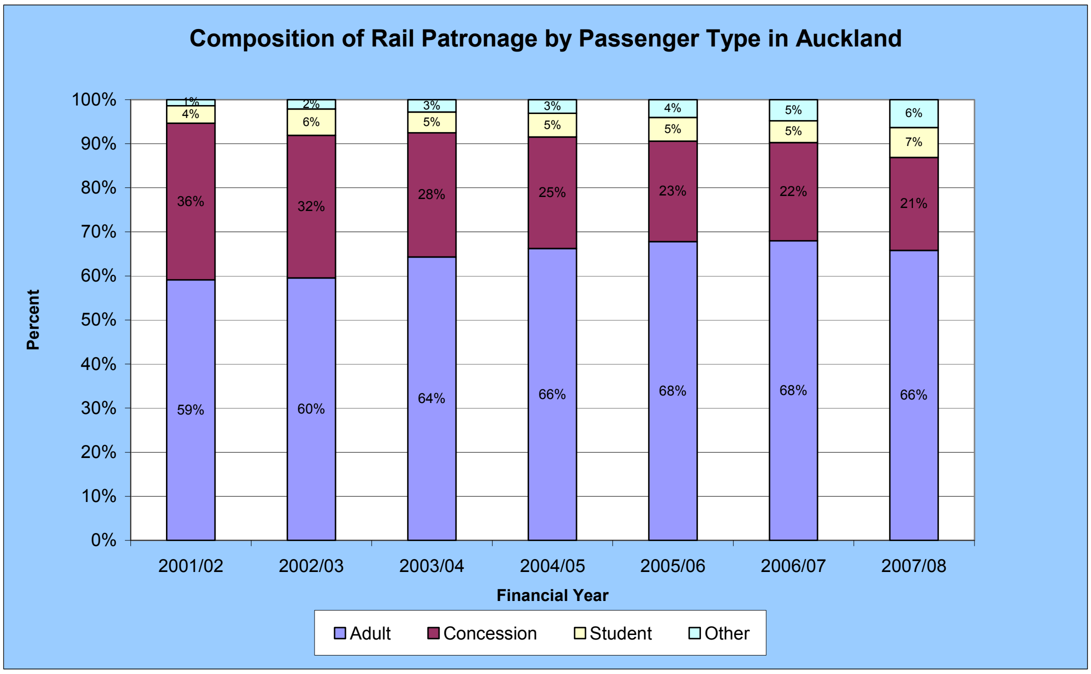
The opening of Britomart and the subsequent significant improvement in rail service induced a change in the characteristics of the rail market.

Figure 5.9 Rail patronage by passenger type in Auckland



Source: ARTA

Figure 5.10 Composition of rail patronage by passenger type in Auckland



Source: ARTA

As shown in figures 5.9 and 5.10, the composition of rail passengers has changed significantly since the opening of Britomart in 2003. Figure 5.9 shows the patronage of adult and other fare categories increased more significantly than that of concession and student categories. The four categories are defined as follows:

1 Adult – adult single ride, adult 10 rides

2 Concession – child single ride, child 10 rides, senior citizens

3 Student – tertiary 10 rides

4 Other – monthly passes (including City Monthly, All Zones Monthly, Regional Monthly and Monthly Discovery Pass), day passes (including Day Rover, Family Pass, Discovery Pass and Bikes)

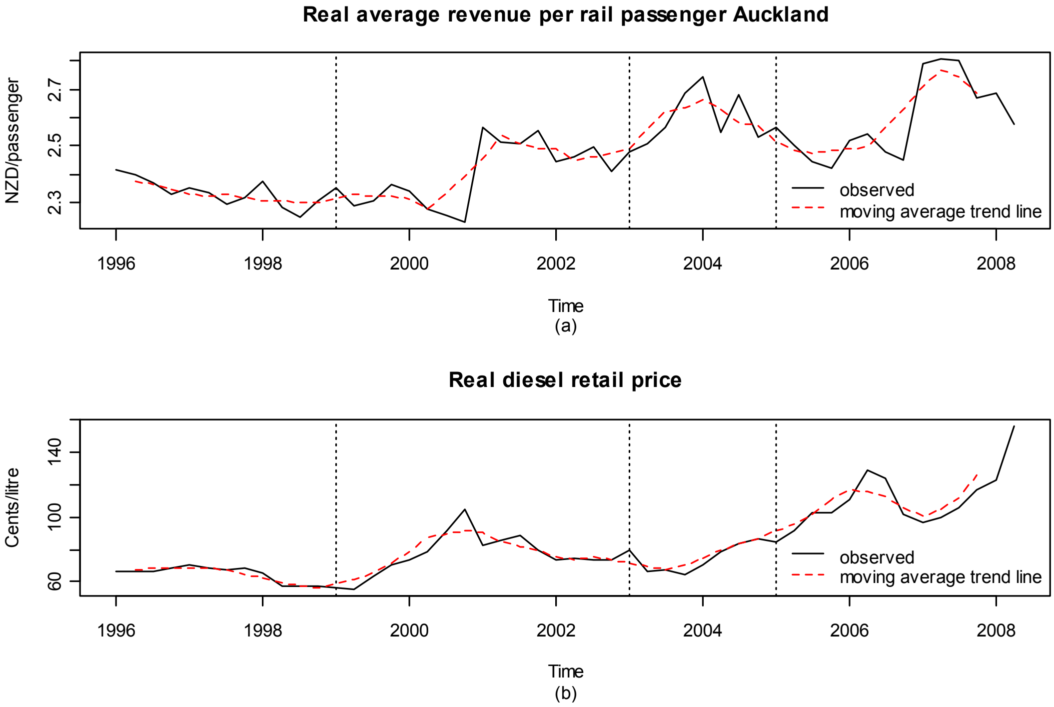
In terms of percentage distribution, as shown in figure 5.10, the pattern after the opening of Britomart in 2003 was quite different from before as follows:

1 The percentage of 'adult' fare users increased from 60% in 2002/03 to a maximum of 68% in 2006/07.

2 The percentage of 'other' fare users increased from 2% in 2002/03 to 6% in 2007/08.

Based on information from ARTA (pers comm), the significant increase of 'other' fare users was associated with the sales of monthly passes, in particular the Monthly Discovery Pass which was introduced as an integrated monthly ticket for all PT services. In other words, it was evident that the rail market had captured a higher proportion of commuters in recent years. Park-and-ride facilities were introduced in numerous locations (eg Panmure, Glen Innes, Papakura, Papatoetoe, Homai) attracting a commuter market.

Figure 5.11 Real revenue per rail passenger in Auckland versus diesel retail price



Figures 5.7(c) and 5.3(c) are repeated in figure 12(a) and (b) to illustrate the relationship between fare and fuel price. Similar to the revenue per passenger of bus services, we believe that the ups and downs of the real revenue per passenger in rail are associated with the fluctuation in diesel price. For example, after a spike in diesel price in early 2000, there was a peak in rail revenue per passenger in 2001. Similarly, after a spike in diesel price in 2006, there was a peak in rail revenue per passenger in 2007.

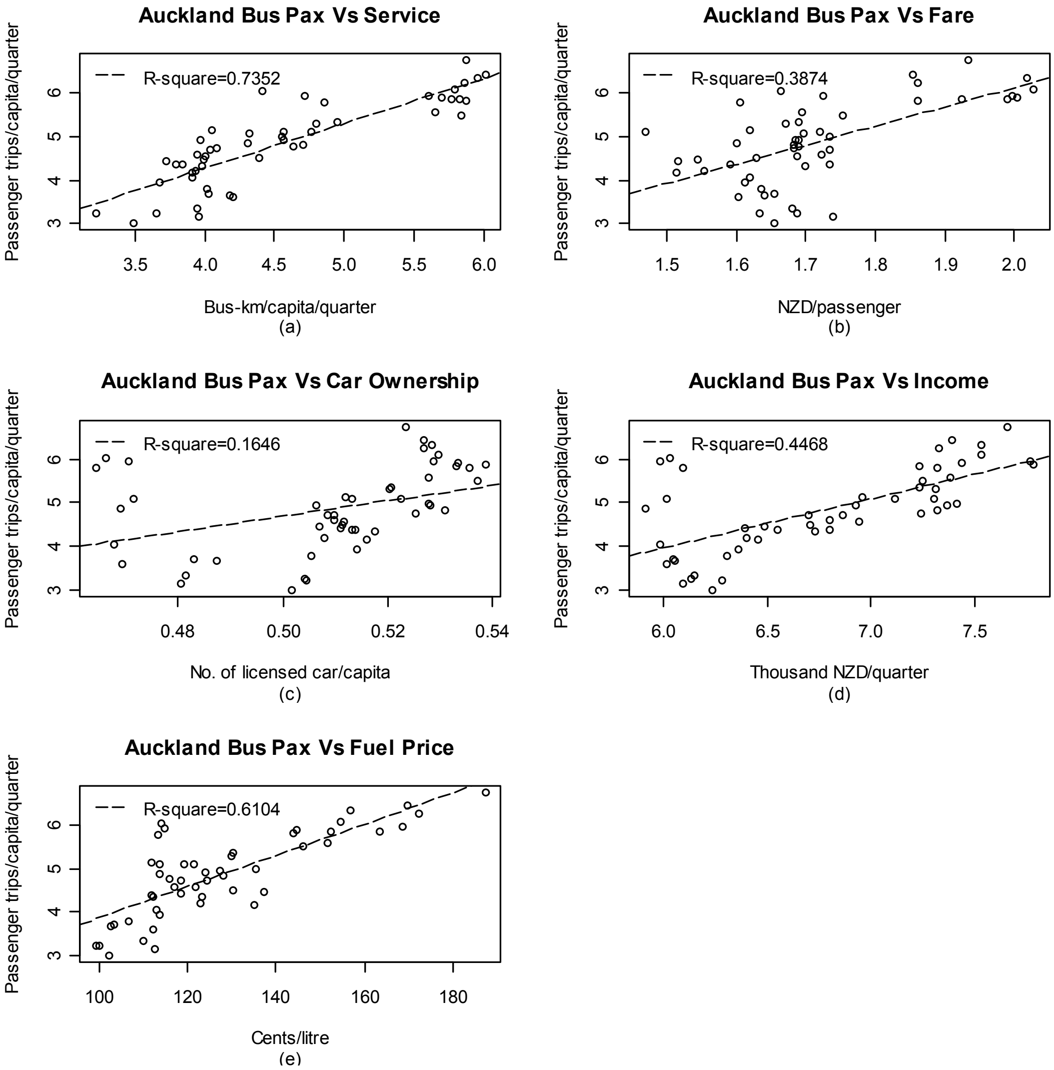
5.1.2 Correlation analysis

In this section, we analyse the correlation between the dependent variable, ie bus/rail patronage per capita and each of the five possible independent variables. Linear regression analysis was applied to each pair of variables. The strength of correlation is illustrated with a scatter plot with the best fitted line and the corresponding R-square value.

5.1.2.1 Auckland bus patronage versus potential influencing factors

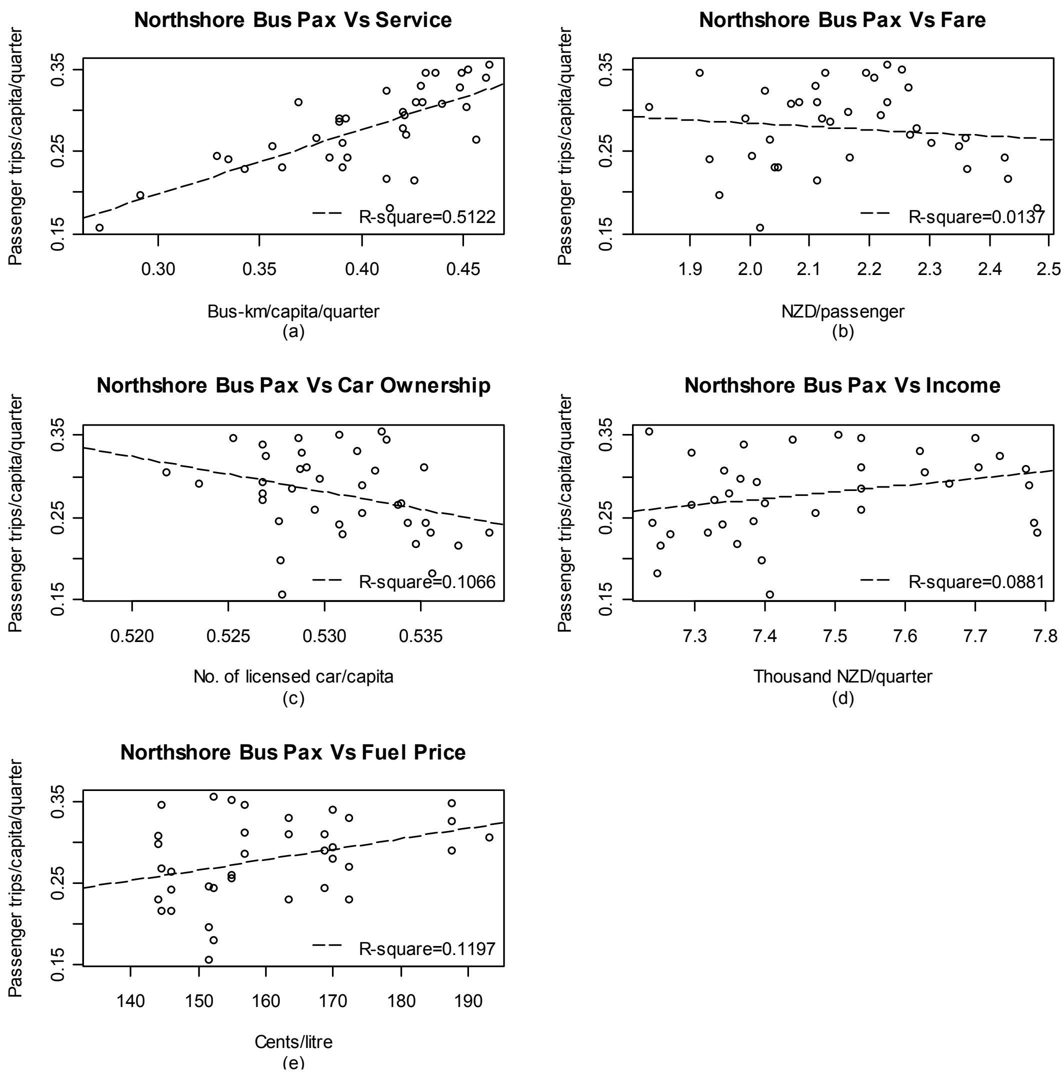
The scatter plots of Auckland bus patronage versus service, fares, car ownership, income and fuel price are depicted in figure 5.12(a) to (e).

Figure 5.12 Scatter plots of Auckland bus patronage versus potential influencing factors



As shown in figure 5.12(a) and (e), with R-square values of 0.7352 and 0.6104, it is clear there was a positive relationship between Auckland's bus patronage with service and fuel price respectively. By intuition, we would expect bus patronage to have a negative relationship with fares, car ownership and income but this is not evident in the plots. As shown in figure 5.12(b), (c) and (d), the best fitted lines all have a positive slope. This means bus patronage apparently increased with the increase in fares, car ownership and income. One possible explanation is that the increase in petrol price in recent years was more significant than the increase in bus fares. It was still 'cheaper' to go on the bus, even though the fare had increased. It is important to note that these plots were produced based on simple linear regression with only one independent variable, ie the regression was not controlled for the influences of the other explanatory variables. As a result, it might have given us a distorted picture.

Figure 5.13 Scatter plots of Northshore bus patronage versus potential influencing factors



5.1.2.2 Northshore bus patronage versus potential influencing factors

The scatter plots of Auckland rail patronage versus service, fares, car ownership, income and fuel price are depicted in figure 5.13(a) to (e). As shown in figure 5.13(b) to (e), all the R-square values are very low. That means North Shore bus patronage did not appear to have a strong correlation with fare, car ownership, income and fuel price. The only possible relationship that can be identified from the plots is a positive one between bus service (in bus-kilometres) and patronage as indicated by a relatively higher R-square value of 0.5122 in figure 5.13(a).

5.1.2.3 Auckland rail patronage versus potential influencing factors

The scatter plots of Auckland rail patronage versus service, fares, car ownership, income and fuel price are depicted in figure 5.14(a) to (e).

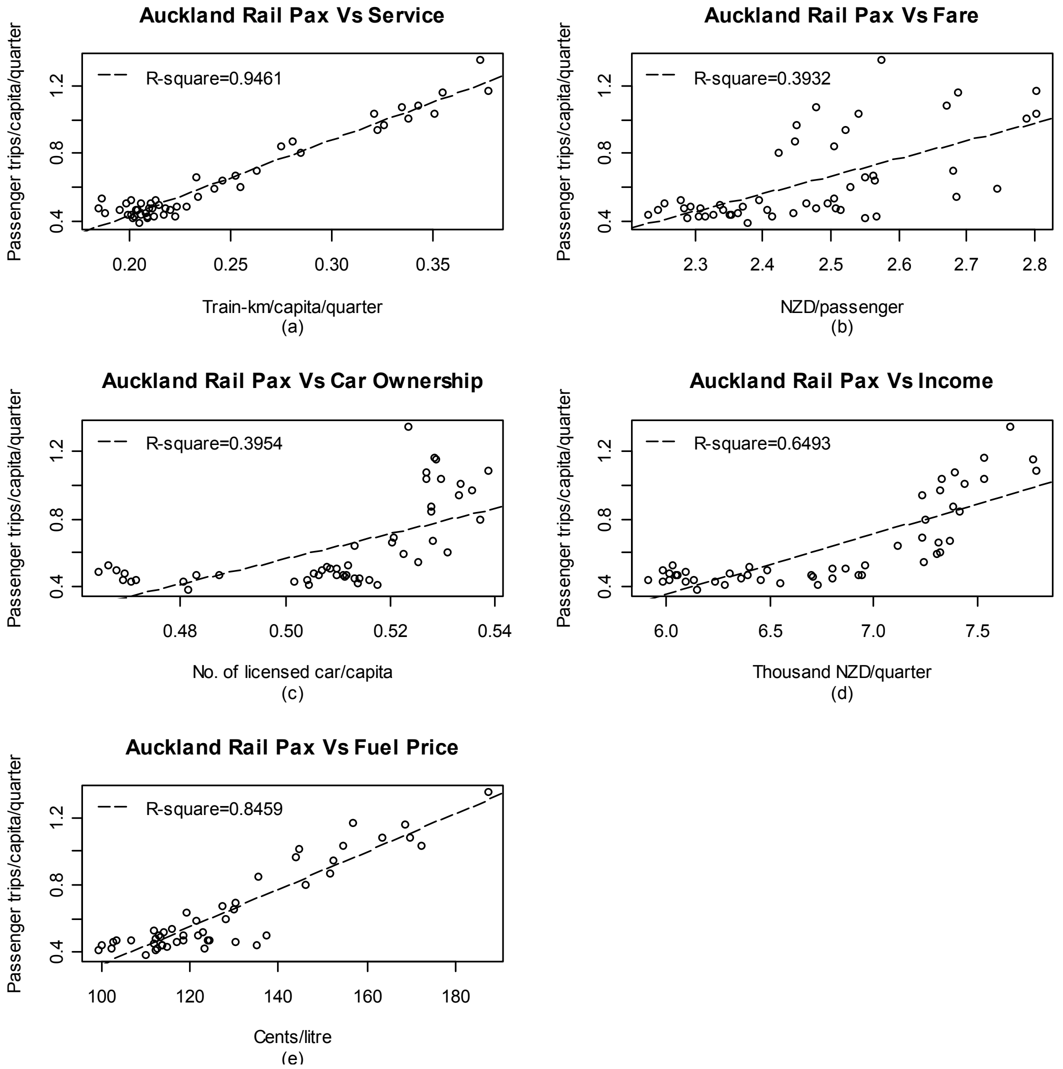


Figure 5.14 Scatter plots of Auckland rail patronage versus potential influencing factors

As shown in figure 5.14(a) and (e), with R-square values of 0.9461 and 0.8459, it is clear that there was a positive relationship between Auckland's rail patronage and service and fuel price. It also appears, from figure 5.14(c) and (d), although not as strong (with R-square values of 0.3954 and 0.6493), that rail patronage was low when car ownership and income were low and a positive relationship only started when they were above a certain level. The positive relationship between patronage and car ownership/income is somewhat counter-intuitive. We believe this could be explained by the change in characteristics of the rail market, as discussed in section 5.1.1.5, as a result of the opening of Britomart and the subsequent significant improvement in rail service. More and more commuters were attracted to use the rail service. Thus the rail market had captured a higher proportion of users with higher income. This is a possible explanation for the observed positive correlation between patronage and car ownership/income and an observed kink in figure 5.14(d). One would expect that the higher income group was the group who travelled on rail by choice and hence would be more sensitive to service and other factors such as travel time reliability, level of comfort and level of congestion (if they drove instead). On the other hand, the lower income group could be a captive market (eg students, senior citizens), ie they might not have a choice because they did not have a driving licence, or they did not have access to a car.

Figure 5.14(b) also indicates Auckland rail patronage was positively correlated to fare as the best fitted line has a positive slope, although the R-square value is low at 0.3932. This is similar to the observation for Auckland bus patronage and could be explained by the higher increase in fuel price relative to the fare increase, ie rail service was still relatively ‘cheaper’ than driving.

5.2 Wellington

We decided to analyse data for Wellington based on annual information rather than quarterly information for the following reasons:

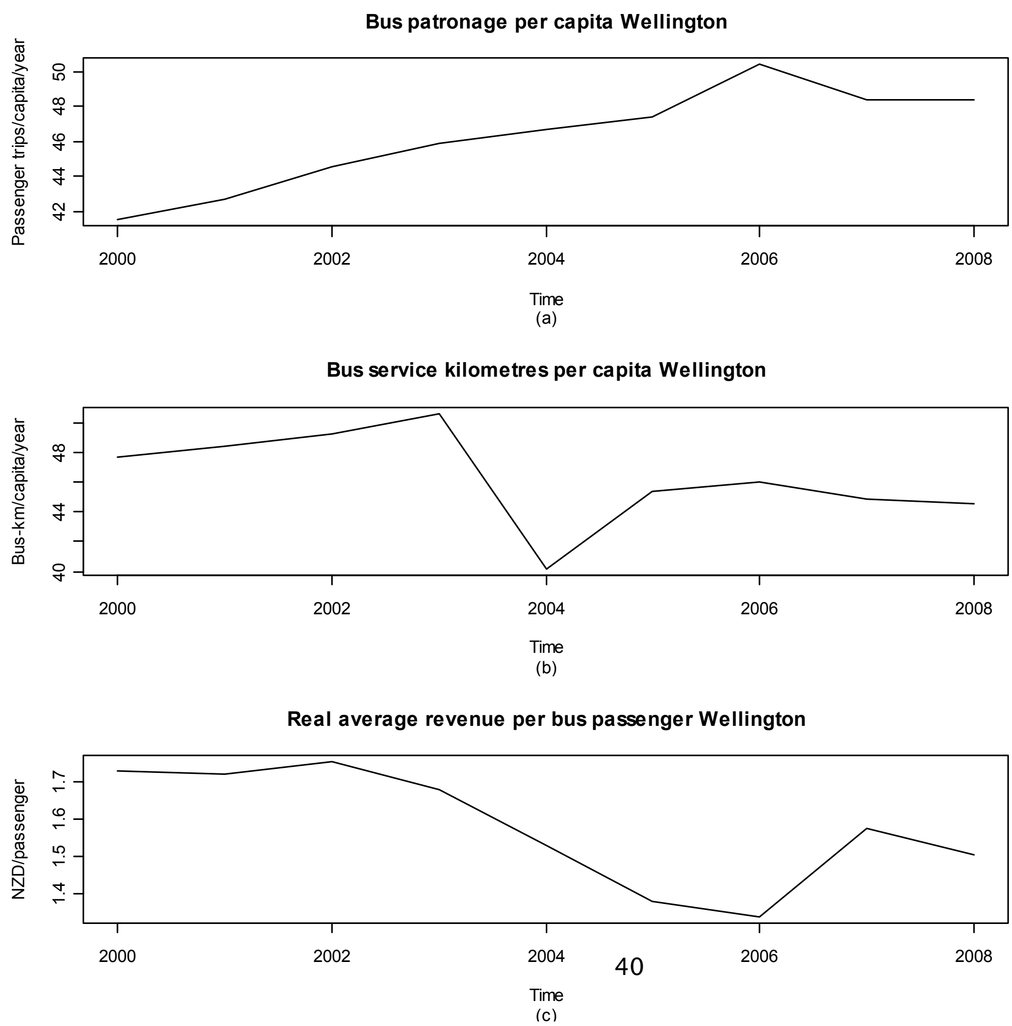
1. The annual data was available for a longer period (from FY 1999/2000 to FY 2007/08) than the quarterly information (from 2002 Q1 to 2008 Q2).
2. Information on average revenue per passenger was only available on an annual basis.

5.2.1 Trend analysis

5.2.1.1 Wellington bus services

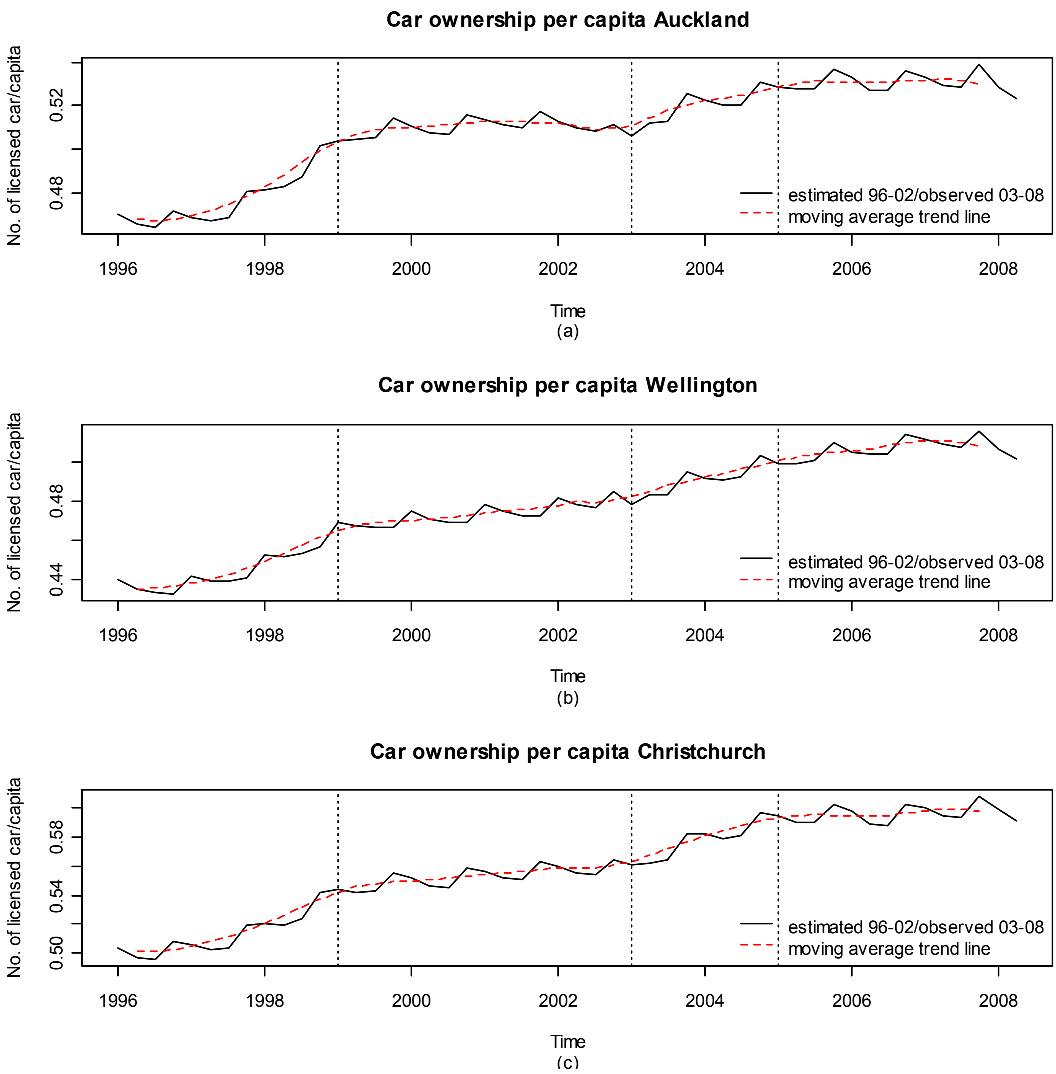
The observed annual patronage, service kilometres and fares for Wellington bus services from FY 1999/2000 to FY 2007/08 are depicted in figure 5.15(a) to (c).

Figure 5.15 Wellington bus patronage, service level and revenue per passenger (FY 1999/2000 to FY 2007/08)



As shown in figure 5.15(a) to (c), the bus patronage per capita in Wellington had been following an increasing trend for the previous eight years but no up or down trend was observed for bus service level and average revenue per passenger. We believe this was due to the fact that Wellington's PT market was quite different from Auckland's and Christchurch's in the sense that it was a mature market.

Figure 5.16 Car ownership per capita by region

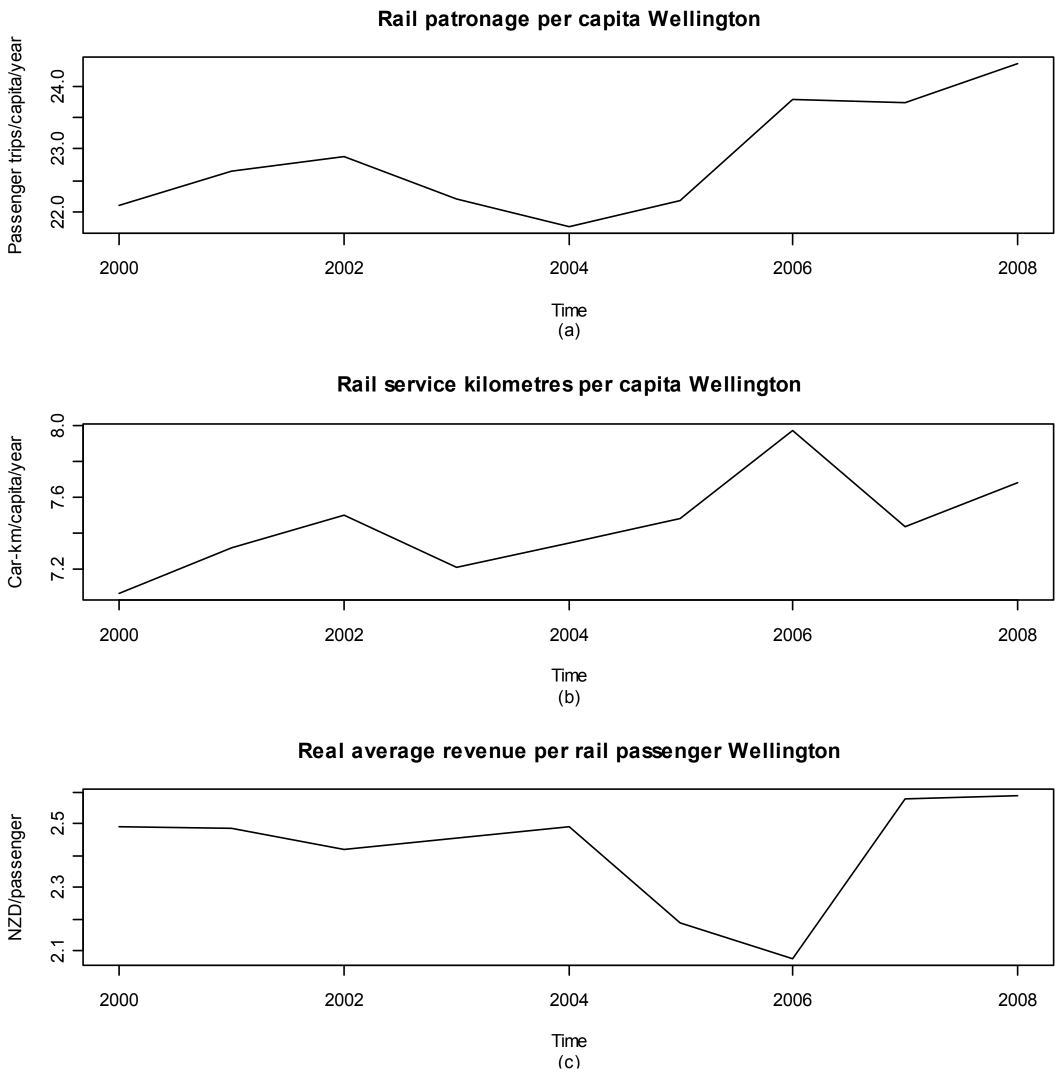


As shown in figure 5.16(a) to (c), Wellington always had the lowest car ownership among the three cities. Wellington also had the highest PT modal share among the three cities. It has a walkable CBD area and a committed parking restraint strategy in the CBD area. There were only three city council car parks in the Wellington city centre and only one of them was available for all day commuter parking. As a result, the characteristics of the Wellington market were quite different from Auckland and Christchurch.

5.2.1.2 Wellington rail services

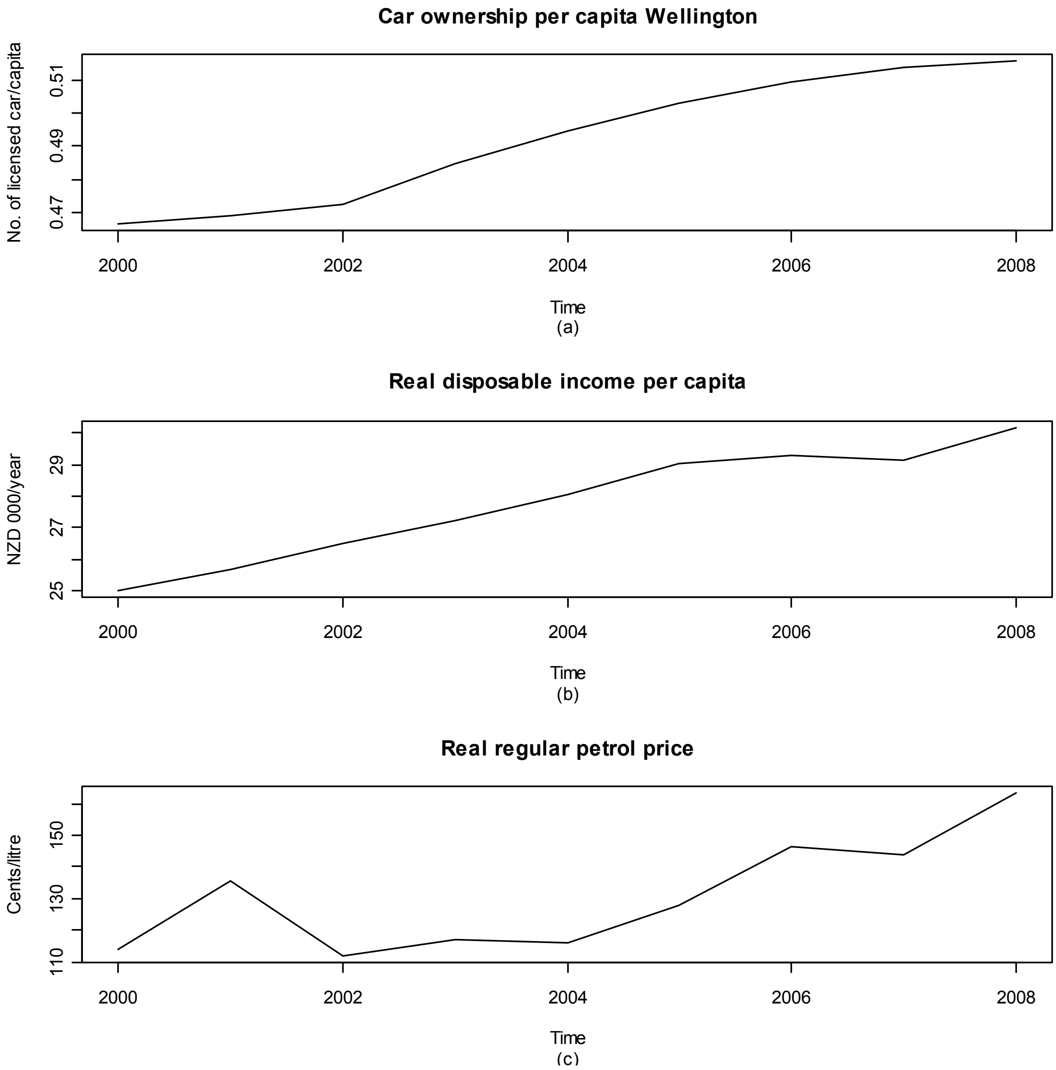
The observed annual patronage, service kilometres and fares for Wellington rail services from FY 1999/2000 to FY 2007/08 are depicted in figure 5.17(a) to (c).

Figure 5.17 Wellington rail patronage, service level and revenue per passenger (FY 1999/2000 to FY 2007/08)



Similar to bus patronage, the rail patronage per capita also followed an increasing trend as shown in figure 5.17(a). The rail service kilometre, as shown in figure 5.17(b), also followed an increasing trend while there was no trend observed for average revenue per passenger in figure 5.17(c).

Figure 5.18 Car ownership in Wellington, national disposable income and petrol price



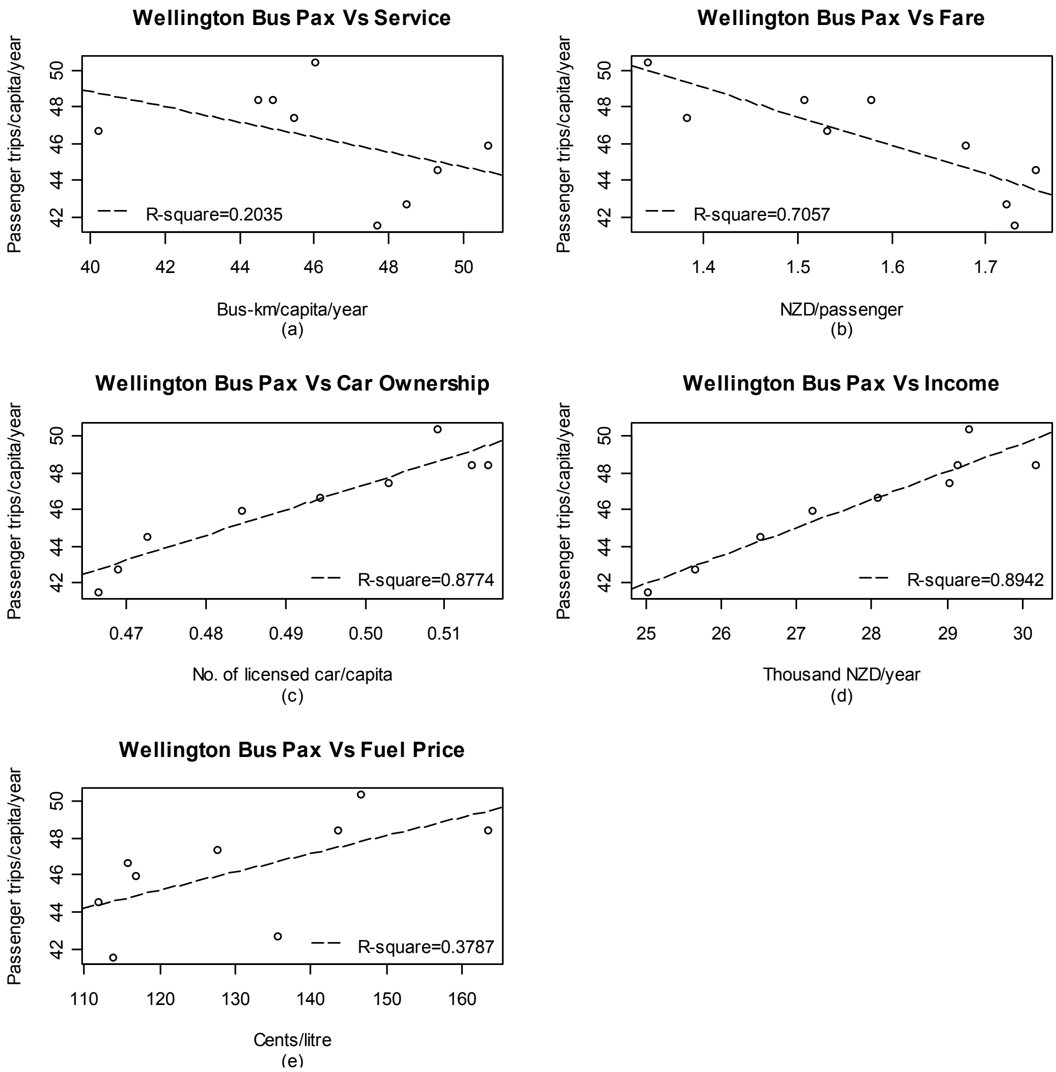
The increasing trend of car ownership is in line with that of income, as shown in figure 5.18(a) and (b), while the increasing trend of both bus and rail patronage is in line with that of petrol price as shown in figure 5.18(c).

5.2.2 Correlation analysis

5.2.2.1 Wellington bus patronage versus potential influencing factors

The scatter plots of Wellington bus patronage versus service, fares, car ownership, income and fuel price with the best fitted lines and R-square values are depicted in figure 5.19(a) to (e).

Figure 5.19 Scatter plots of Wellington bus patronage versus potential influencing factors

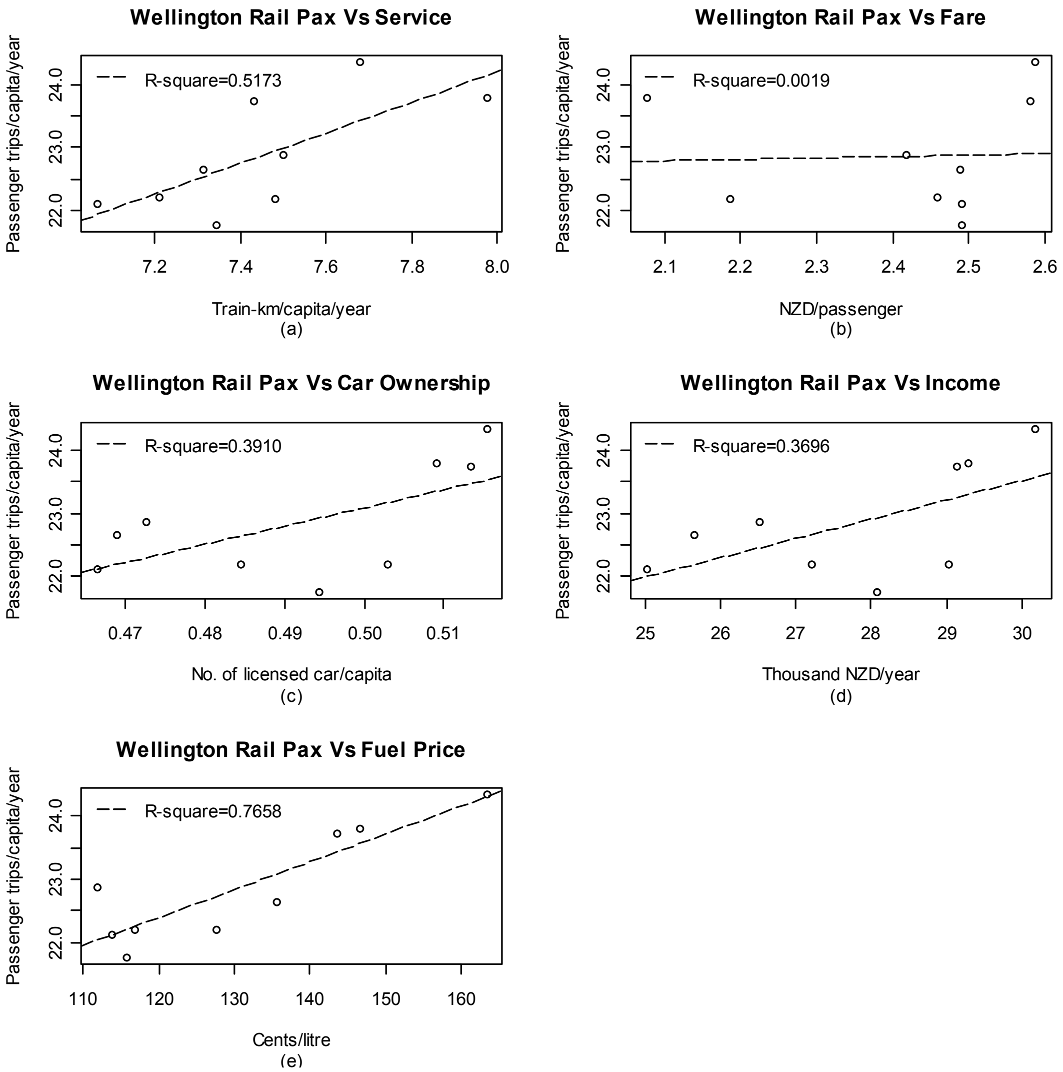


As shown in figure 5.19(b), with an R-square value of 0.7057, it appears there was a negative relationship between bus patronage and fare. It also appears in figure 5.19(c) and (d) that, with R-square values of 0.8774 and 0.8942, bus patronage had a strong positive relationship with car ownership and income. As discussed in the Auckland example, although the positive relationship to car ownership and income was counter intuitive, this could be explained by the characteristics of different segments in the bus market, or it might be due to the effect of regression with a single independent variable. Bus patronage did not appear to have any observable relationship with bus service and fuel price, as indicated by the low R-square values of 0.2035 and 0.3787 in figure 5.20(a) and (e).

5.2.2.2 Wellington rail patronage versus potential influencing factors

The scatter plots of Wellington rail patronage versus service, fares, car ownership, income and fuel price with the best fitted lines and R-square values are depicted in figure 5.20(a) to (e).

Figure 5.20 Scatter plots of Wellington rail patronage versus potential influencing factors



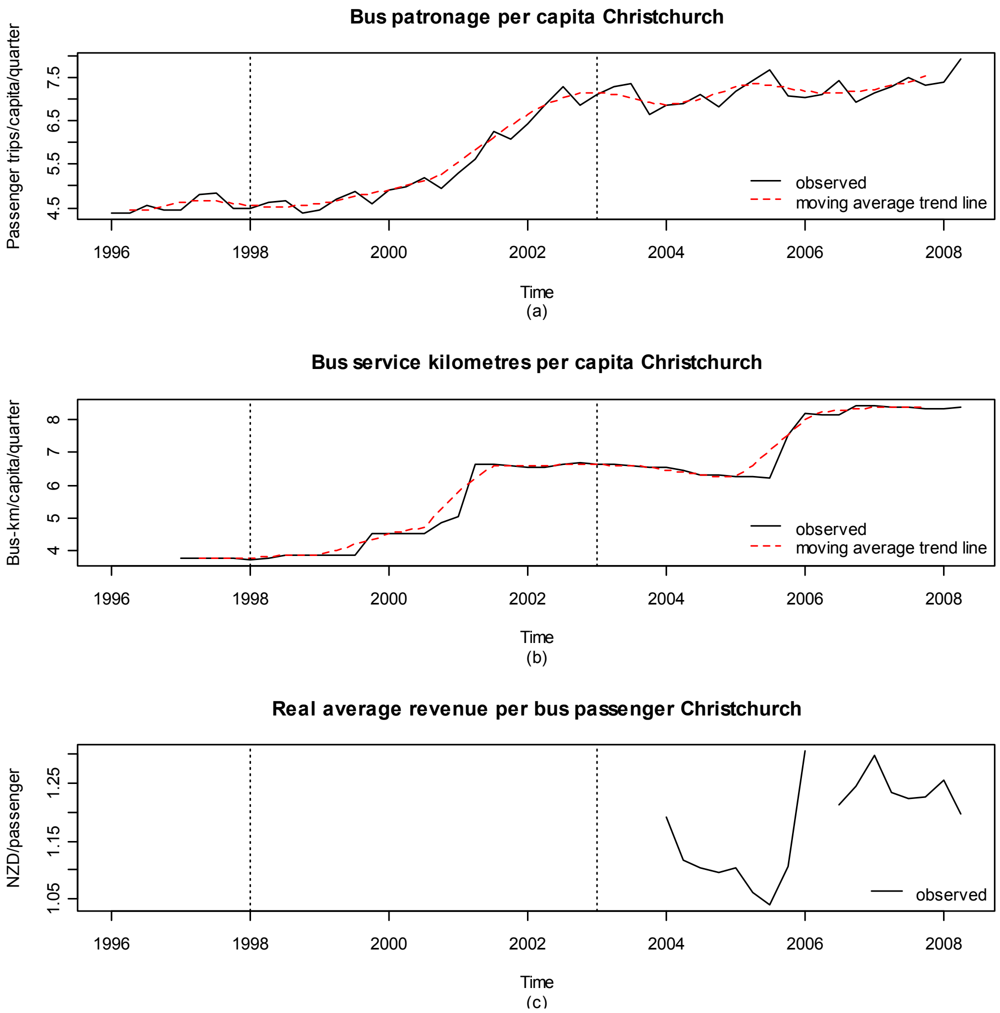
As shown in figure 5.20(e), with a high R-square value of 0.7658, rail patronage in Wellington appeared to have a strong positive relationship with fuel price. It also appeared that rail patronage had a positive relationship with service, with an R-square value of 0.5173 as shown in figure 5.20(a). The relationships to car ownership and income also appeared to be positive, as shown in figure 5.20(c) and (d), as they were for Wellington bus. There was no relationship between rail patronage and fare, as the R-square of 0.0019 as shown in figure 5.20(b) is nearly zero and the best fitted line is almost horizontal.

5.3 Christchurch

5.3.1 Trend analysis

The observed patronage, fares and service level and their trend lines of Christchurch contract bus services are depicted in figure 5.21(a) to (c).

Figure 5.21 Trends in Christchurch bus patronage, service level and revenue per passenger (1996–2008)



As shown in figure 5.21(a), bus patronage per capita in Christchurch also followed an increasing trend over the years as in Auckland. However, the rate of increase had a different pattern. It had the steepest climb during the period 1998 to 2003. From 2003, the trend was still increasing but the rate of increase was slower than in previous years.

As discussed earlier, the bus service information was only an approximation based on bus timetables and by multiplying the estimated weekly total bus-kilometres by the number of weeks per quarter. Therefore, as shown in figure 5.21(b), the bus-kilometres appeared to have had a few ‘step-ups’ over the years.

Based on information from ECan, we believe that the improvement of bus services or the introduction of new bus services induced the growth of bus patronage over the years. The Orbiter ring route was first introduced in 1999. According to information from ECan, the service proved so popular that the full ‘orbit’ or ring route around the city was completed ahead of schedule in 2000, and by 2001 the frequency of the weekday timetable was increased from a 15-minute headway to a 10-minute headway. In November 2004 a second purely cross-suburban service (the Metrostar) was introduced, and starting from November 2006 the services to the Waimakariri District were upgraded with rebranded ‘Northern Star’ buses, frequency improvements and more direct services. The success of the Orbiter route in particular was notable. It is the biggest single route in Canterbury in terms of passenger trips made, contributing over two million trips per year since 2005/06. The Orbiter recorded phenomenal patronage growth in the period from 2000 to 2004. The Metrostar and Northern Star also attracted good patronage numbers and contributed significantly to the rise in bus use in recent years.

However, as shown in figure 5.21(c), information on average revenue per passenger was available only from 2004 Q1 and the observation for 2006 Q2 had data issues. Therefore, the analysis of the effect of the fare would be very limited.

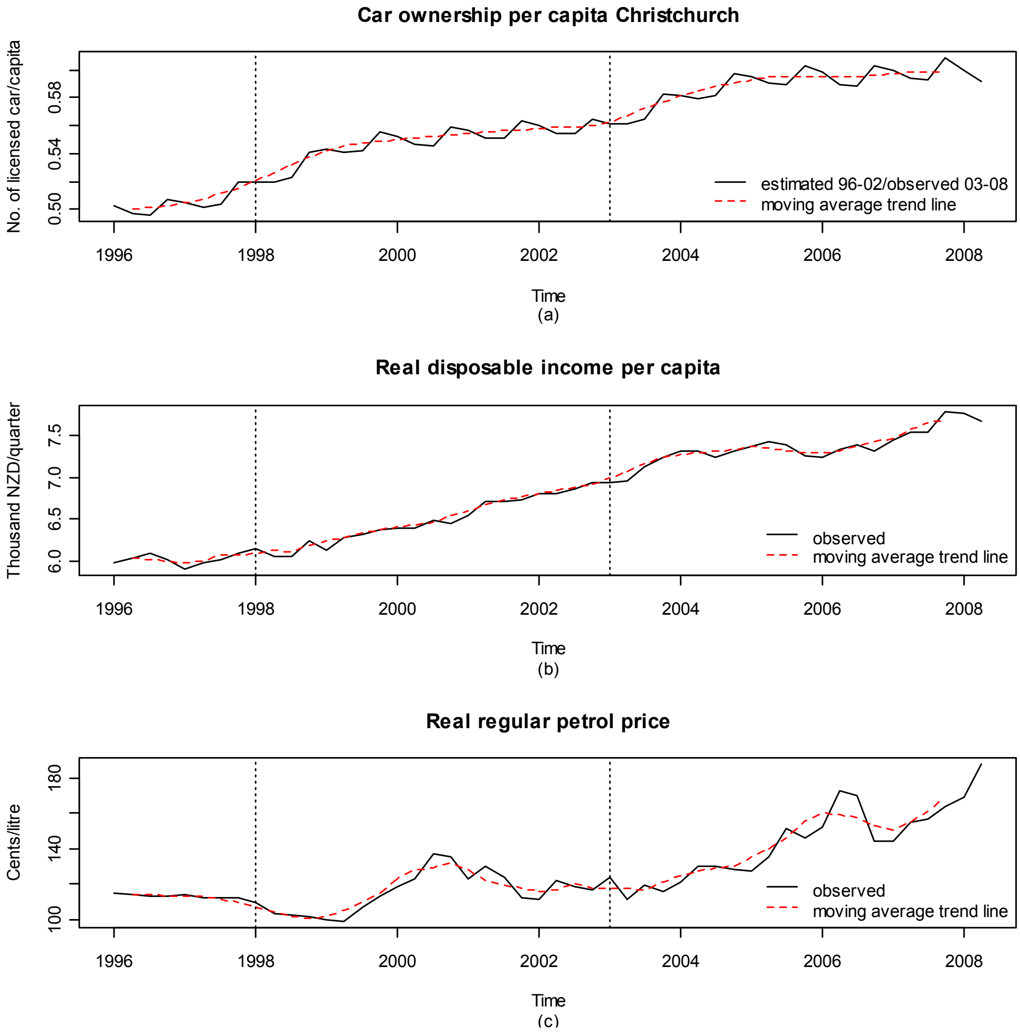
The two possible breaks identified for the analysis for Christchurch are shown as vertical dotted lines in figure 5.22(a) to (e) and all the subsequent time series plots. These two breaks represent the events as follows:

1 In 1998, to ensure that fares were simple and affordable, ECan introduced gold coin cash fares – a one to two section fare system with a $2 zone fare for adults and $1 for children. We believe this policy had a positive effect on bus patronage and hence the steep climb of patronage from 1998 to 2003.

2 In 2003, the electronic Metrocard was introduced which provided an opportunity to further simplify the fare structure – the one to two section fare was simplified to a flat one zone fare across the whole city.

The trends in car ownership in Christchurch, national real disposable income per capita and petrol price are depicted in figure 5.22(a) to (c).

Figure 5.22 Car ownership in Christchurch, national disposable income and petrol price

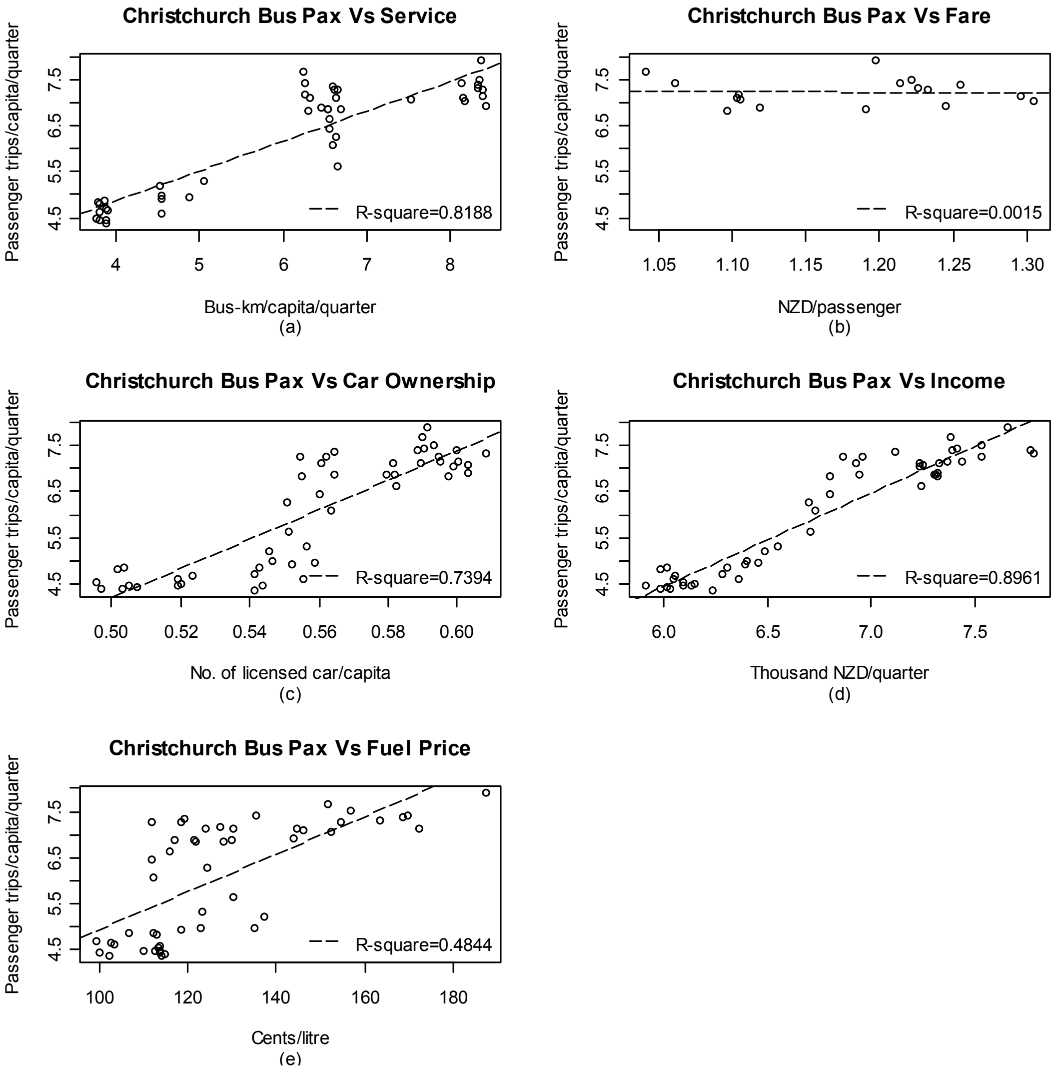


In terms of car ownership, as shown in figure 5.16(a) to (c), Christchurch had the highest car ownership per capita among the three cities. As shown in figure 5.22(a) and (c), car ownership in Christchurch flattened after the steep climb in petrol price in 2000. After that, as shown in figure 5.22(a), (b) and (c), car ownership started to climb again in 2003 as income increased but flattened again with another sharp increase in petrol price from 2005 to 2006.

5.3.2 Correlation analysis

The scatter plots of Christchurch bus patronage versus service, fares, car ownership, income and fuel price are depicted in figure 5.23(a) to (e) with the best fitted lines and R-square values.

Figure 5.23 Christchurch bus patronage versus potential influencing factors



With a high R-square value of 0.8188, as shown in figure 5.23(a), bus patronage in Christchurch appeared to have a strong positive relationship with service. Although the R-square value of 0.4844 is not as high in figure 5.23(e), bus patronage also appeared to have a positive relationship with fuel price. Car ownership and income, as shown by high R-square values of 0.7394 and 0.8961 in figure 5.23(c) and (d), was observed to be similar to Auckland, with a positive relationship between patronage and car ownership/income when they were above a certain level. As indicated by an extremely low R-square value of 0.0015 and an almost horizontal best fitted line in figure 5.23(b), no relationship was observed between bus patronage and fare.

5.4 Summary

This section summarises the findings from our preliminary analysis and information from ARTA, GWRC and ECan on events or changes that might have caused the observed trends.

5.4.1 Auckland

A summary of historical changes and possible effect on PT patronage in Auckland is depicted in table 5.1.

5.4.2 Wellington

A summary of historical changes and possible effect on PT patronage in Auckland is depicted in table 5.2.

5.4.3 Christchurch

A summary of historical changes and possible effect on PT patronage in Auckland is depicted in table 5.3.

Table 5.1 Summary of historic changes and possible effect on public transport patronage in Auckland

Date/period	Events/changes	Possible effect
1996 Q1 to 1999 Q2	• Car prices fell by 50% as Japanese second hand imports entered the market. • 1998 tariffs removed from cars. • Real fuel price declining. • Bus fare increased in 1997.	• Car ownership had a steep climb from 1996 to mid-1999. • Bus patronage declined from 1996 Q1to 1999 Q2. • Rail patronage was stationary.
1999 Q3 to 2003 Q2	• Real fuel priced increased sharply and then dropped to 2000 level. • Bus fare increased in 2000 Q4.	• Car ownership flattened and then continued to climb at a slower rate. • Bus patronage started to climb sharply. • Rail patronage still not moving.
2003 Q3 to 2005 Q4	• Real fuel price was increasing. • Opening of Britomart in Jul 2003. No need for bus transfer to CBD from The Strand. • Rail operator changed from Tranzmetro to Connex in Aug 2004. • Double tracking completed and frequency of train services improved in Feb 2005. • Bus fare increased 2004 Q3and 2005 Q3.	• Rail patronage started climbing and continued sharply. • Car ownership increased at a higher rate compared with previous period. • Bus patronage continued to increase.
2006 Q1 to 2008 Q2	• Rail service further improved as rolling stock refurbished and sufficient train drivers trained. • Stage one Northern Busway opened in Nov 2005. • Fuel price increased significantly in Apr 2006. • Four finance companies specialised in used car loans failed between May and Aug 2006. • Rail fare increased in Feb 2006 and Jan 2007. • Bus fare increased 2007 Q1. • Busway second stage opened in Feb 2008. • Real income dropped just after 2006 but increased afterwards in 2006 Q3.	• Bus patronage continued to climb and increased at a higher rate. • Rail patronage also increased sharply. • Both bus and rail patronage increased at a higher rate when the fuel price increased significantly. • Consistent with the observations in patronage and fuel price, the rate of increase in car ownership decreased and flattened in 2008.

Table 5.2 Summary of historic changes and possible effect on public transport patronage in Wellington

Date/period	Events/changes	Possible effect
Sep 2002	• Rail fare increase	• Negative impact on rail patronage
Jan 2003 – Jul 2003	• Rail track speed restrictions	• Negative impact on rail patronage
Apr and Oct 2003	• Hutt Valley bus service changes	• Increase in bus patronage
July 2006	• Fuel price spike	• Increase in both bus and rail patronage
Sep 2006	• Fare increase and new zonal system	• Negative impact both bus and rail patronage
Feb 2007	• Wellington bus re-scheduling disruptions	• Negative impact on bus patronage
Sep 2008	• Fare increase	• Negative impact on both bus and rail patronage
Oct 2008	• Fuel price spike	• Positive impact on both bus and rail patronage

Table 5.3 Summary of historic changes and possible effect on public transport patronage in Christchurch

Period	Events/changes	Possible effect
1998 to 2003	• Car prices fell by 50% as Japanese second-hand imports entered the market. • 1998 tariffs removed from cars. • Orbiter ring route first introduced in 1999, completed in 2000 and frequency increased from 15-minute headway to 10-minute headway in 2001. • Real fuel price reached a peak in 2003 Q3 and declined afterwards.	• Car ownership had a steep climb from 1996 to mid-1999 and flattened afterwards. • Bus patronage started to climb since 1998. The rate of increase steadily increased since 2000.
2003 to 2008	• Electronic Metrocard introduced in 2003. • The Metrostar bus service introduced in Nov 2004. • The Northern Star bus service introduced in Nov 2006. • Fuel price continued to increase and reached its peak in 2006.	• Growth of bus patronage slowed down but still followed an increasing trend with lower rate of increase.

6 The model

As discussed in section 3.2, two dynamic econometric modelling techniques were identified for time series analysis, namely, the partial adjustment model (PAM) and vector error correction model (VECM). Both PAM and VECM are well developed approaches to estimate relationships between explanatory and dependent variables over time. PAM assumes a geometrically declining adjustment process while VECM assumes a structural model behind the dynamics of interactions between the variables. Both models also enable estimation of short-run and long-run elasticities at the same time. However, VECM is very data intensive as it involves the estimation of a system of equations representing the inter-relationships between the endogenous and exogenous variables over time. As discussed in section 4.1, data availability varied among the three cities. Auckland had the most comprehensive set of data of the three cities. In fact, given the data set collected, Auckland was the only city that had enough data to support the development of a VECM. Stationarity tests[14] on the data collected for Auckland and other potential influencing factors collected on a national basis were performed as detailed in appendix B. The results indicated it would be feasible to deduce a VECM for Auckland. For Wellington and Christchurch, we simply did not have enough information to build a VECM model. As one of our objectives for this study was to identify the differences in characteristics of the three cities, a VECM for Auckland on its own would not be appropriate for this purpose. Therefore, we decided to adopt the alternative approach, ie PAM, for all three cities.

6.1 Variable definitions

As discussed in section 4.1, after consolidation, each regional data set contains six variables. The definition of the variables considered are summarised in table 6.1.

Table 6.1 Definition of variables

Variable	Definition	Unit
$Q_t^{X(M)}$	Patronage per capita in city X on mode M at time t	Passenger trips per capita per period
$S_t^{X(M)}$	Service kilometres per capita in city X on mode M at time t	Bus-km/train-km per capita per period
$F_t^{X(M)}$	Fare (real average revenue per passenger) in city X on mode M at time t	NZD per passenger
C_t^X	Car ownership per capita in city X on mode M at time t	Number of vehicles per capita
I_t	Income (real gross disposable income per capita) at time t	NZD per capita
O_t	Fuel price (real regular petrol price) at time t	Cents per litre

where $X = A$ for Auckland; $X = W$ for Wellington; $X = C$ for Christchurch; $M = B$ for bus; and $M = R$ for rail.

As discussed in section 4.2, we collected information on bus and rail services in Auckland and Wellington, and bus services in Christchurch. Moreover, the Auckland and Christchurch models were built on quarterly data while the Wellington model was built on annual data. Dummy variables were introduced for Auckland

[14] Stationary tests were implemented in JMulTi version 4.23 developed by Lutkepohl and Kratzig (2004)

and Christchurch in order to capture the seasonal effect. Therefore, there would be five equations with the following variable set:

1 Auckland bus services: $\left\{Q_t^{A(B)}, S_t^{A(B)}, F_t^{A(B)}, C_t^A, I_t, O_t, D_{2t}, D_{3t}, D_{4t}\right\}$ (Equation 6.1)

2 Auckland rail services: $\left\{Q_t^{A(R)}, S_t^{A(R)}, F_t^{A(R)}, C_t^A, I_t, O_t, D_{2t}, D_{3t}, D_{4t}\right\}$ (Equation 6.2)

3 Wellington bus services: $\left\{Q_t^{W(B)}, S_t^{W(B)}, F_t^{W(B)}, C_t^W, I_t, O_t\right\}$ (Equation 6.3)

4 Wellington rail services: $\left\{Q_t^{W(R)}, S_t^{W(R)}, F_t^{W(R)}, C_t^W, I_t, O_t\right\}$ (Equation 6.4)

5 Christchurch bus services: $\left\{Q_t^{C(B)}, S_t^{C(B)}, F_t^{C(B)}, C_t^C, I_t, O_t, D_{2t}, D_{3t}, D_{4t}\right\}$ (Equation 6.5)

where $D_{jt}^{X(M)}$ represents the dummy variable for the seasonal effect of the j th quarter in city X on mode M at time t. The seasonal effect represents the impact of the imbalance of public holidays and school holidays in the four quarters throughout the year. Note that only three dummy variables were needed to represent the four quarters as shown in table 6.2.

Table 6.2 Dummy variables to model seasonal effect

Case	D_{2t}	D_{3t}	D_{4t}
Time t in 1st quarter	0	0	0
Time t in 2nd quarter	1	0	0
Time t in 3rd quarter	0	1	0
Time t in 4th quarter	0	0	1

6.2 Model specification

As discussed in section 3.2, PAM is a special form of autoregressive model that simply includes the lagged dependent (or endogenous) variables as one of the regressors. In this case, our dependent variable is the long-run equilibrium demand for PT services, or more specifically $Q_t^{X(M)}$, as defined in table 6.1.

We assume that the long-run equilibrium demand, in passenger trips per capita, can be expressed as a function f of the service $S_t^{X(M)}$, fare $F_t^{X(M)}$, car ownership C_t^X, per capita disposable income I_t and petrol price O_t.

$$Q_t^{X(M)*} = f\left(S_t^{X(M)}, F_t^{X(M)}, C_t^X, I_t, O_t\right)$$ (Equation 6.6)

In estimating the demand model, we assume that all explanatory variables are given or determined exogenously. We further assume a geometrically declining adjustment process, which results in a typical lagged dependent variable model.

$$Q_t^{X(M)*} = f\left(S_t^{X(M)}, F_t^{X(M)}, C_t^X, I_t, O_t\right) + \theta^{X(M)} Q_{t-1}^{X(M)}$$ (Equation 6.7)

where $0 \le \theta^{X(M)} < 1$; and $1-\theta^{X(M)}$ is the adjustment coefficient which indicates the proportion of the gap between equilibrium and actual patronage that is closed each year.

The idea behind the PAM is that an individual's travel behaviour to a certain extent is a habit. One's choices today have an effect on one's future decisions. This is modelled by introducing a lagged independent variable on the right-hand side of the equation and the associated adjustment coefficient.

Assuming that f is in linear form, if all variables are transformed in logarithmic forms, the equation can be expressed as follows: [Type a quote from the document or the summary of an interesting point. You can position the text box anywhere in the document. Use the Text Box Tools tab to change the formatting of the pull quote text box.]

$$\begin{aligned}\ln Q_t^{X(M)} &= \alpha^{X(M)} + \beta_S^{X(M)} \ln S_t^{X(M)} + \beta_F^{X(M)} \ln F_t^{X(M)} + \beta_C^{X(M)} \ln C_t^{X} \\ &+ \beta_I^{X(M)} \ln I_t + \beta_O^{X(M)} \ln O_t + \theta^{X(M)} \ln Q_{t-1}^{X(M)} \\ &+ \gamma_2^{X(M)} D_{2t} + \gamma_3^{X(M)} D_{3t} + \gamma_4^{X(M)} D_{4t}\end{aligned} \qquad \text{(Equation 6.8a)}$$

where $X = A$ for Auckland; $X = C$ for Christchurch; $M = B$ for bus; and $M = R$ for rail.

As discussed earlier, dummy variables $D_{jt}^{X(M)}$ were created to capture the seasonal effects for the Auckland and Christchurch models, which would not be necessary for Wellington as the model for Wellington was developed based on annual data. The general form of the model for Wellington is as follows:

$$\begin{aligned}\ln Q_t^{X(M)} &= \alpha^{X(M)} + \beta_S^{X(M)} \ln S_t^{X(M)} + \beta_F^{X(M)} \ln F_t^{X(M)} + \beta_C^{X(M)} \ln C_t^{X} \\ &+ \beta_I^{X(M)} \ln I_t + \beta_O^{X(M)} \ln O_t + \theta^{X(M)} \ln Q_{t-1}^{X(M)}\end{aligned} \qquad \text{(Equation 6.8b)}$$

where $X = W$ for Wellington; $M = B$ for bus; and $M = R$ for rail.

With this model, as represented by equations 6.8a and 6.8b, we can estimate both the short-run and long-run elasticities at the same time; the short-run elasticities as the coefficients of the independent variables and the long-run elasticities as the short-run elasticities divided by the adjustment coefficient $1-\theta^{X(M)}$.

Elasticity of demand, eg price elasticity of demand, is defined as the measure of responsiveness in the quantity demanded for a commodity as a result of change in price of the same commodity. It is a measure of how consumers react to a change in price. The elasticities of demand to be estimated from equations 6.8a and 6.8b are defined in table 6.3 with the corresponding estimation coefficients.

Table 6.3 Definitions of elasticity of demand and the corresponding estimation coefficients

Elasticity	Definition	Estimation coefficients	
		Short-run	Long-run
Service elasticity of demand	$\dfrac{\text{Percentage change in Patronage per capita}}{\text{Percentage change in Service}}$	$\beta_S^{X(M)}$	$\dfrac{\beta_S^{X(M)}}{1-\theta^{X(M)}}$
Fare elasticity of demand	$\dfrac{\text{Percentage change in Patronage per capita}}{\text{Percentage change in Fare}}$	$\beta_F^{X(M)}$	$\dfrac{\beta_F^{X(M)}}{1-\theta^{X(M)}}$
Car ownership elasticity of demand	$\dfrac{\text{Percentage change in Patronage per capita}}{\text{Percentage change in Car ownership per capita}}$	$\beta_C^{X(M)}$	$\dfrac{\beta_C^{X(M)}}{1-\theta^{X(M)}}$
Income elasticity of demand	$\dfrac{\text{Percentage change in Patronage per capita}}{\text{Percentage change in Real income per capita}}$	$\beta_I^{X(M)}$	$\dfrac{\beta_I^{X(M)}}{1-\theta^{X(M)}}$

Elasticity	Definition	Estimation coefficients	
		Short-run	Long-run
Fuel price elasticity of demand	$\frac{\text{Percentage change in Patronage per capita}}{\text{Percentage change in Fuel price}}$	$\beta_O^{X(M)}$	$\frac{\beta_O^{X(M)}}{1-\theta^{X(M)}}$

Equation 6.8a or 6.8b was applied to individual modes (M) in individual cities (X) so that mode-specific elasticities $\left(\beta_j^{X(M)}\right)$ for each significant factor j could be obtained for each city. The influencing factors considered were included in the equation only if the $t-$statistics associated with the corresponding coefficient $\left(\beta_j^{X(M)}\right)$ showed that the null hypothesis of $\beta_j^{X(M)} = 0$ could be rejected at an acceptable level of significance. A minimum of 5% level of significant was applied in general.

6.3 Model estimation[15]

As discussed in chapter 5, each time series revealed a different story behind the trends. Therefore, we developed models for different periods as appropriate based on the history of each city for each mode. We introduced dummy variables to represent seasonal effects and possible influences of policy changes/ events identified that created the observed breaks and change of trend. If the null hypothesis that the coefficient of a dummy variable equals zero was rejected, the effect represented by the dummy variable would be considered to be significant. We would then include the dummy variable in the equation or change the model time period accordingly.

For each model time period, we applied a procedure following the concept of stepwise regression analysis. Stepwise regression analysis is a well established statistical procedure to determine the most appropriate set of independent variables. The general idea is either to start with a full set of variables and eliminate insignificant variables until the best fit is achieved or to start with one variable and add significant variable(s) until the best fit is achieved.

We first included all five independent variables, ie service, fare, car ownership, income and fuel price. We then eliminated variables with insignificant effect until all the remaining variables were significant. If, however, none of the variables appeared to be significant when we considered them all together, we then applied this concept in the opposite direction. Instead of including all variables, we started with only the PT-related variables, ie service and fare. We then continued with the strategy of keeping the significant variables and adding a variable if it was significant.

In this section, we describe how we applied this model estimation procedure to each mode in each of the three cities to obtain the best fit model(s). The corresponding R-outputs are detailed in appendix C.

6.3.1 Auckland

6.3.1.1 Auckland bus patronage forecast model

Auckland bus patronage model (1999Q3–2008Q2)

Step 1: Full period, all variables, all indicators

As discussed in the trend analysis for Auckland in section 5.1.1, the first observed break was in 1999 Q3. We created a dummy variable to represent the break, D_t^{99Q3} where

15 Model estimation is implemented with the Dynamic Linear Regression R-package developed by Zeileis (2009).

$$D_t^{99Q3} = \begin{cases} 1 & \text{if } t \text{ is in or after 1999 Q3} \\ 0 & \text{otherwise} \end{cases}$$

We included this new dummy variable in equation 6.8a to form a new equation:

Service + fare + car ownership+ fuel price + income + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A \\ &+ \beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+ \gamma_2^{A(B)} D_{2t} + \gamma_3^{A(B)} D_{3t} + \gamma_4^{A(B)} D_{4t} + \gamma^{99Q3} D_t^{99Q3} \end{aligned} \qquad \text{(Equation 6.9)}$$

where $t \in [1996Q1, 2008Q2]$

With dynamic regression, we then determined the best estimates for the coefficients in equation 6.9 and their respective levels of significance. It was observed that only two dummy variables were significant, the Q4 dummy variable, $D_{4t}^{A(B)}$, and the dummy variable for the break in 1993 Q3, D_t^{99Q3}. The R-square value for the best fitted equation 6.9 was 0.9561.

Step 2: Period changed, take out insignificant dummy variables

Since the dummy variable for the break in 1993 Q3 was significant, we changed our model period to 1993 Q3 – 2008 Q2. We also took out the seasonal dummy variables for Q2 and Q3 as they were not significant. The amended equation is as follows:

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

$$\begin{aligned} \ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A \\ &+ \beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+ \gamma_4^{A(B)} D_{4t} \end{aligned} \qquad \text{(Equation 6.10)}$$

where $t \in [1999Q3, 2008Q2]$

Again, we obtained the best estimates for the coefficients in equation 6.10 and their respective levels of significance with dynamic regression. It was observed that fare and fuel price were not significant, ie only three of the independent variables were significant: service, car ownership and income. The R-square value for the best-fitted equation 6.10 was 0.9636. However, the estimated coefficient of car ownership was negative while that of income was positive which was counter intuitive. As discussed in the trend analysis for Auckland in section 5.1.1, the trends of car ownership and income were correlated. It would not be appropriate to include both variables in the equation.

Step 3: Take out insignificant variables, try different combinations of significant variables

We then excluded the two insignificant variables, bus fare and fuel price. For the remaining three variables, as car ownership and income should not be included at the same time, there would be only two combinations as shown in equations 6.11 and 6.12:

Service + income+ Q4 seasonal dummy

$$\ln Q_t^{A(B)} = \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_I^{A(B)} \ln I_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t} \qquad \text{(Equation 6.11)}$$

where $t \in [1999Q3, 2008Q2]$

Only service and seasonal effect of Q4 were significant. The R-square value for the best-fitted equation 6.11 was 0.9519.

Service + car ownership+ Q4 seasonal dummy

$$\ln Q_t^{A(B)} = \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A + \theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t} \quad \text{(Equation 6.12)}$$

where $t \in [1999Q3, 2008Q2]$

All three independent variables were significant with an R-square value of 0.9562. Equation 6.12 is the most satisfactory forecast model for Auckland bus patronage per capita. The forecast model is elaborated with the best estimated coefficients as follows:

Service + car ownership+ Q4 seasonal dummy

$$\begin{aligned} \ln Q_t^{A(B)} &= -0.8046 + 0.4977 \ln S_t^{A(B)} - 1.3239 \ln C_t^A \\ &\quad +0.5071 \ln Q_{t-1}^{A(B)} - 0.0661 D_{4t} \end{aligned} \quad \text{(Equation 6.13)}$$

where $t \in [1999Q3, 2008Q2]$

Auckland bus patronage model (2003Q3–2008Q2)

As discussed in section 5.1.1, the opening of the Britomart Transport Centre in July 2003 marked the beginning of a new trend in railway patronage in Auckland. For consistency in modelling for Auckland, we should also develop a model for bus patronage for the same period.

Step 1: Period 2003 Q3 to 2008 Q2, all variables, all indicators

We included all five independent variables plus a dummy variable to model a possible change of trend due to the opening of Northern Busway in July 2005. We created a dummy variable to represent the break, D_t^{05Q3} where

$$D_t^{05Q3} = \begin{cases} 1 & \text{if } t \text{ is in or after 2005 Q3} \\ 0 & \text{otherwise} \end{cases}$$

We included this new dummy variable in equation 6.8a to form a new equation:

Service + fare + car ownership + income + fuel price + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A \\ &\quad +\beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &\quad +\gamma_2^{A(B)} D_{2t} + \gamma_3^{A(B)} D_{3t} + \gamma_4^{A(B)} D_{4t} + \gamma^{05Q3} D_t^{05Q3} \end{aligned} \quad \text{(Equation 6.14)}$$

where $t \in [2003Q3, 2008Q2]$

We determined the best estimates for the coefficients in equation 6.14 and their respective levels of significance. The R-square value for the best fitted equation 6.14 was 0.9893. It was observed that only one dummy variable was significant, the Q4 dummy variable, D_{4t}.

Step 2: Period unchanged, take out insignificant dummy variables

Since the dummy variable for the break in 2005 Q3 was insignificant, we excluded this dummy variable and the model period remained unchanged as 2003 Q3 – 2008 Q2. We also took out the seasonal dummy variables for quarter 2 and 3 as they were not significant. The amended equation is as follows:

Service + fare + car ownership + income + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} \\ &+\beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+\gamma_4^{A(B)} D_{4t}\end{aligned} \qquad \text{(Equation 6.15)}$$

where $t \in [2003Q3, 2008Q2]$

Again, we obtained the best estimates for the coefficients in equation 6.15 and their respective levels of significance with dynamic regression. It was observed that fare and income were not significant, ie only three of the independent variables were significant, including, service, car ownership and fuel price. The R-square value for the best fitted equation 6.15 was 0.9857.

Step3: Take out insignificant variables, include all significant variables

We then excluded the two insignificant variables, bus fare and income. The resulting model form is as follows:

Service + car ownership + fuel price+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} + \beta_O^{A(B)} \ln O_t \\ &+\theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t}\end{aligned} \qquad \text{(Equation 6.16)}$$

where $t \in [2003Q3, 2008Q2]$

The best fitted model with equation 6.16 had a very high R-square value of 0.9829 and all the three independent variables, service car ownership and fuel price were significant. The equation with the estimated coefficients can be elaborated as follows:

Service + car ownership + fuel price+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= -1.9306 + 0.4603 \ln S_t^{A(B)} - 1.9618 \ln C_t^{A} + 0.2040 \ln O_t \\ &+0.3677 \ln Q_{t-1}^{A(B)} - 0.0478 D_{4t}\end{aligned} \qquad \text{(Equation 6.17)}$$

where $t \in [2003Q3, 2008Q2]$

6.3.1.2 Auckland rail patronage forecast model

As discussed in the trend analysis in section 5.1.1, the opening of the Britomart Transport Centre in July 2003 clearly marked the beginning of a new trend in rail patronage in Auckland. We, therefore, developed our forecast model based on the observations from the period 2003Q–2008Q2.

Auckland rail patronage model (2003Q3–2008Q2)

Step1: Full period, all variables, all indicators

We started with a model including all the variables, including the five independent variables plus the seasonal indicators as follows:

Service + fare + car ownership+ fuel price + income + seasonal dummies

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} \\ &+\beta_I^{A(R)} \ln I_t + \beta_O^{A(R)} \ln O_t + \theta^{A(R)} \ln Q_{t-1}^{A(R)} \\ &+\gamma_2^{A(R)} D_{2t} + \gamma_3^{A(R)} D_{3t} + \gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation 6.18)}$$

where $t \in [2003Q3, 2008Q2]$

The purpose of this step was to find out the significance of the dummy variables so that we could include the significant ones. The results indicated that, as for Auckland bus, only the dummy variable for Q4 was significant. The R-square value for the best fitted equation 6.18 was 0.9904.

Step 2: Take out insignificant dummy variables

We excluded the two dummy variables for Q2 and Q3, ie the model form is as follows:

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} \\ &+\beta_I^{A(R)} \ln I_t + \beta_O^{A(R)} \ln O_t + \theta^{A(R)} \ln Q_{t-1}^{A(R)} \\ &+\gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation 6.19)}$$

where $t \in [2003Q3, 2008Q2]$

The results indicated that car ownership and fuel price were not significant. The R-square value for the best fitted equation 6.19 was 0.9871. However, although income appeared to be a significant factor, the estimated coefficient was positive which was counter intuitive.

Step 3: Take out insignificant variables, try different combinations of significant variables

Based on observations in step 2, we created different possible combinations of variables as follows:

Service + fare + car ownership+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} \\ &+\theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation 6.20)}$$

where $t \in [2003Q3, 2008Q2]$

The best estimates for equation 6.20 indicated that the influence of car ownership was not significant. The R-square value for the best fitted equation 6.20 was 0.9786.

Service + fare + income+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_I^{A(R)} \ln I_t \\ &+\theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation 6.21)}$$

where $t \in [2003Q3, 2008Q2]$

For equation 6.21, income appeared to have significant influence instead of car ownership as in the bus model but the coefficient was positive. The R-square value for the best fitted equation 6.21 was 0.9867.

Service + fare+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t}$$ (Equation 6.22)

where $t \in \left[2003Q3, 2008Q2\right]$

The R-square value for the best fitted equation 6.22 was 0.9785. Both service and fare were significant.

Equations 6.21 and 6.22 are both satisfactory with all independent variables statistically significant and satisfactory R-square values. They are elaborated with the best estimated coefficients as follows:

Model 1 – service + fare + income+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = -1.1448 + 0.9946 \ln S_t^{A(R)} - 0.9672 \ln F_t^{A(R)} + 1.606 \ln I_t + 0.2957 \ln Q_{t-1}^{A(R)} - 0.1325 D_{4t}$$ (Equation 6.23)

where $t \in \left[2003Q3, 2008Q2\right]$

Model 2 – service + fare+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = 1.6913 + 0.8827 \ln S_t^{A(R)} - 0.6755 \ln F_t^{A(R)} + 0.4584 \ln Q_{t-1}^{A(R)} - 0.1478 D_{4t}$$ (Equation 6.24)

where $t \in \left[2003Q3, 2008Q2\right]$

6.3.2 Wellington

As discussed in section 5.2, we developed the model for Wellington based on annual data as information on fares was not available on a quarterly basis. As a result, seasonal factors were not considered in the Wellington models.

6.3.2.1 Wellington bus patronage forecast model

Wellington bus patronage model (FY1999/2000–FY2007/08)

Step 1: Include all variables

We applied the model estimation methodology as discussed earlier. We started with all five independent variables included. The model form is as follows:

Service + fare + car ownership+ fuel price + income

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_C^{W(B)} \ln C_t^{W} + \beta_I^{W(B)} \ln I_t + \beta_O^{W(B)} \ln O_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)}$$ (Equation 6.25)

where $t \in \left[FY99/00, FY07/08\right]$

However, the results indicated that none of the five variables was significant. The R-square value for the best fitted equation 6.25 was 0.9273.

Step 2: Test significance of PT-related variables variables

Because none of the variables was significant in the previous step, we changed our strategy. We tested the significance of both the PT-related variables variables, ie service and fare, and their significance individually.

Service + fare

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.26)}$$

Only bus fare had significant impact on bus patronage. The R-square value for the best fitted equation 6.26 was 0.9170.

Service only

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.27)}$$

Service did not have significant impact on bus patronage. The R-square value for the best fitted equation 6.26 was 0.7881.

Fare only

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.28)}$$

Bus fare had significant impact on bus patronage. The R-square value for the best fitted equation 6.28 was 0.9006.

Step 3: Test significance of other variables

We then tested the significance of other variables, including car ownership, fuel price and income.

Fare + car ownership

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_C^{W(B)} \ln C_t^{W} + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.29)}$$

The results indicated that none of the factors had significant impact on bus patronage in Wellington. The R-square value for the best fitted equation 6.29 was 0.9011.

Fare + fuel price

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_O^{W(B)} \ln O_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.30)}$$

Only bus fare had significant impact on bus patronage. The R-square value for the best fitted equation 6.26 was 0.9015.

Fare + income

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_I^{W(B)} \ln I_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.31)}$$

The results indicated that none of the factors had significant impact on bus patronage in Wellington. The R-square value for the best fitted equation 6.31 was 0.9042.

Therefore, equation 6.28 was the only satisfactory model identified for Wellington. The best fitted forecast model for Wellington bus patronage forecast can be elaborated as follows:

Fare only

$$\ln Q_t^{W(B)} = 2.0590 - 0.2312 \ln F_t^{W(B)} + 0.4934 \ln Q_{t-1}^{W(B)} \quad \text{(Equation 6.32)}$$

6.3.2.2 Wellington rail patronage forecast model

Wellington rail patronage model (FY1999/2000–FY2007/08)

Step 1: Include all variables

We applied the model estimation methodology as discussed earlier. We started with all the five independent variables included. The model form is as follows:

Service + fare + car ownership+ fuel price + income

$$\begin{aligned}\ln Q_t^{W(R)} &= \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^{W} \\ &+\beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)}\end{aligned} \qquad \text{(Equation 6.33)}$$

where $t \in \left[FY99/00, FY07/08\right]$

However, the results indicated that none of the five variables was significant. The R-square value for the best fitted equation 6.33 was 0.9913.

Step 2: Test significance of PT-related variables variables

Because none of the variables was significant in the previous step, we changed our strategy. We tested the significance of both the PT-related variables variables, ie service and fare, and their significance individually.

Service + fare

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \qquad \text{(Equation 6.34)}$$

The results indicated that only service had significant influence. The R-square value for the best fitted equation 6.34 was 0.8987.

Service only

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \qquad \text{(Equation 6.35)}$$

The results indicated that service had significant influence. The R-square value for the best fitted equation 6.35 was 0.8938.

Fare only

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \qquad \text{(Equation 6.36)}$$

Fare was just marginally significant in equation 6.36. The R-square value for the best fitted equation 6.36 was 0.6881.

Step-3 Test significance of other variables

We then tested the significance of other variables, including car ownership, fuel price and income as additional variables to service or fare.

$$\begin{aligned}\ln Q_t^{W(R)} &= \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^{W} \\ &+\beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)}\end{aligned} \qquad \text{(Equation 6.37)}$$

Fare + car ownership + income + fuel price

Only fuel price was significant. The R-square value for the best fitted equation 6.37 was 0.9418.

Fare + car ownership

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^W + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.38)}$$

Both fare and car ownership were not significant. The R-square value for the best fitted equation 6.38 was 0.6881.

Fare + income

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_I^{W(R)} \ln I_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.39)}$$

Both fare and income were not significant. The R-square value for the best fitted equation (6.39) was 0.6882.

Fare + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.40)}$$

Only fuel price was significant equation 6.40. The R-square value for the best fitted equation 6.40 was 0.8810.

As shown above, the level of significance of the variables was not satisfactory when we considered fare with car ownership, fuel price and income as independent variables.

Service + car ownership + income + fuel price

$$\begin{aligned}\ln Q_t^{W(R)} = \alpha^{W(R)} &+ \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_C^{W(R)} \ln C_t^W \\ &+ \beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)}\end{aligned} \quad \text{(Equation 6.41)}$$

Only service and fuel price were significant in equation 6.41. The R-square value for the best fitted equation 6.41 was 0.9873.

Service + car ownership + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_C^{W(R)} \ln C_t^W + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.42)}$$

All the three factors considered were significant with an R-square value of 0.9846. Equation 6.42 is a satisfactory model.

Service + income + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.43)}$$

All the three factors considered were significant with an R-square value of 0.9873. Equation 6.43 is a satisfactory model.

As shown above, two models were identified with satisfactory performance. They are elaborated as follows:

Model 1 – service + car ownership + fuel price

$$\ln Q_t^{W(R)} = -1.3759 + 0.7424 \ln S_t^{W(R)} - 0.3246 \ln C_t^{W} + 0.1303 \ln O_t + 0.6889 \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.44)}$$

Model 2 – service + income + fuel price

$$\ln Q_t^{W(R)} = -0.4663 + 0.7847 \ln S_t^{W(R)} - 0.2176 \ln I_t + 0.1212 \ln O_t + 0.6900 \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.45)}$$

In equation 6.44, the rail patronage in Wellington appears to be affected by the same factors affecting bus patronage in Auckland. The effects of service, car ownership and fuel price are statistically significant. The signs of the coefficients are also intuitively correct, with positive effect from service and fuel price and negative effect from car ownership. Equation 6.45 is similar to equation 6.44 except that the effect of car ownership is replaced with income effect. The income effect was also negative.

6.3.3 Christchurch

Christchurch fitted bus patronage forecast model (1998Q1–2008Q2)

Step 1: Full period, all variables, all indicators

For Christchurch bus services, we first tested the effect of 'gold coin' cash fare by introducing a dummy variable (an indicator) for the observed change of trend in 1998 Q1

$$D_t^{98Q1} = \begin{cases} 1 & \text{if } t \text{ is in or after 1998 Q1} \\ 0 & \text{otherwise} \end{cases}$$

As discussed in section 4.1, available bus fare information covers only a very short period of time from 2004. Thus it was not possible to include fares in our analysis. We included the remaining four independent variables.

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t} + \gamma^{98Q1} D_t^{98Q1} \quad \text{(Equation 6.46)}$$

Service + car ownership+ fuel price + income + seasonal dummies + break

where $t \in \left[1997Q1, 2008Q2\right]$

The results indicated that the seasonal effect of Q4 and the effect of the introduction of gold coin cash fare were both significant. The R-square value for the best fitted equation 6.46 was 0.9899.

Step 2: Period changed, take out insignificant dummy variables

As the effect of gold coin cash fare was found to be significant, we changed the model period to 1998 Q1 to 2008 Q2. As in the Auckland case, we also found the seasonal effect of Q4 was significant. The dummy variables for seasonal effect of Q2 and Q3 were excluded.

Service + car ownership+ fuel price + income + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation 6.47)

where $t \in [1998Q1, 2008Q2]$

The results indicated the influence of service and car ownership was significant. The R-square value for the best fitted equation 6.47 was 0.9875.

Step 3: Take out insignificant variables, try different combinations of significant variables

As service was a significant variable, we tested combinations of service with other variables, including car ownership, income and fuel. We also tested service as the only independent variable.

Service + car ownership+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ Equation 6.48)

where $t \in [1998Q1, 2008Q2]$

The results indicated service was the only variable with significant influence in Christchurch. The R-square value for the best fitted equation 6.48 was 0.9865.

Service +income+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_I^{C(B)} \ln I_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation 6.49)

where $t \in [1998Q1, 2008Q2]$

The results indicated service was the only variable with significant influence in Christchurch. The R-square value for the best fitted equation 6.49 was 0.9861.

Service + fuel+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation 6.50)

where $t \in [1998Q1, 2008Q2]$

The results indicated service was the only variable with significant influence in Christchurch. The R-square value for the best fitted equation 6.50 was 0.9862.

Service only+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation 6.51)

where $t \in [1998Q1, 2008Q2]$

The only satisfactory model was equation 6.51. Service was the only independent variable and had a high R-square value of 0.9861. The forecast model with the estimated coefficients was elaborated as follows:

Service only+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = 0.1147 + 0.0708 \ln S_t^{C(B)} - 0.0898 \ln Q_{t-1}^{C(B)} + 0.8863 D_{4t}$$ (Equation 6.52)

where $t \in [1998Q1, 2008Q2]$

Christchurch fitted bus patronage forecast model (2004Q1–2008Q2)

Step 1: Full period, all variables, all indicators

To test the influence of the introduction of the Metrocard in July 2003, we introduced a new indicator (dummy) variable D_t^{03Q3} as follows:

$$D_t^{03Q3} = \begin{cases} 1 & \text{if } t \text{ is in or after 2003 Q3} \\ 0 & \text{otherwise} \end{cases}$$

Service + car ownership+ fuel price + income + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t} + \gamma^{03Q3} D_t^{03Q3} \end{aligned} \quad \text{(Equation 6.53)}$$

where $t \in [1998Q1, 2008Q2]$

The results indicated that the introduction of the Metrocard had significant impact on bus patronage. The R-square value for the best fitted equation 6.53 was 0.9919.

Step 2: Period changed, take out insignificant dummy variables

Since the impact of the introduction of the Metrocard was significant, theoretically we should build a model based on the data from 2003 Q3 to 2008 Q2. However, bus fare information was available only from 2004 Q1. Therefore, we could only build a model for the period 2004 Q1 to 2008 Q2. Again, we started with all five independent variables and seasonal dummy variables.

Service + fare + car ownership+ fuel price + income + seasonal dummies

$$\begin{aligned} \ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C \\ &+ \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} \\ &+ \gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t} \end{aligned} \quad \text{(Equation 6.54)}$$

where $t \in [2004Q1, 2008Q2]$

The results indicated only three variables were significant: car ownership, fuel price and seasonal effect of Q4. The R-square value for the best fitted equation 6.54 was 0.9477.

We excluded seasonal dummy variables for Q2 and Q3 and tested again including all five independent variables.

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

$$\begin{aligned} \ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C \\ &+ \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} \\ &+ \gamma_4^{C(B)} D_{4t} \end{aligned} \quad \text{(Equation 6.55)}$$

where $t \in [2004Q1, 2008Q2]$

The results indicated that we had two variables with significant effect: fare and fuel price. The R-square value for the best fitted equation 6.55 was 0.9212.

Fare + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \quad \text{(Equation 6.56)}$$

where $t \in [2004Q1, 2008Q2]$

The results indicated that with fare and fuel as independent variables, the R-square value for the best fitted equation 6.56 was only 0.8717. We tried to include additional variables with a view to improve the model performance.

Step 3: Try different combinations of variables

Service + fare + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation 6.57)}$$

where $t \in [2004Q1, 2008Q2]$

Only fuel price was a significant factor. The R-square value for the best fitted equation 6.57 was 0.8738.

Fare + car ownership + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation 6.58)}$$

where $t \in [2004Q1, 2008Q2]$

Only fare and fuel price were significant. The R-square value for the best fitted equation 6.58 was 0.9029.

Service + fare + car ownership + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation 6.59)}$$

where $t \in [2004Q1, 2008Q2]$

Three of the four factors considered, fare, car ownership and fuel, were significant. The R-square value for the best fitted equation 6.59 was 0.9111. However, in this case, the estimated coefficient for the lagged variable was -0.17379 which was negative. This model was not acceptable.

Service + fare + income + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation 6.60)}$$

where $t \in [2004Q1, 2008Q2]$

Only fuel price was significant and the R-square value for the best fitted equation 6.60 was 0.9029.

The results indicated that none of the four models above were totally satisfactory in terms of the R-square value, plus all independent variables being significant. We believe this was due to the quality of the information collected for Christchurch. As discussed in section 4.2 and section 5.3.1, the service data was only an approximation from bus service timetables and there were also issues with the fare information with missing data. Two models were selected for further consideration as follows:

Model 1 – fare + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = 0.1731 - 0.2620 \ln F_t^{C(B)} + 0.2785 \ln O_t + 0.2373 \ln Q_{t-1}^{C(B)} - 0.0429 D_{4t} \qquad \text{(Equation 6.61)}$$

where $t \in [2004Q1, 2008Q2]$

Model 2 – service + fare +income + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= -0.2646 + 0.0767 \ln S_t^{C(B)} - 0.3697 \ln F_t^{C(B)} + 0.3904 \ln I_t \\ &+ 0.1954 \ln O_t + 0.2039 \ln Q_{t-1}^{C(B)} - 0.0459 D_{4t}\end{aligned} \qquad \text{(Equation 6.62)}$$

where $t \in [2004Q1, 2008Q2]$

6.4 Forecast models

6.4.1 Auckland

As discussed in section 4.6.1, the model period for Auckland was identified as 1996 Q1 to 2008 Q2 and the models were developed based on quarterly information.

6.4.1.1 Auckland bus patronage forecast model

As illustrated in section 6.3.1.1, two forecast models were estimated for forecasting Auckland bus patronage per capita. The forecast models are:

Auckland bus patronage model (1999Q3–2008Q2)

Service + car ownership+ Q4 seasonal dummy – equation 6.13 in section 6.3.1.1

$$\begin{aligned}\ln Q_t^{A(B)} &= -0.8046 + 0.4977 \ln S_t^{A(B)} - 1.3239 \ln C_t^{A} \\ &+ 0.5071 \ln Q_{t-1}^{A(B)} - 0.0661 D_{4t}\end{aligned} \qquad \text{(Equation 6.63)}$$

where $t \in [1999Q3, 2008Q2]$

Auckland bus patronage model (2003Q3–2008Q2)

Service + car ownership + fuel price+ Q4 seasonal dummy – equation 6.17 in section 6.3.1.1

$$\begin{aligned}\ln Q_t^{A(B)} &= -1.9306 + 0.4603 \ln S_t^{A(B)} - 1.9618 \ln C_t^{A} + 0.2040 \ln O_t \\ &+ 0.3677 \ln Q_{t-1}^{A(B)} - 0.0478 D_{4t}\end{aligned} \qquad \text{(Equation 6.64)}$$

where $t \in [2003Q3, 2008Q2]$

As shown in equations 6.63 and 6.64, seasonal effect of the 4th quarter was found to be significant and had a negative coefficient, ie the bus patronage in the 4th quarter was significantly less than the rest of the year. This would be the combined effect of school, university and public holidays. Two breaks were identified as described in section 5.1.1: 1) 1999Q3 due to the effect of changes in policy on second-hand imported cars; and 2) 2003Q3 due to the opening of Britomart train station. Thus we applied the dynamic model to the two periods: 1) 1999Q3 to 2008Q2; and 2) 2003Q3 to 2008Q2, with a dummy variable for the Q4 effect. Equation 6.63 as illustrated above is the best fitted model for period 1 while equation 6.64 is the best fitted model for period 2.

Table 6.4 Fitted bus patronage forecast model details and statistics for Auckland

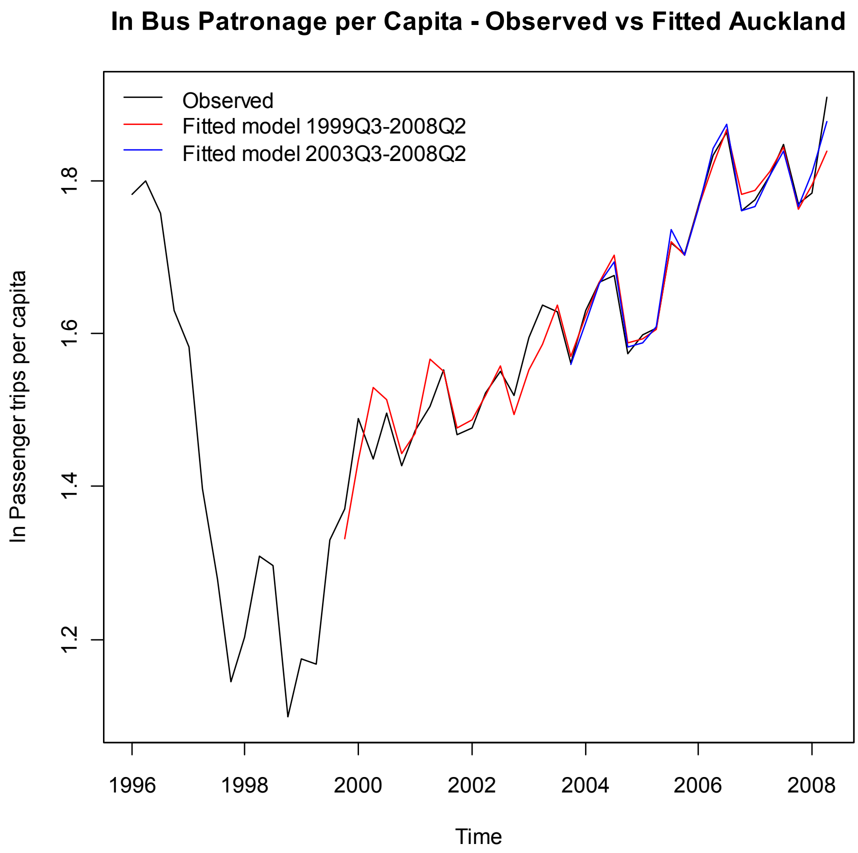
Description		Model parameters	
Variable	**Unit**	**99–08 Equation 6.63**	**03–08 Equation 6.64**
Intercept	Passengers per capita	-0.8046	-1.9306***
Bus service, $\ln S_t^{A(B)}$	Bus-km per capita	0.4977***	0.4603***
Car ownership, $\ln C_t^{A}$	Car ownership per capita	-1.3239^	-1.9618**
Fuel price, $\ln O_t$	Real fuel price	n/a	0.2040**
Seasonal effect Q4, $D_{4t}^{A(B)}$		-0.0661***	-0.0478**
Lagged patronage per capita, $\ln Q_{t-1}^{A(B)}$	Passengers per capita	0.5071***	0.3677***
R-square		0.9562	0.9829
Adjusted R-squared		0.9504	0.9763
Residual standard error		0.0318	0.0159
F-statistic		163.8	149.6
P-value (F-statistic)		<2.2E-16	5.15E-11

Note: Significant codes: '***' 0.001'**' 0.01 '*' 0.05 '^' 0.1

The model parameters and performance statistics are summarised in table 6.4. For the period of 1999Q3 to 2008Q2, it was evident that bus service kilometres had a positive influence on bus patronage while car ownership had a negative effect. The influence of fuel price was found to be significant only in the five-year period 2003Q3 to 2008Q2, and the effect was positive. Car ownership was the only elastic factor among the three factors identified. The signs of the estimated coefficients were all intuitively correct. We concluded that both models were satisfactory.

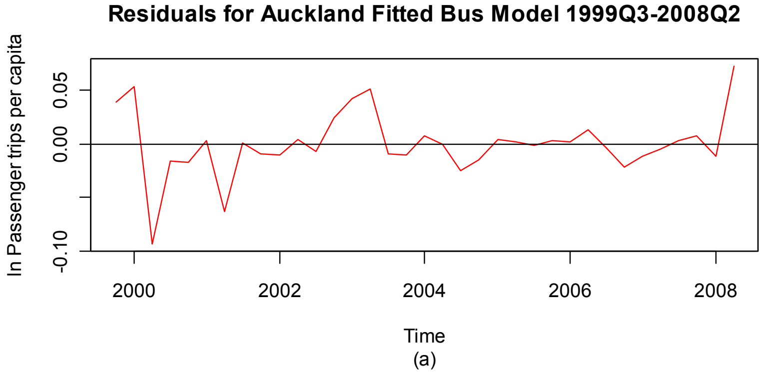
In this case, the model developed based on the later period, ie 2003 Q3 to 2008 Q2 had a better fit as the R-square value of 0.9829 was higher than that of 1999 Q3 to 2008 Q3 which was 0.9562.

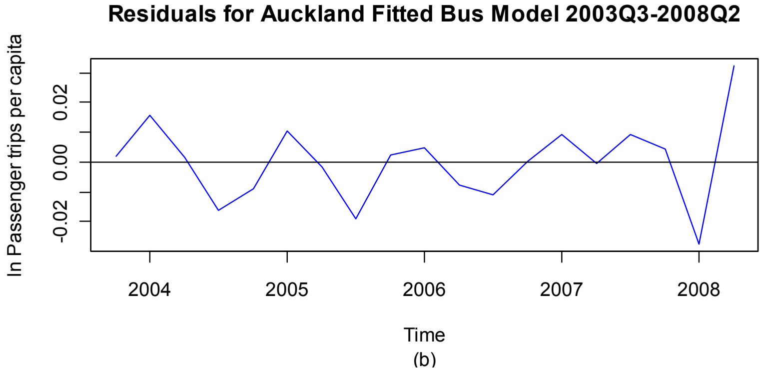
Figure 6.1 Observed versus fitted values – Auckland bus patronage forecast model



The fitted/predicted values from the two models identified for bus patronage forecast in Auckland are plotted with the observed values as shown in figure 6.1. The observation from figure 6.1 shows that the model developed based on data from the later period 2003 Q3–2008 Q2 has a better fit.

Figure 6.2 Residuals – Auckland bus patronage forecast model





The residual values are plotted in figure 6.2 (a) and (b) and they appear to follow a random pattern. This is consistent with our conclusion that both models were satisfactory.

No satisfactory forecast model was identified based on the monthly data for Northern Busway. We believe there were two main reasons:

1 We had only 18 months of data collected during phase one of the busway being open and six months of data collected after full opening of the busway. We would expect there to be a significant behaviour change during this period.

2 Seasonal factors (monthly variation) should be considered in the model. Only two observations for each month, 2006 and 2007, were available which was not sufficient to derive a meaningful monthly index.

6.4.1.2 Auckland rail patronage forecast model

As discussed in section 6.3.1.2, we developed our forecast model based on the observations from the period 2003Q3–2008Q2 due to the observed significant effect of the opening of the Britomart Transport Centre in July 2003 in the trend analysis in section 5.1.1. As illustrated in section 6.3.1.2, two forecast models were estimated for forecasting Auckland rail patronage per capita. The forecast models are:

Model 1 – service + fare + income+ Q4 seasonal dummy – equation 6.23 in section 6.3.1.2

$$\begin{aligned} \ln Q_t^{A(R)} &= -1.1448 + 0.9946 \ln S_t^{A(R)} - 0.9672 \ln F_t^{A(R)} + 1.606 \ln I_t \\ &\quad +0.2957 \ln Q_{t-1}^{A(R)} - 0.1325 D_{4t} \end{aligned} \qquad \text{(Equation 6.65)}$$

where $t \in [2003Q3, 2008Q2]$

Model 2 – service + fare+ Q4 seasonal dummy– equation 6.24 in section 6.3.1.2

$$\ln Q_t^{A(R)} = 1.6913 + 0.8827 \ln S_t^{A(R)} - 0.6755 \ln F_t^{A(R)} + 0.4584 \ln Q_{t-1}^{A(R)} - 0.1478 D_{4t} \quad \text{(Equation 6.66)}$$

where $t \in [2003Q3, 2008Q2]$

As shown in equations 6.65 and 6.66, both service and fare had significant influence on rail patronage in Auckland. As shown in the correlation analysis in section 5.1.2, it was observed that service had a strong positive correlation with patronage. The models estimated in equations 6.65 and 6.66 were consistent with our observations. The positive coefficients of $\ln S_t^{A(R)}$ showed that service had a positive effect on the demand. On the other hand, the fare appeared to have a negative effect on the demand as the coefficients of $\ln F_t^{A(R)}$ were negative. In addition, from equation 6.66, it appeared that income effect was positive rather than negative, which is also consistent with our observation in the correlation analysis in section 5.1.2. As discussed earlier, one possible explanation to this positive relationship is the significant increase in the share of commuters with higher income in the rail market.

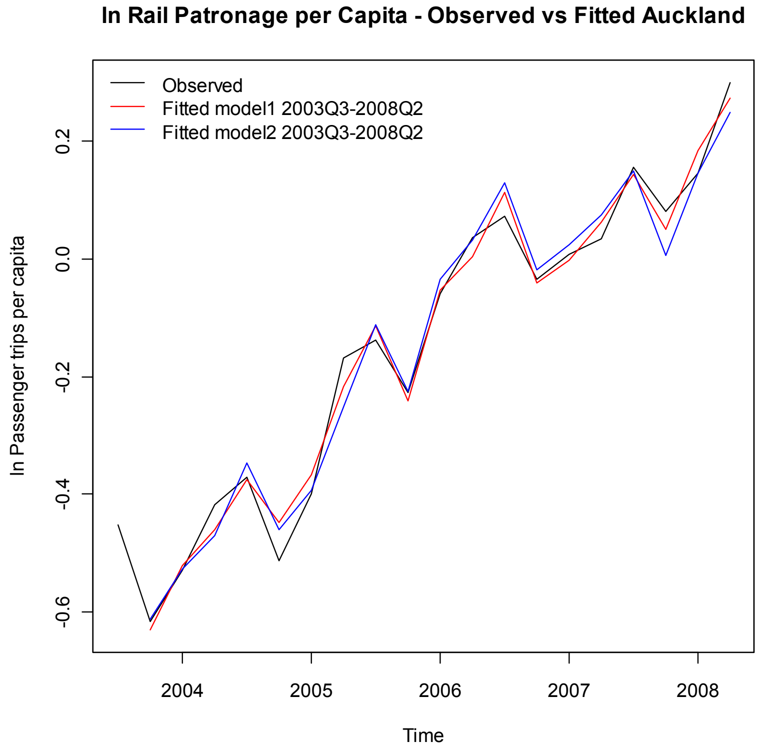
Table 6.5 Fitted rail patronage forecast model details and statistics for Auckland

Description		Model parameters	
Variable	**Unit**	**Model 1 Equation 6.65**	**Model 2 Equation 6.66**
Intercept	Passengers per capita	-1.1448	1.6913**
Rail service, $\ln S_t^{A(R)}$	Train-km per capita	0.9946***	0.8827**
Rail fare, $\ln F_t^{A(R)}$	Real average revenue per passenger	-0.9672***	-0.6755*
Income, $\ln I_t$	Real disposable income per capita	1.6060*	n/a
Seasonal effect Q4, $D_{4t}^{A(B)}$		-0.1325***	-0.1478***
Lagged patronage per capita, $\ln Q_{t-1}^{A(B)}$	Passengers per capita	0.2957^	0.4584*
R-square		0.9867	0.9785
Adjusted R-squared		0.9816	0.9724
Residual standard error		0.0361	0.0442
F-statistic		193.2	159.5
P-value (F-statistic)		1.007e-11	1.653e-11

Note: Significant codes: '***' 0.001'**' 0.01 '*' 0.05 '^' 0.1

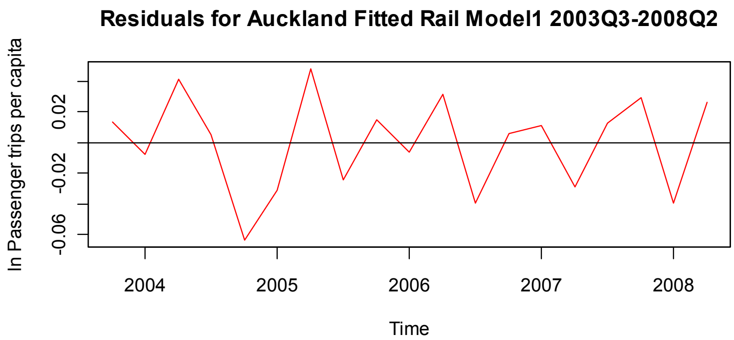
The model parameters and performance statistics are summarised in table 6.5. As shown by the R-squares of 0.9867 and 0.9785, model 1 appeared to have a slightly better fit compared with model 2.

Figure 6.3 Observed versus fitted values – Auckland rail patronage forecast model

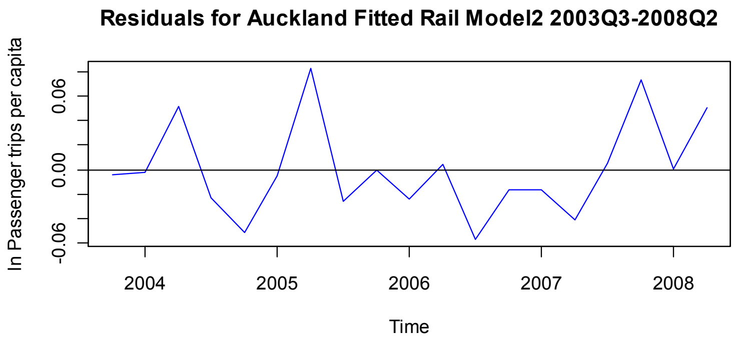


The fitted /predicted values are plotted with the observed values as shown in figure 6.3. Figure 6.3 also shows that model 1 had a slightly better fit than model 2.

Figure 6.4 Residuals – Auckland rail patronage forecast model



(a)



(b)

The residual values are plotted in figure 6.4 (a) and (b) and they appeared to follow a random pattern. Both model 1 and 2 were considered satisfactory.

6.4.2 Wellington

As discussed in section 4.6.2, the model period for Wellington was identified as FY 1999/2000 to FY 2007/08 and the models were developed based on annual information.

6.4.2.1 Wellington bus patronage forecast model

As illustrated in section 6.3.2.1, the forecast model for bus patronage in Wellington is as follows:

Fare only – equation 6.32 in section 6.3.2.1

$$\ln Q_t^{W(B)} = 2.0590 - 0.2312 \ln F_t^{W(B)} + 0.4934 \ln Q_{t-1}^{W(B)} \qquad \text{(Equation 6.67)}$$

where $t \in [FY99/00, FY07/08]$

As shown in equation 6.67, only bus fare was identified to have significant effect on bus patronage in Wellington. The effect was negative with an elasticity value of -0.2312, ie it was not elastic.

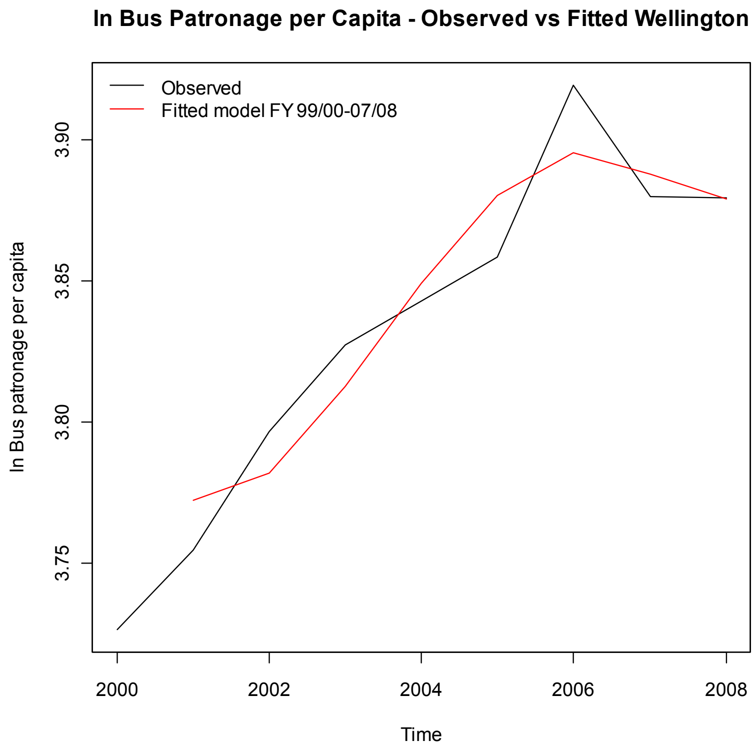
Table 6.6 Fitted bus patronage forecast model details and statistics for Wellington

Variable	Unit	Model parameters FY1999/2000–FY2007/08 Equation 6.67
Intercept	Passengers per capita	2.0590*
Bus fare, $\ln F_t^{W(B)}$	Real revenue per passenger	-0.2312^
Lagged patronage per capita, $\ln Q_{t-1}^{W(B)}$	Passengers per capita	0.4934*
R-square		0.9006
Adjusted R-squared		0.8609
Residual standard error		0.0194
F-statistic		22.65
P-value (F-statistic)		0.0031

Note: Significant codes: '***' 0.001'**' 0.01 '*' 0.05 '^' 0.1

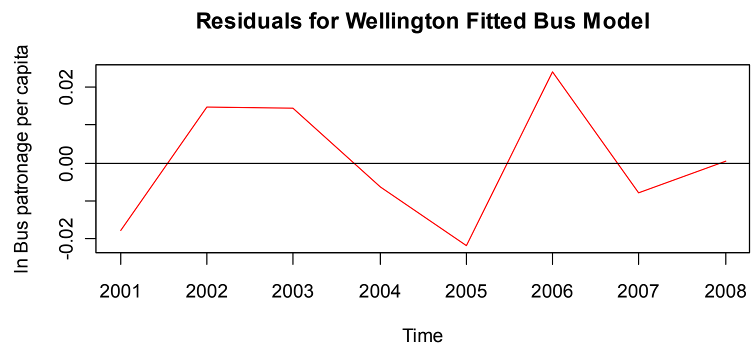
The model parameters and the detailed statistics are summarised in table 6.6. The R-square value of the fitted model was 0.9006.

Figure 6.5 Observed versus fitted values – Wellington bus patronage forecast model



The fitted /predicted values for bus patronage forecast in Wellington are plotted with the observed values as shown in figure 6.5. It appeared that the predicted values followed the trend generally but the goodness of fit was not as high as desired.

Figure 6.6 Residuals – Wellington bus patronage forecast model



The residual values are plotted in figure 6.6 and they appeared to follow a random pattern. We concluded that the model was satisfactory.

6.4.2.2 Wellington rail patronage forecast model

As illustrated in section 6.3.2.2, two forecast models were estimated for forecasting Wellington rail patronage per capita. The forecast models are:

Model 1 – service + car ownership + fuel price – equation 6.44 in section 6.3.2.2

$$\ln Q_t^{W(R)} = -1.3759 + 0.7424 \ln S_t^{W(R)} - 0.3246 \ln C_t^{W} + 0.1303 \ln O_t + 0.6889 \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.68)}$$

where $t \in \left[FY99/00, FY07/08\right]$

Model 2 – service + income + fuel price – equation (6.45) in section 6.3.2.2

$$\ln Q_t^{W(R)} = -0.4663 + 0.7847 \ln S_t^{W(R)} - 0.2176 \ln I_t + 0.1212 \ln O_t + 0.6900 \ln Q_{t-1}^{W(R)} \quad \text{(Equation 6.69)}$$

where $t \in \left[FY99/00, FY07/08\right]$

As shown in equations 6.68 and 6.69, service and fuel price had significant influence on rail patronage in Wellington and the impact was positive. On the other hand, car ownership and income also had significant influence but the effect was negative. As car ownership and income were correlated, they should not be in the same equation. Therefore, they were included in the two models respectively in addition to the other two significant factors, service and fuel price.

It is interesting to note the coefficient for income effect in equation 6.69 is negative whereas for Auckland, in equation 6.65, the coefficient is positive. Interpretations and comparison with international experience are discussed in chapter 7 of this report.

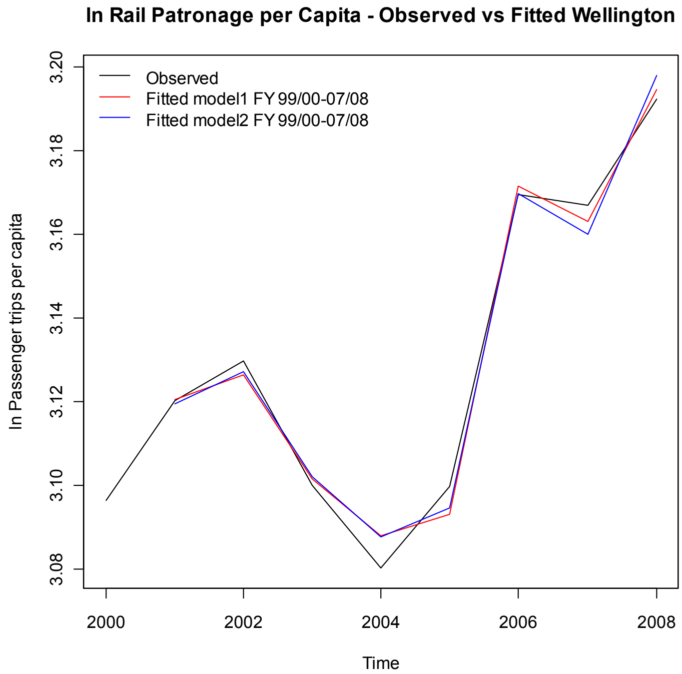
Table 6.7 Fitted bus and rail patronage forecast model details and statistics for Wellington

Description		Estimated elasticities	
Variable	**Unit**	**Model 1 Equation 6.68**	**Model 2 Equation 6.69**
Intercept	Passengers per capita	-1.3759^	-0.4663
Rail service, $\ln S_t^{W(R)}$	Bus-km per capita	0.7424*	0.7847**
Car ownership, $\ln C_t^{W}$	Car ownership per capita	-0.3246^	n/a
Income, $\ln I_t$	Real disposable income per capita	n/a	-0.2176*
Fuel price, $\ln O_t$	Real fuel price	0.1303*	0.1212*
Lagged patronage per capita, $\ln Q_{t-1}^{A(B)}$	Passengers per capita	0.6889**	0.6900**
R-square		0.9846	0.9873
Adjusted R-squared		0.9641	0.9704
Residual standard error		0.0076	0.0069
F-statistic		48.0	58.44
P-value (F-statistic)		0.0047	0.0035

Note: Significant codes: '***' 0.001'**' 0.01 '*' 0.05 '^' 0.1

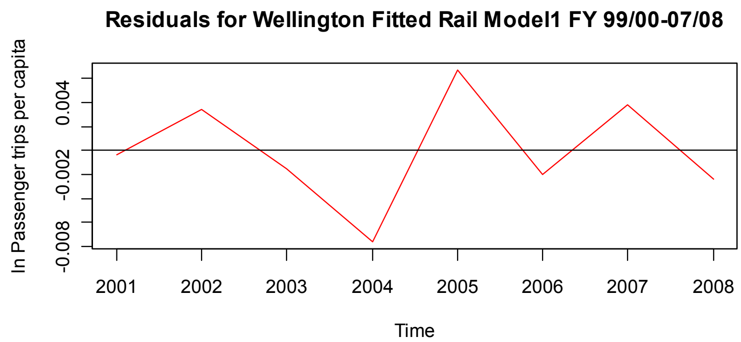
The model parameters and statistics are summarised in table 6.7. Both models had satisfactory performance with very high R-square values of 0.9846 and 0.9873 respectively. Model 2 had a slightly better fit as compared with model 1, ie when income rather than car ownership was considered as an independent variable.

Figure 6.7 Observed versus fitted values – Wellington rail patronage forecast models

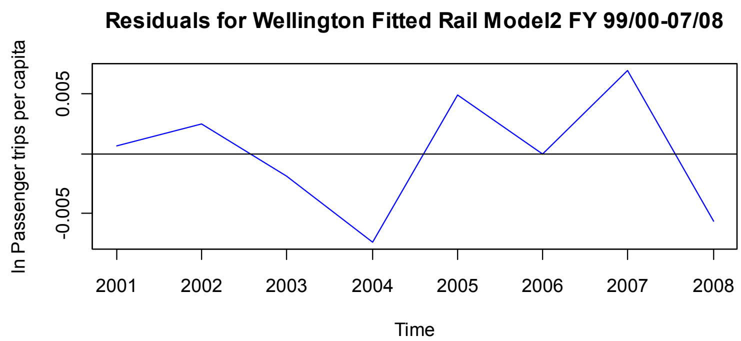


The fitted/predicted values are plotted with the observed values as shown in figure 6.7. The difference in goodness of fit is hardly noticeable.

Figure 6.8 Residuals – Wellington rail patronage forecast models



(a)



(b)

The residuals are plotted in figure 6.8 (a) and (b) and they appear to follow a random pattern. We concluded that both models were satisfactory.

6.4.3 Christchurch

As discussed in section 4.6.3, the model period for Christchurch was identified as 1997 Q1 to 2008 Q2. Information on fares, however, was available only from 2004 Q1 to 2008 Q2. Therefore, two model periods were considered: 1) 1997 Q1 to 2008; and 2) 2004 Q1 to 2008 Q2. The models were developed based on quarterly information. As illustrated in section 6.3.3, three forecast models were estimated for forecasting Christchurch bus patronage per capita, one based on period 1 and two based on period 2. The forecast models are as follows:

Christchurch fitted bus patronage forecast model (1998Q1–2008Q2)

Service only+ Q4 seasonal dummy – equation 6.52 in section 6.3.3

$$\ln Q_t^{C(B)} = 0.1147 + 0.0708 \ln S_t^{C(B)} + 0.8863 \ln Q_{t-1}^{C(B)} - 0.0898 D_{4t} \qquad \text{(Equation 6.70)}$$

where $t \in \left[1998Q1, 2008Q2\right]$

Christchurch fitted bus patronage forecast model (2004Q1–2008Q2)

Model 1 – fare + fuel price + Q4 seasonal dummy – equation 6.61 in section 6.3.3

$$\ln Q_t^{C(B)} = 0.1731 - 0.2620 \ln F_t^{C(B)} + 0.2785 \ln O_t + 0.2373 \ln Q_{t-1}^{C(B)} - 0.0429 D_{4t} \qquad \text{(Equation 6.71)}$$

where $t \in \left[2004Q1, 2008Q2\right]$

Model 2 – service + fare +income + fuel price + Q4 seasonal dummy – equation 6.62 in section 6.3.3

$$\begin{aligned}\ln Q_t^{C(B)} &= -0.2646 + 0.0767 \ln S_t^{C(B)} - 0.3697 \ln F_t^{C(B)} + 0.3904 \ln I_t \\ &+ 0.1954 \ln O_t + 0.2039 \ln Q_{t-1}^{C(B)} - 0.0459 D_{4t}\end{aligned} \qquad \text{(Equation 6.72)}$$

where $t \in \left[2004Q1, 2008Q2\right]$

As shown in equations 6.70 to 6.72, as in the Auckland case, seasonal effect of Q4 was found to be significant and had a negative coefficient, ie the bus patronage in Q4 was significantly less than the rest of the year. Two breaks were identified as described in section 5.3.1: 1) 1998 Q1 due to the introduction of 'gold coin' cash fare; and 2) 2003 Q3 due to the introduction of the Metrocard. As illustrated in section 6.3.3, both changes were found to have a significant impact on bus patronage. We, therefore, developed the first model, equation 6.70, based on a model period starting from 1998 Q1. The second model period, however, had to start from 2004 Q1 rather than from 2003 Q3 as fare data was available only from 2004 Q1. Two models, equations 6.71 and 6.72, were identified for this period for further investigation.

As shown in equation 6.70, service was found to be the only factor which had a significant effect on patronage with a positive coefficient of 0.0708, ie inelastic. We believe that bus fare should also have had a significant effect since it was observed that the introduction of a gold coin fare had caused a change of trend in 1998. Unfortunately, fare information was not available and it became an important 'missing' variable in equation 6.70.

In equation 6.71, bus fare and fuel price were found to be a significant factor while fuel price was the only significant factor in equation 6.72.

Table 6.8 Fitted bus patronage forecast model details and statistics for Christchurch 1998Q1–2008Q2

Variable	Unit	Model parameters		
		1998Q1–2008Q2 Equation 6.70	2004Q1–2008Q2	
			Model 1 Equation 6.71	Model 2 Equation 6.72
Intercept	Passengers per capita	0.1147**	0.1731	-0.2646
Bus service, $\ln S_t^{C(B)}$	Bus-km per capita	0.0708*	n/a	0.0767
Bus fare, $\ln F_t^{C(B)}$	Real average revenue per passenger	n/a	-0.2620*	-0.3697
Income, $\ln I_t$	Real disposable income per capita	n/a	n/a	0.3904
Fuel price, $\ln O_t$	Real fuel price	n/a	0.2785**	0.1954^
Seasonal effect Q4, $D_{4t}^{C(B)}$		-0.0898***	-0.0429*	-0.0459*
Lagged patronage per capita, $\ln Q_{t-1}^{W(B)}$	Passengers per capita	0.8863***	0.2373	0.2039
R-square		0.9861	0.8717	0.9029
Adjusted R-squared		0.9850	0.8251	0.8382
Residual standard error		0.0230	0.0167	0.0161
F-statistic		875.5	18.69	13.95
P-value (F-statistic)		<2.2E-16	7.21E-05	4.20E-04

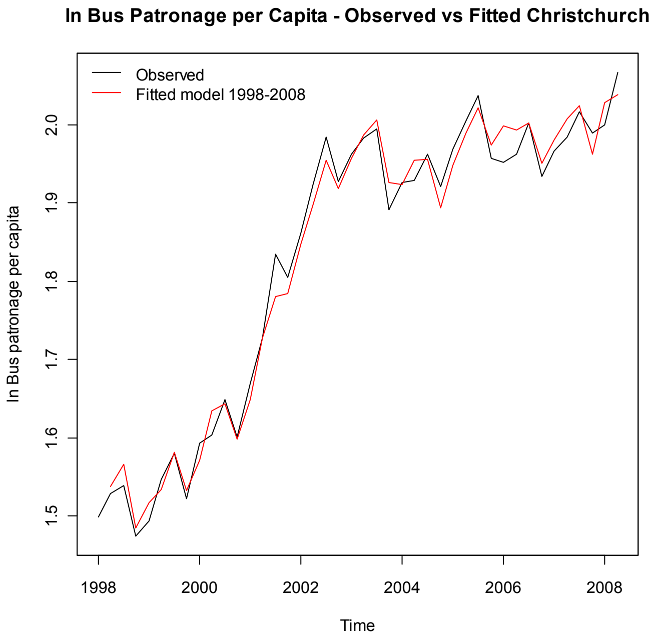
Note: Significant codes: '***' 0.001'**' 0.01 '*' 0.05 '^' 0.1

The model parameters and statistics are summarised in table 6.8. Equation 6.70 had the highest R-square value of 0.9861 despite fare being a missing variable as discussed. This implies that in the long run, service had a significant impact on bus patronage in Christchurch. This is consistent with the results from the trend analysis in section 5.3.1.

The two models identified for the more recent model period were not 'ideal' due to different reasons. Nevertheless, given the quality of the information collected, they were the best we could obtain. The R-square value of equation 6.71) was only 0.8717 with all independent variables, fare and fuel price, being significant. From both equation 6.70 and the trend analysis in section 5.3.1, we expected that service should have a significant impact. Therefore, we selected equation 6.72 for further investigation. We included service and income in addition to fare and fuel price and the R-square value improved to 0.9029. Although all the coefficients were intuitively correct, ie the coefficients of both bus service and fuel price were positive while that of bus fare was negative, only fuel price appeared to be a significant factor. We believe the real cause was the quality of data on service which was only an approximation from the bus timetable rather than actual figures.

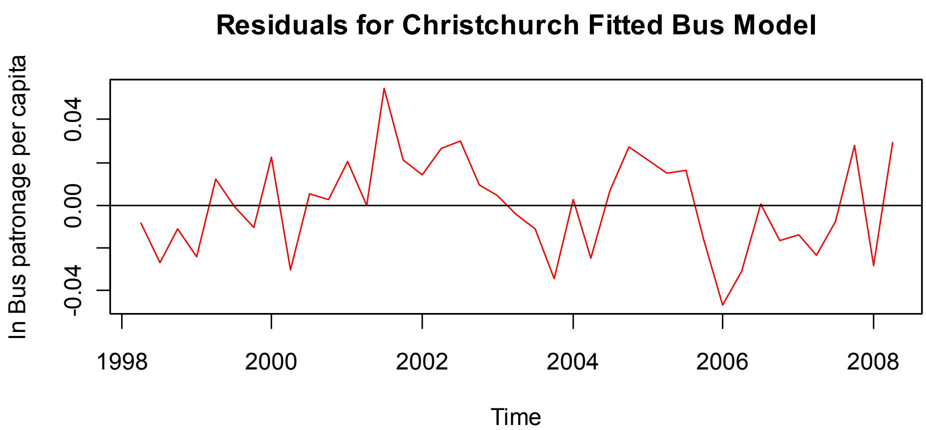
To decide which would be the best model for Christchurch, the plots of predicted values, observed values and residuals were investigated as follows.

Figure 6.9 Observed versus fitted values – Christchurch bus patronage forecast model 1998–2008



The fitted/predicted values of the 1998Q1–2008Q2 model are plotted with the observed values in figure 6.9.

Figure 6.10 Residuals – Christchurch bus patronage forecast model 1998–2008



The residuals are plotted in figure 6.10 and they appeared to follow a random pattern.

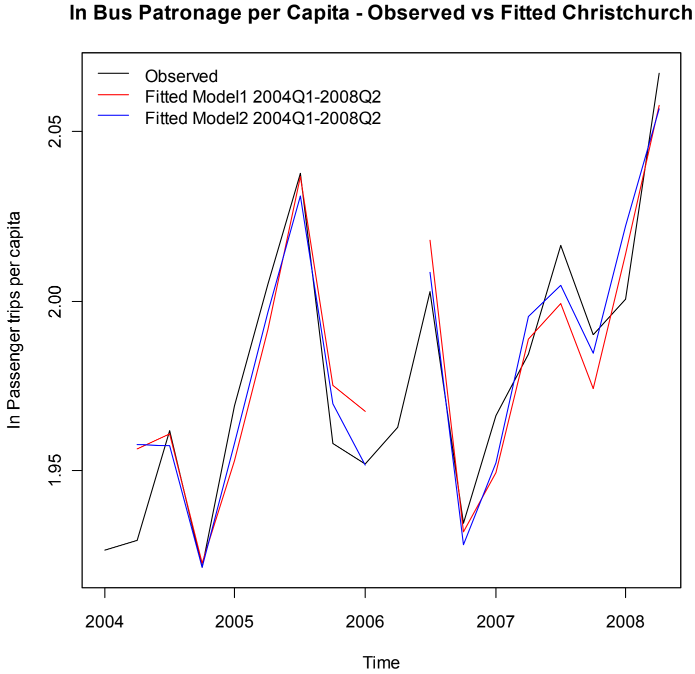
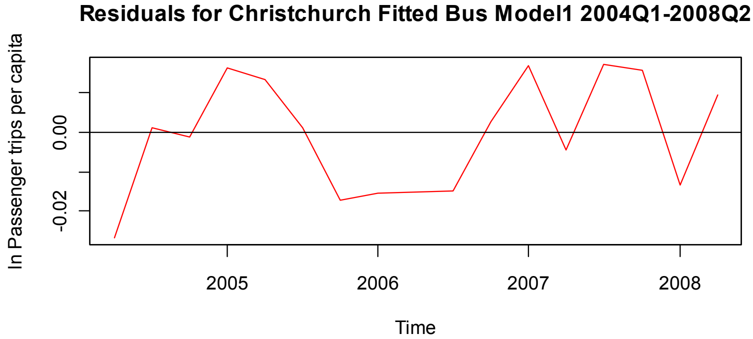


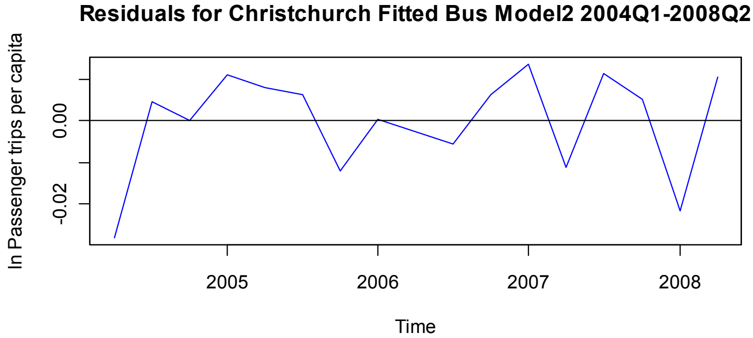
Figure 6.11 Observed versus fitted values – Christchurch bus patronage forecast model 2004–2008

The fitted/predicted values of the 2004Q1–2008Q2 models are plotted with the observed values in figure 6.11. It appeared because of its higher R-square values, that model 2 had a better fit than model 1.

Figure 6.12 Residuals – Christchurch bus patronage forecast model 2004–2008



(a)



(b)

The residuals are plotted in figure 6.12(a) and (b). As shown in the residual plots, the residuals of model 2 in figure 6.12(b) appeared to be more random and smaller in magnitude compared with the residuals of model 1 in figure 6.12(a).

This was not surprising as more factors were considered in model 2 than in model 1. Nevertheless, as discussed earlier, the only significant factor in model 2 was fuel price. Based on our criteria of acceptable R-square value and all independent variables being significant, model 2 would not be an appropriate forecast model.

We concluded that model 1, equation 6.71, although not ideal, should be a recommended forecast model for Christchurch as it was developed based on a more recent period and took into account the impact of fuel price fluctuations in recent years. On the other hand, equation 6.70 which was developed based on longer-term time series should be applied to assess the impact on changes in service since service was not a variable in equation 6.71.

6.5 Summary

Based on the discussion above, the recommended forecast models for the three cities are summarised in this section. The corresponding R-outputs are depicted in appendix D.

6.5.1 Auckland

6.5.1.1 Auckland bus patronage forecast model

Auckland bus patronage model (2003Q3–2008Q2)

Service + car ownership + fuel price+ Q4 seasonal dummy – equation 6.64 in section 6.4.1.1

$$\ln Q_t^{A(B)} = -1.9306 + 0.4603 \ln S_t^{A(B)} - 1.9618 \ln C_t^{A} + 0.2040 \ln O_t \\ +0.3677 \ln Q_{t-1}^{A(B)} - 0.0478 D_{4t} \qquad \text{(Equation 6.73)}$$

where $t \in \left[2003Q3, 2008Q2\right]$

6.5.1.2 Auckland rail patronage forecast model

Model 1 – service + fare + income+ Q4 seasonal dummy – equation 6.65 in section 6.4.1.2

$$\ln Q_t^{A(R)} = -1.1448 + 0.9946 \ln S_t^{A(R)} - 0.9672 \ln F_t^{A(R)} + 1.606 \ln I_t \\ +0.2957 \ln Q_{t-1}^{A(R)} - 0.1325 D_{4t} \qquad \text{(Equation 6.74)}$$

where $t \in \left[2003Q3, 2008Q2\right]$

Model 2 – service + fare+ Q4 seasonal dummy – equation 6.66 in section 6.4.1.2

$$\ln Q_t^{A(R)} = 1.6913 + 0.8827 \ln S_t^{A(R)} - 0.6755 \ln F_t^{A(R)} \\ +0.4584 \ln Q_{t-1}^{A(R)} - 0.1478 D_{4t} \qquad \text{(Equation 6.75)}$$

where $t \in \left[2003Q3, 2008Q2\right]$

6.5.1.3 Influencing factors for Auckland

The factors influencing PT patronage in Auckland, as modelled in equations 6.73 to 6.75, are summarised in table 6.9.

Table 6.9 Summary of factors influencing Auckland public transport patronage

	Auckland bus equation 6.73	Auckland rail	
		Model 1 equation 6.74	Model 2 equation 6.75
Service	positive	positive	positive
Fare	n/a	negative	negative
Car ownership	negative	n/a	n/a
Income	n/a	positive	n/a
Fuel price	positive	n/a	n/a
4th quarter effect	negative	negative	negative

6.5.2 Wellington

6.5.2.1 Wellington bus patronage forecast model

Fare only – equation 6.67 in section 6.4.2.1

$$\ln Q_t^{W(B)} = 2.0590 - 0.2312 \ln F_t^{W(B)} + 0.4934 \ln Q_{t-1}^{W(B)} \qquad \text{(Equation 6.76)}$$

where $t \in [FY99/00, FY07/08]$

6.5.2.2 Wellington rail patronage forecast model

Model 1 – service + car ownership + fuel price – equation 6.68 in section 6.4.2.2

$$\begin{aligned}\ln Q_t^{W(R)} &= -1.3759 + 0.7424 \ln S_t^{W(R)} - 0.3246 \ln C_t^{W} + 0.1303 \ln O_t \\ &\quad +0.6889 \ln Q_{t-1}^{W(R)}\end{aligned} \qquad \text{(Equation 6.77)}$$

where $t \in [FY99/00, FY07/08]$

Model 2 – service + income + fuel price – equation 6.69 in section 6.4.2.2

$$\begin{aligned}\ln Q_t^{W(R)} &= -0.4663 + 0.7847 \ln S_t^{W(R)} - 0.2176 \ln I_t + 0.1212 \ln O_t \\ &\quad +0.6900 \ln Q_{t-1}^{W(R)}\end{aligned} \qquad \text{(Equation 6.78)}$$

where $t \in [FY99/00, FY07/08]$

6.5.2.3 Influencing factors for Wellington

The factors influencing PT patronage in Wellington, as modelled in equations 6.76 to 6.78, are summarised in table 6.10.

Table 6.10 Summary of factors influencing Wellington public transport patronage

	Wellington bus Equation 6.76	Wellington rail	
		Model 1 Equation 6.77	Model 2 Equation 6.78
Service	n/a	positive	positive
Fare	negative	n/a	n/a
Car ownership	n/a	negative	n/a
Income	n/a	n/a	negative
Fuel price	n/a	positive	positive

6.5.3 Christchurch

6.5.3.1 Christchurch bus patronage forecast model

Christchurch fitted bus patronage forecast model (1998Q1–2008Q2)

Service only + Q4 seasonal dummy – equation 6.52 in section 6.3.3

$$\ln Q_t^{C(B)} = 0.1147 + 0.0708 \ln S_t^{C(B)} + 0.8863 \ln Q_{t-1}^{C(B)} - 0.0898 D_{4t} \qquad \text{(Equation 6.79)}$$

where $t \in [1998Q1, 2008Q2]$

Christchurch fitted bus patronage forecast model (2004Q1–2008Q2)

Model 1 – fare + fuel price + Q4 seasonal dummy – equation 6.71 in section 6.5.3

$$\ln Q_t^{C(B)} = 0.1731 - 0.2620 \ln F_t^{C(B)} + 0.2785 \ln O_t + 0.2373 \ln Q_{t-1}^{C(B)} - 0.0429 D_{4t} \qquad \text{(Equation 6.80)}$$

where $t \in [2004Q1, 2008Q2]$

6.5.3.2 Influencing factors for Christchurch

The factors influencing bus patronage in Christchurch, as modelled in equations 6.79 and 6.80, are summarised in table 6.11.

Table 6.11 Summary of factors influencing Christchurch bus patronage

	Model 1 (1998Q1–2008Q2)	Model 2 (2004Q1–2008Q2)
Service	positive	n/a
Fare	n/a	negative
Car ownership	n/a	n/a
Income	n/a	n/a
Fuel price	n/a	positive
Q4 effect	negative	negative

6.6 Estimated elasticities

We obtained the best estimates of demand elasticities from the recommended models as summarised in section 6.5.

Table 6.12 Bus demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland 2003Q3–2008Q2		Wellington		Christchurch	
	short-run	long-run	short-run	long-run	short-run	long-run
Bus service	0.46	0.73	n/a	n/a	0.07[a]	0.62[a]
Bus fare	n/a	n/a	-0.23	-0.46	-0.26[b]	-0.34[b]
Car ownership	-1.96	-3.10	n/a	n/a	n/a	n/a
Income	n/a	n/a	n/a	n/a	n/a	n/a
Fuel price	0.20	0.32	n/a	n/a	0.28[b]	0.37[b]

Note: [a] Estimates based on 1998Q1–2008Q2 model, [b] Estimates based on 2004Q1–2008Q2 model 1

Table 6.13 Rail demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Rail service	0.99	0.88	1.41	1.63	0.74	0.78	2.39	2.53
Rail fare	-0.97	-0.68	-1.37	-1.25	n/a	n/a	n/a	n/a
Car ownership	n/a	n/a	n/a	n/a	-0.32	n/a	-1.04	n/a
Income	1.61	n/a	2.28	n/a	n/a	-0.22	n/a	-0.70
Fuel price	n/a	n/a	n/a	n/a	0.13	0.12	0.42	0.39

The best elasticity estimates for bus demand and rail demand are depicted in tables 6.12 and 6.13. Our estimates are compared with international experiences in chapter 7.

7 Comparison with international experiences

As the objectives of this study were to identify the factors influencing PT patronage in New Zealand and estimate the demand elasticities for the factors identified, the comparison with international experiences was conducted in a two-fold manner:

1 To compare the factors influencing PT patronage

2 To compare the estimated elasticity values.

7.1 Influencing factors

Table 7.1 Summary of factors influencing public transport patronage in New Zealand

	Auckland		Wellington		Christchurch bus
	Bus	Rail	Bus	Rail	
Service	positive	positive	n/a	positive	positive
Fare	n/a	negative	negative	n/a	negative
Car ownership	negative	n/a	n/a	negative	n/a
Income	n/a	positive	n/a	negative	n/a
Fuel price	positive	n/a	n/a	positive	positive

The influencing factors identified in the recommended forecast models in section 6.5 are depicted in table 7.1.

7.1.1 Service and fare

Litman (2004) found in his review that PT fares, service quality (service speed, frequency, coverage and comfort), and parking pricing tend to have the greatest impact on PT patronage. It appeared from our findings that both service and fare did have a significant effect in all three cities although not on all modes. It is important to note that fare was not a significant factor influencing bus demand in Auckland. Litman (2004) found that large cities tend to have lower price elasticities than suburbs and smaller cities as they have a greater portion of PT-dependent users. Auckland has always been the largest city in New Zealand with more than one third of the country's population. We would expect Auckland to have a bigger proportion of PT-dependent users, ie international students and new immigrants. This could be a reason why fare was found to be an insignificant factor affecting the bus patronage which has always been the major PT mode in Auckland.

7.1.2 Differences between modes and cities

Litman (2004) also found that bus and rail often had different elasticities because they served different markets. In our case, it appeared that bus and rail had different influencing factors for the same reason. For example, in Wellington, fuel price was found to be an influencing factor for rail but not for bus. On the contrary, in Auckland, fuel price was found to be an influencing factor for bus but not for rail. We believe this could be evidence of a strong substitution effect between car and bus in Auckland; and between car and rail in Wellington. This hypothesis was also supported by the observation that car ownership was found to have a significant effect on Auckland bus and Wellington rail patronage. On the other hand, the influence of car ownership on bus patronage in Christchurch was found to be insignificant. We believe a possible explanation is that Christchurch has the highest car ownership level among the three cities. As mentioned in Dargay and Hanly (2002b), the analysis of bus patronage in England showed the overall picture could be quite different as car ownership became saturated. If the car ownership in Christchurch was saturated, the decision of whether to take the bus or not would more likely to be influenced by factors other than car ownership. Thus, our study results indicated that service, fare and fuel price were the factors that had significant influence.

7.1.3 Car ownership and income

Among the five influencing factors we considered – service, fare, car ownership, income and fuel price – two of the five factors, car ownership and income, did not enter into any models for the three cities at the same time. This was similar to the experience in France, described in Bresson et al (2004), where a model was first estimated using the entire set of explanatory variables, including disposable income, a structural indicator (representing the effect of changes in population characteristics, urban sprawl and car ownership), the PT fare, three service measures (seat-kilometres per capita, service frequency and network density) and fuel price. The income variable was dropped as it was found to have an insignificant effect on patronage. It was not surprising that income was a major determinant of car ownership. One would expect that car ownership and income are positively correlated and hence should not be included in the equation at the same time.

In our case, the car ownership effect was found to have a negative significant effect for Auckland bus and Wellington rail. On the other hand, income was found to have positive effect on Auckland rail. It was found generally in the literature, such as Bresson et al (2003) and Dargay and Hanly (2002a; 2002b) that the income effect was negative rather than positive. Nevertheless, as stressed in Dargay and Hanly (2002b), negative income elasticity pertained to a period of rising car ownership and use. It is likely that the income effect could become positive rather than negative when the use of private vehicles approached saturation.

7.2 Elasticities

Litman (2004) provides a comprehensive review on elasticities for PT planning in the literature including the classic reviews by Goodwin (1992) and Oum et al (1992). In this section, we compare our results to the relevant studies in the literature.

7.2.1 Service elasticity

Table 7.2 Bus service elasticity estimates in Auckland, Wellington and Christchurch

	Auckland 2003Q3–2008Q2		Wellington		Christchurch 1998Q1–2008Q2	
	short-run	long-run	short-run	long-run	short-run	long-run
Bus service	0.46	0.73	n/a	n/a	0.07	0.62

Table 7.3 Rail service elasticity estimates in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Rail service	0.99	0.88	1.41	1.63	0.74	0.78	2.39	2.53

Table 7.4 Estimated elasticities for bus services in England

Variable	Estimated elasticity
Journeys (-1)	0.52
Income	-0.39
Service	0.49
Motoring costs	0.32
Percent pensioners	-0.08
Fare	-0.33

Source: Dargay and Hanly (2002a), table 1

Table 7.5 Public transport service level elasticities in England and France

	Estimated elasticities using shrinkage estimators in a log-log model	
	England[a]	France[b]
Short-run	0.57	0.29
Long-run	0.77	0.57

Note:[a] The main objective was to estimate bus elasticities based on time series data for different regions and for a panel of 46 English counties (urban and rural).

[b] The French study includes all PT modes for urban travel and covers a sample of 62 out of about 100 urban areas.

Source: Bresson et al (2003)

The estimated bus and rail service elasticities in Auckland, Wellington and Christchurch are depicted in tables 7.2 and 7.3. Our estimate of short-run service level elasticity of 0.46 for bus services in Auckland was comparable to the 0.46 for bus services in England, as given in Dargay and Hanly (2002a) and shown in table 7.4. Our estimate of a long-run service level elasticity of 0.73 was also comparable to the 0.77 for England, as given in Bresson et al (2003) and shown in table 7.5. On the other hand, our short-run estimate of 0.07 for Christchurch bus services appeared to be much lower than the estimates of 0.57 and 0.29 for England and France respectively, as shown in table 7.5. On the other hand, our long-run estimate of 0.62 was similar to the long-run estimate of 0.57 in France as shown in table 7.2. We believe that one possible explanation is the high car ownership in Christchurch, which was the highest among all three cities. Whether to take the bus was more of a long-term decision rather than a short-term one and was taken concurrently with a car ownership decision. Thus the service elasticity in Christchurch appeared to be lower than in international experience.

Both Auckland and Wellington rail had very high short-run and long-run service elasticities; short-run in the range of 0.88–0.99 in Auckland and 0.74–0.78 in Wellington; long-run in the range of 1.41–1.63 in Auckland and 2.39–2.53 in Wellington. These were much higher than the estimates of 0.29 (short-run) and 0.57 (long-run) in France as shown in table 7.5. Moreover, only the long-run estimates were consistent with the expectation of higher elasticities for smaller cities. We believe that this could be due to the service kilometres of rail service in Auckland being measured in train-kilometres which was not a precise measure of service level as the train composition could vary from 2-car to 6-car depending on time of the day and day of the week. Thus the estimates for Auckland could be overestimated. The fact that bus and rail served different markets could also be another reason.

7.2.2 Fare elasticity

Table 7.6 Estimated bus fare elasticities in Auckland, Wellington and Christchurch

	Auckland 2003Q3–2008Q2		Wellington		Christchurch 2004Q1–2008Q2 Model 2	
	short-run	long-run	short-run	long-run	short-run	long-run
Bus fare	n/a	n/a	-0.23	-0.46	-0.26	-0.34

Table 7.7 Estimated rail fare elasticities in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Rail fare	-0.97	-0.68	-1.37	-1.25	n/a	n/a	n/a	n/a

Table 7.8 Estimated national and regional resident population

	Estimated population	As percentage of national population
NZ	4,268,650	
Auckland	1,435,645	33.63%
Wellington	475,944	11.15%
Canterbury	554,442	12.99%

The estimated bus and rail fare elasticities in Auckland, Wellington and Christchurch are depicted in tables 7.6 and 7.7. Litman (2004) found that large cities tend to have lower price elasticities than suburbs and smaller cities as they have a greater portion of PT-dependent users. The estimated resident population in the three regions in 2008 based on our data imputation model are shown in table 7.8.

Table 7.9 Bus fare elasticities versus size of city

	Large cities (more than 1 million population)	Smaller cities (less than 1 million population)
Average for all hours	-0.36	-0.43
Peak hour	-0.18	-0.27
Off-peak	-0.39	-0.46
Off-peak average	-0.42	
Peak hour average	-0.23	

Source: As cited in Litman (2004) from Pham and Linsalata (1991)

Comparing the estimated long-run bus fare elasticity of -0.46 in Wellington and -0.34 in Christchurch with international statistics as shown in table 7.9, it appeared our estimated value of -0.46 for Wellington was comparable with the estimate for smaller cities with less than 1 million population as in Pham and Linsalata (1991), but the estimated value of -0.34 in Christchurch was lower than expected. Nevertheless, both values were within the range of most typical elasticity estimates of -0.1 to -0.6 in Oum et al (1992) but slightly less than the estimate of -0.55 in table 3 of Goodwin (1992).

Our model structure was very similar to the model in Dargay and Hanly (2002a). The major difference was in how we applied the model. Dargay and Hanly (2002a) applied the model to annual bus data in England by county. Because the availability of data per county was very limited – only 10 annual observations – they pooled the data and developed a pooled model for all counties. The results for the model, with all coefficients constrained to be the same across counties and assuming constant fare elasticity, are as shown in table 7.4 in section 7.2.1.

Table 7.10 Public transport fare level elasticities in England and France

	Estimated elasticities using shrinkage estimators in a log-log model	
	England[a]	France[b]
Short-run	-0.51	-0.32
Long-run	-0.69	-0.61

Note:[a] The main objective was to estimate bus elasticities based on time series data for different regions and for a panel of 46 English counties (urban and rural).

[b] The French study includes all PT modes for urban travel and covers a sample of 62 out of about 100 urban areas.

Source: Bresson et al (2003)

Our estimate of the short-run bus fare elasticity in Wellington of -0.23 was less elastic than the overall short-run fare elasticity of -0.33 of England in Dargay and Hanly (2002a) as shown in table 7.4. On the other hand, the short-run bus fare elasticity in Christchurch of -0.34 was comparable with the values of -0.33 in England in Dargay and Hanly (2002a) and -0.32 in France in Bresson et al (2003), as shown in table 7.10.

In terms of rail fare elasticities, our short-run and long-run estimates for Auckland were in the range of -0.68 to -0.97 and -1.25 to -1.37, respectively. Compared with the French study of Bresson et al (2003) of -0.32 and -0.61, our estimates for Auckland were much higher. We believe that this could be the effect of significant increase in rail fare of 10% in February 2006 plus another 15% in January 2007.

7.2.3 Car ownership elasticity

Table 7.11 Car ownership elasticity of bus demand in Auckland, Wellington and Christchurch

	Auckland 2003Q3–2008Q2		Wellington		Christchurch 2004Q1–2008Q2 Model 2	
	short-run	long-run	short-run	long-run	short-run	long-run
Car ownership	-1.96	-3.10	n/a	n/a	n/a	n/a

Table 7.12 Car ownership elasticity of rail demand in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Car ownership	n/a	n/a	n/a	n/a	-0.32	n/a	-1.04	n/a

Table 7.13 Elasticities estimated on basis of structural model of bus and automobile use, national UK data

	Bus passenger-km	Bus journeys	Car ownership	Car passenger-km
Car ownership				
Short-run	0	0		0.94
Long-run	-0.73	-0.64		0.81
Bus fare				
Short-run	-0.31	-0.52	0.19	0.18
Long-run	-0.94	-1.08	0.42	0.34
Income				
Short-run	0.14	0.38	0.37	0.14
Long-run	0.07	-0.26	0.56	0.70
Motoring costs				
Short-run	0	0	-0.38	-0.44
Long-run	0.37	0.33	-0.51	-0.96

Source: Dargay and Hanly (2002b)

The estimated car ownership elasticities of bus and rail demand in Auckland, Wellington and Christchurch are depicted in tables 7.11 and 7.12.

Dargay and Hanly (2002b) developed a structural model to examine the influences of income, bus fares, car ownership and use and motoring costs on bus patronage in England. As shown in table 7.13, the estimated short-run and long-run elasticities of bus journeys with respect to car ownership are 0 and -0.64 respectively. We could not make a direct comparison of our results with these estimates because both car ownership and car passenger-kilometres were considered in Dargay and Hanly (2002b), while we considered only car ownership. In other words, we were trying to measure car ownership and its use with only one variable. We believe that this is why our estimated values of -1.96 (short-run) and -3.10 (long-run) for Auckland bus, and -0.32 (short-run) and -1.04 (long-run) for Wellington rail, were much higher in magnitude, especially for Auckland bus, compared with their estimates.

7.2.4 Income elasticity

Table 7.14 Income elasticity of rail demand in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Income	1.61	n/a	2.28	n/a	n/a	-0.22	n/a	-0.70

Table 7.15 Public transport income level elasticities in England and France

	Estimated elasticities using shrinkage estimators in a log-log model	
	England[a]	France[b]
Short-run	-0.67	-0.05
Long-run	-0.90	-0.09

Note:[a] The main objective was to estimate bus elasticities based on time series data for different regions and for a panel of 46 English counties (urban and rural).

[b] The French study includes all PT modes for urban travel and covers a sample of 62 out of about 100 urban areas.

Source: Bresson et al (2003)

The estimated car ownership elasticities of rail demand in Auckland, Wellington and Christchurch are depicted in table 7.14 but there were no reliable estimates of the income elasticities of bus demand from our results.

As discussed in section 7.1.3, the income effect was found to be negative in most studies in the literature, such as Bresson et al (2003) and Dargay and Hanly (2002a; 2002b). In this study, the estimated income elasticities of rail demand in Auckland were positive, 1.61 (short-run) and 2.28 (long-run), while the estimates for Wellington rail were negative, -0.22 (short-run) and -0.70 (long-run). We believe that this was due to the difference between the two markets. Auckland rail went through tremendous improvement during the study period while Wellington rail was a well-established mature system. In Auckland, as analysed in section 5.1.1, the share of commuters in the rail market increased significantly after the opening of Britomart put the CBD within walking distance of the new station. The newly attracted commuters would have characteristics such as higher income. Thus the elasticity estimates for Auckland were positive. On the other hand, the negative income elasticities of Wellington came from the increase in car ownership as a result of increase in income.

7.2.5 Fuel price elasticity

Table 7.16 Fuel price elasticity of bus demand in Auckland, Wellington and Christchurch

	Auckland 2003Q3-2008Q2		Wellington		Christchurch 2004Q1-2008Q2 Model 2	
	short-run	long-run	short-run	long-run	short-run	long-run
Fuel price	0.20	0.32	n/a	n/a	0.28	0.37

Table 7.17 Fuel price elasticity of rail demand in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Fuel price	n/a	n/a	n/a	n/a	0.13	0.12	0.42	0.39

Table 7.18 Public transport patronage elasticities with respect to fuel price

Trip purpose	Estimated elasticity
Commuting	0.20
Business	0.24
Education	0.01
Other	0.15
Total	0.13

Source: As cited in Litman (2004) from TRACE (1999).

The estimated fuel price elasticities of bus and rail demand in Auckland, Wellington and Christchurch are depicted in tables 7.16 and 7.17.

It is important to note that the fluctuations in fuel price in recent years have been at much higher sustained levels than historical trends and this recent pattern compares to what happened in 1970s. As a result, we would expect our estimates to be higher than the estimates from international experience.

In the short run, the fuel price elasticities, for Auckland bus of 0.20 and for Wellington rail of 0.12-0.13 were relatively less elastic than Christchurch, and both comparable with the estimated elasticities for PT in Europe, as shown in table 7.18, of 0.20 for commuting trips and 0.13 for all trips. Bresson et al (2004) also had similar estimates of short-run and long-run fuel price elasticities of 0.08 and 0.14 respectively in France. On the other hand, Christchurch's short-run estimate of 0.28 was higher but comparable with the elasticity of bus patronage with respect to motoring cost in England of 0.32, as shown in table 7.4. A higher positive fuel price elasticity represents a higher level of substitution between car and the alternative mode. Our results indicated that the increase in fuel price in recent years had a more substantial effect in Christchurch and Auckland and much less so in Wellington. We believe this was again due to the fact that Wellington had a much more mature market compared with Auckland and Christchurch.

8 Discussion

8.1 Results

Table 8.1 Bus demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland 2003Q3-2008Q2		Wellington		Christchurch	
	short-run	long-run	short-run	long-run	short-run	long-run
Bus service	0.46	0.73	n/a	n/a	0.07[a]	0.62[a]
Bus fare	n/a	n/a	-0.23	-0.46	-0.26[b]	-0.34[b]
Car ownership	-1.96	-3.10	n/a	n/a	n/a	n/a
Income	n/a	n/a	n/a	n/a	n/a	n/a
Fuel price	0.20	0.32	n/a	n/a	0.28[b]	0.37[b]

Note: [a] Estimates based on the 1998Q1–2008Q2 model, [b] Estimates based on the 2004Q1–2008Q2 model 1

Table 8.2 Rail demand elasticity estimates in Auckland, Wellington and Christchurch

	Auckland				Wellington			
	short-run		long-run		short-run		long-run	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Rail service	0.99	0.88	1.41	1.63	0.74	0.78	2.39	2.53
Rail fare	-0.97	-0.68	-1.37	-1.25	n/a	n/a	n/a	n/a
Car ownership	n/a	n/a	n/a	n/a	-0.32	n/a	-1.04	n/a
Income	1.61	n/a	2.28	n/a	n/a	-0.22	n/a	-0.70
Fuel price	n/a	n/a	n/a	n/a	0.13	0.12	0.42	0.39

Colour key:

Colour	Meaning
(cyan)	The estimated value was lower than international experience
(yellow)	The estimated value was compatible with international experience
(red)	The estimated value was higher than international experience

From our econometric results, it was evident that all five factors – service, fare, car ownership, income and fuel price – had significant influences in at least one of the three major cities in New Zealand. The estimated elasticities are summarised and colour coded in tables 8.1 and 8.2 to show whether they were lower, compatible or higher than the estimates from international experience. As discussed in section 7.2, most of our elasticity estimates were comparable to international experiences. The following key observations were made:

1 Service and fare had significant influence in all three cities although not on all modes.

2 Service had influence in all cities and on almost all modes except Wellington bus.

3 Rail service and fare elasticities were higher than the corresponding estimates for bus demand in Auckland.

4 Bus fare was a significant influencing factor in both Wellington and Christchurch but not in Auckland.

5 Car ownership was an influencing factor for Auckland bus and Wellington rail but fare was not an influencing factor in these cases.

6 Car ownership elasticities were far higher than the estimates from international experience.

7 Income elasticities were negative for Wellington rail but positive for Auckland rail, while international estimates were negative.

8 The fluctuations in fuel price in recent years had an impact in all three cities, although not on all modes. Fuel price elasticity was the lowest in Wellington and highest in Christchurch. The influence of fuel price was found to be significant only in recent years in both Auckland (from 2003Q3) and Christchurch (from 2004Q1) bus demand.

9 It was evident that the increase in patronage was influenced by the increase in fuel price but not influenced by fare for Auckland bus and Wellington rail.

The above observations can be explained as follows:

1 Service was the key driving factor among all five factors considered and had influence on almost every mode in all cities except on Wellington bus. This is a very encouraging result as it means the investment in PT infrastructure, new services and service improvement in the last decade did have positive influences on PT patronage in all cities.

2 The market in Wellington was quite different from Auckland and Christchurch, as Wellington had a more mature market. Wellington had the highest PT use among the three cities. Wellington also had the highest walk modal share as the CBD area is more compact and walkable. On top of that, the council also had a committed parking restraint policy in the CBD. Hence the elasticity estimates for Wellington were lower (less elastic) than those for Auckland and Christchurch in general.

3 Auckland's PT system, especially the rail system, had gone through tremendous improvement over the last decade. As a result, the service and fare elasticities were higher (more elastic) than in other cities.

4 Auckland has always been the biggest city in New Zealand with a higher proportion of PT-dependent population. As a result, it appeared that fare did not have a significant influence on bus patronage in Auckland.

5 Car ownership was the most elastic among all the factors identified and fuel price was found to have a significant influence on Auckland and Christchurch bus patronage but only in recent years. That means in the long run, the most effective policy to encourage use of PT could be by controlling car ownership or its use. Fuel price, as part of motoring cost, could have an influence in the short run as well.

6 The observed unusual positive income effect on Auckland rail could be explained by the change in rail market as a result of the tremendous investment in infrastructure and service improvement. The opening of Britomart in 2003 induced this effect as the CBD became within walking distance from the new station. A higher proportion of commuters with higher income were attracted to use the rail service. This group also appeared to be more sensitive to quality of service and hence rail had higher service elasticities.

7 Christchurch had the highest car ownership per capita of the three cities but relatively cheaper bus fares and a more convenient ticketing system. As a result, the estimated fuel price elasticity in Christchurch was higher than in Auckland and Wellington. This implied a higher substitution effect between bus and car in Christchurch compared with the other cities.

8 Despite the increase in fares as a result of increases in diesel price, the increase in fuel price was more significant. In other words, PT was still relatively 'cheap' compared with driving. As a result, the

increase in patronage was influenced by the increase in fuel price but not influenced by fare for Auckland bus and Wellington rail.

8.2 Data limitation and its implications

Our methodology and analysis in this study was limited by data availability. There were a few important issues in the data collected. For Auckland, detailed bus information was available just for contract services which represented only part of the picture in Auckland. The rail service information was measured in train-kilometres which was not accurate. For Wellington, the annual information available covered a period of nine years only. For Christchurch, service information was only estimated from timetables and fare information was only available from 2004.

Given the limitations imposed by data availability and the issues described above, PAM would be the best model to apply. Due to the differences in data format and information availability for the three cities, separate models were developed for each of the cities. The existence of endogenous explanatory variables could be an issue in a PAM in this case. Endogenous variables are the ones that can be explained by our system. In our formulation, only patronage was considered as an endogenous variable and all the structural and economic determinants (service, fare, income, car ownership and fuel price) were considered to be exogenous. One would expect that this might not be true. As patronage increases/ decreases over time, PT operators might adjust the fares and service accordingly. The structural relationships between the endogenous and exogenous variables can be modelled by VECM but not PAM. Unfortunately, it was impossible to apply VECM to all three cities.

Models can only be as good as the quality of the information they are built on. The missing or inaccurate information could have implications on the quality of the models estimated but we have already made the most out of the information collected. As shown in chapter 6, the model results are consistent with the observations in our preliminary analysis in chapter 5. The signs of estimated coefficients are intuitively correct and the significance is also consistent with the preliminary analysis. The quality of the forecast models was assured by the performance statistics and graphical inspection of the residuals in section 6.4. Furthermore, as shown in chapter 7, our study results are also comparable with international experience.

8.3 Technical issues and their implications

Three technical issues were identified during the peer review process, namely, ‘transport services offered variable’, spurious regressions and series length. The referee report where these issues were raised is depicted in appendix E.1; and the responses to the referee report are detailed in appendix E.2. In summary, these technical issues were addressed as follows:

Table 8.3 Summary of how the technical issues were addressed

Issue	Additional tests/responses
Issue 1: Transport services offered variable	We applied a technique commonly used to handle/test for possible endogeneity in time series data. We introduced lagged service variable(s), ie S_{t-1} and S_{t-2}, as instrument(s). We then applied the Breusch-Godfrey LM test (Baum and Wiggins 1999) for autocorrelation and the Wu-Hausman F-test to test the hypothesis that the variables were exogenous (Baum et al 2002).
Issue 2: Spurious regressions	We performed the Johansen trace test (Lutkepohl and Kratzig 2004) for cointegration for each set of variables in the models developed. In addition to the cointegration test, the Breusch-Godfrey LM test (Baum and Wiggins 1999) for autocorrelation was performed to ensure the residuals of the original PAM equation were not autocorrelated.
Issue 3: Series	As discussed earlier, like any modelling approach, PAM comes with its limitations. Ideally the

Issue	Additional tests/responses
length	changes of trend should be modelled as structural breaks in a VECM model. Since structural breaks cannot be handled properly with a PAM modelling approach, we can only be aware of its limitations and try to use the models that were developed based on longer time series for forecasting purposes.

The test results are depicted in table E.1 in appendix E.2. In summary, only the Auckland models passed both the endogeneity tests and cointegration tests; the Wellington data set was not sufficient to support either test; and the Christchurch models passed the endogeneity tests but the variables appeared to be not cointegrated. Some final comments from the peer reviewer on the implications of these technical issues are detailed in appendix E.3.

8.4 Further research directions

This study has only made a start in econometric analysis of PT information in New Zealand. It is clear that the Auckland data set has the quality to support the development of better models. Some suggestions were made in the final referee report by the peer reviewer who raised the technical issues described above, as shown in appendix E.3. In addition, further research is recommended in the following directions.

8.4.1 Updates of econometric analysis on a regular basis

The model developed in this study should be applied on a regular basis so authorities and PT operators will be able to monitor the changes in the market and hence respond accordingly. For example, the fuel price was at its peak at the end of the study period. Since then it has gone down and up again. It would be meaningful, for instance, to update the models in one year's time to see how the market has changed in response to the change in fuel price in the opposite direction. We would expect that what happened until the second quarter of 2008 was only a transition. Updating the model with a longer time series could certainly improve the accuracy of the model. The observed positive income effect on Auckland rail patronage could possibly be transitional as well. Further investigation would be worthwhile.

8.4.2 Fare policy in response to fuel price changes

It was evident from this study that fuel price has had a positive effect on PT patronage in recent years. At the same time, operators also justified fare increases to allow for diesel price increases, quite often at the same level of increase. One would expect this could be justified if the load factor was 100% which would not be true especially during off-peak periods. In other words, fare increases could only be partially justified by an increase in diesel price. Allowance needs to be made for the increase in revenue from the expected increase in patronage. The extent would be dependent on the fuel price elasticity of PT demand. The forecast models developed in this study could be further developed to serve this purpose.

8.4.3 System performance analysis

With the advance in technology, eg forthcoming integrated ticketing in Auckland, the electronic Metrocard introduced in Christchurch and GPS information on bus services, we should be prepared to fully utilise the information that can be captured in the future. For example, in this study, service was measured in bus-kilometres per capita which was a very crude approximation. In the future, we can capture other meaningful information, for example, delays or travel time reliability measures for buses. Other relevant variables, such as car use, parking availability and price, should also be considered.

With a more comprehensive data set, we can apply more advanced econometric analysis techniques and hence improve the quality of forecasts. For example, VECM was considered in this study as a more appropriate modelling technique but only the Auckland data set had sufficient information for this model to be developed.

8.4.4 Integrated performance analysis and forecasting

We have found in this study that econometric analysis can be a powerful planning tool for PT patronage forecast but we do need to be able to capture good quality information. A good system should be designed to capture as much information as possible in the future to support not only system performance analysis but also updating and further development of appropriate econometric forecast models, and it should also ultimately support policy decision making.

For example, an integrated ticketing system should be able to capture automatically comprehensive information on demand, origins, destinations, fare, travel time, etc. With an appropriate link, this information can be fed into an econometric forecast model to obtain updated forecasts automatically on a periodic basis. Management decisions on policy changes, such as fare adjustments and service improvements, can then be made based on the information provided.

9 References

Asteriou, D and SG Hall (2006) Applied econometrics – a modern approach using EViews and Microfit. New York: Palgrave Macmillan.

Baum, CF, ME Schaffer and S Stillman (2007) *IVENDOG: stata module to calculate Durbin-Wu-Hausman endogeneity test after ivreg*. Revised version. Boston: Boston College Department of Economics.

Baum, CF, and V Wiggins (2002) *BGTEST: stata module to calculate Breusch-Godfrey test for serial correlation.* Revised version. Boston: Boston College Department of Economics.

Bresson, G and A Pirotte (1999) *Modelling the demand for public transportation* (in French). Report VIA GTI-ERMES Université Panthéon-Assas Paris II.

Bresson, G, J Dargay, J-L Madre and A Pirotte (2003) The main determinants of the demand for public transport: a comparative analysis of England and France using shrinkage estimators. *Transportation Research Part A 37:* 605–627.

Bresson, G, J Dargay, J-L Madre and A Pirotte (2004) Economic and structural determinants of the demand for public transport: an analysis on a panel of French urban areas using shrinkage estimators. *Transportation Research Part A 38(4)*: 269–285.

Dargay, J and M Hanly (2002a) The demand for local bus services in England. *Journal of Transport Economics and Policy 36*, no.1: 73–91.

Dargay, J and M Hanly (2002b) Bus patronage in Great Britain – econometric analysis. *Transportation Research Record 1799*: 97–106.

Espino, R, J de D Ortuzar and C Roman (2007) Understanding suburban travel demand: flexible modelling with revealed and stated choice data. *Transportation Research Part A 41*: 899–912.

García-Ferrer, A, M Bujosa, A de Juan and P Poncela (2006) Demand forecast and elasticities estimation of public transport. *Journal of Transport Economics and Policy 40,* no.1: 45–67.

Greer, M (2008) *Report on the determinants of bus ridership by commuters to work in Auckland.* A research report prepared for Auckland Regional Transport Authority. Auckland: University of Auckland.

Goodwin, P (1992) A review of new demand elasticities with special reference to short and long run effects of price changes. *Journal of Transport Economics and Policy 26*, no.2: 155–163.

Hensher, DA (2008) Assessing systematic sources of variation in public transport elasticities: some comparative warnings. *Transportation Research Part A 42*: 1031–1042.

Hensher, DA and J King (1998) Establishing fare elasticity regimes for urban passenger transport: time-based fares for concession and non-concession markets segmented by trip length. *Journal of Transportation and Statistics 1*, no.1: 43–61.

Kennedy, D and I Wallace (2007) Impacts of transport fuel price changes on New Zealand transport. *Land Transport NZ research report no.331*.

Litman, T (2004) Transit price elasticities and cross-elasticities. *Journal of Public Transportation 7,* no.2: 37–58.

Lutkepohl, H and K Kratzig (2004) *Applied time series econometrics*. Cambridge: Cambridge University Press.

Madre, J-L and M Boulahbal (1999. Long term forecasting of mobility in relation to urban sprawl and demographic evolution, *Presented at UITP Congress*, Toronto, May 1999.

Oum, TH, WG Waters and J Yong (1992) Concepts of price elasticities of transport demand and recent empirical estimates – an interpretative survey. *Journal of Transport Economics and Policy 26,* no.2: 139–154.

Paulley, N, R Balcombe, R Mackett, H Titheridge, J Preston, M Wardman, J Shires and P White (2006) The demand for public transport: the effect of fares, quality of service, income, and car ownership. *Transport Policy 13*, no.4: 295–306

Pham, L and J Linsalata (1991) *Effects of fare changes on bus ridership*. American Public Transit Association.

Stock, JH and MW Watson (2007) *Introduction to econometrics*. 2nd ed. Addison Wesley.

TRACE (1999) *Elasticity handbook: elasticities for prototypical contexts.* Prepared for the European Commission, Directorate-General for Transport, Contract No. RO-97-SC.2035.

Zeileis, A (2009) dynlm: Dynamic Linear Regression. R package version 0.2-2. *http://CRAN.R-project.org/package=dynlm*

Appendix A: Data imputation

Car ownership

Step 1: Derive seasonal index – based on the quarterly licensed cars statistics available from 2002 Q4 to 2008 Q2, the seasonal index by region are derived as shown in table A.1.

Table A.1 Seasonal index for licensed cars by region

	Auckland	Wellington	Christchurch
Quarter 1	1.0031	1.0113	1.0054
Quarter 2	0.9946	1.0002	0.9923
Quarter 3	0.9923	0.9947	0.9897
Quarter 4	1.0100	0.9938	1.0126

Step 2: De-seasonalise quarterly licensed cars statistics

$$V_i^{X\prime} = V_i^X \big/ \tau_i^X$$

where V_i^X is the observed number of licensed cars in city X in quarter i; τ_i^X is the seasonal index of quarter i in city X; and $V_i^{X\prime}$ is the de-seasonalised number of licensed cars in city X in quarter i.

Step 3: Regression – regression analysis by region is performed based on the following equations:

$$V_t^{X\prime} = \beta_0^X + \beta_1^X V_t^{NZ} + \varepsilon$$

where $Q_i^{X\prime}$ is the de-seasonalised number of licensed cars in city X at time t; Q_t^{NZ} is the observed number of licensed cars in New Zealand at time t; and β_0^X and β_1^X are the regression coefficients.

The outputs from the regression analysis are shown in R-output 1 to R-output 3. The plots are shown in figure A.1.

R-output 1: regression results of regional licensed cars in Auckland versus national licensed cars

```
Call:
lm(formula = CarAKL ~ CarNZ)
Residuals:
     30      34      38       42      46      50
 -178.5   545.2   171.3 -1276.1   515.8   222.2

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.523e+04   7.622e+03  -3.311  0.0296 *
CarNZ        3.394e-01   3.477e-03  97.627  6.6e-08 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 758.7 on 4 degrees of freedom
 (44 observations deleted due to missingness)
Multiple R-squared: 0.9996,   Adjusted R-squared: 0.9995
F-statistic: 9531 on 1 and 4 DF, p-value: 6.6e-08
```

R-output 2: regression results of regional licensed cars in Wellington versus national licensed cars

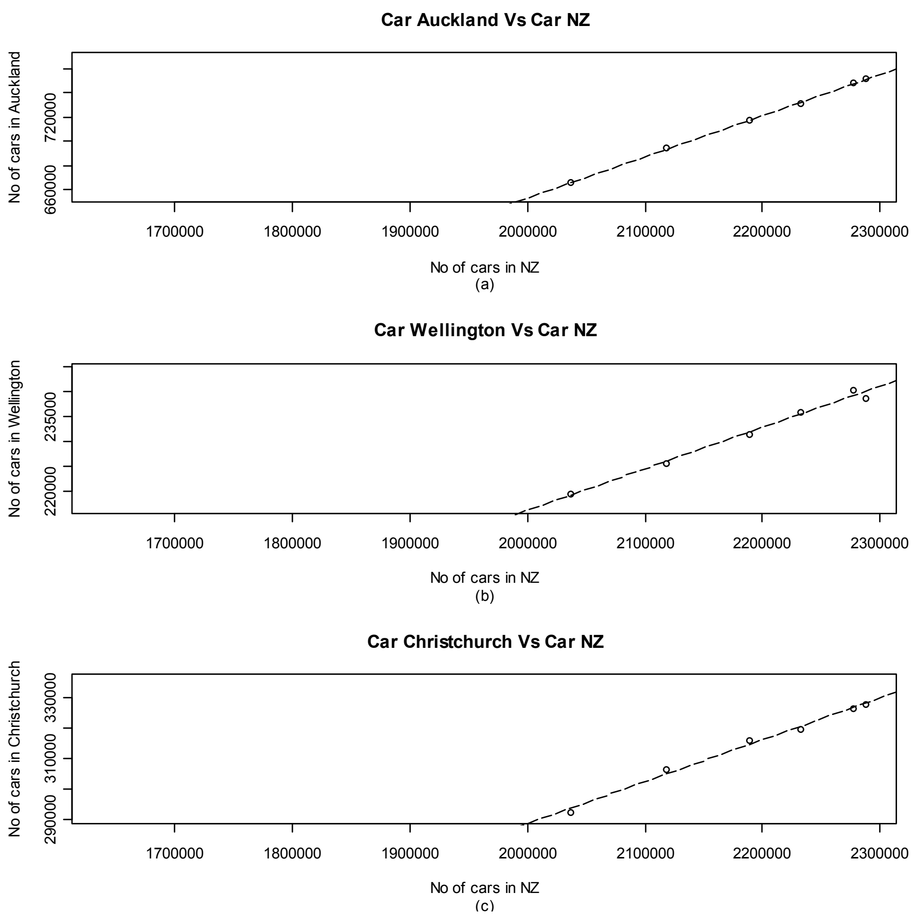
```
Call:
lm(formula = CarWLT ~ CarNZ)
Residuals:
    30      34      38      42      46      50
 261.0 -437.3 -291.1  426.0 1329.8 -1288.4
```

```
Coefficients:
       Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.972e+04  9.988e+03   4.978  0.00761 **
CarNZ       8.317e-02  4.556e-03  18.254  5.3e-05 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 994.2 on 4 degrees of freedom
 (44 observations deleted due to missingness)
Multiple R-squared: 0.9881,   Adjusted R-squared: 0.9852
F-statistic: 333.2 on 1 and 4 DF, p-value: 5.298e-05
```

R-output 3: regression results of regional licensed cars in Christchurch versus national licensed cars

```
Call:
lm(formula = CarCHC ~ CarNZ)

Residuals:
     30      34      38      42      46      50
-1373.2  1476.8  1485.2  -862.2  -457.9  -268.7

Coefficients:
             Estimate   Std. Error t value   Pr(>|t|)
(Intercept) 1.394e+04   1.357e+04    1.027   0.362
CarNZ       1.373e-01   6.190e-03   22.185   2.44e-05 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1351 on 4 degrees of freedom
 (44 observations deleted due to missingness)
Multiple R-squared: 0.9919,   Adjusted R-squared: 0.9899
F-statistic: 492.2 on 1 and 4 DF, p-value: 2.444e-05
```

Figure A.1 Plots of regional car ownership versus national car ownership



Step 4: Estimate quarterly regional car ownership – the equations obtained in step 3 and the seasonal indices obtained in step 1 are then applied to the annual statistics to obtain the estimated quarterly car ownership for the period 1996 Q1 to 2002 Q3.

Regional resident population

Step 1: Regression analysis – regress the regional resident population as the dependent variable with the national resident population information available.

$$P_t^X = \beta_0^X + \beta_1^X P_t^{NZ} + \varepsilon$$

where P_t^X is the resident population in city X at time t; and P_t^{NZ} is the national resident population at time t.

The regression outputs are shown in R-output 4 to R-output 6. The plots are shown in figure A.2.

R-output 4: regression results of resident population Auckland versus resident population NZ

```
Call:
lm(formula = PopAKL ~ PopNZ)

Residuals:
   Min    1Q Median      3Q  Max
 -9631  -6744 -2161   3450 16433

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.514e+05  7.681e+04  -12.39 5.88e-07 ***
PopNZ        5.592e-01  1.938e-02   28.86 3.51e-10 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9511 on 9 degrees of freedom
 (39 observations deleted due to missingness)
Multiple R-squared: 0.9893,   Adjusted R-squared: 0.9881
F-statistic: 832.7 on 1 and 9 DF, p-value: 3.512e-10
```

R-output 5: regression results of resident population Wellington versus resident population NZ

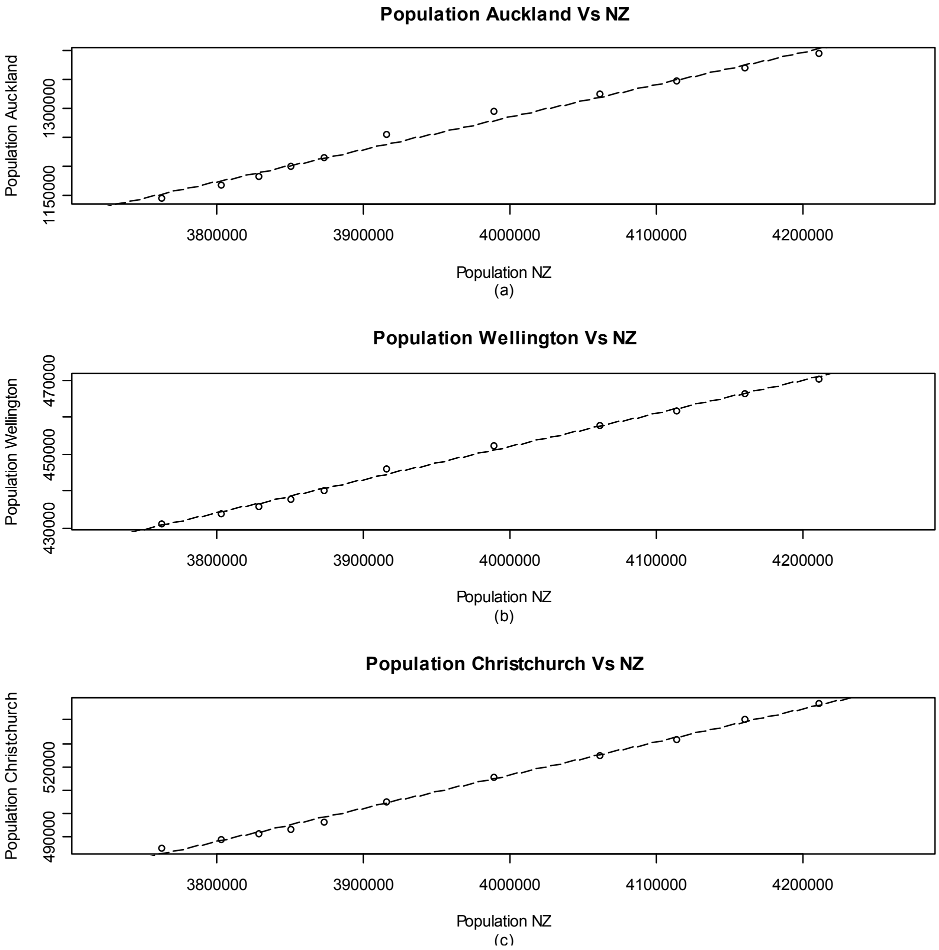
```
Call:
lm(formula = PopWLT ~ PopNZ)

Residuals:
    Min     1Q Median      3Q   Max
-1085.9 -529.4 -348.7  474.5 1394.3

Coefficients:
         Estimate     Std. Error t value  Pr(>|t|)
(Intercept) 9.398e+04  7.040e+03   13.35  3.09e-07 ***
PopNZ       8.948e-02  1.776e-03   50.38  2.40e-12 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 871.6 on 9 degrees of freedom
 (39 observations deleted due to missingness)
Multiple R-squared: 0.9965,   Adjusted R-squared: 0.9961
F-statistic: 2539 on 1 and 9 DF, p-value: 2.398e-12
```

R-output 6: regression results of resident population Christchurch versus resident population NZ

```
Call:
lm(formula = PopCHC ~ PopNZ)

Residuals:
    Min     1Q  Median     3Q    Max
-1698.8 -879.1   376.9  582.0 2698.0

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5.034e+04  1.093e+04  -4.606 0.00128 **
```

```
PopNZ        1.417e-01  2.757e-03  51.388 2.01e-12 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1353 on 9 degrees of freedom
 (39 observations deleted due to missingness)
Multiple R-squared: 0.9966,   Adjusted R-squared: 0.9962
F-statistic: 2641 on 1 and 9 DF, p-value: 2.010e-12
```

Figure A.2 Plots of regional resident population versus national resident population



Appendix B: Stationarity tests

Table B1 Auckland patronage and potential influencing factors

Variable	Description	Stationarity test results[16]	Variable type
$\ln Q_t^{A(B)}$	Bus patronage per capita in Auckland	ADF Test for series: lnPaxBusPerCapAKL sample range: [1998 Q2, 2008 Q2], T = 41 lagged differences: 8 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -3.9818	Null hypothesis of non-stationarity is rejected at 1% level of significance. $\ln Q_t^{A(B)}$ is stationary, ie $I(0)$
$\ln Q_t^{A(R)}$	Rail patronage per capita in Auckland	ADF Test for series: lnPaxRailPerCapAKL sample range: [1997 Q2, 2008 Q2], T = 45 lagged differences: 4 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -1.1705	Null hypothesis of non-stationarity is not rejected. $\ln Q_t^{A(R)}$ is non-stationary. Test on first difference to proceed
$\Delta \ln Q_t^{A(R)}$	First difference of rail patronage per in Auckland	ADF Test for series: lnPaxRailPerCapAKL_d1 sample range: [1997 Q2, 2008 Q2], T = 45 lagged differences: 3 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -3.7994	Null hypothesis of non-stationarity is rejected at 5% level of significance. First difference $\Delta \ln Q_t^{A(R)}$ is stationary, ie levels variable $\ln Q_t^{A(R)}$ is $I(1)$
$\ln S_t^{A(B)}$	Bus-km per capita in Auckland	ADF Test for series: lnBusKmPerCapAKL sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend	Null hypothesis of non-stationarity is not rejected. $\ln S_t^{A(B)}$ is non-stationary. Test on first difference

[16] ADF tests are performed with JMulTi version 4.23 developed by Lutkepohl and Kratzig (2004). The lagged differences are optimised based on the four criteria (or the most common suggestion out of the four) in JMulTi, ie Akaike Info Criterion, Final Prediction Error, Hannan-Quinn Criterion and Schwarz Criterion.

Variable	Description	Stationarity test results[16]	Variable type
		asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -2.6721	to proceed
$\Delta \ln S_t^{A(B)}$	First difference of bus-km per capita in Auckland	ADF Test for series: lnBusKmPerCapAKL_d1 sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -9.6731	Null hypothesis of non-stationarity is rejected at 1% level of significance. First difference $\Delta \ln S_t^{A(B)}$ is stationary, ie levels variable $\ln S_t^{A(B)}$ is $I(1)$
$\ln S_t^{A(R)}$	Train-km per capita in Auckland	ADF Test for series: lnRailKmPerCapAKL sample range: [1997 Q3, 2008 Q2], T = 44 lagged differences: 4 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -1.1784	Null hypothesis of non-stationarity is not rejected. $\ln S_t^{A(B)}$ is non-stationary. Test on first difference to proceed
$\Delta \ln S_t^{A(R)}$	First difference of train-km per capita in Auckland	ADF Test for series: lnRailKmPerCapAKL_d1 sample range: [1997 Q1, 2008 Q2], T = 46 lagged differences: 2 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -6.8570	Null hypothesis of non-stationarity is rejected at 1% level of significance. First difference $\Delta \ln S_t^{A(R)}$ is stationary, ie levels variable $\ln S_t^{A(R)}$ is $I(1)$
$\ln F_t^{A(B)}$	Real average revenue per bus passenger in Auckland	ADF Test for series: lnRealFareBusAKL sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -2.8867	Null hypothesis of non-stationarity is not rejected. $\ln F_t^{A(B)}$ is non-stationary. Test on first difference to proceed

Variable	**Description**	**Stationarity test results**[16]	**Variable type**
$\Delta \ln F_t^{A(B)}$	First difference of real average revenue per bus passenger in Auckland	ADF Test for series: lnRealFareBusAKL_d1 sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -7.8690	Null hypothesis of non-stationarity is rejected at 1% level of significance. First difference $\Delta \ln F_t^{A(B)}$ is stationary, ie levels variable $\ln F_t^{A(B)}$ is $I(1)$
$\ln F_t^{A(R)}$	Real average revenue per rail passenger in Auckland	ADF Test for series: lnRealFareRailAKL sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -3.9023	Null hypothesis of non-stationarity is rejected at 5% level of significance. $\ln F_t^{A(R)}$ is stationary, ie $I(0)$
$\ln C_t^A$	Car ownership per capita in Auckland	ADF Test for series: lnCarOwnPerCapAKL sample range: [1998 Q3, 2008 Q2], T = 40 lagged differences: 8 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -3.8546	Null hypothesis of non-stationarity is rejected at 5% level of significance. $\ln C_t^A$ is stationary, ie $I(0)$
$\ln O_t$	Real fuel price (regular petrol)	ADF Test for series: lnRealFuelP sample range: [1996 Q3, 2008 Q2], T = 48 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -2.1174	Null hypothesis of non-stationarity is not rejected. $\ln O_t$ is non-stationary. Test on first difference to proceed
$\Delta \ln O_t$	First difference of real fuel price (regular petrol)	ADF Test for series: lnRealFuelP_d1 sample range: [1996 Q4, 2008 Q1], T = 46 lagged differences: 1 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10%	Null hypothesis of non-stationarity is rejected at 1% level of significance. First difference $\Delta \ln O_t$ is stationary, ie levels variable $\ln O_t$ is $I(1)$

Variable	Description	Stationarity test results[16]	Variable type
		-3.96 -3.41 -3.13 value of test statistic: -6.2044	
$\ln I_t$	Real gross disposable income per capita	ADF Test for series: lnRealIncome sample range: [1996 Q3, 2008 Q1], T = 47 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -2.9094	Null hypothesis of non-stationarity is not rejected. $\ln I_t$ is non-stationary. Test on first difference to proceed
$\Delta \ln I_t$	First difference of real gross disposable income per capita	ADF Test for series: lnRealIncome_d1 sample range: [1996 Q3, 2008 Q1], T = 47 lagged differences: 0 intercept, time trend asymptotic critical values reference: Davidson and MacKinnon (1993) p708, table 20.1. 1% 5% 10% -3.96 -3.41 -3.13 value of test statistic: -7.7677	Null hypothesis of non-stationarity is rejected at 1% level of significance. First difference $\Delta \ln I_t$ is stationary, ie levels variable $\ln I_t$ is $I(1)$

Reference

Davidson, R and J MacKinnon (1993) *Estimation and inference in econometrics*. London: Oxford University Press.

Appendix C: Model estimation outputs

Auckland bus patronage model (1999Q3–2008Q2)

Step 1: full period, all variables, all indicators

We first started with all the variables for the full period, ie 1996 Q1-2008 Q2, plus a dummy variable (an indicator) for the observed change of trend in 1999 Q3, D_t^{99Q3}.

$$D_t^{99Q3} = \begin{cases} 1 & \text{if } t \text{ is in or after 1999 Q3} \\ 0 & \text{otherwise} \end{cases}$$

Service + fare + car ownership+ fuel price + income + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} \\ &+ \beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+ \gamma_2^{A(B)} D_{2t} + \gamma_3^{A(B)} D_{3t} + \gamma_4^{A(B)} D_{4t} + \gamma^{99Q3} D_t^{99Q3} \end{aligned} \quad \text{(Equation C.1)}$$

where $t \in \left[1996Q1, 2008Q2\right]$

```
Time series regression with 'ts' data:
Start = 1996(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts96 ~ lnBusKmPerCapAKL.ts96 +
  lnRealBusFareAKL.ts96 + lnCarOwnPerCapAKL.ts96 + lnRealIncomePerCap.ts96 +
  lnRealFuelP.ts96 + IndQtr2.ts96 + IndQtr3.ts96 + IndQtr4.ts96 +
  L(lnBusPaxPerCapAKL.ts96) + IndTrend99Q3)

Residuals:
      Min        1Q    Median       3Q      Max
-0.114183 -0.020165 -0.002953 0.024500 0.107902

Coefficients:
                          Estimate Std. Error t value  Pr(>|t|)
(Intercept)              -0.072934   0.979450  -0.074   0.94103
lnBusKmPerCapAKL.ts96     0.377445   0.123009   3.068   0.00396 **
lnRealBusFareAKL.ts96    -0.072008   0.171840  -0.419   0.67754
lnCarOwnPerCapAKL.ts96   -0.464153   0.611459  -0.759   0.45248
lnRealIncomePerCap.ts96  -0.087114   0.319712  -0.272   0.78673
lnRealFuelP.ts96         -0.034851   0.107960  -0.323   0.74860
IndQtr2.ts96             -0.003783   0.020147  -0.188   0.85208
IndQtr3.ts96             -0.026719   0.021292  -1.255   0.21718
IndQtr4.ts96             -0.102896   0.022060  -4.664  3.77e-05 ***
L(lnBusPaxPerCapAKL.ts96) 0.680145   0.081125   8.384  3.60e-10 ***
IndTrend99Q3              0.134162   0.040884   3.282   0.00222 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04876 on 38 degrees of freedom
Multiple R-squared: 0.9561,   Adjusted R-squared: 0.9446
F-statistic: 82.84 on 10 and 38 DF, p-value: < 2.2e-16
```

As shown above, the coefficients of two dummy variables were significantly not equal to zero. The first one was the dummy variable for quarter 4 with a level significance of 3.77e-05 and the second one was the dummy variable created for the observed change of trend in 1999 Q3 with a level of significance of 0.00222.

Step 2: period changed, take out insignificant dummy variables

Since the dummy variable created for the change of trend in 1999 Q3 was significant, we modified the model period to 1999 Q3 to 2008 Q2. We also took out the dummy variables for seasons 2 and 3 as they were not significant.

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} \\ &+\beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+\gamma_4^{A(B)} D_{4t}\end{aligned} \quad \text{(Equation C.2)}$$

where $t \in [1999Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1999(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts99 ~ lnBusKmPerCapAKL.ts99 +
  lnRealBusFareAKL.ts99 + lnCarOwnPerCapAKL.ts99 + lnRealIncomePerCap.ts99 +
  lnRealFuelP.ts99 + IndQtr4.ts99 + L(lnBusPaxPerCapAKL.ts99))

Residuals:
      Min        1Q    Median       3Q      Max
-0.069796 -0.012427 -0.002373 0.009343 0.065664

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -2.08376    0.83179   -2.505  0.01857 *
lnBusKmPerCapAKL.ts99      0.42998    0.13702    3.138  0.00408 **
lnRealBusFareAKL.ts99      0.08670    0.12631    0.686  0.49831
lnCarOwnPerCapAKL.ts99    -1.78433    0.77876   -2.291  0.02998 *
lnRealIncomePerCap.ts99    0.42353    0.23002    1.841  0.07660 .
lnRealFuelP.ts99           0.08780    0.08013    1.096  0.28292
IndQtr4.ts99              -0.04940    0.01691   -2.922  0.00695 **
L(lnBusPaxPerCapAKL.ts99) 0.36582    0.12257    2.985  0.00597 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03051 on 27 degrees of freedom
Multiple R-squared: 0.9636,   Adjusted R-squared: 0.9542
F-statistic: 102.1 on 7 and 27 DF, p-value: < 2.2e-16
```

As shown above, out of the five independent variables, only three were significant: service at level of significance of 0.00408; car ownership at 0.02998; and income at 0.07660. The estimated coefficient of car ownership was negative while that of income was positive which was counter intuitive. As discussed in the trend analysis for Auckland in section 5.1.1, the trends of car ownership and income were pretty much correlated. It would not be appropriate to include both variables in the equation.

Step 3: take out insignificant variables, try different combinations of significant variables

We then excluded the two insignificant variables, bus fare and fuel price. For the remaining three variables, there were only two combinations to try as car ownership should not be included at the same time:

Service + income+ Q4 seasonal dummy

$$\ln Q_t^{A(B)} = \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_I^{A(B)} \ln I_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t} \quad \text{(Equation C.3)}$$

where $t \in [1999Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1999(4), End = 2008(2)
Call:
dynlm(formula = lnBusPaxPerCapAKL.ts99 ~ lnBusKmPerCapAKL.ts99 +
  lnRealIncomePerCap.ts99 + IndQtr4.ts99 + L(lnBusPaxPerCapAKL.ts99))
```

```
Residuals:
      Min        1Q    Median       3Q      Max
-0.077430 -0.017937 -0.005086 0.010924 0.086771

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.02635    0.30712    0.086 0.932203
lnBusKmPerCapAKL.ts99     0.34891    0.08869    3.934 0.000457 ***
lnRealIncomePerCap.ts99   0.15847    0.21503    0.737 0.466866
IndQtr4.ts99             -0.08095    0.01365   -5.929  1.7e-06 ***
L(lnBusPaxPerCapAKL.ts99) 0.48018    0.12191    3.939 0.000452 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03328 on 30 degrees of freedom
Multiple R-squared: 0.9519,   Adjusted R-squared: 0.9455
F-statistic: 148.4 on 4 and 30 DF, p-value: < 2.2e-16
```

Service + car ownership+ Q4 seasonal dummy

$$\ln Q_t^{A(B)} = \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A + \theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t} \quad \text{(Equation C.4)}$$

where $t \in [1999Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1999(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts99 ~ lnBusKmPerCapAKL.ts99 +
  lnCarOwnPerCapAKL.ts99 + IndQtr4.ts99 + L(lnBusPaxPerCapAKL.ts99))

Residuals:
       Min         1Q     Median        3Q       Max
-0.0933966 -0.0108164 -0.0008438 0.0057223 0.0721826

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -0.80458    0.56163   -1.433 0.162318
lnBusKmPerCapAKL.ts99      0.49768    0.10998    4.525 8.87e-05 ***
lnCarOwnPerCapAKL.ts99    -1.32387    0.70225   -1.885 0.069124 .
IndQtr4.ts99              -0.06607    0.01561   -4.234 0.000200 ***
L(lnBusPaxPerCapAKL.ts99)  0.50706    0.10205    4.969 2.54e-05 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03175 on 30 degrees of freedom
Multiple R-squared: 0.9562,   Adjusted R-squared: 0.9504
F-statistic: 163.8 on 4 and 30 DF, p-value: < 2.2e-16
```

Only in the second of the above two combinations were both independent variables (service and car ownership) significant. Therefore, this is an acceptable model for Auckland based on the period from 1999 Q3 to 2008 Q2.

Auckland bus patronage model (2003Q3–2008Q2)

As discussed in section 5.1.1, the opening of the Britomart Transport Centre in July 2003 marked the beginning of a new trend in railway patronage in Auckland. For consistency in modelling for Auckland, we also needed to develop a model for bus patronage for the same period.

Step 1: Period 2003 Q3 to 2008 Q2, all variables, all indicators

We included all five independent variables plus an indicator for a possible change of trend due to the opening of the Northern Busway in July 2005. We created a dummy variable to represent the break, D_t^{05Q3} where

Service + fare + car ownership + income + fuel price + seasonal dummies + break

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} \\ &+\beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+\gamma_2^{A(B)} D_{2t} + \gamma_3^{A(B)} D_{3t} + \gamma_4^{A(B)} D_{4t} + \gamma^{05Q3} D_t^{05Q3}\end{aligned} \qquad \text{(Equation C.5)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts03 ~ lnBusKmPerCapAKL.ts03 +
  lnRealBusFareAKL.ts03 + lnCarOwnPerCapAKL.ts03 + lnRealIncomePerCap.ts03 +
  lnRealFuelP.ts03 + IndQtr2.ts03 + IndQtr3.ts03 + IndQtr4.ts03 +
  L(lnBusPaxPerCapAKL.ts03) + IndNSH.ts03)

Residuals:
      Min        1Q    Median       3Q      Max
-0.022831 -0.005161 -0.001235 0.005818 0.023197

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -2.451099   0.749921  -3.268  0.0114 *
lnBusKmPerCapAKL.ts03      0.858893   0.297363   2.888  0.0202 *
lnRealBusFareAKL.ts03      0.145832   0.191159   0.763  0.4674
lnCarOwnPerCapAKL.ts03    -1.868938   0.872764  -2.141  0.0646 .
lnRealIncomePerCap.ts03    0.040536   0.368688   0.110  0.9152
lnRealFuelP.ts03           0.215517   0.105375   2.045  0.0751 .
IndQtr2.ts03               0.002391   0.012707   0.188  0.8554
IndQtr3.ts03              -0.014483   0.015872  -0.913  0.3882
IndQtr4.ts03              -0.040520   0.021339  -1.899  0.0941 .
L(lnBusPaxPerCapAKL.ts03)  0.212387   0.138943   1.529  0.1649
IndNSH.ts03               -0.080576   0.060341  -1.335  0.2185
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01602 on 8 degrees of freedom
Multiple R-squared: 0.9893,   Adjusted R-squared: 0.976
F-statistic: 74.08 on 10 and 8 DF, p-value: 8.812e-07
```

As shown above, the dummy variable created for the impact of the opening of the Northern Busway was not significant. Moreover, the dummy variables for Q2 and Q3 were also insignificant.

Step 2: Period unchanged, take out insignificant dummy variables

We excluded the three insignificant dummy variables and re-ran the model period which remained unchanged.

Service + fare + car ownership + income + fuel price + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} \\ &+\beta_I^{A(B)} \ln I_t + \beta_O^{A(B)} \ln O_t + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+\gamma_4^{A(B)} D_{4t}\end{aligned} \qquad \text{(Equation C.6)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts03 ~ lnBusKmPerCapAKL.ts03 +
  lnRealBusFareAKL.ts03 + lnCarOwnPerCapAKL.ts03 + lnRealIncomePerCap.ts03 +
```

```
  lnRealFuelP.ts03 + IndQtr4.ts03 + L(lnBusPaxPerCapAKL.ts03))

Residuals:
       Min         1Q     Median        3Q       Max
-0.0301333 -0.0080909 -0.0008048 0.0063041 0.0273053

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              -1.95438    0.61949   -3.155 0.00916 **
lnBusKmPerCapAKL.ts03     0.43022    0.10704    4.019 0.00202 **
lnRealBusFareAKL.ts03     0.21424    0.15179    1.411 0.18579
lnCarOwnPerCapAKL.ts03   -2.50441    0.68943   -3.633 0.00394 **
lnRealIncomePerCap.ts03  -0.20375    0.32258   -0.632 0.54054
lnRealFuelP.ts03          0.24094    0.08447    2.852 0.01573 *
IndQtr4.ts03             -0.02949    0.01810   -1.630 0.13140
L(lnBusPaxPerCapAKL.ts03) 0.26178    0.10963    2.388 0.03599 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01581 on 11 degrees of freedom
Multiple R-squared: 0.9857,   Adjusted R-squared: 0.9766
F-statistic: 108.2 on 7 and 11 DF, p-value: 2.938e-09
```

As shown above, fare and income were not significant. We excluded these variables.

Step3: Take out insignificant variables, include all significant variables

We included all the remaining significant variables and obtained a satisfactory model for the period 2003Q3–2008Q2.

Service + car ownership + fuel price+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_C^{A(B)} \ln C_t^A + \beta_O^{A(B)} \ln O_t \\ &\quad + \theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t}\end{aligned} \qquad \text{(Equation C.7)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts03 ~ lnBusKmPerCapAKL.ts03 +
  lnCarOwnPerCapAKL.ts03 + lnRealFuelP.ts03 + IndQtr4.ts03 +
  L(lnBusPaxPerCapAKL.ts03))

Residuals:
      Min        1Q   Median       3Q      Max
-0.027448 -0.008362 0.001468 0.006955 0.032450

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.93063    0.45113   -4.280 0.000897 ***
lnBusKmPerCapAKL.ts03      0.46028    0.08629    5.334 0.000136 ***
lnCarOwnPerCapAKL.ts03    -1.96184    0.57928   -3.387 0.004865 **
lnRealFuelP.ts03           0.20400    0.06155    3.314 0.005592 **
IndQtr4.ts03              -0.04783    0.01299   -3.682 0.002763 **
L(lnBusPaxPerCapAKL.ts03)  0.36767    0.08198    4.485 0.000614 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01589 on 13 degrees of freedom
Multiple R-squared: 0.9829,   Adjusted R-squared: 0.9763
F-statistic: 149.6 on 5 and 13 DF, p-value: 5.145e-11
```

Auckland rail patronage model (2003Q3–2008Q2)

Step1: Full period, all variables, all indicators

We started with including all the variables, including the five independent variables plus the seasonal indicators.

Service + fare + car ownership+ fuel price + income + seasonal dummies

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} \\ &+\beta_I^{A(R)} \ln I_t + \beta_O^{A(R)} \ln O_t + \theta^{A(R)} \ln Q_{t-1}^{A(R)} \\ &+\gamma_2^{A(R)} D_{2t} + \gamma_3^{A(R)} D_{3t} + \gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation C.8)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + lnCarOwnPerCapAKL.ts03 + lnRealIncomePerCap.ts03 +
  lnRealFuelP.ts03 + IndQtr2.ts03 + IndQtr3.ts03 + IndQtr4.ts03 +
  L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q   Median       3Q      Max
-0.064183 -0.011345 0.009056 0.015605 0.030001

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -0.681657  1.662629   -0.410  0.6914
lnRailKmPerCapAKL.ts03     1.010085  0.279308    3.616  0.0056 **
lnRealRailFareAKL.ts03    -0.870139  0.337085   -2.581  0.0296 *
lnCarOwnPerCapAKL.ts03     0.489137  1.907331    0.256  0.8034
lnRealIncomePerCap.ts03    1.452767  0.641808    2.264  0.0499 *
lnRealFuelP.ts03           0.011887  0.263095    0.045  0.9650
IndQtr2.ts03               0.045111  0.028974    1.557  0.1539
IndQtr3.ts03               0.008976  0.032170    0.279  0.7865
IndQtr4.ts03              -0.115305  0.043902   -2.626  0.0275 *
L(lnRailPaxPerCapAKL.ts03 )0.273500  0.194100    1.409  0.1924
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03679 on 9 degrees of freedom
Multiple R-squared: 0.9904,   Adjusted R-squared: 0.9809
F-statistic: 103.7 on 9 and 9 DF, p-value: 5.239e-08
```

As shown above, only one dummy variable, Q4, was significant.

Step 2: Take out insignificant dummy variables

We excluded the dummy variables for quarter 2 and 3 and reran the regression.

$$\begin{aligned}\ln Q_t^{A(R)} &= \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} \\ &+\beta_I^{A(R)} \ln I_t + \beta_O^{A(R)} \ln O_t + \theta^{A(R)} \ln Q_{t-1}^{A(R)} \\ &+\gamma_4^{A(R)} D_{4t}\end{aligned} \qquad \text{(Equation C.9)}$$

$t \in [2003Q3, 2008Q2]$

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

where

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + lnCarOwnPerCapAKL.ts03 + lnRealIncomePerCap.ts03 +
  lnRealFuelP.ts03 + IndQtr4.ts03 + L(lnRailPaxPerCapAKL.ts03))

Residuals:
```

```
     Min       1Q  Median      3Q     Max
-0.06644 -0.02189 0.00994 0.01845 0.05283

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.78291    1.62304   -1.099 0.29545
lnRailKmPerCapAKL.ts03     1.02415    0.29259    3.500 0.00497 **
lnRealRailFareAKL.ts03    -0.91755    0.34820   -2.635 0.02320 *
lnCarOwnPerCapAKL.ts03    -0.46583    1.70184   -0.274 0.78937
lnRealIncomePerCap.ts03    1.53042    0.65781    2.327 0.04011 *
lnRealFuelP.ts03           0.09372    0.26699    0.351 0.73219
IndQtr4.ts03              -0.11491    0.04366   -2.632 0.02332 *
L(lnRailPaxPerCapAKL.ts03) 0.24590    0.18893    1.302 0.21967
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0387 on 11 degrees of freedom
Multiple R-squared: 0.9871,   Adjusted R-squared: 0.9789
F-statistic: 120.1 on 7 and 11 DF, p-value: 1.676e-09
```

As shown above, car ownership and fuel price were not significant. Although income appeared to be a significant factor, the estimated coefficient was positive which was counter intuitive.

Step 3: Take out insignificant variables, try different combinations of significant variables

Based on observations in step-2, we created different possible combinations of variables as follows:

Service + fare + car ownership+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_C^{A(R)} \ln C_t^{A} + \theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t} \quad \text{(Equation C.10)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + lnCarOwnPerCapAKL.ts03 + IndQtr4.ts03 +
  L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q    Median       3Q      Max
-0.058347 -0.022494 -0.002960 0.002811 0.083644

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                 1.44430    1.05458    1.370  0.1940
lnRailKmPerCapAKL.ts03      0.92338    0.32798    2.815  0.0146 *
lnRealRailFareAKL.ts03     -0.69811    0.26951   -2.590  0.0224 *
lnCarOwnPerCapAKL.ts03     -0.49034    1.87335   -0.262  0.7976
IndQtr4.ts03               -0.13898    0.04774   -2.911  0.0121 *
L(lnRailPaxPerCapAKL.ts03)  0.44256    0.19032    2.325  0.0369 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04578 on 13 degrees of freedom
Multiple R-squared: 0.9786,   Adjusted R-squared: 0.9704
F-statistic: 119.1 on 5 and 13 DF, p-value: 2.189e-10
```

As shown above, car ownership was not a significant factor.

Service + fare + income

Service + fare + income+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \beta_I^{A(R)} \ln I_t + \theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t}$$

(Equation C.11)

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + lnRealIncomePerCap.ts03 + IndQtr4.ts03 +
  L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q   Median       3Q      Max
-0.063648 -0.026550 0.006022 0.020504 0.048103

Coefficients:
                          Estimate Std. Error  t value Pr(>|t|)
(Intercept)               -1.14479    1.06804   -1.072 0.303277
lnRailKmPerCapAKL.ts03     0.99461    0.23105    4.305 0.000856 ***
lnRealRailFareAKL.ts03    -0.96723    0.22613   -4.277 0.000900 ***
lnRealIncomePerCap.ts03    1.60601    0.56710    2.832 0.014138 *
IndQtr4.ts03              -0.13253    0.02734   -4.848 0.000318 ***
L(lnRailPaxPerCapAKL.ts03) 0.29570    0.15344    1.927 0.076105 .
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03609 on 13 degrees of freedom
Multiple R-squared: 0.9867,   Adjusted R-squared: 0.9816
F-statistic: 193.2 on 5 and 13 DF, p-value: 1.007e-11
```

As shown above, real income was a significant factor. However, the estimated coefficient had a positive sign which was counter intuitive.

$$\ln Q_t^{A(R)} = \alpha^{A(R)} + \beta_S^{A(R)} \ln S_t^{A(R)} + \beta_F^{A(R)} \ln F_t^{A(R)} + \theta^{A(R)} \ln Q_{t-1}^{A(R)} + \gamma_4^{A(R)} D_{4t}$$

(Equation C.12)

Service + fare+ Q4 seasonal dummy

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + IndQtr4.ts03 + L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q    Median       3Q      Max
-0.056704 -0.023389 -0.004208 0.004447 0.082838

Coefficients:
                           Estimate Std. Error  t value Pr(>|t|)
(Intercept)                 1.69131    0.45480    3.719 0.002290 **
lnRailKmPerCapAKL.ts03      0.88265    0.27893    3.164 0.006892 **
lnRealRailFareAKL.ts03     -0.67551    0.24667   -2.739 0.016001 *
IndQtr4.ts03               -0.14775    0.03285   -4.498 0.000501 ***
L(lnRailPaxPerCapAKL.ts03)  0.45840    0.17434    2.629 0.019811 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04423 on 14 degrees of freedom
Multiple R-squared: 0.9785,   Adjusted R-squared: 0.9724
F-statistic: 159.5 on 4 and 14 DF, p-value: 1.653e-11
```

We obtained two satisfactory models in this case:

- service + fare + income
- service + fare.

Wellington bus patronage model (FY99/00–FY07/08)

Step 1: Include all variables

We started with all the five independent variables included.

Service + fare + car ownership+ fuel price + income

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_C^{W(B)} \ln C_t^{W} + \beta_I^{W(B)} \ln I_t + \beta_O^{W(B)} \ln O_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation C.13)}$$

where $t \in [FY99/00, FY07/08]$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ lnBusKmPerCapWLT.ts00 +
  lnRealBusFareWLT.ts00 + lnCarOwnPerCapWLT.ts00 + lnRealIncomePerCap.ts00 +
  lnRealFuelP.ts00 + L(lnBusPaxPerCapWLT.ts00))

Residuals:
      2001      2002       2003      2004       2005      2006       2007
-0.0080429 0.0105728 -0.0003453 0.0054221 -0.0255946 0.0225043 -0.0031128
      2008
-0.0014036

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                5.210737   32.430067  0.161  0.899
lnBusKmPerCapWLT.ts00      0.137068    0.258228  0.531  0.689
lnRealBusFareWLT.ts00     -0.152981    0.545408 -0.280  0.826
lnCarOwnPerCapWLT.ts00     1.505717   12.018560  0.125  0.921
lnRealIncomePerCap.ts00   -0.100920    3.076152 -0.033  0.979
lnRealFuelP.ts00          -0.080974    0.357719 -0.226  0.858
L(lnBusPaxPerCapWLT.ts00) -0.008537    3.484986 -0.002  0.998

Residual standard error: 0.03714 on 1 degrees of freedom
Multiple R-squared: 0.9273,   Adjusted R-squared: 0.4909
F-statistic: 2.125 on 6 and 1 DF, p-value: 0.4817
```

As shown above, none of the variables was significant.

Step 2: Test significance of PT related variables variables

Because none of the variables was significant in the previous step, we changed our strategy. We tested the significance of both the PT-related variables variables, ie service and fare, and their significance individually.

Service + fare

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation C.14)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ lnBusKmPerCapWLT.ts00 +
  lnRealBusFareWLT.ts00 + L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003     2004      2005     2006      2007     2008
-0.016539 0.013006 0.006767 0.008685 -0.023603 0.019282 -0.008879 0.001281
```

```
Coefficients:
                          Estimate Std. Error   t value Pr(>|t|)
(Intercept)                 1.4623     0.9051     1.616  0.1815
lnBusKmPerCapWLT.ts00       0.1123     0.1263     0.890  0.4240
lnRealBusFareWLT.ts00      -0.2533     0.1016    -2.493  0.0673 .
L(lnBusPaxPerCapWLT.ts00)   0.5395     0.1600     3.373  0.0280 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01983 on 4 degrees of freedom
Multiple R-squared: 0.917,   Adjusted R-squared: 0.8548
F-statistic: 14.74 on 3 and 4 DF, p-value: 0.01255
```

Service only

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_S^{W(B)} \ln S_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)}$$ (Equation C.15)

```
Time series regression with ts data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ lnBusKmPerCapWLT.ts00 +
  L(lnBusPaxPerCapWLT.ts00))

Residuals:
      2001      2002      2003      2004      2005      2006       2007
-0.0185068 0.0019641 0.0006343 0.0016121 0.0016142 0.0503206 -0.0334067
      2008
-0.0042318

Coefficients:
                        Estimate Std. Error  t value  Pr(>|t|)
(Intercept)              0.87778    1.24951    0.703  0.5137
lnBusKmPerCapWLT.ts00    0.03509    0.17501    0.201  0.8490
L(lnBusPaxPerCapWLT.ts00) 0.74041    0.19749    3.749  0.0133*
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02835 on 5 degrees of freedom
Multiple R-squared: 0.7881,   Adjusted R-squared: 0.7033
F-statistic: 9.298 on 2 and 5 DF, p-value: 0.02067
```

Fare only

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \theta^{W(B)} \ln Q_{t-1}^{W(B)}$$ (Equation C.16)

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ +lnRealBusFareWLT.ts00 +
  L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003      2004      2005     2006      2007     2008
-0.017702 0.014627 0.014371 -0.006231 -0.021781 0.023986 -0.007862 0.000592

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                 2.05903    0.59481    3.462  0.0180 *
lnRealBusFareWLT.ts00      -0.23115    0.09643   -2.397  0.0618 .
L(lnBusPaxPerCapWLT.ts00)   0.49344    0.14816    3.330  0.0208 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01941 on 5 degrees of freedom
Multiple R-squared: 0.9006,   Adjusted R-squared: 0.8609
F-statistic: 22.65 on 2 and 5 DF, p-value: 0.003114
```

As shown above, only fare had significant influence on bus patronage in Wellington.

Step 3: Test significance of other variables

We then tested the significance of other variables, including car ownership, fuel price and income.

Fare + car ownership

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_C^{W(B)} \ln C_t^W + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation C.17)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ +lnRealBusFareWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003      2004      2005     2006      2007      2008
-0.018113 0.014969 0.014456 -0.005580 -0.021205 0.024178 -0.006682 -0.002023

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                2.5774     3.7374    0.690  0.528
lnRealBusFareWLT.ts00     -0.2114     0.1764   -1.199  0.297
lnCarOwnPerCapWLT.ts00     0.2117     1.5023    0.141  0.895
L(lnBusPaxPerCapWLT.ts00)  0.3946     0.7205    0.548  0.613

Residual standard error: 0.02165 on 4 degrees of freedom
Multiple R-squared: 0.9011,   Adjusted R-squared: 0.8269
F-statistic: 12.15 on 3 and 4 DF, p-value: 0.01772
```

Fare + fuel price

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_O^{W(B)} \ln O_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation C.18)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ +lnRealBusFareWLT.ts00 +
  lnRealFuelP.ts00 + L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003      2004      2005     2006      2007     2008
-0.015567 0.013673 0.013307 -0.008108 -0.022779 0.024624 -0.007941 0.002790

Coefficients:
                         Estimate      Std. Error t value Pr(>|t|)
(Intercept)               2.07442       0.66721    3.109   0.0359 *
lnRealBusFareWLT.ts00    -0.23390       0.10834   -2.159   0.0970 .
lnRealFuelP.ts00         -0.01420       0.07586   -0.187   0.8606
L(lnBusPaxPerCapWLT.ts00) 0.50784       0.18201    2.790   0.0493 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02161 on 4 degrees of freedom
Multiple R-squared: 0.9015,   Adjusted R-squared: 0.8276
F-statistic: 12.2 on 3 and 4 DF, p-value: 0.01759
```

Fare + income

$$\ln Q_t^{W(B)} = \alpha^{W(B)} + \beta_F^{W(B)} \ln F_t^{W(B)} + \beta_I^{W(B)} \ln I_t + \theta^{W(B)} \ln Q_{t-1}^{W(B)} \quad \text{(Equation C.19)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ +lnRealBusFareWLT.ts00 +
  lnRealIncomePerCap.ts00 + L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003      2004      2005     2006      2007      2008
-0.015669 0.012751 0.014323 -0.005283 -0.021782 0.025501 -0.003944 -0.005898

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                1.8461     0.8530    2.164  0.0964 .
lnRealBusFareWLT.ts00     -0.1924     0.1456   -1.322  0.2568
```

```
lnRealIncomePerCap.ts00    0.2034     0.5244     0.388  0.7179
L(lnBusPaxPerCapWLT.ts00)  0.3673     0.3637     1.010  0.3697
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02131 on 4 degrees of freedom
Multiple R-squared: 0.9042,   Adjusted R-squared: 0.8324
F-statistic: 12.59 on 3 and 4 DF, p-value: 0.01664
```

As shown above, none of the three factors, car ownership, fuel price and income, had a significant impact on bus patronage in Wellington.

Wellington rail patronage model (FY99/00–FY07/08)

Step 1: Include all variables

Service + fare + car ownership+ fuel price + income

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^{W} + \beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.20)}$$

where $t \in \left[FY99/00, FY07/08\right]$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnRealRailFareWLT.ts00 + lnCarOwnPerCapWLT.ts00 + lnRealIncomePerCap.ts00 +
  lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001      2002       2003       2004      2005       2006      2007
 0.0003839 0.0037294 -0.0053622 -0.0028431 0.0046541 -0.0030636 0.0037111
      2008
-0.0012096

Coefficients:
                          Estimate   Std. Error t value Pr(>|t|)
(Intercept)              -0.90309    1.94134   -0.465   0.723
lnRailKmPerCapWLT.ts00     0.64980    0.27183    2.390   0.252
lnRealRailFareWLT.ts00    -0.08338    0.12283   -0.679   0.620
lnCarOwnPerCapWLT.ts00    -0.13226    0.65889   -0.201   0.874
lnRealIncomePerCap.ts00   -0.15846    0.41691   -0.380   0.769
lnRealFuelP.ts00           0.13529    0.05179    2.612   0.233
                  L(lnRailPaxPerCapWLT.ts00) 0.82545  0.24418  3.381  0.183

Residual standard error: 0.009858 on 1 degrees of freedom
Multiple R-squared: 0.9913,   Adjusted R-squared: 0.9393
F-statistic: 19.05 on 6 and 1 DF, p-value: 0.1736
```

As shown above, none of the variables was significant.

Step 2: Test significance of PT related variables variables

Because none of the variables was significant in the previous step, we changed our strategy. We tested the significance of both the PT-related variables variables, ie service and fare, and their significance individually.

Service + fare

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.21)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
```

```
  lnRealRailFareWLT.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001       2002       2003       2004      2005      2006      2007
 2.399e-02 -4.522e-03 -3.445e-03 -2.267e-02 1.483e-03 6.796e-04 4.573e-03
      2008
-8.779e-05

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -0.91564    0.75419   -1.214  0.2915
lnRailKmPerCapWLT.ts00        0.96343    0.33630    2.865  0.0457 *
lnRealRailFareWLT.ts00        0.07228    0.18306    0.395  0.7131
L(lnRailPaxPerCapWLT.ts00)    0.65536    0.35601    1.841  0.1395
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01692 on 4 degrees of freedom
Multiple R-squared: 0.8978,   Adjusted R-squared: 0.8211
F-statistic: 11.71 on 3 and 4 DF, p-value: 0.01891
```

Service only

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.22)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001       2002       2003       2004       2005       2006      2007
 0.0267826 -0.0038900 -0.0070313 -0.0198016 -0.0007324 -0.0007614 0.0030580
      2008
 0.0023762

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.0071     0.6543    -1.539   0.18438
lnRailKmPerCapWLT.ts00     0.8570     0.1835     4.671   0.00548 **
L(lnRailPaxPerCapWLT.ts00) 0.7737     0.1753     4.412   0.00694 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01543 on 5 degrees of freedom
Multiple R-squared: 0.8938,   Adjusted R-squared: 0.8513
F-statistic: 21.04 on 2 and 5 DF, p-value: 0.003675
```

Fare only

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.23)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ +lnRealRailFareWLT.ts00 +
  L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001       2002       2003       2004       2005      2006       2007
 0.0321724 -0.0002565 -0.0374858 -0.0125893 -0.0115148 0.0142161 -0.0071494
      2008
 0.0226074

Coefficients:
                            Estimate Std. Error t value Pr(>|t|)
(Intercept)                 -0.7725     1.1758   -0.657  0.5403
lnRealRailFareWLT.ts00      -0.3479     0.1712   -2.033  0.0978 .
L(lnRailPaxPerCapWLT.ts00)   1.3492     0.4077    3.310  0.0213 *
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02644 on 5 degrees of freedom
Multiple R-squared: 0.6881,   Adjusted R-squared: 0.5633
F-statistic: 5.515 on 2 and 5 DF, p-value: 0.05434
```

As shown above, service had a more significant influence on rail patronage in Wellington. Fare did not appear to have a significant influence on rail patronage when both service and fare were considered but when it was considered as the only independent variable, it was just marginally significant.

Step 3: Test significance of other variables

We then tested the significance of other variables, including car ownership, fuel price and income as additional variables to service or fare.

Fare + car ownership + income + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^{W} + \beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.24)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ +lnRealRailFareWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + lnRealIncomePerCap.ts00 + lnRealFuelP.ts00 +
  L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001      2002      2003      2004      2005      2006      2007      2008
-0.001112  0.012208 -0.018899  0.007362 -0.006828  0.004764  0.004281 -0.001776

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -2.64409    3.29718   -0.802   0.5067
lnRealRailFareWLT.ts00    -0.30019    0.15177   -1.978   0.1865
lnCarOwnPerCapWLT.ts00    -0.85261    1.07359   -0.794   0.5104
lnRealIncomePerCap.ts00    0.29562    0.67998    0.435   0.7062
lnRealFuelP.ts00           0.21474    0.07279    2.950   0.0982 .
L(lnRailPaxPerCapWLT.ts00) 1.09159    0.39818    2.741   0.1113
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01806 on 2 degrees of freedom
Multiple R-squared: 0.9418,   Adjusted R-squared: 0.7962
F-statistic: 6.471 on 5 and 2 DF, p-value: 0.1393
```

Fare + car ownership

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_C^{W(R)} \ln C_t^{W} + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.25)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ +lnRealRailFareWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001      2002      2003      2004      2005      2006      2007      2008
 0.032016 -0.000667 -0.037703 -0.012283 -0.011288  0.014193 -0.007047  0.022779

Coefficients:
                            Estimate    Std. Error t value Pr(>|t|)
(Intercept)                -0.806542    2.027819  -0.398   0.711
lnRealRailFareWLT.ts00     -0.351249    0.244057  -1.439   0.223
lnCarOwnPerCapWLT.ts00     -0.009393    0.425418  -0.022   0.983
L(lnRailPaxPerCapWLT.ts00)  1.358982    0.634570   2.142   0.099 .
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02956 on 4 degrees of freedom
Multiple R-squared: 0.6881,   Adjusted R-squared: 0.4542
F-statistic: 2.942 on 3 and 4 DF, p-value: 0.1621 Fare +Income
```

Fare + income

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_I^{W(R)} \ln I_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.26)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ +lnRealRailFareWLT.ts00 +
  lnRealIncomePerCap.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001      2002      2003      2004      2005     2006      2007     2008
 0.031812 -0.000816 -0.037842 -0.012075 -0.011089 0.014121 -0.007144 0.023034

Coefficients:
                           Estimate   Std. Error t value Pr(>|t|)
(Intercept)               -0.78217    1.34013   -0.584   0.5908
lnRealRailFareWLT.ts00    -0.35369    0.24656   -1.434   0.2248
lnRealIncomePerCap.ts00   -0.01032    0.27758   -0.037   0.9721
L(lnRailPaxPerCapWLT.ts00) 1.36501    0.62260    2.192   0.0934 .
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02956 on 4 degrees of freedom
Multiple R-squared: 0.6882,   Adjusted R-squared: 0.4543
F-statistic: 2.943 on 3 and 4 DF, p-value: 0.1621
```

Fare + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_F^{W(R)} \ln F_t^{W(R)} + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.27)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ +lnRealRailFareWLT.ts00 +
  lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001     2002      2003      2004      2005     2006      2007     2008
 0.009778 0.025409 -0.016093 -0.007350 -0.014322 0.006526 -0.005417 0.001467

Coefficients:
                           Estimate   Std. Error t value Pr(>|t|)
(Intercept)               -0.05310    0.85967   -0.062   0.9537
lnRealRailFareWLT.ts00    -0.20649    0.13058   -1.581   0.1890
lnRealFuelP.ts00           0.16354    0.06421    2.547   0.0635 .
L(lnRailPaxPerCapWLT.ts00) 0.82316    0.34917    2.358   0.0779 .
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01826 on 4 degrees of freedom
Multiple R-squared: 0.881,   Adjusted R-squared: 0.7918
F-statistic: 9.873 on 3 and 4 DF, p-value: 0.02547
```

As shown above, the level of significance of the variables was not satisfactory when we considered fare with car ownership, fuel price and income as independent variables.

Service + car ownership + income + fuel price

$$\begin{aligned}\ln Q_t^{W(R)} &= \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_C^{W(R)} \ln C_t^W \\ &+\beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)}\end{aligned} \quad \text{(Equation C.28)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + lnRealIncomePerCap.ts00 + lnRealFuelP.ts00 +
  L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001      2002       2003       2004      2005       2006      2007       2008
-0.0003664 0.0034129 -0.0015615 -0.0076412 0.0066948 -0.0020126 0.0038506 -0.0023765

Coefficients:
                           Estimate   Std. Error t value Pr(>|t|)
(Intercept)               -0.42850    1.54783   -0.277   0.8079
lnRailKmPerCapWLT.ts00     0.78606    0.15667    5.017   0.0375 *
lnCarOwnPerCapWLT.ts00     0.01350    0.53238    0.025   0.9821
lnRealIncomePerCap.ts00   -0.22614    0.34597   -0.654   0.5805
lnRealFuelP.ts00           0.12080    0.04033    2.995   0.0957 .
L(lnRailPaxPerCapWLT.ts00) 0.68986    0.12004    5.747   0.0290 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008425 on 2 degrees of freedom
Multiple R-squared: 0.9873,   Adjusted R-squared: 0.9557
F-statistic: 31.18 on 5 and 2 DF, p-value: 0.03137
```

Service + car ownership + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_C^{W(R)} \ln C_t^W + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C29)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001      2002       2003       2004      2005       2006      2007   2008
 6.433e-04 2.544e-03 -1.902e-03 -7.435e-03 4.911e-03 -6.492e-05 7.036e-03  -5.733e-03

Coefficients:
                           Estimate   Std. Error t value Pr(>|t|)
(Intercept)               -1.37586    0.48855   -2.816   0.0669 .
lnRailKmPerCapWLT.ts00     0.74242    0.12748    5.824   0.0101 *
lnCarOwnPerCapWLT.ts00    -0.32464    0.11304   -2.872   0.0640 .
lnRealFuelP.ts00           0.13026    0.03386    3.848   0.0310 *
L(lnRailPaxPerCapWLT.ts00) 0.68886    0.10797    6.380   0.0078 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.007578 on 3 degrees of freedom
Multiple R-squared: 0.9846,   Adjusted R-squared: 0.9641
F-statistic: 48.04 on 4 and 3 DF, p-value: 0.004721
```

Service + income + fuel price

$$\ln Q_t^{W(R)} = \alpha^{W(R)} + \beta_S^{W(R)} \ln S_t^{W(R)} + \beta_I^{W(R)} \ln I_t + \beta_O^{W(R)} \ln O_t + \theta^{W(R)} \ln Q_{t-1}^{W(R)} \quad \text{(Equation C.30)}$$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008
```

```
Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnRealIncomePerCap.ts00 + lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001      2002      2003      2004      2005      2006      2007      2008
-0.000359 0.003357 -0.001572 -0.007615 0.006651 -0.001940 0.003984 -0.002506

Coefficients:
                           Estimate   Std. Error t value Pr(>|t|)
(Intercept)               -0.46634    0.33435   -1.395  0.25741
lnRailKmPerCapWLT.ts00     0.78471    0.12032    6.522  0.00732 **
lnRealIncomePerCap.ts00   -0.21761    0.06669   -3.263  0.04703 *
lnRealFuelP.ts00           0.12123    0.02987    4.058  0.02697 *
L(lnRailPaxPerCapWLT.ts00) 0.69004    0.09787    7.050  0.00587 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.00688 on 3 degrees of freedom
Multiple R-squared: 0.9873,   Adjusted R-squared: 0.9704
F-statistic: 58.44 on 4 and 3 DF, p-value: 0.003539
```

As shown above, two models were identified with satisfactory performance as follows:

- service + car ownership + fuel price
- service + income + fuel price.

Christchurch fitted bus patronage forecast model (1998Q1–2008Q2)

Step 1: Full period, all variables, all indicators

We first started with all the variables for the full period, ie 1996 Q1 to 2008 Q2, plus a dummy variable (an indicator) for the observed change of trend in 1998 Q1, D_t^{98Q1} as a result of the introduction of gold coin cash fare.

$$D_t^{98Q1} = \begin{cases} 1 & \text{if } t \text{ is in or after 1998 Q1} \\ 0 & \text{otherwise} \end{cases}$$

Service + car ownership+ fuel price + income + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t} + \gamma^{98Q1} D_t^{98Q1} \end{aligned} \quad \text{(Equation C.31)}$$

where $t \in [1997Q1, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1997(1), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts96 ~ lnBusKmPerCapCHC.ts96 +
  lnCarOwnPerCapCHC.ts96 + lnRealIncomePerCap.ts96 + lnRealFuelP.ts96 +
  IndQtr2.ts96 + IndQtr3.ts96 + IndQtr4.ts96 + L(lnBusPaxPerCapCHC.ts96) +
  IndGCoin.ts96)

Residuals:
      Min        1Q    Median        3Q      Max
-0.036596 -0.013401 -0.002787 0.013599 0.054207

Coefficients:
                             Estimate   Std. Error t value  Pr(>|t|)
(Intercept)                 -1.028549   0.597780  -1.721   0.0939 .
lnBusKmPerCapCHC.ts96        0.070676   0.043493   1.625   0.1129
lnCarOwnPerCapCHC.ts96      -0.433320   0.265623  -1.631   0.1115
lnRealIncomePerCap.ts96      0.571411   0.274154   2.084   0.0443 *
lnRealFuelP.ts96            -0.019699   0.044633  -0.441   0.6616
IndQtr2.ts96                 0.004165   0.009913   0.420   0.6769
IndQtr3.ts96                 0.012110   0.010504   1.153   0.2566
IndQtr4.ts96                -0.075189   0.010999  -6.836   5.39e-08 ***
L(lnBusPaxPerCapCHC.ts96)    0.815036   0.072303  11.273   2.30e-13 ***
```

```
IndGCoin.ts96                0.022024   0.010262   2.146   0.0387 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02277 on 36 degrees of freedom
Multiple R-squared: 0.9899,   Adjusted R-squared: 0.9873
F-statistic: 390.5 on 9 and 36 DF, p-value: < 2.2e-16
```

As shown above, the seasonal effect of Q4 and the effect of the introduction of gold coin cash fare were both significant.

Step 2: Period changed, take out insignificant dummy variables

As the effect of gold coin cash fare was found to be significant, we changed the model period to 1998 Q1 to 2008 Q2. The dummy variables for seasonal effect of Q2 and Q3 were also excluded. The output from R is as follows:

Service + car ownership+ fuel price + income + Q4 seasonal dummy + break

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \qquad \text{(Equation C.32)}$$

where $t \in \left[1998Q1, 2008Q2\right]$

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  lnCarOwnPerCapCHC.ts98 + lnRealIncomePerCap.ts98 + lnRealFuelP.ts98 +
  IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
      Min        1Q   Median       3Q      Max
-0.042880 -0.015833 0.001684 0.017431 0.046459

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                -0.88306    0.60191   -1.467  0.1515
lnBusKmPerCapCHC.ts98       0.07985    0.03938    2.028  0.0505 .
lnCarOwnPerCapCHC.ts98     -0.55937    0.31251   -1.790  0.0824 .
lnRealIncomePerCap.ts98     0.44363    0.27056    1.640  0.1103
lnRealFuelP.ts98           -0.01878    0.04376   -0.429  0.6705
IndQtr4.ts98               -0.07886    0.01053   -7.491  1.08e-08 ***
L(lnBusPaxPerCapCHC.ts98)   0.82940    0.06928   11.971  9.62e-14 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02278 on 34 degrees of freedom
Multiple R-squared: 0.9875,   Adjusted R-squared: 0.9853
F-statistic:  447 on 6 and 34 DF, p-value: < 2.2e-16
```

As shown above, the influence of service and car ownership was found to be significant.

Step 3: Take out insignificant variables, try different combinations of significant variables

As service was a significant variable, we tested combinations of service with other variables, including car ownership, income and fuel. We also tested service as the only independent variable.

Service + car ownership+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \qquad \text{(Equation C.33)}$$

where $t \in \left[1998Q1, 2008Q2\right]$

```
Time series regression with 'ts' data:
```

```
Start = 1998(2), End = 2008(2)
Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  lnCarOwnPerCapCHC.ts98 + IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
       Min         1Q    Median        3Q       Max
-0.0433897 -0.0149798 0.0005365 0.0192871 0.0495277

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -0.042894  0.160402 -0.267  0.7907
lnBusKmPerCapCHC.ts98         0.083427  0.035371  2.359  0.0239 *
lnCarOwnPerCapCHC.ts98       -0.181363  0.179336 -1.011  0.3186
IndQtr4.ts98       -0.087728  0.008802 -9.967 6.79e-12 ***
L(lnBusPaxPerCapCHC.ts98)     0.904029  0.049012 18.445 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.023 on 36 degrees of freedom
Multiple R-squared: 0.9865,   Adjusted R-squared: 0.985
```

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_I^{C(B)} \ln I_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation C.34)

```
F-statistic: 657.3 on 4 and 36 DF, p-value: < 2.2e-16
```

Service + income+ Q4 seasonal dummy

where $t \in [1998Q1, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  lnRealIncomePerCap.ts98 + IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
       Min         1Q    Median        3Q       Max
-0.0465878 -0.0158942 0.0002645 0.0161129 0.0548082

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.099416  0.208028  0.478  0.6356
lnBusKmPerCapCHC.ts98     0.069768  0.036495  1.912  0.0639 .
lnRealIncomePerCap.ts98   0.011478  0.153920  0.075  0.9410
IndQtr4.ts98             -0.089677  0.008851 -10.132 4.37e-12 ***
L(lnBusPaxPerCapCHC.ts98) 0.883550  0.059629 14.817 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02332 on 36 degrees of freedom
Multiple R-squared: 0.9861,   Adjusted R-squared: 0.9846
F-statistic:  639 on 4 and 36 DF, p-value: < 2.2e-16
```

Service + fuel+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$ (Equation C.35)

where $t \in [1998Q1, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  lnRealFuelP.ts98 + IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
       Min         1Q    Median        3Q       Max
-0.0466109 -0.0150181 0.0007507 0.0180222 0.0524055

Coefficients:
```

```
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                0.195377   0.154231   1.267  0.2134
lnBusKmPerCapCHC.ts98      0.082779   0.040079   2.065  0.0461 *
lnRealFuelP.ts98          -0.019608   0.036276  -0.541  0.5922
IndQtr4.ts98              -0.090299   0.008692 -10.389  2.22e-12 ***
L(lnBusPaxPerCapCHC.ts98) 0.882528   0.046778  18.866 <2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02323 on 36 degrees of freedom
Multiple R-squared: 0.9862,   Adjusted R-squared: 0.9847
F-statistic: 644.2 on 4 and 36 DF, p-value: < 2.2e-16
```

Service only+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \quad \text{(Equation C.36)}$$

where $t \in \left[1998Q1, 2008Q2\right]$

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
       Min         1Q    Median        3Q       Max
-0.0468237 -0.0161091 0.0002652 0.0162674 0.0545626

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                0.114658   0.038194   3.002 0.00478 **
lnBusKmPerCapCHC.ts98      0.070835   0.033117   2.139 0.03911 *
IndQtr4.ts98              -0.089807   0.008561 -10.490 1.23e-12 ***
L(lnBusPaxPerCapCHC.ts98) 0.886340   0.045799  19.353 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.023 on 37 degrees of freedom
Multiple R-squared: 0.9861,   Adjusted R-squared: 0.985
F-statistic: 875.5 on 3 and 37 DF, p-value: < 2.2e-16
```

As shown above, service was the only variable with significant influence in Christchurch. The only satisfactory model has service as the only independent variable with a high R-square value of 0.9861.

Christchurch fitted bus patronage forecast model (2004Q1–2008Q2)

Step: 1 Full period, all variables, all indicators

To test the influence of the introduction of the Metrocard in July 2003, we introduced a new indicator (dummy) variable D_t^{03Q3} as follows:

$$D_t^{03Q3} = \begin{cases} 1 & \text{if } t \text{ is in or after 2003 Q3} \\ 0 & \text{otherwise} \end{cases}$$

Service + car ownership+ fuel price + income + seasonal dummies + break

$$\begin{aligned} \ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t \\ &+ \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t} + \gamma^{03Q3} D_t^{03Q3} \end{aligned} \quad \text{(Equation C.37)}$$

where $t \in \left[1998Q1, 2008Q2\right]$

The output from R is as follows:

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)
```

```
Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  lnCarOwnPerCapCHC.ts98 + lnRealIncomePerCap.ts98 + lnRealFuelP.ts98 +
  IndQtr2.ts98 + IndQtr3.ts98 + IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98) +
  IndTKTG.ts98)

Residuals:
      Min        1Q   Median       3Q      Max
-0.039283 -0.012639 0.003184 0.009165 0.033768

Coefficients:
                            Estimate  Std. Error t value  Pr(>|t|)
(Intercept)                -0.761162    0.568332  -1.339  0.19021
lnBusKmPerCapCHC.ts98       0.031606    0.038087   0.830  0.41298
lnCarOwnPerCapCHC.ts98      0.133602    0.354918   0.376  0.70916
lnRealIncomePerCap.ts98     0.712856    0.239590   2.975  0.00563 **
lnRealFuelP.ts98           -0.038525    0.041669  -0.925  0.36234
IndQtr2.ts98                0.003494    0.010175   0.343  0.73364
IndQtr3.ts98                0.020703    0.011339   1.826  0.07752 .
IndQtr4.ts98               -0.081135    0.009796  -8.283 2.35e-09 ***
L(lnBusPaxPerCapCHC.ts98)   0.805004    0.062588  12.862 5.74e-14 ***
IndTKTG.ts98               -0.063930    0.016566  -3.859  0.00054 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01918 on 31 degrees of freedom
Multiple R-squared: 0.9919,   Adjusted R-squared: 0.9896
F-statistic: 422.3 on 9 and 31 DF, p-value: < 2.2e-16
```

As shown above, the introduction of the Metrocard had significant impact on bus patronage.

Step 2: Period changed, take out insignificant dummy variables

Since the impact of the introduction of the Metrocard was significant, theoretically we had to build a model based on the data from 2003 Q3 to 2008 Q2. However, bus fare information was available only from 2004 Q1. Therefore, we could only build a model for the period 2004 Q1 to 2008 Q2. Again, we started with all the independent variables and seasonal dummy variables included for this period.

Service + fare + car ownership+ fuel price + income + seasonal dummies

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} \\ &+\beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} \\ &+\gamma_2^{C(B)} D_{2t} + \gamma_3^{C(B)} D_{3t} + \gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation C.38)}$$

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnBusKmPerCapCHC.ts04 +
  lnRealBusFareCHC.ts04 + lnCarOwnPerCapCHC.ts04 + lnRealIncomePerCap.ts04 +
  lnRealFuelP.ts04 + IndQtr2.ts04 + IndQtr3.ts04 + IndQtr4.ts04 +
  L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q    Median       3Q      Max
-0.016225 -0.006346 -0.001301 0.006723 0.018690

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                   1.64484    1.05003    1.566  0.1683
lnBusKmPerCapCHC.ts04         0.01970    0.16291    0.121  0.9077
lnRealBusFareCHC.ts04        -0.39468    0.27862   -1.417  0.2064
lnCarOwnPerCapCHC.ts04        2.03286    0.99925    2.034  0.0881 .
lnRealIncomePerCap.ts04       0.21895    0.25391    0.862  0.4216
lnRealFuelP.ts04              0.26041    0.09894    2.632  0.0390 *
```

```
IndQtr2.ts04                  0.02630    0.02176    1.209  0.2722
IndQtr3.ts04                  0.03578    0.02146    1.667  0.1465
IndQtr4.ts04                 -0.04442    0.01635   -2.716  0.0348 *
L(lnBusPaxPerCapCHC.ts04)   -0.16681    0.34947   -0.477  0.6500
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01443 on 6 degrees of freedom
Multiple R-squared: 0.9477,   Adjusted R-squared: 0.8693
F-statistic: 12.08 on 9 and 6 DF, p-value: 0.003336
```

As shown above, only two variables were significant, fuel price and seasonal effect of Q4. We excluded seasonal dummy variables for Q2 and Q3 and tested again with all five independent variables included.

Service + fare + car ownership+ fuel price + income + Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{C(B)} &= \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^C \\ &+\beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} \\ &+\gamma_4^{C(B)} D_{4t}\end{aligned} \quad \text{(Equation C.39)}$$

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnBusKmPerCapCHC.ts04 +
  lnRealBusFareCHC.ts04 + lnCarOwnPerCapCHC.ts04 + lnRealIncomePerCap.ts04 +
  lnRealFuelP.ts04 + IndQtr4.ts04 + L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q   Median       3Q      Max
-0.020162 -0.008118 0.002287 0.008171 0.015772

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                 0.67574    0.82764    0.816  0.4379
lnBusKmPerCapCHC.ts04       0.12528    0.14732    0.850  0.4198
lnRealBusFareCHC.ts04      -0.55192    0.26593   -2.075  0.0716 .
lnCarOwnPerCapCHC.ts04      0.88361    0.64809    1.363  0.2099
lnRealIncomePerCap.ts04     0.25190    0.24890    1.012  0.3412
lnRealFuelP.ts04            0.24962    0.10188    2.450  0.0399 *
IndQtr4.ts04               -0.05200    0.01655   -3.142  0.0138 *
L(lnBusPaxPerCapCHC.ts04) -0.07049    0.36651   -0.192  0.8523
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01534 on 8 degrees of freedom
Multiple R-squared: 0.9212,   Adjusted R-squared: 0.8523
F-statistic: 13.36 on 7 and 8 DF, p-value: 0.0007689
```

As shown above, now we had two variables with significant effect, fare and fuel price.

Fare + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \quad \text{(Equation C.40)}$$

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnRealBusFareCHC.ts04 +
  lnRealFuelP.ts04 + IndQtr4.ts04 + L(lnBusPaxPerCapCHC.ts04))

Residuals:
```

```
      Min        1Q   Median       3Q      Max
-0.026842 -0.013884 0.001090 0.014043 0.017157

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                 0.17309    0.29959    0.578  0.5751
lnRealBusFareCHC.ts04      -0.26196    0.08780   -2.984  0.0124 *
lnRealFuelP.ts04            0.27846    0.08637    3.224  0.0081 **
IndQtr4.ts04               -0.04290    0.01722   -2.491  0.0300 *
L(lnBusPaxPerCapCHC.ts04)   0.23730    0.28728    0.826  0.4264
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0167 on 11 degrees of freedom
Multiple R-squared: 0.8717,   Adjusted R-squared: 0.8251
F-statistic: 18.69 on 4 and 11 DF, p-value: 7.211e-05
```

As shown above, with fare and fuel as independent variables, the R-square value was 0.8717 which was not ideal. We tried to include additional variables with a view to improving the model performance.

Step 3: Try different combinations of variables

Service + fare + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$

(Equation C.41)

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnBusKmPerCapCHC.ts04 +
  lnRealBusFareCHC.ts04 + lnRealFuelP.ts04 + IndQtr4.ts04 +
  L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q   Median       3Q      Max
-0.027188 -0.011340 0.001618 0.013866 0.017123

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                0.26393    0.38326    0.689  0.5067
lnBusKmPerCapCHC.ts04      0.06582    0.16164    0.407  0.6924
lnRealBusFareCHC.ts04     -0.36115    0.26015   -1.388  0.1952
lnRealFuelP.ts04           0.26500    0.09573    2.768  0.0198 *
IndQtr4.ts04              -0.04282    0.01792   -2.390  0.0380 *
L(lnBusPaxPerCapCHC.ts04) 0.16642    0.34584    0.481  0.6407
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01737 on 10 degrees of freedom
Multiple R-squared: 0.8738,   Adjusted R-squared: 0.8107
F-statistic: 13.85 on 5 and 10 DF, p-value: 0.0003178
```

Fare + car ownership + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t}$$

(Equation C.42)

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnRealBusFareCHC.ts04 +
```

```
  lnCarOwnPerCapCHC.ts04 + lnRealFuelP.ts04 + IndQtr4.ts04 +
  L(lnBusPaxPerCapCHC.ts04))

Residuals:
       Min         1Q     Median        3Q       Max
-0.0181076 -0.0115086 -0.0001819 0.0101889 0.0223514

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.960507   0.517327   1.857 0.09302 .
lnRealBusFareCHC.ts04    -0.384909   0.105450  -3.650 0.00446 **
lnCarOwnPerCapCHC.ts04    1.022532   0.570365   1.793 0.10326
lnRealFuelP.ts04          0.325077   0.082979   3.918 0.00288 **
IndQtr4.ts04             -0.051338   0.016404  -3.130 0.01070 *
L(lnBusPaxPerCapCHC.ts04) 0.000516   0.293508   0.002 0.99863
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01523 on 10 degrees of freedom
Multiple R-squared: 0.9029,   Adjusted R-squared: 0.8544
F-statistic: 18.6 on 5 and 10 DF, p-value: 8.915e-05
```

Service + fare + car ownership + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_C^{C(B)} \ln C_t^{C} + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \qquad \text{(Equation C.43)}$$

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnBusKmPerCapCHC.ts04 +
  lnRealBusFareCHC.ts04 + lnCarOwnPerCapCHC.ts04 + lnRealFuelP.ts04 +
  IndQtr4.ts04 + L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q   Median       3Q      Max
-0.021581 -0.010190 0.001160 0.010231 0.019009

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                1.24488    0.60803    2.047 0.07091 .
lnBusKmPerCapCHC.ts04      0.13415    0.14725    0.911 0.38605
lnRealBusFareCHC.ts04     -0.60257    0.26153   -2.304 0.04669 *
lnCarOwnPerCapCHC.ts04     1.15138    0.59241    1.944 0.08383 .
lnRealFuelP.ts04           0.30352    0.08697    3.490 0.00683 **
IndQtr4.ts04              -0.05223    0.01657   -3.151 0.01171 *
L(lnBusPaxPerCapCHC.ts04) -0.17379    0.35249   -0.493 0.63379
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01537 on 9 degrees of freedom
Multiple R-squared: 0.9111,   Adjusted R-squared: 0.8519
F-statistic: 15.38 on 6 and 9 DF, p-value: 0.0002861
```

As shown above the estimated coefficient for the lagged variable is -0.17379 which is negative. This model is not acceptable.

Service + fare +income + fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = \alpha^{C(B)} + \beta_S^{C(B)} \ln S_t^{C(B)} + \beta_F^{C(B)} \ln F_t^{C(B)} + \beta_I^{C(B)} \ln I_t + \beta_O^{C(B)} \ln O_t + \theta^{C(B)} \ln Q_{t-1}^{C(B)} + \gamma_4^{C(B)} D_{4t} \qquad \text{(Equation C.44)}$$

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
```

```
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnBusKmPerCapCHC.ts04 + 
  lnRealBusFareCHC.ts04 + lnRealIncomePerCap.ts04 + lnRealFuelP.ts04 + 
  IndQtr4.ts04 + L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q    Median        3Q       Max 
-0.028112 -0.007045 0.004977 0.008737 0.013829 

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)  
(Intercept)              -0.26464    0.47879   -0.553  0.5939  
lnBusKmPerCapCHC.ts04     0.07671    0.14961    0.513  0.6205  
lnRealBusFareCHC.ts04    -0.36968    0.24061   -1.536  0.1588  
lnRealIncomePerCap.ts04   0.39044    0.23781    1.642  0.1350  
lnRealFuelP.ts04          0.19535    0.09816    1.990  0.0778 .
IndQtr4.ts04             -0.04586    0.01667   -2.751  0.0224 *
L(lnBusPaxPerCapCHC.ts04) 0.20389    0.32060    0.636  0.5406  
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.01606 on 9 degrees of freedom
Multiple R-squared: 0.9029,   Adjusted R-squared: 0.8382 
F-statistic: 13.95 on 6 and 9 DF, p-value: 0.0004199
```

Appendix D: Dynamic linear regression results – recommended forecast models

Auckland bus patronage model (2003Q3–2008Q2)

Service + car ownership + fuel price+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(B)} &= -1.9306 + 0.4603 \ln S_t^{A(B)} - 1.9618 \ln C_t^{A} + 0.2040 \ln O_t \\ &+0.3677 \ln Q_{t-1}^{A(B)} - 0.0478 D_{4t}\end{aligned} \quad \text{(Equation D.1)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapAKL.ts03 ~ lnBusKmPerCapAKL.ts03 +
  lnCarOwnPerCapAKL.ts03 + lnRealFuelP.ts03 + IndQtr4.ts03 +
  L(lnBusPaxPerCapAKL.ts03))

Residuals:
      Min        1Q   Median       3Q      Max
-0.027448 -0.008362 0.001468 0.006955 0.032450

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.93063    0.45113   -4.280 0.000897 ***
lnBusKmPerCapAKL.ts03      0.46028    0.08629    5.334 0.000136 ***
lnCarOwnPerCapAKL.ts03    -1.96184    0.57928   -3.387 0.004865 **
lnRealFuelP.ts03           0.20400    0.06155    3.314 0.005592 **
IndQtr4.ts03              -0.04783    0.01299   -3.682 0.002763 **
L(lnBusPaxPerCapAKL.ts03) 0.36767    0.08198    4.485 0.000614 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01589 on 13 degrees of freedom
Multiple R-squared: 0.9829,   Adjusted R-squared: 0.9763
F-statistic: 149.6 on 5 and 13 DF, p-value: 5.145e-11
```

Auckland rail patronage forecast model

Model 1 – service + fare + income+ Q4 seasonal dummy

$$\begin{aligned}\ln Q_t^{A(R)} &= -1.1448 + 0.9946 \ln S_t^{A(R)} - 0.9672 \ln F_t^{A(R)} + 1.606 \ln I_t \\ &+0.2957 \ln Q_{t-1}^{A(R)} - 0.1325 D_{4t}\end{aligned} \quad \text{(Equation D.2)}$$

where $t \in [2003Q3, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + lnRealIncomePerCap.ts03 + IndQtr4.ts03 +
  L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q   Median       3Q      Max
-0.063648 -0.026550 0.006022 0.020504 0.048103

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              -1.14479    1.06804   -1.072 0.303277
lnRailKmPerCapAKL.ts03     0.99461    0.23105    4.305 0.000856 ***
```

```
lnRealRailFareAKL.ts03    -0.96723    0.22613   -4.277 0.000900 ***
lnRealIncomePerCap.ts03    1.60601    0.56710    2.832 0.014138 *
IndQtr4.ts03              -0.13253    0.02734   -4.848 0.000318 ***
L(lnRailPaxPerCapAKL.ts03) 0.29570    0.15344    1.927 0.076105 .
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03609 on 13 degrees of freedom
Multiple R-squared: 0.9867,   Adjusted R-squared: 0.9816
F-statistic: 193.2 on 5 and 13 DF, p-value: 1.007e-11
```

Model 2 – service + fare+ Q4 seasonal dummy

$$\ln Q_t^{A(R)} = 1.6913 + 0.8827 \ln S_t^{A(R)} - 0.6755 \ln F_t^{A(R)} + 0.4584 \ln Q_{t-1}^{A(R)} - 0.1478 D_{4t} \qquad \text{(Equation D.3)}$$

where $t \in \left[2003Q3, 2008Q2\right]$

```
Time series regression with 'ts' data:
Start = 2003(4), End = 2008(2)

Call:
dynlm(formula = lnRailPaxPerCapAKL.ts03 ~ lnRailKmPerCapAKL.ts03 +
  lnRealRailFareAKL.ts03 + IndQtr4.ts03 + L(lnRailPaxPerCapAKL.ts03))

Residuals:
      Min        1Q    Median       3Q      Max
-0.056704 -0.023389 -0.004208 0.004447 0.082838

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)                1.69131    0.45480    3.719 0.002290 **
lnRailKmPerCapAKL.ts03     0.88265    0.27893    3.164 0.006892 **
lnRealRailFareAKL.ts03    -0.67551    0.24667   -2.739 0.016001 *
IndQtr4.ts03              -0.14775    0.03285   -4.498 0.000501 ***
L(lnRailPaxPerCapAKL.ts03) 0.45840    0.17434    2.629 0.019811 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Wellington bus patronage forecast model

Fare only

$$\ln Q_t^{W(B)} = 2.0590 - 0.2312 \ln F_t^{W(B)} + 0.4934 \ln Q_{t-1}^{W(B)} \qquad \text{(Equation D.4)}$$

where $t \in \left[FY99/00, FY07/08\right]$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnBusPaxPerCapWLT.ts00 ~ +lnRealBusFareWLT.ts00 +
  L(lnBusPaxPerCapWLT.ts00))

Residuals:
     2001     2002     2003      2004      2005     2006      2007     2008
-0.017702 0.014627 0.014371 -0.006231 -0.021781 0.023986 -0.007862 0.000592

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               2.05903    0.59481    3.462  0.0180 *
lnRealBusFareWLT.ts00    -0.23115    0.09643   -2.397  0.0618 .
L(lnBusPaxPerCapWLT.ts00) 0.49344    0.14816    3.330  0.0208 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01941 on 5 degrees of freedom
Multiple R-squared: 0.9006,   Adjusted R-squared: 0.8609
F-statistic: 22.65 on 2 and 5 DF, p-value: 0.003114
```

Wellington rail patronage forecast model

Model 1 – service + car ownership + fuel price

$$\ln Q_t^{W(R)} = -1.3759 + 0.7424 \ln S_t^{W(R)} - 0.3246 \ln C_t^{W} + 0.1303 \ln O_t + 0.6889 \ln Q_{t-1}^{W(R)} \qquad \text{(Equation D.5)}$$

where $t \in \left[FY99/00, FY07/08\right]$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnCarOwnPerCapWLT.ts00 + lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
      2001      2002       2003       2004      2005       2006      2007        2008
 6.433e-04 2.544e-03 -1.902e-03 -7.435e-03 4.911e-03 -6.492e-05 7.036e-03  -5.733e-03

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -1.37586    0.48855   -2.816  0.0669 .
lnRailKmPerCapWLT.ts00     0.74242    0.12748    5.824  0.0101 *
lnCarOwnPerCapWLT.ts00    -0.32464    0.11304   -2.872  0.0640 .
lnRealFuelP.ts00           0.13026    0.03386    3.848  0.0310 *
L(lnRailPaxPerCapWLT.ts00) 0.68886    0.10797    6.380  0.0078 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.007578 on 3 degrees of freedom
Multiple R-squared: 0.9846,   Adjusted R-squared: 0.9641
F-statistic: 48.04 on 4 and 3 DF, p-value: 0.004721
```

Model 2 – service + income + fuel price

$$\ln Q_t^{W(R)} = -0.4663 + 0.7847 \ln S_t^{W(R)} - 0.2176 \ln I_t + 0.1212 \ln O_t + 0.6900 \ln Q_{t-1}^{W(R)} \qquad \text{(Equation D.6)}$$

where $t \in \left[FY99/00, FY07/08\right]$

```
Time series regression with 'ts' data:
Start = 2001, End = 2008

Call:
dynlm(formula = lnRailPaxPerCapWLT.ts00 ~ lnRailKmPerCapWLT.ts00 +
  lnRealIncomePerCap.ts00 + lnRealFuelP.ts00 + L(lnRailPaxPerCapWLT.ts00))

Residuals:
     2001     2002      2003      2004     2005      2006     2007      2008
-0.000359 0.003357 -0.001572 -0.007615 0.006651 -0.001940 0.003984 -0.002506

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)               -0.46634  0.33435   -1.395   0.25741
lnRailKmPerCapWLT.ts00     0.78471  0.12032    6.522   0.00732 **
lnRealIncomePerCap.ts00   -0.21761  0.06669   -3.263   0.04703 *
lnRealFuelP.ts00           0.12123  0.02987    4.058   0.02697 *
L(lnRailPaxPerCapWLT.ts00) 0.69004  0.09787    7.050   0.00587 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.00688 on 3 degrees of freedom
Multiple R-squared: 0.9873,   Adjusted R-squared: 0.9704
F-statistic: 58.44 on 4 and 3 DF, p-value: 0.003539
```

Christchurch fitted bus patronage forecast model (1998Q1–2008Q2)

Service only+ Q4 seasonal dummy

$$\ln Q_t^{C(B)} = 0.1147 + 0.0708 \ln S_t^{C(B)} - 0.0898 \ln Q_{t-1}^{C(B)} + 0.8863 D_{4t}$$ (Equation D.7)

where $t \in [1998Q1, 2008Q2]$

```
Time series regression with 'ts' data:
Start = 1998(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts98 ~ lnBusKmPerCapCHC.ts98 +
  IndQtr4.ts98 + L(lnBusPaxPerCapCHC.ts98))

Residuals:
       Min         1Q    Median        3Q       Max
-0.0468237 -0.0161091 0.0002652 0.0162674 0.0545626

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.114658   0.038194   3.002 0.00478 **
lnBusKmPerCapCHC.ts98     0.070835   0.033117   2.139 0.03911 *
IndQtr4.ts98             -0.089807   0.008561  -10.490 1.23e-12 ***
L(lnBusPaxPerCapCHC.ts98) 0.886340   0.045799  19.353 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.023 on 37 degrees of freedom
Multiple R-squared: 0.9861,   Adjusted R-squared: 0.985
F-statistic: 875.5 on 3 and 37 DF, p-value: < 2.2e-16
```

Christchurch fitted bus patronage forecast model (2004Q1–2008Q2)

Model 1 – Fare fuel price + Q4 seasonal dummy

$$\ln Q_t^{C(B)} = 0.1731 - 0.2620 \ln F_t^{C(B)} + 0.2785 \ln O_t + 0.2373 \ln Q_{t-1}^{C(B)} - 0.0429 D_{4t}$$ (Equation D.8)

where $t \in [2004Q1, 2008Q2]$

```
Time series regression with 'zooreg' data:
Start = 2004(2), End = 2008(2)

Call:
dynlm(formula = lnBusPaxPerCapCHC.ts04 ~ lnRealBusFareCHC.ts04 +
  lnRealFuelP.ts04 + IndQtr4.ts04 + L(lnBusPaxPerCapCHC.ts04))

Residuals:
      Min        1Q   Median       3Q      Max
-0.026842 -0.013884 0.001090 0.014043 0.017157

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               0.17309    0.29959    0.578  0.5751
lnRealBusFareCHC.ts04    -0.26196    0.08780   -2.984  0.0124 *
lnRealFuelP.ts04          0.27846    0.08637    3.224  0.0081 **
IndQtr4.ts04             -0.04290    0.01722   -2.491  0.0300 *
L(lnBusPaxPerCapCHC.ts04) 0.23730    0.28728    0.826  0.4264
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0167 on 11 degrees of freedom
Multiple R-squared: 0.8717,   Adjusted R-squared: 0.8251
F-statistic: 18.69 on 4 and 11 DF, p-value: 7.211e-05
```

Appendix E: Referee reports on technical issues and responses

Appendix E.1: Referee report from Andrew Coleman, Motu Economics, 18 November 2009

Comment: Appraisal of factors influencing public transport patronage.

This comment is an econometric assessment of the paper 'Appraisal of factors influencing public transport patronage', written for the NZ Transport Agency. The paper is intended to investigate the major factors that influence public transport patronage in Auckland, Wellington and Christchurch. The paper develops a model of patronage and estimates a series of regressions that allow the estimation of demand elasticities for bus and rail services.

The focus of this comment is these regressions, starting with the econometric equations describing public transport demand in Auckland. The lengthiest comment, that outlines many of the reoccurring issues, concerns Auckland buses. Additional issues concerning the equations for Wellington and Christchurch are discussed at the end.

Overall evaluation

This study is ambitious. It aims to find causal links between a variety of factors including infrastructure development, public transport service levels, fares, motoring costs, and income on public transport patronage (section 4.1). In attempting to do this, however, it faces considerable difficulties.

The first difficulty is the problem of identifying which changes in the level of public transport services are caused by exogenous factors, and which reflect decisions to increase service provision because of changes in demand or other factors. In my view, insufficient attention is directed to this problem, and until it is addressed doubts should remain about the meaning of the estimated coefficients. In particular, it is plausible that many of the elasticity estimates are substantially biased because they do not adequately take into account the possibility that service levels are determined endogenously.

Second, the sample periods used in the paper are short. The problem is most acute for Wellington, but it is serious for all of the regions, particularly as the authors restrict much of the analysis to sub-samples comprising fewer than five years. While it is sometimes possible to estimate parameters accurately using small amounts of data, the serial correlation in the data used in this study make it unlikely in this case. In addition, the use of dummy variables to model structural breaks or the estimation of models over length-restricted subsamples mean that authors have not really done enough to install confidence that the estimated coefficients are likely to be representative of the coefficients that would be estimated on longer samples, should they become available.

The authors can do little about data availability. However, it would be useful if they addressed more clearly the potential problems that can arise from the use of variables that describe public transport service provision. They need to establish whether fluctuations in these variables are largely caused by exogenous factors (such as a government decision to make large infrastructure investments) or whether they are a response to endogenous factors such as demand by patrons, whether or not the latter is caused by measured variables (such as the price of fuel) or other factors. If they can demonstrate that service provision is largely exogenous, the paper will be much improved, in part because they should then be able to use the full sample periods rather than the restricted samples whose results they emphasise. At the moment, however, doubts must remain as to how their estimated elasticities should be interpreted, as it is not clear whether the reported elasticities make proper allowances for the way independent factors such as fuel prices affect final demand indirectly through changes in the level of services provided by transport operators.

The authors should provide more details about the serial correlation structure of the estimated residuals. Readers need to be convinced that the residuals are serially uncorrelated, and thus that the estimated standard errors are appropriately calculated. This is particularly important given the short samples, and the widespread inclusion of lagged dependent variables. In addition, the use of 'stepwise-regression' methodology to eliminate potentially insignificant variables will be flawed if the standard errors of the coefficient estimates are not properly adjusted for the serial correlation in the error structure.

Overall, I am unsure what weight to place on their estimated elasticities. I find it difficult to take the Wellington estimates seriously because the sample periods are so short – indeed, most of the preferred estimates for Auckland and Christchurch as well as Wellington are based on very short samples. On the whole, I am unconvinced about the accuracy of the estimates for variables such as income, fares and the fuel price, because of concern over the exogeneity of the service provision variable. The estimates I have most sympathy for are the estimates concerning rail passengers in Auckland, even these are very large. I like these estimates because it seems likely that the public transport service provision variable for Auckland rail largely reflects exogenous factors.

In terms of their discussion and conclusions (section 8), I have sympathy with their complaint that the study was limited by data availability, and that further updates should be attempted as more data becomes available. Econometrically, any future studies need to be focused on the extent to which public transport service provision is exogenously or endogenously determined, however, and there should be considerable attention paid to creating variables that measure the exogenous components of public transport provision as a means of identifying other parameters.

Auckland

The analysis for Auckland is the strongest. It comprises an analysis of rail and bus services, using quarterly patronage data between 1996 and 2008.

Auckland buses

The core regression for Auckland bus [A(B)] service is (section 6.3.1.1, equation 6.9)

$$\begin{aligned}\ln Q^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln S_t^{A(B)} + \beta_F^{A(B)} \ln F_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A(B)} \\ &+ \beta_I^{A(B)} \ln I_t^{A(B)} + \beta_o^{A(B)} \ln O_t^{A(B)} + \theta^{A(B)} \ln Q_{t-1}^{A(B)} \\ &+ \gamma_2^{A(B)} D_{2t} + \gamma_3^{A(B)} D_{3t} + \gamma_4^{A(B)} D_{4t} + \gamma^{99Q3} D_t^{99Q3}\end{aligned}$$

where

Q is the number of passenger trips per period (per capita)

S is the km of services of transport offered (per capita)

F is the price of the fare

C is the number of cars per capita

I is real income per capita

O is the oil price per capita.

D are dummies representing seasonality or the post-1999 period.

The regression model was estimated, and then variables whose coefficients were statistically insignificant were dropped in a stepwise procedure. Rather than include a dummy variable for the post-1999 period, the regression was shortened to only include post-1999 or post-2003 variables. The final regression was (section 6.3.1.1, equation 6.17):

$$\ln Q_t^{A(B)} = -1.93 + 0.46 \ln S_t^{A(B)} - 1.96 \ln C_t^{A(B)} + 0.20 \ln O_t^{A(B)} + 0.37 \ln Q_{t-1}^{A(B)} - 0.05 D_{4t} + e_t$$

$$R^2 = 0.98 \quad df = 13$$

Issue1: Transport services offered variable

The regressions estimate how the number of passenger trips per period (per capita), *Q*, depends on various demand variables such as the price of oil or the transport fare and at least one supply-side variable, the distance of transport services offered (per capita), *S*. The variable S is potentially problematic, and the paper would be more convincing if the authors discussed the reasons for including it, and the potential problems that stem from including it, at greater length.

There are two problems with the variable S.

1 First, it may be acting as a proxy for the other factors. If so, this will affect the estimates of how much these other factors affect passenger trips (Q). Suppose we call the other factors X. Further suppose that the number of services provided, S depends on a set of factors Z as well as X. Then mathematically we have a system:

$$\begin{aligned} Q &= f(S,X)+e_1 \\ S &= g(Z,X)+e_2 \end{aligned} \qquad \Rightarrow Q = f\big(g(Z,X)+e_2, X\big)+e_1$$

If X is a major factor determining S, then the effect of X on Q has two components: the direct effect; and the indirect effect through S. Estimates of the effect of X on Q that do not take into account the effect on S will then be potentially misleading.

Before discussing this further, to give an idea of how potentially important this problem is, consider the following exaggerated example. Suppose Q was the number of bus trips by males, and X are the price variables included in the above regression. Now suppose S was the number of bus trips taken by females. A regression of Q on (S,X) will have an almost perfect fit, because the economic factors determining female trips are very similar to the factors determining male trips. Furthermore, the additional explanatory power of X in the regression for Q is likely to be minor, as its effects are already apparent in the number of female patrons. Consequently, this regression would have a very good explanatory power for the number of male patrons, given the number of female patrons, but it would be unable to tell very much at all about the fundamental role of how X affects the number of male patrons. To work this out, we would need to estimate how X affects the number of female patrons, and incorporate this indirect information into our estimates of the effect of X on Q.

In the paper, S is the number of bus services put in place by the provider. It seems plausible that the number of services provided in part reflects some of the economic factors X such as the oil price or the innate demand stemming from congestion etc (This will be true if the provider increases services in response to demand.) If this is the case, S and Q are likely to be correlated because they both reflect common factors X. Consequently, the effect of X on Q is likely to be incorrectly estimated, as it ignores any indirect effect on Q through S. Obviously, if the coefficients are incorrectly estimated, the elasticities measuring how Q depends on X will also be incorrectly calculated.

Figure 5.12(a) in section 5.1.2 shows that S and Q are highly correlated. It would be worthwhile examining the extent to which S depends on the other variables X. If there is a relationship, this needs to be taken into account when calculating how Q depends on X.

2 The second problem is the standard simultaneous variables problem. There are some similarities to the problem analysed above, although some differences.

The variable S – the supply of bus trip-miles – has all the hallmarks of a supply side variable. If so, it needs to be treated within a simultaneous equation framework where the price for bus services is affected by both the supply and demand factors. Typically, if one wishes to estimate both the demand and supply equations, a set of variables Z_1 and Z_2 that affect either demand or supply but not both is needed, in addition to the variables (X) that affect both:

$$Q = f(S, X, Z_1) + e_1$$
$$S = g(Q, X, Z_2) + e_2$$

While it is not straightforward to model a transport network in a simple supply and demand framework, it would be useful for the authors to discuss the assumptions that would be necessary to estimate the equation for Q as a function of S, X and Z_1 without having unbiased coefficient estimates. In particular, what assumptions about the relationship between S and Q are necessary? It would also be useful to consider an instrumental variables approach if there are variables Z_2 that have a direct affect on the transport network but not on demand.

I am certainly sympathetic to an approach that wishes to incorporate information about how *exogenous* changes in the transport network affect passenger numbers change. In general, one wishes to believe that if you build a rail network, this will lead to an expansion in passenger numbers. At the moment, however, the paper doesn't really explain why the size of the network (in terms of bus-miles, S) changes from quarter to quarter. If the changes in S are driven by the variables X such as prices, these changes will not be exogenous, and then there will be a simultaneity bias in the coefficient estimates. At a minimum, the authors should regress S against the other variables in the original regression, to establish that changes in S are not a function of common variables. An exogeneity test would be useful. The ideal solution would be to find an instrument that directly affects the size of the transport network, but not demand, and estimate a simultaneous equation system. This might be a lot easier said than done, however.

Recommendation:

The authors need to explain the role of the variable S in the regression. It would be useful to see how S depends on the other variables, and to know the correlation between S and Q. If S does depend on the other variables X, this dependence should be taken into account when calculating elasticities.

Response: Please refer to appendix E.2.

Issue 2 Spurious regressions.

In appendix B, the authors conducted adjusted Dickey Fuller tests for stationarity and establish:

1 $Q_t^{A(B)}$ is stationary around a time trend (at a 1% critical level) (but the first order regression coefficient is not reported)

2 $S_t^{A(B)}$ is not stationary around a time trend but the difference is stationary (again, the first order regression coefficients are not reported)

3 $C_t^{A(B)}$ is stationary around a time trend (at 5% critical level)

4 $O_t^{A(B)}$ is not stationary around a time trend but the difference is stationary.

These tests should not be considered definitive: the time series are simply much too short, and thus the critical values should not be taken seriously. (The coefficients (and standard errors) in these stationarity regressions should be reported, as the formal test statistics have little meaning in series this short.) Nonetheless, the tests are suggestive that all of the series have considerable autoregressive persistence: that the coefficient ρ in the regression

$$X_t = \alpha + \rho X_{t-1} + e_t$$

is close to 1.

When time series are persistent, particularly if they have time trends, a regression between two or more of the series can be very misleading because of the problem of spurious regression. This problem arises because in order to minimise the sum of squared errors, a simple linear regression will choose coefficients that reflect the trends in the series. Consider the following regression between two variables Y and Z, each which has a trend: $Y_t = \alpha_0 + \alpha_1 Z_t + u_t$. For example, Y could be Auckland bus patronage, which increased steadily after 1999, and Z could be Australian coal exports, which also steadily increased over the period. A regression estimated using these two variables could have a very high R^2, and a positive and *ostensibly* statistically significant coefficient that reflects the fact that both had a common time trend. The regression would be misleading because the coefficients are estimated conditional on the assumption that the error terms in the regression not only are mean zero and are uncorrelated with the independent variables, but have no serial correlation either. When a spurious regression occurs, the residuals u_t have a high degree of serial correlation. The variables Y and Z *ostensibly* appear to be significantly related because the standard error of the coefficient α_1 is calculated without taking into account the serial correlation in u_t. If the serial correlation in u_t were taken into account, the standard error would probably suggest the two variables were not related.

Technically, a spurious regression is said to occur if at least one of the variables Y and Z are non-stationary integrated processes, but the estimated residual series is not I(0) (Hamilton chapter 18.) Given that the variables used in the report are persistent, the authors should report the serial correlation of the estimated residual processes u_t in more or less all of their regressions, and a test as to whether or not the residual is serially correlated. A simple Durbin-Watson test or a Durbin-h test would suffice to test whether the residuals are serially correlated. I would recommend the following procedure

1 Estimate the regression coefficients and the residuals $\hat{u}_t$

2 Estimate the degree of serial correlation in the residuals, and conduct a test such as the Durbin-Watson test to test whether this serial correlation is zero or not.

3 If it is zero, fine.

4 If not, use a unit root test to test whether the residuals are stationary or not.

   a If non-stationary, there is a formal spurious regression problem – basically the top level regression is meaningless

   b If stationary, but the residuals are serially correlated, a correction procedure such as generalized least squares or Cochrane Orcutt should be used to estimate the parameters of the top-level equation in a manner that takes into account the serial correlation of the errors.

Response: Please refer to appendix E.2.

I would be surprised if there is not serial correlation of some sort – but even if there is not, the authors need to report the test statistics so the reader has confidence that there is no serial correlation. If there is serial correlation that is fixable (ie not a spurious regression) the equations need re-estimation to take into account serial correlation. This may alter the calculated elasticities even though the parameters estimated are unbiased.

Most importantly, the standard errors of the coefficients need to be calculated correctly. If the standard errors have been calculated using standard classical methods, they will be wrong if there is serial correlation in the residuals. This is important for two reasons. First, the variables in the regression may not be statistically significant at all – it is just they appear statistically significant because the standard errors aren't corrected for the serial correlation in the error structure. Secondly, the authors have used a step-wise regression method to calculate their final regression. They may have excluded or included some variables based on incorrectly calculated standard errors. If this is the case, they may want to reassess the final form of their equation.

The potential problem of spurious regressions is made more difficult in this case by the relationship between Q and S. If both S and Q are highly correlated because they both are determined by the same driving variables, the residuals will have little serial correlation (as would be the case in an equation that regresses male bus trips against female bus trips, to use the previous example.) If both S and Q were I(1) variables, and the residuals were I(0), the variables would be cointegrated, which is fine. Nonetheless, there would still be a problem with interpreting the coefficients on X in a regression of Q on S and X, because X determine S and Q jointly, and the marginal effect of X on Q conditional on S does not reflect the total effect of X on Q.

I suspect the Auckland bus equation is not too badly behaved, in part because the term on the lagged dependent variable is not too high. Moreover, in section 6.4.1.1 the plot of the residuals in figure 6.2 for the final equation appears well behaved – although the series is very short so it is not clear whether this reflects over-fitting, or the correlation between S and Q. This needs to be investigated more fully.

However, there could be problems with other equations. As a general rule, tests for serial correlation should be included in time series work, particularly if the variables are as persistent as these ones are.

Recommendation

1 *All the regressions should report a test for serial correlation in the residuals, and correct standard errors for any serial correlation*

2 *The tests for stationarity, appendix B should report coefficients and standard errors: the length of the series is so short that it is difficult to interpret only the reported test statistics.*

3 *It would be useful to report the results from a regression of Q against S: the autocorrelation of both Q and S, and the autocorrelation of the residuals from a regression between Q and S. Ideally, this should be done on the whole period as well as the sub-periods.*

Response: Please refer to appendix E.2.

Issue 3: Series length

The series length for some of these regressions is extremely short. The recommended best equation for Auckland buses has 20 data points and 6 variables, leaving only 13 degrees of freedom. One needs to be very careful about inference with this level of degrees of freedom, particularly if there are serial correlation issues, for it is possible that the data from the sample period is not particularly representative of the underlying data generating process.

In appendix C, the model is 1) first estimated with a dummy variable for the post 1999 period and then 2) estimated over the shortened 1999–2008 period. This leads to a significant change in many of the coefficients. But it doesn't seem clear to me why it is necessary to drop the 1996–1999 observations. Obviously something changed in 1999 which lead to an increased use of bus transport. This could have a level increase (the model with a dummy variable) or it could have been a change in the underlying structural parameters (the estimated change in the coefficients.) The authors don't suggest why one model is preferred over the other – the R^2 are similar in both cases – or even whether the change in the parameters is statistically significant (a Chow test). It would be useful to discuss why they prefer the model on the shorter time period, and why they believe the parameters should have changed pre and post 1999. If there isn't a good economic reason for the change, one needs to be suspicious about the future usefulness of the model because it is possible that another set of parameter changes could occur.

The authors further shorten the period to get to their preferred equation, estimated only on the post-2003 period. While the rationale for this period is better – it is to be consistent with rail data after the opening of the Britomart centre – the shortness of the data period rather undermines the exercise. A regression with 13 degrees of freedom lacks credibility. That said, as shown in appendix C, it is worth stressing that the coefficients estimated post-2003 and post-1999 are similar. Perhaps the authors should emphasise the longer regression, since the two appear reasonably similar.

Recommendation: The authors

1 *explain why the estimated coefficients might have changed after 1999*

2 *if they continue to use the post-1999 period, they should use the whole of the period as the post 2003 period really is too short .*

Response: Please refer to appendix E.2.

Auckland rail

The above three criticisms can be applied to some extent to each of the econometric models. The Auckland rail model is potentially the most interesting, however, because the supply of rail miles variable (S) seems most likely to be exogenous. There was a large increase in rail services in 2003, and a further increase in 2005, and quarterly changes in rail services don't seem driven by the supplier altering services in response to quarterly changes in demand.

To some extent the recommendations for the Auckland bus model should be followed in general here, including the treatment of serial correlation. Nonetheless there are specific features of the Auckland rail network that should be considered.

1 *Data coverage*

The equation is estimated for the post 2003 period, because of the increase in rail services. It doesn't strike me as obvious that the equation couldn't have been estimated over the whole period, however. If rail service provision (S) was largely exogenous, the additional rail services should provide the variation in the data necessary to generate accurate estimates. It will be nice to see if the parameter estimates are similar over the short and long periods; if so, the long period estimates are to be preferred, to avoid the problems of inference over short periods.

2 *Transport services offered variable'*

It would be useful to see if S is as dependent on the other variables (X) as in the case of Auckland buses. If it is dependent, this needs to be taken into account, given that there is a very large coefficient between Q and S; if not the contrast with buses should be emphasized for it would suggest that structurally the rail equation is better.

Wellington

The Wellington analysis has too little data to be taken seriously. It doesn't make much sense to estimate models with only 5 to 6 degrees of freedom and several included independent variables. The coefficient on the lagged dependent variable is also quite high (0.5) suggesting important information is missing from the equation.

Christchurch

The regressions for Christchurch buses are estimated over two different periods: 1998–2008 and 2004–2008. The results are quite different.

The first equation is very problematic. The lagged dependent variable has a coefficient of 0.89 and is probably not significantly different to 1. This suggests that the equation that is being estimated is the change in Christchurch bus passengers, not the level of Christchurch bus passengers:

$$\ln Q_t = 0.11 + 0.07 \ln S_t - 0.09 D_t^4 + 0.89 \ln Q_{t-1} + e_t$$
$$=> \Delta \ln Q_t = 0.11 + 0.07 \ln S_t - 0.09 D_t^4 - 0.11 \ln Q_{t-1} + e_t$$

If so, the independent variables should also be in differenced form, not level form.

Recommendation: this equation should be re-estimated in difference form.

Response: This was a typo. Equations 6.70 and D.7 should be:

$$\ln Q_t^{C(B)} = 0.1147 + 0.0708 \ln S_t^{C(B)} + 0.8863 \ln Q_{t-1}^{C(B)} - 0.0898 D_{4t}$$

The second equation is considerably better behaved in levels. However, there are only 17 observations and 11 degrees of freedom, which rather undermines confidence in making an inference. Since this equation does not include a variable for service levels, many of the problems alluded to in the Auckland study are circumvented. In order to install greater confidence in the results, given the short sample period, more information about the serial correlation properties of the residuals should be reported. If these results are going to be used, there needs to be confidence that there is no serial correlation in the errors, and thus there is no reason to doubt that the sample period includes a large and representative range of data.

Appendix E.2 Response to referee report dated 18 November from Andrew Coleman, 5 January 2010

Overview

In summary, the following issues were identified by the third peer reviewer:

Region	Issues
Auckland	*Issue1: Transport services offered variable* *Issue 2: Spurious regressions* *Issue 3: Series length*
Wellington	The Wellington data series is simply too short. It doesn't make much sense to estimate models with only 5–6 degrees of freedom and several included independent variables.
Christchurch	The regressions for Christchurch buses are estimated over two different periods: 1998–2008 and 2004–2008. Different problems were identified with the two equations.

Perhaps we should start with an explanation of the background of the study. The client, NZTA, specifically requested an investigation of at least the three major regions in New Zealand, with a view to comparing the different characteristics of the three regions. During the inception stage, two econometric modelling approaches were identified in the literature, namely, partial adjustment model (PAM) and vector error correction model (VECM). As the study progressed, due to data availability and limitations, PAM was identified as the only 'feasible' model for the three cities. Please refer to chapter 3 of the draft report for detailed explanation.

Like any modelling approach, PAM comes with its limitations. The assumptions behind the PAM are stated clearly in section 6.1 of the draft report. In particular, as stated in section 6.1, we assume that all explanatory variables are given or determined exogenously. As discussed in section 3.3, PAM has been applied to model demand for bus services in England (Dargay and Hanly 2002b) based on the same assumptions and service-km was also considered as one of the independent variables. We were aware that this assumption could be a problem and might not be valid. We, therefore, pointed out in section 8.2, that '[..] although VECM was considered to be a more ideal model, it was impossible to apply this method to all three cities'. In fact, we believe that the three issues for Auckland as stated above might not be a problem at all if a VECM was developed instead of the PAM approach. We, therefore, as explained in section 6 , performed the stationarity tests as shown in appendix B of the draft report and illustrated that it would be feasible to deduce a VECM for Auckland. In fact, VECM has also been applied to model demand for bus services in England as well by the same authors (Dargay and Hanly 2002a).

Nevertheless, we do agree that it is important to assess the validity of the models developed. We, therefore, performed a series of tests, following the recommendations from the third reviewer, wherever appropriate for all the three regions, in order to address the issues identified.

Issue 1: 'Transport services offered variable'.

We applied a technique commonly used to handle/test for possible endogeneity in time series data. We introduced lagged service variable(s), ie S_{t-1} and S_{t-2}, as instrument(s). For example, the Auckland bus forecast model as described with equation 6.73, ie:

$$\begin{aligned}\ln Q_t^{A(B)} &= -1.9306 + 0.4603 \ln S_t^{A(B)} - 1.9618 \ln C_t^{A} + 0.2040 \ln O_t \\ &+0.3677 \ln Q_{t-1}^{A(B)} - 0.0478 D_{4t}\end{aligned}$$

where $t \in \left[2003Q3, 2008Q2\right]$

We introduced two possible sets of equations with instrument variables (IV) as follows.

<u>*IV Set 1*</u>

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln \hat{S}_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} + \beta_O^{A(B)} \ln O_t \\ &+\theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t}\end{aligned} \quad \text{(IV 1a)}$$

$$\ln \hat{S}_t^{A(B)} = \hat{\pi}_0^{A(B)} + \hat{\pi}_{t-1}^{A(B)} \ln S_{t-1}^{A(B)} \quad \text{(IV 1b)}$$

<u>*IV Set 2*</u>

$$\begin{aligned}\ln Q_t^{A(B)} &= \alpha^{A(B)} + \beta_S^{A(B)} \ln \hat{S}_t^{A(B)} + \beta_C^{A(B)} \ln C_t^{A} + \beta_O^{A(B)} \ln O_t \\ &+\theta^{A(B)} \ln Q_{t-1}^{A(B)} + \gamma_4^{A(B)} D_{4t}\end{aligned} \quad \text{(IV 2a)}$$

$$\ln \hat{S}_t^{A(B)} = \hat{\pi}_0^{A(B)} + \hat{\pi}_{t-1}^{A(B)} \ln S_{t-1}^{A(B)} + \hat{\pi}_{t-2}^{A(B)} \ln S_{t-2}^{A(B)} \quad \text{(IV 2b)}$$

The performance and validity of the sets of equations (IV 1) or (IV 2) were assessed based on:

1 Breusch-Godfrey LM test, as described in Baum and Wiggins (1999), for autocorrelation on, eg equations (IV 1b) and (IV 2b)

H_0 : No serial correlation

H_a : There is autocorrelation

2 Wu-Hausman F-test, as described in Baum et al (2002), for the sets, eg (IV 1) & (IV 2)

H_0 : Variables are exogenous

The tests were carried out in STATA and the test results are provided in appendices E.2.1 and E.2.2.

H_a : Variables are NOT exogenous

Issue 2: Spurious regressions

Spurious regressions indeed could be a problem in a PAM if the variables considered are not cointegrated. For example, equation 6.73 would have a spurious regression problem if the endogenous variables, ie $\ln Q_t^{A(B)}, \ln S_t^{A(B)}$ and $\ln C_t^A$, are not cointegrated, assuming that fuel price, $\ln O_t$, is exogenous. Apart from stationarity tests that were performed before, we, therefore, performed the Johansen trace test, as described in Lutkepohl and Kratzig (2004), to test for cointegration for each set of variables in the models developed with the structural breaks identified for the Auckland and Christchurch data sets respectively. Detailed printouts from JMulTi for stationarity are provided in appendix E.2.3, as per the third peer reviewer's request and the cointegration test results are detailed in appendix E.2.4.

In addition to cointegration test, Breusch-Godfrey LM test for autocorrelation was performed to ensure that the residuals of the original PAM equation are not autocorrelated. The tests were carried out in STATA, as described in Baum and Wiggins (1999), and the test results are provided in appendices E.2.1 and E.2.2.

Issue 3: Series length.

As discussed earlier, like any modelling approach, PAM comes with its limitations. Ideally the changes of trend should be modelled as structural breaks in a VECM model. Since structural breaks cannot be handled properly with a PAM modelling approach, we can only be aware of its limitations and try to use the models that were developed based on longer time series for forecasting purposes.

Summary of test results

The results of the tests are summarised in table E.1. The observations are summarised as follows:

Auckland

Only the Auckland models passed both the endogeneity tests and cointegration tests. Therefore, there is no evidence of the issues of endogeneity and spurious regression in the Auckland models.

Wellington

The Wellington data set was not sufficient to support the endogeneity tests and cointegration tests. The residuals of the Wellington rail models appeared to have serial autocorrelation.

Christchurch

There is no evidence that the residuals of the Christchurch models have serial autocorrelation. However, it appeared that the variables considered were not cointegrated. This is not a surprise as we were aware that the service information was only estimated from timetables (section 8.2).

Conclusions

The results from the further tests do not change our conclusion that '[..] the missing or inaccurate information could have implications on the quality of the models estimated but we have already made the most out of the information collected' (section 8.2), as we were well aware that '[..] the power of a forecast model could only be as good as the data that it was built on' (section 4.5).

It is clear that the Auckland data set has the quality to support the development of better models. There is no evidence of the issues of endogeneity and spurious regression in the Auckland PAM models. To better utilise the Auckland data set, it is recommended to develop more advanced models for Auckland (section 8.3.3), for example, a VECM with the structural breaks incorporated. In this way, the information in this data set can be utilised for the full period, ie 1996Q1–2008Q2, and it is anticipated that more accurate forecast model(s) for Auckland can be developed.

References

Baum, CF, ME Schaffer and S Stillman (2007) *IVENDOG: stata module to calculate Durbin-Wu-Hausman endogeneity test after ivreg*. Revised version. Boston: Boston College Department of Economics.

Baum, CF, ME Schaffer and S Stillman (2007) *IVENDOG: stata module to calculate Durbin-Wu-Hausman endogeneity test after ivreg*. Revised version. Boston: Boston College Department of Economics.

Dargay, J and M Hanly (2002a) The demand for local bus services in England. *Journal of Transport Economics and Policy 36*, no.1: 73–91.

Dargay, J and M Hanly (2002b) Bus patronage in Great Britain – econometric analysis. *Transportation Research Record 1799*: 97–106.

Lutkepohl, H and K Kratzig (2004) *Applied time series econometrics*. Cambridge: Cambridge University Press.

Table E.1 Summary of test results

Model	Independent variables	Modelled period	Endogeneity test - Ho: variables are exogenous (Wu-Hausman p-value)		Bgodfrey test on PAM – H0: no serial correlation (p-value)	Cointegration – Johansen trace test	Model acceptance
			IV – S(t-1)	IV – S(t-1), S(t-2)			
AKL bus model 1	Service + car ownership + Q4 seasonal dummy	[1999Q3,2008Q2]	0.1127	0.1157	0.8151	Cointegrated	√
AKL bus model 2	Service + car ownership + fuel price + Q4 seasonal dummy	[2003Q3,2008Q2]	0.4105	0.9338	0.0947	Cointegrated	√
AKL rail model 1	Service + fare + income + Q4 seasonal dummy	[2003Q3,2008Q2]	0.6500	0.4043	0.1151	Cointegrated	√
AKL rail model 2	Service + fare + Q4 seasonal dummy	[2003Q3,2008Q2]	0.6479	0.9527	0.9164	Cointegrated	√
WLT bus model	Fare only	[FY1999/2000,FY2007/08]	n/a	n/a	0.1604	n/a	X
WLT rail model 1	Service + car ownership + fuel price	[FY1999/2000,FY2007/08]	n/a	n/a	0.0279	n/a	X
WLT rail model 2	Service + income + fuel price	[FY1999/2000,FY2007/08]	n/a	n/a	0.0334	n/a	X
CHC bus model 1	Service + Q4 seasonal dummy	[1998Q1,2008Q2]	n/a	n/a	0.1110	Not cointegrated	X
CHC bus model 2	Fare + fuel price + Q4 seasonal dummy	[2004Q1,2008Q2]	n/a	n/a	0.5590	Not cointegrated	X

Appendix E.2.1: Instrument variable (IV) regression results with STATA for Auckland (Breusch-Godfrey LM and endogeneity tests)

```
. *
```

. *AKL bus model 1–1999Q3

```
. regress lnBusKmPerCapAKL lnBusKmPerCapAKLlag1 if tin(1999q3,2008q2)

      Source |     SS       df       MS              Number of obs =      36
-------------+------------------------------        F( 1,  34)    =  222.54
       Model | .915342616    1 .915342616           Prob > F      =  0.0000
    Residual | .139848072   34 .004113179           R-squared     =  0.8675
-------------+------------------------------        Adj R-squared =  0.8636
       Total | 1.05519069   35 .030148305           Root MSE      =  .06413

------------------------------------------------------------------------------
lnBusKmPer~L |      Coef.   Std. Err.    t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~1 |   .9003044   .0603513   14.92  0.000      .7776559  1.022953
       _cons |   .1685912   .0925703    1.82  0.077    -.0195343  .3567168
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |      chi2         df          Prob > chi2
-------------+-------------------------------------------------------------
      1      |      0.265        1           0.6069
---------------------------------------------------------------------------
                H0: no serial correlation

. *
. regress lnBusKmPerCapAKL lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2 if tin(1999q3,2008q2)

    Source |     SS        df      MS           Number of obs =     36
-------------+------------------------------ F( 2,  33)    = 119.26
     Model | .926942292   2  .463471146       Prob > F      = 0.0000
  Residual | .128248397   33 .003886315       R-squared     = 0.8785
-------------+------------------------------ Adj R-squared = 0.8711
     Total | 1.05519069   35 .030148305       Root MSE      = .06234

------------------------------------------------------------------------------
lnBusKmPer~L |     Coef.    Std. Err.     t    P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~1 |   .6533785   .1544971   4.23   0.000    .3390517   .9677053
lnBusKmPer~2 |   .2661442   .1540505   1.73   0.093   -.0472739   .5795622
     _cons |     .1424819   .0912415   1.56   0.128   -.0431504   .3281142
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |      chi2         df          Prob > chi2
-------------+-------------------------------------------------------------
     1   |       2.631        1           0.1048
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. regress lnBusPaxPerCapAKL lnBusKmPerCapAKL lnCarOwnPerCapAKL lnBusPaxPerCapAKLlag1 IndQtr4 if
tin(1999q3,2008q2)

      Source |      SS      df      MS            Number of obs =     36
-------------+------------------------------  F( 4,  31)    = 183.46
       Model | .745813587    4 .186453397        Prob > F      = 0.0000
    Residual | .031505151   31 .001016295        R-squared     = 0.9595
-------------+------------------------------  Adj R-squared = 0.9542
       Total |  .777318738   35 .022209107         Root MSE     = .03188

------------------------------------------------------------------------------
```

```
lnBusPaxPe~L |      Coef.     Std. Err.   t  P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~L |     .4233067   .0878845   4.82  0.000     .244065   .6025484
lnCarOwnPe~L |   -1.137092   .6847746  -1.66  0.107  -2.533699   .2595155
lnBusPaxPe~1 |      .587144   .0729092   8.05  0.000    .4384448   .7358432
   IndQtr4 |     -.0700429   .0152546  -4.59  0.000  -.1011548  -.0389309
     _cons |     -.6965566    .555452  -1.25  0.219  -1.829408   .4362952
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |   chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |     0.055        1          0.8151
---------------------------------------------------------------------------
              H0: no serial correlation

. ivregress 2sls lnBusPaxPerCapAKL lnCarOwnPerCapAKL lnBusPaxPerCapAKLlag1 IndQtr4 (
lnBusKmPerCapAKL = lnBusKmPerCapAKLlag1) if tin(1999q3,2008q2)

Instrumental variables (2SLS) regression        Number of obs =      36
                                                Wald chi2(4)  = 596.31
                                                Prob > chi2   = 0.0000
                                                R-squared     = 0.9430
                                                Root MSE      = .03509

------------------------------------------------------------------------------
lnBusPaxPe~L |   Coef.     Std. Err.     z   P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~L |  .7353922   .2364917    3.11  0.002     .271877   1.198907
lnCarOwnPe~L | -2.747615   1.344729   -2.04  0.041  -5.383236  -.1119941
lnBusPaxPe~1 |  .4370741     .13118    3.33  0.001    .1799661   .6941821
     IndQtr4 | -.0422957   .0254958   -1.66  0.097  -.0922665   .0076752
       _cons | -1.996584   1.087142   -1.84  0.066  -4.127344   .1341759
------------------------------------------------------------------------------
Instrumented: lnBusKmPerCapAKL
Instruments:  lnCarOwnPerCapAKL lnBusPaxPerCapAKLlag1 IndQtr4
              lnBusKmPerCapAKLlag1

. estat endogenous

 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)     = 2.94205 (p = 0.0863)
 Wu-Hausman F(1,30)       =  2.6699 (p = 0.1127)

. ivregress 2sls lnBusPaxPerCapAKL lnCarOwnPerCapAKL lnBusPaxPerCapAKLlag1 IndQtr4 (
lnBusKmPerCapAKL = lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2) if tin(1999q3,2008
> q2)
Instrumental variables (2SLS) regression    Number of obs =      36
                                            Wald chi2(4)  = 605.04
                                            Prob > chi2   = 0.0000
                                            R-squared     = 0.9438
                                            Root MSE      = .03484

------------------------------------------------------------------------------
lnBusPaxPe~L |     Coef.   Std. Err.     z  P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~L |  .7275733  .2313934     3.14  0.002     .2740506     1.181096
lnCarOwnPe~L | -2.707265   1.31917    -2.05  0.040     -5.29279    -.1217401
lnBusPaxPe~1 |  .4408339  .1288232     3.42  0.001      .188345     .6933229
   IndQtr4 |   -.0429909   .025064    -1.72  0.086    -.0921154     .0061337
     _cons |    -1.964013  1.066518    -1.84  0.066    -4.054351     .1263244
------------------------------------------------------------------------------
Instrumented: lnBusKmPerCapAKL
Instruments:  lnCarOwnPerCapAKL lnBusPaxPerCapAKLlag1 IndQtr4
        lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2

. estat endogenous

 Tests of endogeneity
```

```
 Ho: variables are exogenous

 Durbin (score) chi2(1)   =  2.8965 (p = 0.0888)
 Wu-Hausman F(1,30)       = 2.62495 (p = 0.1157)
```

. *AKL bus model 2 - 2003Q3

```
. regress lnBusKmPerCapAKL lnBusKmPerCapAKLlag1 if tin(2003q3,2008q2)
   Source |       SS    df      MS            Number of obs =     20
-------------+------------------------------  F( 1,  18)    =  69.28
    Model |  .199210458   1 .199210458        Prob > F      = 0.0000
  Residual | .051756372  18 .002875354        R-squared     = 0.7938
-------------+------------------------------  Adj R-squared = 0.7823
    Total |  .25096683  19 .013208781         Root MSE      = .05362
------------------------------------------------------------------------------
lnBusKmPer~L |   Coef.     Std. Err.   t  P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~1 |  .8026494  .0964307   8.32  0.000   .6000559  1.005243
    _cons |     .3449033   .159871   2.16  0.045   .0090267  .6807799
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |     chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |      0.244         1            0.6212
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. regress lnBusKmPerCapAKL lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2 if tin(2003q3,2008q2)

   Source |    SS        df     MS               Number of obs =     20
-------------+------------------------------     F( 2,  17)    =  33.29
    Model |  .199921802   2 .099960901           Prob > F      = 0.0000
  Residual | .051045028  17 .003002649           R-squared     = 0.7966
-------------+------------------------------     Adj R-squared = 0.7727
    Total |  .25096683   19 .013208781           Root MSE      =  .0548


------------------------------------------------------------------------------
lnBusKmPer~L |    Coef.    Std. Err.   t     P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~1 |  .9091283   .239934     3.79  0.001   .4029119  1.415345
lnBusKmPer~2 | -.1062633  .2183212    -0.49  0.633  -.5668807  .3543541
       _cons |    .34271  .1634337     2.10  0.051  -.0021049  .6875249
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |     chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |      0.331         1           0.5650
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. regress lnBusPaxPerCapAKL lnBusKmPerCapAKL lnCarOwnPerCapAKL lnRealFuelP lnBusPaxPerCapAKLlag1
IndQtr4 if tin(2003q3,2008q2)

    Source |     SS      df     MS            Number of obs =     20
-------------+------------------------------  F( 5,  14)    = 169.41
     Model | .198611765   5 .039722353        Prob > F      = 0.0000
  Residual | .003282663  14 .000234476        R-squared     = 0.9837
-------------+------------------------------  Adj R-squared = 0.9779
     Total | .201894428  19 .010626023        Root MSE      = .01531

------------------------------------------------------------------------------
lnBusPaxPe~L |    Coef.   Std. Err.   t    P>|t|   [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
lnBusKmPer~L |  .4608974  .0818964   5.63  0.000    .285247   .6365478
lnCarOwnPe~L | -1.951866  .5078065  -3.84  0.002  -3.041003  -.8627299
 lnRealFuelP |   .204324  .0588237   3.47  0.004   .0781597   .3304883
lnBusPaxPe~1 |  .3665044  .0741789   4.94  0.000   .2074064   .5256024
     IndQtr4 | -.0478434  .0125178  -3.82  0.002  -.0746914  -.0209954
       _cons | -1.924969  .4143493  -4.65  0.000   -2.81366  -1.036278
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |      chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |      2.793         1          0.0947
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. ivregress 2sls lnBusPaxPerCapAKL lnCarOwnPerCapAKL lnRealFuelP lnBusPaxPerCapAKLlag1 IndQtr4 (
lnBusKmPerCapAKL = lnBusKmPerCapAKLlag1) if tin(2003q3,2008q2)

Instrumental variables (2SLS) regression    Number of obs =     20
                                            Wald chi2(5)  = 617.50
                                            Prob > chi2   = 0.0000
                                            R-squared     = 0.9692
                                            Root MSE      = .01765

------------------------------------------------------------------------------
lnBusPaxPe~L |   Coef.    Std. Err.    z    P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~L |  .7511417  .4006561   1.87  0.061  -.0341299   1.536413
lnCarOwnPe~L | -2.832896  1.318893  -2.15  0.032  -5.417879  -.2479126
 lnRealFuelP |  .0902761  .1673473   0.54  0.590  -.2377186   .4182707
lnBusPaxPe~1 |  .2442737  .1849243   1.32  0.187  -.1181713   .6067188
     IndQtr4 | -.0379087    .01964  -1.93  0.054  -.0764023   .0005849
       _cons | -2.199237  .6028153  -3.65  0.000  -3.380734  -1.017741
------------------------------------------------------------------------------
Instrumented: lnBusKmPerCapAKL
Instruments:  lnCarOwnPerCapAKL lnRealFuelP lnBusPaxPerCapAKLlag1 IndQtr4
        lnBusKmPerCapAKLlag1

. estat endogenous

 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)     = 1.05409 (p = 0.3046)
 Wu-Hausman     F(1,13      = .723281 (p = 0.4105)

. ivregress 2sls lnBusPaxPerCapAKL lnCarOwnPerCapAKL lnRealFuelP lnBusPaxPerCapAKLlag1 IndQtr4 (
lnBusKmPerCapAKL = lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2) if ti
> n(2003q3,2008q2)

Instrumental variables (2SLS) regression      Number of obs =      20
                                              Wald chi2(5)  = 1182.52
                                              Prob > chi2   =  0.0000
                                              R-squared     =  0.9837
                                              Root MSE      =  .01282

------------------------------------------------------------------------------
lnBusPaxPe~L |     Coef.   Std. Err.  z    P>|z|  [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~L |  .4697478  .1087119   4.32  0.000   .2566762   .6828193
lnCarOwnPe~L | -1.978731  .4962431  -3.99  0.000   -2.95135  -1.006113
 lnRealFuelP |  .2008464  .0593584   3.38  0.001    .084506   .3171868
lnBusPaxPe~1 |  .3627772  .0715375   5.07  0.000   .2225663   .5029882
     IndQtr4 | -.0475404  .0108683  -4.37  0.000  -.0688418   -.026239
       _cons | -1.933332  .3558613  -5.43  0.000  -2.630808  -1.235857
------------------------------------------------------------------------------
Instrumented: lnBusKmPerCapAKL
Instruments:  lnCarOwnPerCapAKL lnRealFuelP lnBusPaxPerCapAKLlag1 IndQtr4
        lnBusKmPerCapAKLlag1 lnBusKmPerCapAKLlag2
```

```
. estat endogenous

 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)      = .011011 (p = 0.9164)
 Wu-Hausman F(1,13)          = .007161 (p = 0.9338)
```

. *AKL rail model 1–2003Q3

```
. regress lnRailKmPerCapAKL lnRailKmPerCapAKLlag1 if tin(2003q3,2008q2)

    Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F( 1, 18)     =   80.00
     Model | .416980584     1 .416980584           Prob > F      = 0.0000
  Residual | .093824966    18 .005212498           R-squared     = 0.8163
-------------+------------------------------           Adj R-squared = 0.8061
     Total | .510805551    19 .026884503           Root MSE      =   .0722

------------------------------------------------------------------------------
lnRailKmPe~L |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~1 |     .80137    .089598     8.94   0.000     .6131317    .9896083
       _cons |   -.213641   .1131995    -1.89   0.075    -.4514644    .0241825
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          1.198               1                   0.2737
---------------------------------------------------------------------------
                        H0: no serial correlation

. *
. regress lnRailKmPerCapAKL lnRailKmPerCapAKLlag1 lnRailKmPerCapAKLlag2 if tin(2003q3,2008q2)

    Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F( 2,   17)   =   50.20
     Model | .436842809     2 .218421405           Prob > F      = 0.0000
  Residual | .073962742    17  .00435075           R-squared     = 0.8552
-------------+------------------------------           Adj R-squared = 0.8382
     Total | .510805551    19 .026884503           Root MSE      =  .06596

------------------------------------------------------------------------------
lnRailKmPe~L |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~1 |   .3920206     .20834     1.88   0.077    -.0475384    .8315795
lnRailKmPe~2 |   .4073726   .1906601     2.14   0.047     .0051149    .8096304
       _cons |  -.2028347   .1035434    -1.96   0.067    -.4212922    .0156228
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          0.528               1                   0.4674
---------------------------------------------------------------------------
                        H0: no serial correlation

. *
. regress lnRailPaxPerCapAKL lnRailKmPerCapAKL lnRealRailFareAKL lnRealIncomePerCap
lnRailPaxPerCapAKLlag1 IndQtr4 if tin(2003q3,2008q2)

   Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F( 5,   14)   =  211.94
    Model | 1.35034074     5 .270068148           Prob > F      = 0.0000
  Residual | .017839749    14 .001274268           R-squared     = 0.9870
-------------+------------------------------           Adj R-squared = 0.9823
```

```
    Total  | 1.36818049  19   .0720095           Root MSE      =  .0357

------------------------------------------------------------------------------
lnRailPaxP~L |   Coef.    Std. Err.   t    P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L |  1.057607  .2159264   4.90  0.000   .5944913   1.520723
lnRealRail~L | -.9929639  .2215545  -4.48  0.001  -1.468151  -.5177768
lnRealInco~p |  1.559281  .5581257   2.79  0.014     .36222   2.756341
lnRailPaxP~1 |  .2505009  .1419538   1.76  0.099  -.0539597   .5549615
     IndQtr4 | -.1292404  .0267546  -4.83  0.000  -.1866233  -.0718574
       _cons | -.9582788   1.03284  -0.93  0.369    -3.1735   1.256942
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |      chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1     |     2.483        1          0.1151
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. ivregress 2sls lnRailPaxPerCapAKL lnRealRailFareAKL lnRealIncomePerCap lnRailPaxPerCapAKLlag1
IndQtr4 ( lnRailKmPerCapAKL = lnRailKmPerCapAKLlag1) if tin(2003q3,2
> 008q2)

Instrumental variables (2SLS) regression   Number of obs =     20
                                           Wald chi2(5)  = 177.26
                                           Prob > chi2   = 0.0000
                                           R-squared     = 0.8912
                                           Root MSE      = .08629

------------------------------------------------------------------------------
lnRailPaxP~L |   Coef.    Std. Err.   z    P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L | -1.132357  11.08659  -0.10  0.919  -22.86168  20.59697
lnRealRail~L | -.0184003  4.957228  -0.00  0.997  -9.734389  9.697589
lnRealInco~p |  .3343127  6.339687   0.05  0.958  -12.09124  12.75987
lnRailPaxP~1 |  1.611806   6.89245   0.23  0.815  -11.89715  15.12076
     IndQtr4 | -.3175528  .9544598  -0.33  0.739   -2.18826  1.553154
       _cons | -1.778249  4.840083  -0.37  0.713  -11.26464   7.70814
------------------------------------------------------------------------------
Instrumented: lnRailKmPerCapAKL
Instruments:  lnRealRailFareAKL lnRealIncomePerCap lnRailPaxPerCapAKLlag1
        IndQtr4 lnRailKmPerCapAKLlag1

. estat endogenous

 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)     = .326432 (p = 0.5678)
 Wu-Hausman F(1,13)         = .215702 (p = 0.6500)

. *
. ivregress 2sls lnRailPaxPerCapAKL lnRealRailFareAKL lnRealIncomePerCap lnRailPaxPerCapAKLlag1
IndQtr4 ( lnRailKmPerCapAKL = lnRailKmPerCapAKLlag1 lnRailKmPerCapAK
> Llag2) if tin(2003q3,2008q2)

Instrumental variables (2SLS) regression     Number of obs =     20
                                             Wald chi2(5)  = 974.98
                                             Prob > chi2   = 0.0000
                                             R-squared     = 0.9802
                                             Root MSE      =  .0368

------------------------------------------------------------------------------
lnRailPaxP~L |    Coef.    Std. Err.   z   P>|z|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L |  .4760046  .7241737   0.66  0.511  -.9433498  1.895359
lnRealRail~L | -.7341428  .3823743  -1.92  0.055  -1.483583   .015297
lnRealInco~p |  1.233958  .6925541   1.78  0.075   -.123423   2.591339
lnRailPaxP~1 |  .6120313  .4526676   1.35  0.176  -.2751809   1.499244
```

```
     IndQtr4 | -.1792517  .0653606  -2.74  0.006  -.3073562  -.0511472
       _cons | -1.176044   1.09557  -1.07  0.283  -3.323322   .9712347
------------------------------------------------------------------------------
Instrumented: lnRailKmPerCapAKL
Instruments:  lnRealRailFareAKL lnRealIncomePerCap lnRailPaxPerCapAKLlag1
        IndQtr4 lnRailKmPerCapAKLlag1 lnRailKmPerCapAKLlag2

. estat endogenous

 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)      = 1.08145 (p = 0.2984)
 Wu-Hausman F(1,13)          = .743124 (p = 0.4043)
```

. *AKL rail model 2–2003Q3

```
. *
. regress lnRailPaxPerCapAKL lnRailKmPerCapAKL lnRealRailFareAKL lnRailPaxPerCapAKLlag1 IndQtr4
if tin(2003q3,2008q2)

    Source |      SS      df     MS            Number of obs =      20
-------------+------------------------------   F( 4,  15)    = 180.90
     Model | 1.34039482    4 .335098704        Prob > F      = 0.0000
  Residual | .027785675   15 .001852378        R-squared     = 0.9797
-------------+------------------------------   Adj R-squared = 0.9743
     Total | 1.36818049   19   .0720095        Root MSE      = .04304

------------------------------------------------------------------------------
lnRailPaxP~L |     Coef.   Std. Err.    t    P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L |  .9270626  .2541707   3.65  0.002   .3853106   1.468815
lnRealRail~L | -.6984519  .2349501  -2.97  0.009  -1.199236  -.1976676
lnRailPaxP~1 |  .4249257  .1537095   2.76  0.014   .0973016   .7525498
     IndQtr4 | -.1452438  .0315097  -4.61  0.000  -.2124051  -.0780825
       _cons |  1.761024  .4165537   4.23  0.001   .8731604   2.648887
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p)  |      chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |     0.011        1          0.9164
---------------------------------------------------------------------------
            H0: no serial correlation

. ivregress 2sls lnRailPaxPerCapAKL lnRealRailFareAKL lnRailPaxPerCapAKLlag1 IndQtr4 (
lnRailKmPerCapAKL = lnRailKmPerCapAKLlag1) if tin(2003q3,2008q2)

Instrumental variables (2SLS) regression     Number of obs =      20
                                             Wald chi2(4)  =   74.73
                                             Prob > chi2   = 0.0000
                                             R-squared     = 0.7426
                                             Root MSE      =  .1327

------------------------------------------------------------------------------
lnRailPaxP~L |    Coef.    Std. Err.     z    P>|z|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L | -2.436441  21.64472  -0.11  0.910  -44.85931  39.98643
lnRealRail~L |  .4990813  7.735273   0.06  0.949  -14.66178  15.65994
lnRailPaxP~1 |   2.39764  12.69528   0.19  0.850  -22.48465  27.27993
     IndQtr4 | -.4284192  1.823677  -0.23  0.814   -4.00276  3.145922
       _cons |  -3.00247  30.66071  -0.10  0.922  -63.09636  57.09142
------------------------------------------------------------------------------
Instrumented: lnRailKmPerCapAKL
Instruments:  lnRealRailFareAKL lnRailPaxPerCapAKLlag1 IndQtr4
        lnRailKmPerCapAKLlag1

. estat endogenous
 Tests of endogeneity
 Ho: variables are exogenous
```

```
 Durbin (score) chi2(1)      = .306467 (p = 0.5799)
 Wu-Hausman F(1,14)          = .217865 (p = 0.6479)

. ivregress 2sls lnRailPaxPerCapAKL lnRealRailFareAKL lnRailPaxPerCapAKLlag1 IndQtr4 (
lnRailKmPerCapAKL = lnRailKmPerCapAKLlag1 lnRailKmPerCapAKLlag2) if tin(2003q3
> ,2008q2)

Instrumental variables (2SLS) regression    Number of obs =     20
                                            Wald chi2(4)  = 943.75
                                            Prob > chi2   = 0.0000
                                            R-squared     = 0.9796
                                            Root MSE      = .03736

------------------------------------------------------------------------------
lnRailPaxP~L |   Coef.  Std. Err.    z  P>|z|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~L |  .9943002  .9592722   1.04  0.300  -.8858388  2.874439
lnRealRail~L |  -.722391   .389963  -1.85  0.064  -1.486705  .0419225
lnRailPaxP~1 |  .3854905  .5635581   0.68  0.494  -.7190631  1.490044
     IndQtr4 |  -.139583  .0832198  -1.68  0.093  -.3026909  .0235248
       _cons |  1.856248  1.370682   1.35  0.176  -.8302392  4.542734
------------------------------------------------------------------------------
Instrumented: lnRailKmPerCapAKL
Instruments:  lnRealRailFareAKL lnRailPaxPerCapAKLlag1 IndQtr4
        lnRailKmPerCapAKLlag1 lnRailKmPerCapAKLlag2

. estat endogenous
 Tests of endogeneity
 Ho: variables are exogenous

 Durbin (score) chi2(1)      = .005212 (p = 0.9424)
 Wu-Hausman F(1,14)          = .003649 (p = 0.9527)
```

Appendix E.2.2: Breusch-Godfrey LM test results for Wellington and Christchurch

. * WLT bus model

```
. *
. regress lnBusPaxPerCapWLT lnRealBusFareWLT lnBusPaxPerCapWLTlag1

    Source |     SS      df      MS            Number of obs =      8
-------------+------------------------------   F( 2,    5)    =  22.65
     Model | .017077266   2 .008538633         Prob > F      = 0.0031
  Residual | .001884543   5 .000376909         R-squared     = 0.9006
-------------+------------------------------   Adj R-squared = 0.8609
     Total | .018961809   7  .00270883         Root MSE      = .01941

------------------------------------------------------------------------------
lnBusPaxPe~T |    Coef.   Std. Err.    t    P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRealBusF~T | -.2311504  .096428   -2.40  0.062  -.4790263  .0167256
lnBusPaxPe~1 |  .4934377  .1481612   3.33  0.021   .1125771  .8742983
       _cons |  2.059026  .5948096   3.46  0.018   .5300192  3.588033
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |     chi2        df        Prob > chi2
-------------+-------------------------------------------------------------
      1    |     1.970        1          0.1604
---------------------------------------------------------------------------
            H0: no serial correlation

. *
```

. * WLT rail model 1

```
. *
. regress lnRailPaxPerCapWLT lnRailKmPerCapWLT lnCarOwnPerCapWLT lnRealFuelP
lnRailPaxPerCapWLTlag1
```

```
   Source  |      SS      df      MS            Number of obs =      8
-------------+------------------------------   F( 4,   3)    =  48.04
    Model  | .011034763   4 .002758691         Prob > F      = 0.0047
  Residual | .000172277   3 .000057426         R-squared     = 0.9846
-------------+------------------------------   Adj R-squared = 0.9641
    Total  |  .01120704   7 .001601006         Root MSE      = .00758

------------------------------------------------------------------------------
lnRailPaxP~T |   Coef.    Std. Err.    t   P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~T |  .7424225  .1274827   5.82  0.010    .3367157  1.148129
lnCarOwnPe~T | -.3246443  .1130417  -2.87  0.064  -.6843933  .0351047
 lnRealFuelP |  .1302622  .0338558   3.85  0.031     .022518  .2380064
lnRailPaxP~1 |  .6888637  .1079693   6.38  0.008    .3452573   1.03247
       _cons | -1.375862  .4885455  -2.82  0.067  -2.930632  .1789073
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |     chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |      4.833         1          0.0279
---------------------------------------------------------------------------
            H0: no serial correlation

. *
. *
```

. * WLT rail model 2

```
. *
. regress lnRailPaxPerCapWLT lnRailKmPerCapWLT lnRealIncomePerCap lnRealFuelP
lnRailPaxPerCapWLTlag1

    Source |      SS      df      MS             Number of obs =      8
-------------+------------------------------    F( 4, 3)      =  58.44
     Model | .011065041   4  .00276626          Prob > F      = 0.0035
  Residual |    .000142   3 .000047333          R-squared     = 0.9873
-------------+------------------------------    Adj R-squared = 0.9704
     Total |  .01120704   7 .001601006          Root MSE      = .00688

------------------------------------------------------------------------------
lnRailPaxP~T |   Coef.    Std. Err.   t    P>|t|   [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRailKmPe~T |  .7847138  .1203201   6.52  0.007    .4018016   1.167626
lnRealInco~p | -.2176122  .0666941  -3.26  0.047  -.4298626  -.0053617
 lnRealFuelP |  .1212285  .0298744   4.06  0.027    .0261548   .2163022
lnRailPaxP~1 |  .6900357  .0978743   7.05  0.006    .3785559   1.001516
       _cons | -.4663425  .3343441  -1.39  0.257  -1.530375   .5976895
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
  lags(p) |     chi2        df         Prob > chi2
-------------+-------------------------------------------------------------
    1   |     4.525        1          0.0334
---------------------------------------------------------------------------
            H0: no serial correlation
```

. * CHC bus model 1

```
. *
. regress lnBusPaxPerCapCHC lnBusKmPerCapCHC lnBusPaxPerCapCHClag1 IndQtr4 if tin(1998q1,2008q2)

    Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  3,    38) =  914.75
     Model |  1.49959065     3  .499863549           Prob > F      =  0.0000
  Residual |  .020765116    38   .00054645           R-squared     =  0.9863
-------------+------------------------------           Adj R-squared =  0.9853
     Total |  1.52035576    41  .037081848           Root MSE      =  .02338

------------------------------------------------------------------------------
lnBusPaxPe~C |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnBusKmPer~C |    .076811    .033407     2.30   0.027     .009182    .1444399
lnBusPaxPe~1 |   .8858974   .0465388    19.04   0.000    .7916846    .9801103
     IndQtr4 |  -.0888294   .0086743   -10.24   0.000   -.1063897   -.0712692
       _cons |   .1036139     .03808     2.72   0.010    .0265249    .1807029
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          2.539               1                   0.1110
---------------------------------------------------------------------------
                        H0: no serial correlation

. *
```

. * CHC bus model 2

```
. *
. regress lnBusPaxPerCapCHC lnRealBusFareCHC lnRealFuelP lnBusPaxPerCapCHClag1 IndQtr4 if
tin(2004q1,2008q2)

    Source |       SS       df       MS              Number of obs =      18
-------------+------------------------------           F(  4,    13) =   15.27
     Model |  .022199858     4  .005549965           Prob > F      =  0.0001
  Residual |  .004724335    13   .00036341           R-squared     =  0.8245
-------------+------------------------------           Adj R-squared =  0.7705
     Total |  .026924193    17  .001583776           Root MSE      =  .01906

------------------------------------------------------------------------------
lnBusPaxPe~C |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnRealBusF~C |   -.206604   .0967713    -2.13   0.052    -.4156657    .0024577
 lnRealFuelP |   .1717965   .0847515     2.03   0.064    -.0112979    .3548909
lnBusPaxPe~1 |   .5056908   .2831901     1.79   0.097    -.1061042    1.117486
     IndQtr4 |  -.0557107   .0184441    -3.02   0.010    -.0955568   -.0158646
       _cons |   .1700198   .2919687     0.58   0.570    -.4607401    .8007798
------------------------------------------------------------------------------

. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          0.341               1                   0.5590
---------------------------------------------------------------------------
                        H0: no serial correlation
```

Appendix E.2.3: Stationary test results with JMulTi[17]

```
ADF Test for series:   lnBusPaxPerCapAKL
sample range:         [1998 Q2, 2008 Q2], T = 41
lagged differences:    8
intercept,             time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993),p708, table 20.1
1%        5%      10%
-3.96   -3.41   -3.13
value of test statistic: -3.9818
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)     -0.8238      -3.9818
dx(-1)      0.5067       2.6116
dx(-2)      0.0404       0.2340
dx(-3)     -0.0282      -0.1782
dx(-4)      0.3033       2.3632
dx(-5)     -0.0193      -0.1395
dx(-6)      0.2147       1.6482
dx(-7)      0.0364       0.2701
dx(-8)      0.3516       2.7022
constant    1.2863       4.0563
trend       0.0114       3.6916
RSS         0.0745

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1998 Q4, 2008 Q2], T = 39

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   8
Final Prediction Error:  8
Hannan-Quinn Criterion:  8
Schwarz Criterion:       8

ADF Test for series:   lnRailPaxPerCapAKL
sample range:       [1997 Q2, 2008 Q2], T = 45
lagged differences:    4
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%       5%     10%
-3.96   -3.41   -3.13
value of test statistic: -1.1705
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.0910       -1.1705
dx(-1)    -0.3175       -1.9563
dx(-2)    -0.4611       -2.8412
dx(-3)    -0.2752       -1.7288
dx(-4)     0.2590        1.7231
constant  -0.0080       -0.1765
trend      0.0049        2.4042
RSS        0.2206

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1998 Q4, 2008 Q2], T = 39

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   4
Final Prediction Error:  4
Hannan-Quinn Criterion:  4
Schwarz Criterion:       3
```

[17] For easy reference, stationarity test results in this section are provided in the same order as in appendix B.

```
ADF Test for series:   lnRailPaxPerCapAKL_d1
sample range:       [1997 Q2, 2008 Q2], T = 45
lagged differences:    3
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%     5%     10%
-3.96   -3.41   -3.13
value of test statistic: -3.7994
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)     -1.9114     -3.7994
dx(-1)      0.5364      1.3710
dx(-2)      0.0375      0.1412
dx(-3)     -0.2524     -1.6722
constant    0.0422      2.8258
trend       0.0029      2.5650
RSS         0.2286

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   3
Final Prediction Error:  3
Hannan-Quinn Criterion:  2
Schwarz Criterion:       2

ADF Test for series:   lnBusKmPerCapAKL
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993),p708, table 20.1
 1%      5%     10%
-3.96   -3.41   -3.13
value of test statistic: -2.6721
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)     -0.2348     -2.6721
constant    0.3586      2.7123
trend       0.0028      2.6122
RSS         0.2089

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnBusKmPerCapAKL_d1
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%     10%
-3.96   -3.41   -3.13
value of test statistic: -9.6731
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)     -1.3456     -9.6731
constant    0.0077      0.7717
```

```
trend       0.0009      1.2813
RSS         0.2129

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnRailKmPerCapAKL
sample range:       [1997 Q3, 2008 Q2], T = 44
lagged differences:    4
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p 708, table 20.1
 1%       5%      10%
-3.96   -3.41   -3.13
value of test statistic: -1.1784
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.0860      -1.1784
dx(-1)    -0.2576      -1.5649
dx(-2)    -0.2617      -1.6009
dx(-3)    -0.2400      -1.4424
dx(-4)     0.2831       1.7872
constant  -0.1036      -0.9740
trend      0.0026       2.2581
RSS        0.1101

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   4
Final Prediction Error:  4
Hannan-Quinn Criterion:  3
Schwarz Criterion:       0

ADF Test for series:   lnRailKmPerCapAKL_d1
sample range:       [1997 Q1, 2008 Q2], T = 46
lagged differences:    2
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p 708, table 20.1
 1%       5%      10%
-3.96   -3.41   -3.13
value of test statistic: -6.8570
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -2.2650      -6.8570
dx(-1)     0.8326       3.4581
dx(-2)     0.3620       2.3980
constant   0.0262       2.8769
trend      0.0023       3.1761
RSS       0.1308

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   2
Final Prediction Error:  2
Hannan-Quinn Criterion:  2
Schwarz Criterion:       2
```

```
ADF Test for series:   lnRealBusFareAKL
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993),p708, table 20.1
 1%     5%     10%
-3.96   -3.41   -3.13
value of test statistic: -2.8867
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.2848      -2.8867
constant   0.1554       2.9364
trend      0.0015       2.6974
RSS        0.0501

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnRealBusFareAKL_d1
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%      10%
-3.96   -3.41   -3.13
value of test statistic: -7.8690
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -1.1667      -7.8690
constant   0.0038       0.7340
trend      0.0003       0.8945
RSS        0.0577

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnRealRailFareAKL
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%       5%     10%
-3.96   -3.41   -3.13
value of test statistic: -3.9023
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.4953      -3.9023
constant   0.4486       3.9120
trend      0.0017       3.0767
RSS        0.0536
```

```
OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnRealRailFareAKL_d1
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%     10%
-3.96   -3.41   -3.13
value of test statistic: -8.2054
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -1.2143      -8.2054
constant   0.0020       0.3608
trend      0.0001       0.2008
RSS        0.0685

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnCarOwnPerCapAKL
sample range:       [1998 Q3, 2008 Q2], T = 40
lagged differences:    8
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%     10%
-3.96   -3.41   -3.13
value of test statistic: -3.8546
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.4227      -3.8546
dx(-1)     0.2871       1.9069
dx(-2)     0.0823       0.5253
dx(-3)     0.0352       0.2237
dx(-4)     0.3762       2.5831
dx(-5)    -0.0878      -0.5769
dx(-6)     0.0132       0.0819
dx(-7)    -0.0307      -0.1920
dx(-8)     0.5031       3.1450
constant  -0.2798      -3.8342
trend      0.0008       3.0295
RSS        0.0008

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   8
Final Prediction Error:  8
Hannan-Quinn Criterion:  8
Schwarz Criterion:       4
```

```
ADF Test for series:   lnRealFuelP
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%      10%
-3.96   -3.41   -3.13
value of test statistic: -2.1174
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.2064      -2.1174
constant   1.0081       2.1399
trend      0.0026       2.4979
RSS        0.1450

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA

sample range:       [1999 Q1, 2008 Q2], T = 38

optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0

ADF Test for series:   lnRealFuelP_d1
sample range:       [1996 Q4, 2008 Q2], T = 47
lagged differences:    1
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%      10%
-3.96   -3.41   -3.13
value of test statistic: -6.0859
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -1.2892      -6.0859
dx(-1)     0.2874       1.9241
constant   0.0129       1.4825
trend      0.0010       1.4877
RSS        0.1468

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1999 Q1, 2008 Q2], T = 38
optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   1
Final Prediction Error:  1
Hannan-Quinn Criterion:  1
Schwarz Criterion:       0
ADF Test for series:   lnRealIncomePerCap
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:    0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%      5%      10%
-3.96   -3.41   -3.13
value of test statistic: -2.8636
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -0.2989      -2.8636
constant   0.5772       2.8885
trend      0.0018       2.8325
RSS        0.0059

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1999 Q1, 2008 Q2], T = 38
```

```
optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0
ADF Test for series:   lnRealIncomePerCap_d1
sample range:       [1996 Q3, 2008 Q2], T = 48
lagged differences:      0
intercept, time trend
asymptotic critical values
reference: Davidson and MacKinnon (1993), p708, table 20.1
 1%       5%     10%
-3.96   -3.41   -3.13
value of test statistic: -7.4326
regression results:
---------------------------------------
variable   coefficient  t-statistic
---------------------------------------
 x(-1)    -1.1312      -7.4326
constant   0.0057       2.9100
trend      0.0000       0.1627
RSS        0.0068

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1999 Q1, 2008 Q2], T = 38
optimal number of lags (searched up to 10 lags of 1. differences):
Akaike Info Criterion:   0
Final Prediction Error:  0
Hannan-Quinn Criterion:  0
Schwarz Criterion:       0
```

Appendix E.2.4: Cointegration test results with JMulTi

#Auckland bus models 1 and 2

```
Johansen trace test for: lnBusKmPerCapAKL lnBusPaxPerCapAKL lnCarOwnPerCapAKL
unrestricted dummies:   D[1999 Q3] D[1999 Q4] D[2000 Q1] D[2000 Q2] D[2000 Q3] D[2000 Q4] D[2003
Q3] D[2003 Q4] D[2004 Q1] D[2004 Q2] D[2004 Q3] D[2004 Q4] S[1999 Q3] S[2003 Q3]
restricted dummies:    t[1999 Q3] t[2003 Q3]
sample range:       [1997 Q3, 2008 Q2], T = 44
included lags (levels):   6
dimension of the process: 3
trend and intercept included
seasonal dummies included
response surface computed:
-----------------------------------------------
 r0 LR       pval     90%    95%    99%
-----------------------------------------------
 0  296.74  0.0000   70.88  74.99  83.10
 1  126.84  0.0000   45.16  48.52  55.23
 2  57.78   0.0000   22.57  25.04  30.12

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1998 Q3, 2008 Q2], T = 40
optimal number of lags (searched up to 10 lags of levels):
Akaike Info Criterion:    6
Final Prediction Error:  10
Hannan-Quinn Criterion:   6
Schwarz Criterion:        6
```

#Auckland rail models

```
Johansen trace test for: lnRailPaxPerCapAKL lnRealRailFareAKL
unrestricted dummies:   D[2003 Q3] D[2003 Q4] D[2004 Q1] D[2004 Q2] D[2004 Q3] D[2004 Q4] D[2005
Q1] D[2005 Q2] D[2005 Q3] D[2005 Q4] S[2003 Q3]
restricted dummies:    t[2003 Q3]
sample range:       [1998 Q3, 2008 Q2], T = 40
included lags (levels):  10
dimension of the process: 2
trend and intercept included
seasonal dummies included
response surface computed:
```

```
-----------------------------------------------
 r0 LR      pval      90%    95%    99%
-----------------------------------------------
 0  107.88  0.0000  34.19  37.14  43.10
 1  32.99   0.0002  16.65  18.81  23.33

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1998 Q3, 2008 Q2], T = 40
optimal number of lags (searched up to 10 lags of levels):
Akaike Info Criterion:   10
Final Prediction Error:   8
Hannan-Quinn Criterion:  10
Schwarz Criterion:       10
```

#Christchurch bus model 1998Q1

```
Johansen trace test for: lnBusKmPerCapCHC lnBusPaxPerCapCHC
unrestricted dummies:   D[1998 Q1] D[2003 Q3] S[1998 Q1] S[2003 Q3]
restricted dummies:    t[1998 Q1] t[2003 Q3]
sample range:       [1997 Q2, 2008 Q2], T = 45
included lags (levels):  1
dimension of the process: 2
trend and intercept included
response surface computed:
-----------------------------------------------
 r0 LR      pval     90%    95%    99%
-----------------------------------------------
 0  34.94  0.2738  40.64  43.93  50.53
 1  7.95   0.8840  20.10  22.54  27.61

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [1997 Q3, 2008 Q2], T = 44
optimal number of lags (searched up to 2 lags of levels, max lag adjusted):
Akaike Info Criterion:   1
Final Prediction Error:  1
Hannan-Quinn Criterion:  1
Schwarz Criterion:       1
```

#Christchurch bus model 2004Q1

```
Johansen trace test for: lnBusPaxPerCapCHC lnRealBusFareCHC
sample range:       [2004 Q4, 2008 Q2], T = 15
included lags (levels):   3
dimension of the process: 2
trend and intercept included
response surface computed:
-----------------------------------------------
 r0   LR    pval     90%    95%    99%
-----------------------------------------------
 0  34.57  0.0024  23.32  25.73  30.67
 1  10.32  0.1145  10.68  12.45  16.22

OPTIMAL ENDOGENOUS LAGS FROM INFORMATION CRITERIA
sample range:       [2006 Q3, 2008 Q2], T = 8
optimal number of lags (searched up to 10 lags of levels):
Akaike Info Criterion:   3
Final Prediction Error:  2
Hannan-Quinn Criterion:  3
Schwarz Criterion:       3
```

Appendix E.3: A second comment. Andrew Coleman, Motu Economic and Public Policy Research, September 2010

Introduction

In an earlier comment, I queried the extent to which various parameter estimates in this report could be relied upon to accurately represent the true underlying values. The basis of the criticisms was twofold:

1 Whether the model was sufficiently well identified that the coefficient estimates could be used to calculate elasticities

2 Whether the data series were sufficiently long to accurately estimate the parameters.

The main issue surrounding identification concerned the 'service provision' variable, *S*. The key results presented in the paper are that the point estimate of the short run elasticity of service provision for Auckland bus services is 0.46, (a 1% increase in the number of miles of bus services leads to a 0.46% increase in the miles driven by passengers) and for rail services is in the range 0.88–0.99. There is a sizeable international literature examining the importance of the service variable on the number of trips taken, much of which is discussed in the original paper. The Auckland estimates are in line with those of the international literature. (For instance, in a British review by Paulley et al (2006), the average estimate is 0.4 in the short run and 0.7 in the long run. In a different setting, a review of substantial bus service expansions in US cities by Pratt (2000) has an elasticity of 0.9.)

The main question raised in the original critique is the extent to which the service provision variable S is exogenous to the number of passenger miles travelled. *If* it is not exogenous, a single equation of the type estimated in the paper will not be well specified, in the sense that elasticities cannot be estimated from the coefficients of the structural equation using a standard formula. The critique suggested two possible problems:

1 The single equation parameter estimates may be biased, because they ignore feedback effects from increased passenger demand to increased service provision (simultaneous bias problem).

2 The elasticity *calculations* for other variables such as fares and car-ownership will be incorrect because they ignore the effect of induced supply provision on final demand quantities.

The original critique probably underplayed the difficulty of this issue, as it focused on the contemporaneous feedback loops. If the service provision variable responds with a lag to demand factors, the analysis is more complex, as changes in demand factors (such as fares) may have an immediate effect on demand quantities, and then a lagged indirect effect as a supply provision changes. One would imagine that fare elasticities are underestimated in these circumstances: a rise in fares might lead to a decrease in patronage (the direct effect) followed by a subsequent decrease in service levels and a subsequent decrease in patronage (the indirect effect). This issue is covered in more detail in the mathematical appendix to this comment: the appendix outlines some of the difficulties of estimating elasticities when the supply of services is responsive to the demand for services, either with or without a lag.

There are two basic classes of difficulties. The first concerns the appropriate treatment of the supply of service variable, if it is not exogenous to demand factors. If it is not exogenous, many of the elasticities such as the elasticity of passenger demand to fares will be calculated incorrectly because they omit the indirect effect of fares on the supply of services, and from the supply of services to passenger numbers. The appendix shows how these calculations need to be modified *if* the supply of service variable is affected by the demand for bus services, or by other factors that affect passenger demand.

Second, the example in the appendix shows that the exact form of the elasticity formula used to calculate elasticities from structural model coefficients depends on the exact structure of the model, including the appropriate lag structure. If the econometric model is not correctly specified, eg because the wrong number of lags is included, or because a contemporary endogenous variable is wrongly treated, then the formula for calculating the elasticities will not be correct. The most appropriate way to deal with this problem is to allow for model uncertainty: that is to estimate a sequence of different models and for each model calculate the appropriate elasticities. This provides a greater range of uncertainty as it allows for not just parameter uncertainty but model uncertainty. For example, the paper could estimate a pair of structural equations with a more general lag structure as well as the model estimated in the paper, and calculate and present the elasticities for this model as well.

Issue 1: Transport services offered variable

IV estimates of the original relationship

In response to questioning, the authors re-estimated the original equations using instrumental variables techniques as well as OLS. The aim was to find an instrument for the supply of services variable that could be used to deal with the simultaneous equation problem. As the appendix makes clear, the formula used to calculate elasticities would only be valid if the service variable was exogenous to the current value and lags of 1) the number of passengers carried and 2) other variables that determine passenger demand such as fares. The authors attempted to establish exogeneity by using lags of the supply of service variables as instruments. This is certainly the correct approach, and if the response of 'service supply' to contemporaneous variables was the main problem, these additional regressions would help.

The authors used the IV for the Auckland bus and Auckland rail service, and estimated the equations over different sub-periods. For each equation, they reported the OLS results and then the IV results using either 1 lag or 2 lags of the service variable as an instrument.

Q is the number of passenger trips per period (per capita)

S is the km of services of transport offered (per capita)

F is the price of the fare

C is the number of cars per capita

I is real income per capita

O is the oil price per capita.

Auckland buses

Auckland buses 1999–2008

The first stage regression is either $S_t = \alpha_0 + \alpha_1 S_{t-1} + e_t$ or $S_t = \alpha_0 + \alpha_1 S_{t-1} + \alpha_2 S_{t-2} + e_t$

Theses regressions suggest that nearly 90% of the variation in the supply of services variable S_t is explained by its first lag, and that the second lag only makes a marginal improvement.

The second stage IV equation (coefficient estimates displayed below) leads to a significant change in the estimated coefficients. There is little difference between IV1 and IV2; consistent with the second lag adding little to the first stage regression, it adds little to the second stage regression when included in IV2. Further, the test for exogeneity can only just be rejected at conventional levels.

	S_t	C_t	Q_{t-1}	
OLS	0.42 (0.09)	-1.14 (0.68)	0.59 (0.07)	
IV1	0.73 (0.23)	-2.74 (1.34)	0.43 (0.13)	
IV2	0.72 (0.23)	-2.71 (1.32)	0.44 (0.13)	

Superficially, the IV regression in this case is broadly consistent with the OLS regression, although the parameter estimate on the supply variable is much higher. This is supportive of the initial OLS approach, although it suggests the standard errors on the initial estimate may be low.

Auckland buses 2003–2008

The analysis is repeated over a short period. The first stage regressions are similar, and again the second lag S_{t-2} adds little to the explanatory power of the regression. The second stage results are quite strange, however:

	S_t	C_t	Q_{t-1}	O_t
OLS	0.46 (0.08)	-1.95 (0.51)	0.37(0.07)	0.20 (0.06)
IV1	0.75 (0.40)	-2.83 (1.31)	0.24 (0.18)	0.09 (0.17)
IV2	0.47 (0.23)	-1.97 (1.32)	0.36 (0.13)	0.20 (0.06)

These results are strange because of the way the second IV estimates are almost identical to the OLS estimates and are so different from the IV 1 estimates. Moreover, the IV 1 estimates over this shortened period (2003–2008) are almost identical to the IV estimates of the longer period (1999–2008). It is not at all clear what this means.

These two sets of estimates both suggest that the point estimate of the supply of service elasticity is in the range 0.42 to 0.75. As noted above, either figure is consistent with international evidence.

The IV regressions suggest contemporaneous simultaneity is not a major problem. I am happy to accept this result. However, it does not necessarily solve the whole problem about how one calculates elasticities from the estimated parameters. There are still two issues, as discussed in the mathematical appendix.

The first issue is that even if there is no lagged response, the elasticity estimate may still need adjustment. Equation 1 in the mathematical appendix outlines the two equation system (variables Q and S are in logarithms; the system has no lags). X is a vector of variables that may affect both Q and S, Z1 are variables that only affect Q directly, and Z2 are variables that only affect S directly.

$$\begin{aligned} Q_t &= \alpha_1 S_t + \beta_1 X_t + \gamma_1 Z_{1t} \quad + \qquad + e_{1t} \\ S_t &= \alpha_2 Q_t + \beta_2 X_t + \qquad + \delta_2 Z_{2t} \quad + e_{2t} \end{aligned}$$

If X is price, the coefficient β_1 measures the direct effect of price on the quantity of passenger services. The reduced form elasticity comprises the direct and indirect effects: by equation 2a, the elasticity is

$$\frac{\beta_1 + \alpha_1\beta_2}{1-\alpha_1\alpha_2} = \beta_1 + \frac{\alpha_1\beta_2 + \alpha_2\beta_1}{1-\alpha_1\alpha_2}$$

If the supply variable S_t is exogenous to Q_t, this means α_2=0. The equation establishes that α_1 is in the range 0.4 to 0.75. It doesn't report whether β_2=0, however: and the way the software package operates, it estimates the coefficient β_2 in the first stage regression. If one wishes to use one of these equations to estimate the elasticities (which is of course the point of the paper), the full equation for S_t should be reported, and it would be useful to report whether the indirect component of the elasticity formula is zero or not.

The second and probably most important issue is that the supply of service variable may respond with a lag to the other variables including demand. This is quite plausible: one can imagine the bus company delaying a quarter or so before deciding to change service levels in response to fluctuations in fares or passenger demand. If this is the case, the lack of a contemporaneous relationship between the supply of service variable and other variables does not mean there is no relationship. Moreover, as indicated in the appendix, the formula for calculating elasticities from the parameter estimates is quite different in this case.

For future work, and with more data, I can imagine a researcher simultaneously estimating equations examining the determinants of both the number of passengers and the supply of services, with a lengthy series of lags or a partial adjustment model. Unless the determinants of the supply of services is known, it is difficult to be sure that the elasticities affecting the number of passengers are properly calculated. In fairness to the current research, the extent that this problem affects elasticity estimation has not been fully recognised in most of the literature. Nonetheless, the determinants of the supply of services are crucial, and if they are not truly exogenous to demand factors, demand elasticities will be biased.

Auckland trains

The paper estimates two different equations over the period 2003–08.

Equation 1

Even though the first stage regression seems to fit reasonably well, the IV estimates (particularly IV1) are all over the place and provide little confidence that the instruments are working particularly well.

	S_t	$Fares_t$	$Income_t$	Q_{t-1}
OLS	1.05 (0.22)	-0.99 (0.22)	1.56 (0.56)	0.25 (.14)
IV1	-1.13 (11)	-0.02 (4.95)	0.33 (6.33)	1.61 (6.8)
IV2	0.47 (0.72)	-0.73 (0.38)	1.23 (0.69)	0.61 (0.06)

Equation 2

These estimates are similar, and also suffer from an extremely poorly performing IV1 equation. Ignoring these equations, the IV2 and OLS equations are consistent and suggest that the supply of service variable is broadly exogenous to the number of passenger variables. This makes sense, given the big increase in rail services over the period.

	S_t	$Fares_t$		Q_{t-1}
OLS	0.93 (0.25)	-0.70 (0.23)		0.42 (.15)
IV1	-2.43 (21)	-0.49 (7.8)		2.4 (12.7)
IV2	0.99 (0.96)	-0.72 (0.38)		0.39 (0.56)

I have less concern about the way the rail elasticities are calculated, because it seems much more likely the increase in rail services occurred for exogenous political reasons. Nonetheless, in future work the relationship between rail service provision and other variables could be explicitly explored.

Issue 2: Spurious regressions

The author has addressed many of my earlier concerns about whether the error processes in the estimated regressions are serially correlated. The short length of the series means that these tests cannot be considered conclusive. Nonetheless, in general the results suggest that serial correlation among the error terms of the regressions is not a major problem for the Auckland results. However, they strengthen the questions raised about the Auckland bus results.

Auckland buses

The Auckland bus result is odd. According to the adjusted Dickey Fuller test results in the appendix E.2.3, bus passenger numbers and cars do not have a unit root (around a trend), but bus services and bus fares do have a unit root.

Variable	Coefficient on unit root test (=0 if unit root)	Test statistic Critical value ~ -3.4	Unit root in levels
Bus passengers Q	-0.82	-3.95	No
Bus services S	-0.23	-2.67	Yes

Variable	Coefficient on unit root test (=0 if unit root)	Test statistic Critical value ~ -3.4	Unit root in levels
Bus fare	-0.28	-2.88	Yes
Cars	-0.42	-3.85	No
Fuel price	-0.20	-2.11	Yes
Income	-0.29	-2.9	Yes

The cointegration test is for bus passengers, bus services and cars. Since only one of these variables is a unit root, the test doesn't make much sense: you need at least two variables with a unit root to test for cointegration. The test will be able to reject the hypothesis that the residuals have a unit root, because the dependent variable does not have a unit root.

The interesting thing from this analysis is that it suggests the bus service provided variable and the bus passenger variable have quite different behaviour. This is difficult to see in the graphs in section 5.1.1.1 figure 5.2, except for the period 1996–1999 when there is a substantial decline in bus passengers. If the results are correct – and with small samples it is very difficult to be convinced that they are correct – there are two issues. First, strictly speaking, it raises questions as to the long-run relationship between bus passengers and bus service provision (because if services is a unit root series it can drift anywhere, but passenger numbers will not). Second, it suggests that the relationship between services and passenger numbers takes place gradually through time. To expand on the latter point, given the short series, it probably is not true that passenger numbers does not have a unit root but service provision does have a unit root. Rather, it is probably that they both have a unit root or both don't have a unit root, but have a quite different serial persistence. It would appear that service provision is much more persistent than passenger demand. If this is the case, changes in one variable may be related to changes in the other variable, but it takes a long time for the effect to take place. Again this means that the formula for calculating elasticities from parameter estimates has to take into account this lagged adjustment structure.

Auckland rail

The results for Auckland rail are more promising. Both the passenger numbers and the services provided variables seem to have unit roots. The errors of the standard regression seem to have little serial correlation, indicating that the two series are cointegrated, and thus linked in the long run. I am reasonably satisfied with this regression.

I don't quite understand the results of the Johansen trace test. I would have thought it would have been between passenger numbers and service provided, and perhaps some other variables. Instead it is between passenger numbers and rail fares. This makes little sense, given that rail fares does not obviously have a unit root. Nonetheless, this is not particularly important, given the results on the residuals of the main regression reported in appendix E.2.1, p12.

Christchurch buses

The results seem to suggest that the Christchurch model has problems with serial correlation in the errors. 1) Over the full sample, the hypothesis that there is first order serial correlation in the errors of the main regression can only just be rejected at conventional levels in appendix E.2.1.) The cointegration test fails to reject the hypothesis that the passenger numbers and service provided variables are cointegrated.

These two results suggest that the main regression for Christchurch buses may have some problems; as much to the point, it raises the likelihood that different calculations are needed to calculate the elasticities from the parameters estimates, because of the lag structure of the results. The authors acknowledge these difficulties.

Conclusion

The author of the paper has responded reasonably to my initial questions by doing some additional econometric tests to try and ascertain the extent to which the supply of services variable is exogenous. But overall, I remain sceptical. The problem is not so much with the test procedures: it is whether the modelling approach can reasonably be expected to deliver accurate estimates of various demand elasticities, given the shortness of the series and the uncertainty about the true underlying economic model. It must be stressed that the formulae to calculate the elasticities are only correct if the econometric model accurately captures the underlying economic mechanisms. If it proves that the supply of services does not depend on the various factors that affect demand, the formula will be correct. While I can believe there is little relationship in the short term – the bus companies take a while to adjust service levels to demand – it seems reasonable to believe that these adjustments are made in the longer term. If so, future research should aim to find a strategy that could tackle this issue head on, either by adjusting for any endogenous response from demand factors to the supply of services, or by seeking data for which the supply of services is more likely to be exogenous.

One possibility for future research is to concentrate on much smaller area units than cities, and to separate the data according to the extent service delivery has changed over time. For instance, it may be possible to find individual bus routes for which frequency and service delivery have not changed over the period (or have only changed once or twice), so that the effects of fare changes can be ascertained without having to worry about changes in service delivery. Or one could concentrate on services where the major changes in service levels have been for reasons that seem unrelated to immediate demand conditions: the extension of the rail network, or the North Shore dedicated bus route seem to be likely candidates. In either case, the aim would be to try and select data for which the econometric issues are minimised.

Mathematical appendix: Identification and parameter estimates

This appendix explores the mathematical links between the variables used in a system of equations that links bus patronage, service levels, and various supply and demand factors.

Variable definitions

Q is the number of passenger trips per period (per capita)

S is the km of services of transport offered (per capita)

X are variables that directly affect both S and Q (eg price of fares, price of oil)

Z_1 are variables that have a direct effect on Q only (eg income per capita)

Z_2 are variables that have a direct effect on S only (eq political decision to expand railways)

(i) Contemporaneous effects

If there were no lagged effects, we could write a simultaneous equation system as

$$\begin{aligned} Q_t &= f(S_t, X_t, Z_{1t}) + e_{1t} \\ S_t &= g(Q_t, X_t, Z_{2t}) + e_{2t} \end{aligned}$$

This pair of equations simply suggests that (a) an increase in service provision leads to more passengers and (b) an increase in demand leads to an increase in bus services (eg more frequent bus trips, or an extension of the service network).

In linear terms, these form the *structural form* equations

$$\begin{aligned} Q_t &= \alpha_1 S_t + \beta_1 X_t + \gamma_1 Z_{1t} \quad + \qquad + e_{1t} \\ S_t &= \alpha_2 Q_t + \beta_2 X_t + \qquad + \delta_2 Z_{2t} \quad + e_{2t} \end{aligned} \qquad \text{(Equation 1)}$$

These can be rearranged in matrix form

$$\begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix} \begin{bmatrix} Q_t \\ S_t \end{bmatrix} = \begin{bmatrix} \beta_1 & \gamma_1 & 0 \\ \beta_2 & 0 & \delta_2 \end{bmatrix} \begin{bmatrix} X_t \\ Z_{1t} \\ Z_{2t} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \qquad \text{(Equation 1a)}$$

When the structural form is estimated, S is included as an explanatory variable for Q

From equation 1 the direct effect of a change in X on Q, conditioning on S, is β_1.

We can invert the left-hand side matrix to get the *reduced form*, in which Q is estimated as an indirect function of all the other variables:

$$\begin{aligned} \begin{bmatrix} Q_t \\ S_t \end{bmatrix} &= \begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \beta_1 & \gamma_1 & 0 \\ \beta_2 & 0 & \delta_2 \end{bmatrix} \begin{bmatrix} X_t \\ Z_{1t} \\ Z_{2t} \end{bmatrix} + \begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \\ &= \frac{1}{1-\alpha_1\alpha_2} \begin{bmatrix} \beta_1 + \alpha_1\beta_2 & \gamma_1 & \alpha_1\delta_2 \\ \beta_2 + \alpha_2\beta_1 & \alpha_2\gamma_1 & \delta_2 \end{bmatrix} \begin{bmatrix} X_t \\ Z_{1t} \\ Z_{2t} \end{bmatrix} + \begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \end{aligned} \qquad \text{(Equation 2)}$$

Equation 2 shows that in addition to the direct effect of X on Q, two indirect effects are possible:

1 X has a direct effect on S, which then affects Q

2 X affects Q, which has an effect on S, which feeds back and affects Q further.

These indirect effects are captured in the reduced form: the total effect of a change in X on Q is

$$\frac{\beta_1 + \alpha_1\beta_2}{1-\alpha_1\alpha_2} = \beta_1 + \frac{\alpha_1\beta_2 + \alpha_2\beta_1}{1-\alpha_1\alpha_2} \qquad \text{(Equation 2a)}$$

Clearly the direct and total (direct plus indirect) effects will rarely be equal.

(ii) Lagged effects

Suppose it takes a while for additional demand Q to affect the supply of bus services S. Further, suppose there is a lagged effect from the supply of services to the demand for those services, and, more generally, there is persistence in the supply and demand for services. Then we can modify the matrix form of the equations (1a) to:

$$\begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix} \begin{bmatrix} Q_t \\ S_t \end{bmatrix} = + \begin{bmatrix} \delta_1 & \phi_1 \\ \delta_2 & \phi_2 \end{bmatrix} \begin{bmatrix} Q_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} \beta_1 & \gamma_1 & 0 \\ \beta_2 & 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} X_t \\ Z_{1t} \\ Z_{2t} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \qquad \text{(Equation 3)}$$

or ....

$\Omega_0 Y_t = \Omega_1 Y_{t-1} + \beta\chi_t + e_t$ where $Y_t = \begin{bmatrix} Q_t \\ S_t \end{bmatrix}$ and $\chi_t = \begin{bmatrix} X_t \\ Z_{1t} \\ Z_{2t} \end{bmatrix}$ (Equation 3a)

Rearranging

$$Y_t = \Omega_0^{-1}\Omega_1 Y_{t-1} + \Omega_0^{-1}\beta\chi_t + \Omega_0^{-1}e_t$$
$$\Rightarrow \left(I - \Omega_0^{-1}\Omega_1 L\right)Y_t = \Omega_0^{-1}\beta\chi_t + \Omega_0^{-1}e_t \qquad \text{(Equation 4)}$$
$$\Rightarrow Y_t = \left(I - \Omega_0^{-1}\Omega_1 L\right)^{-1}\left(\Omega_0^{-1}\beta\chi_t + \Omega_0^{-1}e_t\right)$$

$$\Rightarrow Y_t = \left(I + \Omega_0^{-1}\Omega_1 L + (\Omega_0^{-1}\Omega_1)^2 L^2 + (\Omega_0^{-1}\Omega_1)^3 L^3 ....\right)\left(\Omega_0^{-1}\beta\chi_t + \Omega_0^{-1}e_t\right) \qquad \text{(Equation 5)}$$
$$= \Omega_0^{-1}\beta\chi_t + \Omega_0^{-1}\Omega_1\Omega_0^{-1}\beta\chi_{t-1} + (\Omega_0^{-1}\Omega_1)^2\Omega_0^{-1}\beta\chi_{t-2} .... + \left(I - \Omega_0^{-1}\Omega_1 L\right)^{-1}\Omega_0^{-1}e_t$$

The first term of equation 5 is the same as the term in the reduced form (equation 2):

$$\Omega_0^{-1}\beta = \begin{bmatrix} 1 & -\alpha_1 \\ -\alpha_2 & 1 \end{bmatrix}^{-1}\begin{bmatrix} \beta_1 & \gamma_1 & 0 \\ \beta_2 & 0 & \lambda_2 \end{bmatrix} = \frac{1}{1-\alpha_1\alpha_2}\begin{bmatrix} \beta_1 + \alpha_1\beta_2 & \gamma_1 & \alpha_1\delta_2 \\ \beta_2 + \alpha_2\beta_1 & \alpha_2\gamma_1 & \lambda_2 \end{bmatrix}$$

This is the contemporaneous effect of χ on Q and S, taking into account the indirect feedback interaction.

The second term is the lagged effect of X on Q and S, one period afterwards:

$$\Omega_0^{-1}\Omega_1\Omega_0^{-1}\beta = \begin{bmatrix} \delta_1 + \alpha_1\delta_2 + \alpha_2\phi_1 + \alpha_1\alpha_2\phi_2 & \phi_1 + \alpha_1(\delta_1 + \phi_2) + \alpha_1^{\,2}\delta_2 \\ \delta_2 + \alpha_2(\delta_1 + \phi_2) + \alpha_2^{\,2}\phi_1 & \phi_2 + \alpha_1\delta_2 + \alpha_2\phi_1 + \alpha_1\alpha_2\delta_1 \end{bmatrix}\begin{bmatrix} \beta_1 & \gamma_1 & 0 \\ \beta_2 & 0 & \lambda_2 \end{bmatrix}$$

Subsequent terms are the lagged effects of χ on Q and S after 2, 3, 4... periods.

In general, the effect of a temporary increase in X at time t on Q and S will decay over time in the way described by equation 5. However, if there is a permanent increase in χ, the effects on Q and S will accumulate through time. The total effect of a permanent increase is $\left(I - \Omega_0^{-1}\Omega_1\right)^{-1}\Omega_0^{-1}\beta\Delta\chi$

This can be expanded to

$$\frac{\theta}{\lambda}\begin{bmatrix} 1-\phi_2 & \alpha_1 + \phi_1 \\ \alpha_2 + \delta_2 & 1-\delta_1 \end{bmatrix}\beta\Delta\chi$$

where $\theta = \dfrac{1}{1-\alpha_1\alpha_2}$ and $\lambda = 1 - \theta(\delta_1 + \phi_2 + \alpha_2\phi_1 + \alpha_1\delta_2 + \phi_1\delta_2 - \delta_1\phi_2)$

(Note that if $\Omega_1 = [0]$, $\lambda = 1$ and the solution is the same as above).

In this case, the long term effect of a change in X on Q is:

$$\Delta Q = \frac{\theta}{\lambda}\left[\beta_1(1-\phi_2) + \beta_2(\alpha_1 + \phi_1)\right]\Delta X$$

The methodology used in the paper calculates a long run elasticity as $\varepsilon = \dfrac{\beta_1}{1-\delta_1}$

Recall that δ_1, δ_2 are the lagged effect of passenger demand on passenger demand and service provision, and ϕ_1, ϕ_2 are the lagged effect of service provision on passenger demand and service provision. Even if ϕ_1, ϕ_2 =0, the long-term effect of a change in X is Q is

$$\Delta Q = \frac{\beta_1 + \beta_2\alpha_1}{1 - \alpha_1\alpha_2 - \delta_1 - \alpha_1\delta_2}\Delta X$$

This is not equal to the long run elasticity calculated in the paper except under very particular circumstances. It follows that if there are additional lags in the true relationship, the long run elasticities

$$=\left[\frac{\beta_1}{1-\delta_1}+\frac{\alpha_1[\beta_2(1-\delta_1)+\beta_1(\delta_2+\alpha_2)]}{(1-\alpha_1\alpha_2-\delta_1-\alpha_1\delta_2)(1-\delta_1)}\right]\Delta X$$

will be incorrectly estimated if the formula in the paper is used. The problem is more acute the more the supply of bus services responds to the demand for services or to the underlying causal factors (such as fares), and less acute the less supply responds to these factors. Indeed, one of the few circumstances where the estimator will not be biased occurs if $\beta_2=\alpha_2=\delta_2=0$; that is, the supply of services is unresponsive to contemporary demand, contemporary factors X that also affect demand (such as fares, or petrol prices), or lagged demand. In short, the estimator is likely to be well behaved only if service provision is independent of contemporary and lagged demand factors.

The above example shows the case if there is a single lag linking the supply of bus services and the demand for bus services. Different formula would need to be calculated if the relationship is more complex, eg if it takes two periods for supply to respond to demand, or if the relationship is based on a partial adjustment process.

Elasticity estimation

There are two basic classes of difficulties. The first concerns the appropriate treatment of the supply of service variable, if it is not exogenous to demand factors. If it is not exogenous, many of the elasticities such as the elasticity of passenger demand to fares will be calculated incorrectly because they omit the indirect effect of fares on the supply of services, and from the supply of services to passenger numbers. The above calculations show how these calculations need to be modified *if* the supply of service variable is affected by the demand for bus services, or by other factors that affect passenger demand, for particular lag structures.

This critique can be expressed differently. When the supply of services variable is related to the passenger demand variable or to other factors that influence passenger demand, there is a difference between the elasticity calculated from a single structural equation of passenger demand versus demand factors and the supply of passenger services, and the reduced form equation that excludes the supply of services variable. The elasticities based on a single passenger demand equation will not be correct if both passenger demand and the supply of services depend on the same variables, or if there are feedback loops from passenger demand to the supply of services. If the supply of services variable is exogenous, the elasticities calculated from the structural and the reduced form equations will (asymptotically) be the same.

Secondly, the example shows that the exact form of the elasticity formula that should be used to calculate elasticities from structural model coefficients depends on the exact structure of the model, including the appropriate lag structure. If the econometric model is not correctly specified – say because the wrong number of lags is included, or because a contemporary endogenous variable is wrongly treated – then the formula for calculating the elasticities will not be correct. The most appropriate way to deal with this problem is to allow for model uncertainty: that is to estimate a sequence of different models and for each model calculate the appropriate elasticities. This provides a greater range of uncertainty as it allows for not just parameter uncertainty but model uncertainty. An example would be to estimate a pair of structural equations with a more general lag structure as well as the model estimated in the paper, and present the elasticity estimates from the two equations.

References

Paulley, N, R Balcombe, R Mackett, H Titheridge, J Preston, M Wardman, J Shires and P White (2006) The demand for public transport: the effect of fares, quality of service, income, and car ownership. *Transport Policy 13*, no.4: 295–306

Pratt, Richard (2000) *Traveller response to transportation system charges*, Transport Research Board web document 12.