Drivers of demand for transport
October 2013

Aleksandra Simic
Robert Bartels
Frontier Economics, Australia

NZ Transport Agency research report 534
Contracted research organisation – Frontier Economics, Australia

ISBN 978-0-478-40783-9 (electronic)
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

Simic, A and R Bartels (2013) Drivers of demand for transport. *NZ Transport Agency research report 534.* 52pp.

Frontier Economics, 280 George Street, Sydney NSW 2000, Australia
Telephone 61 2 8224 9710; facsimile 61 3 9620 4499
Email aleksandra.simic@frontier-economics.com.au , www.frontier-economics.com

Frontier Economics was contracted by the NZ Transport Agency in 2012 to carry out this research.

Keywords: econometric analysis, road transport demand forecasting, time series data analysis

An important note for the reader

Acknowledgements

The authors would like to thank NZ Transport Agency, Ministry of Transport (New Zealand) and Statistics NZ for providing the data and their support throughout this project. We would also like to thank Paul Clark of the NZ Transport Agency; Professor David Hensher of the Institute of Transport and Logistics Studies at the University of Sydney; Tantri Tantirigama and Joanne Leung of the Ministry of Transport, and John Créquer, Andrew Richens, and Alistair Smith of Statistics NZ for their invaluable advice and guidance.

Contents

Executive summary

Frontier Economics (Frontier) was engaged by the New Zealand Transport Agency ('Transport Agency') to assist in identifying a 'best fit' methodology for assessing the historical relationship in New Zealand between: 1) economic activity and road freight activity; and 2) income growth and passenger vehicle travel. The Transport Agency would like to use these models to develop long-term road transport forecasts.

The present report represents an initial step in this modelling exercise. Our objective was to inform the development of demand models by exploring both data availability and analytical steps that should be considered when developing demand models. The actual development of the models and the estimation of the key parameters are outside the scope and will be undertaken under a separate research project.

Our findings, mainly based on literature reviews, are organised in terms of the key analytical steps we recommend the Transport Agency consider when developing the models: data preparation, selection of candidate demand drivers, data investigation and model selection, and model testing and validation. We limited our review to recent studies which could be classified into three categories: 1) studies that investigate the degree of (de)coupling of transport demand (freight in particular) and economic activity; 2) studies that investigate the use of different modelling techniques to forecast aggregate transport demand; and 3) recent models developed by equivalent organisations to the Transport Agency in other countries.

To model freight demand, we recommend the Transport Agency develop models by vehicle type using RUC-based vehicle kilometres travelled (VKT) data. The reason is that different types of trucks are used for different purposes and their use may, at least to some extent, proxy freight transport demand by different industries and hence reflect the changing structure of the economy (ie shift from goods-producing sectors to service sectors).[1] McKinnon (2007) states it is now widely accepted that the increase in industrial output, rather than gross domestic product growth, is the main driving force behind freight transport.

To model passenger vehicle travel, we recommend the Transport Agency investigate developing a panel model using its regional VKT data. The reason is that VKT data on passenger vehicle travel is available for only a relatively short period of time (at best 11 years) which severely limits any analysis of how the effect of income on travel demand has changed over time. Income levels, however, differ across regional council areas in New Zealand, and this difference can provide some insight into the relationship between income and travel demand.

Our preliminary data investigations indicate that VKT data is volatile. Therefore, consideration needs to be given to whether observed fluctuations and level shifts are due to factors such as missing telemetric data or additions of new road segments/bypasses, and how best to address such data issues.

Based on our literature reviews, we provide a summary of the most frequently used econometric techniques for modelling road transport demand. All the models belong to the same broad 'family' of models and hence one cannot recommend *ex ante* which of them should be used. Model selection is an integral part of model development and should be guided by data properties and diagnostic checks.

Some recent studies have explicitly recognised that there are possibly feedback loops between transport demand and some of its drivers, and have dealt with this issue by using, what is known in the econometrics

[1] The Transport Agency does not have time series data on tonne kilometres travelled by commodity type.

literature as, the two-stage least squares (2SLS) approach. To our knowledge, very few studies have used a more complex modelling framework in which the 2SLS method is combined with the seemingly unrelated regression method to attempt to improve the efficiency of the estimated parameters.

Developing econometric models will not be a trivial exercise as there are a number of technical issues discussed in the report that will need to be considered. Furthermore, it should be recognised that developing the models will be an iterative process – some analytical steps will need to be repeated several times before a preferred model is selected. Even after the preferred models have been selected, we recommend the Transport Agency continue to check its forecasting performance on a regular basis.

At the end of the report, we provide two suggestions for research projects that could complement the Transport Agency's work in developing econometric demand modes.

Abstract

Frontier Economics (Frontier) was engaged by the New Zealand Transport Agency (Transport Agency) to assist in identifying a 'best fit' methodology for assessing the historical relationship in New Zealand between: 1) economic activity and road freight activity; and 2) income growth and passenger vehicle travel. The Transport Agency would like to use these models to develop long-term road transport forecasts.

The present report represents an initial step in this modelling exercise. The objective was to inform the development of demand models by exploring both data availability and analytical steps that should be considered when developing demand models. The actual development of the models and the estimation of the key parameters are outside the scope and will be undertaken under a separate research project.

The findings, mainly based on literature reviews, are organised in terms of the key analytical steps we recommend the Transport Agency consider when developing the models: data preparation, selection of candidate demand drivers, data investigation and model selection, and model testing and validation.

1 Introduction

Frontier Economics (Frontier) was engaged by the New Zealand Transport Agency (Transport Agency) to assist in identifying a 'best fit' methodology, and appropriate data sources and variables, for assessing the historical relationship in New Zealand between:

- economic activity and road freight activity, and
- income growth and passenger vehicle travel.

1.1 Background

The motivation for this research project came out of recent work undertaken by the New Zealand Institute for Economic Research (NZIER) on behalf of the Transport Agency.

The Transport Agency commissioned the development of a National Land Transport Demand Model (NLTDM) as part of the 2011/12 research programme. The model projects future freight and passenger vehicle demand based on a set of scenario-derived assumptions.

As a first step towards specifying the model's parameters, ie determining which demand relationships matter and the strengths of these relationships, the NZIER undertook a literature review. They found that a large number of parameter estimates in the reviewed studies were either poorly estimated, of limited relevance to New Zealand, or, in some cases, absent.

One of the issues raised by the NZIER was the apparent lack of statistical rigour in the models used to estimate the effect of income and economic activity on transport demand. For example, the NZIER noted that the demand models commonly assumed a constant elasticity of demand without any statistical testing of the appropriate functional form.

The outcome of the NZIER's review was that, although the reviewed literature pointed to the existence of a relationship between economic activity and freight demand, as well as between income growth and passenger vehicle travel, there seemed to be little or no empirical evidence of how these might have changed over time.

1.2 Scope of work

The objective of the current research project was to inform the development of econometric (ie regression-based) models that the Transport Agency could use to estimate the extent to which the relationship between economic activity and road freight activity, and between income and passenger vehicle travel, has changed over time, the magnitude of those changes and possible causes. The actual development of those demand models and the estimation of the key parameters would be undertaken under a separate research project.

During the inception meeting in January 2013, Frontier received additional instructions/clarifications:

- In addition to estimating the historical relationship between variables of interest, the Transport Agency was also interested in using the econometric models to produce long-term road freight activity and passenger vehicle travel forecasts.

- The demand models should estimate the relationships of interest at the national level (rather than the regional or origin-destination level).
- The Transport Agency viewed the development of econometric models as a separate/independent research activity from the NLTDM (ie our research and recommendations should not be guided or constrained in any way by the NLTDM). The NLTDM was to be used by the Transport Agency for scenario analysis at regional/industry level, while any econometric models developed on the basis of our research would be used by the Transport Agency to investigate long-term trends at the aggregate national level.
- The main criterion in selecting a 'best fit' methodology for developing demand models should be data availability. This will allow the Transport Agency to implement the recommended methodology with the currently available data collected on a regular basis by different New Zealand government departments.
- Additional criteria included:
  - Accuracy and lack of bias. This criterion relates to: a) model selection (eg if the data are non-stationary, selecting a model that addresses this issue); b) functional form selection (eg log vs linear); and c) variable selection (ie models should take into consideration all relevant demand drivers).
  - Validation and testing. Models should allow for assessment of statistical significance of demand drivers, diagnostic checking and assessment of forecasting performance.
  - Transparency and reproducibility. It should be possible for a third party to implement and review models without a major effort.

1.3 How the report is organised

The report is organised as follows. In chapter 2 we first describe different methods for estimating transport demand. We then review several recent studies which examine the relationship between transport demand and economic activity in New Zealand. These studies were a starting point in our investigations as they provided an insight into data availability and the existing models. We then review available data sources and data characteristics. In chapter 3 we review the relevant literature, focusing on the analytical approaches recently used to investigate the relationship between economic activity and transport demand (freight and passenger), and the analytical issues that researchers have had to address. In chapter 4 we further explore the characteristics of the datasets that could be used to develop demand models and also investigate possible drivers of transport demand in New Zealand. We provide our recommendations for demand modelling in chapter 5.

2 Scoping analysis

One of the Transport Agency's research agendas is to use currently available data to:

- understand what are the key economic/income drivers of road transport demand and whether and how their impact on demand has changed over time
- obtain long-run income elasticities of freight and passenger road transport demand
- develop road transport demand forecasts.

The present report represents an initial step in this broad research agenda. Its objective is to inform the development of demand models by exploring both data availability and analytical steps that should be considered when developing demand models. The actual development of the models and the estimation of the key parameters will be undertaken as a separate research project.

We start by briefly describing different types of demand models and comment on the data needed to estimate them. We then discuss a few recent research projects undertaken by the Transport Agency which consider some of the same analytical issues as the present analysis. We conclude by considering data availability.

2.1 Modelling demand

Broadly speaking, transport demand studies can be categorised into three groups:

- **Econometric models using continuous variables.** These models try to explain changes in demand by looking at what happens to the total quantity of demand for travel conditional on a number of drivers. Drivers are selected on the basis of general economic theory. These models generally do not fully specify the economic process that gives rise to demand; hence they are also referred to as 'reduced form' models. Implicit in this type of model are insights from economic theory that demand is a function of own price, price of substitutes and an economic activity variable. The popularity of this type of model is to some extent attributable to the low data requirements. These models are usually estimated using panel or time series data, with data aggregated over passengers, shippers, commodities, etc. Depending on one's research objective and data availability, the level of spatial aggregation may be national, regional or origin-destination.[2]
- **Discrete choice models.** These models focus on individual decision makers (passengers, firms, shippers). They investigate agents' travel-choice behaviour (ie the probability of making a particular choice) based on the attributes of various modes of transport and agents' characteristics. These models capture the situation in which an agent is making discrete (ie all or nothing) decisions. To undertake this type of modelling, comprehensive survey data is required (see, for example, Hensher et al 2013).

[2] A recent example of origin-destination demand modelling is a passenger rail demand model developed for the UK Department of Transport, Transport Scotland and the Passenger Demand Forecasting Council. Using panel data on travel between distinct origin and destination points, rail demand is modelled as a function of train fares, fuel price, car driving time between origin and destination points, proportion of households without a car at the origin and socio-economic variables (Oxera and Arup 2010).

- **Hybrid models.** A hybrid model combines a continuous demand and a discrete choice model to form, what is considered to be, a more complete and accurately specified demand system. An example of this type of model is a so-called discrete/continuous model for car ownership and car use. These models simultaneously explain: 1) whether a household will own a vehicle and the vehicle type; and 2) conditional on the vehicle ownership choice, the number of kilometres driven by year. De Jong et al (2004) explain that the basic idea behind this modelling approach is that decisions of households on car ownership and car use are strongly interrelated and should therefore be studied together.[3]

Commenting on the choice between aggregate and disaggregate models, Oum et al (2000) note that the choice depends largely on the purpose of the study and the availability and cost of obtaining the data. They state:

> *When the purpose of the model is to forecast aggregate traffic volumes, it is natural and even preferable to use aggregate data ... if the purpose of the model is to simulate how decision makers such as shippers or travellers would respond to changes in policy or managerial control variables, a disaggregate discrete choice model is more attractive. A disaggregate model requires an extensive and expensive data base. The prudent course of action, therefore, is to use an aggregate demand model to generate the preliminary results before undertaking a larger discrete choice study if necessary.*

2.2 Recent New Zealand transport models

Over the past few years, the Transport Agency has commissioned a number of research projects (some of them peer reviewed) on modelling the relationship between transport demand and economic activity. These studies were a starting point in our investigation as they provided an insight into data availability and the existing transport demand models in New Zealand. The reviewed studies all developed econometric models using data aggregated at the national or regional levels.

Conder (2009) developed a car ownership model for New Zealand using aggregate national time series data, with gross domestic product (GDP) per capita, an index of car-purchase prices, and a time trend as explanatory variables. The model was estimated using the ordinary least squares (OLS) regression technique on annual data going back to 1973.[4] According to the authors, a preferred modelling approach would be to develop what they refer to as a ‘static disaggregate car ownership and use model’. Such a model would have several modules, each estimated separately, with most modules requiring household level data. One module would use household level data (eg household income, household structure) to predict the car-ownership pattern. Another module would use household level data to predict car use by household type (eg number of employed household members, proximity to public transport, etc) and car type (eg first or second car). Another module would use time series data to estimate the elasticity of car use with respect to fuel price. However, the authors expressed reservations as to whether the data for estimating such a model was readily available, and noted that the time and resource requirements would be considerable.

---

[3] For a recent discussion of the ‘hybrid’ modelling approach and its application, see Truong and Hensher (2012).

[4] There are two parts to the model: car-ownership saturation level and path to saturation. The car-ownership saturation level is based on the observed car-owning rate of the population with relatively high income. The path to saturation is estimated using the OLS model.

Wang (2011) undertook an analysis of factors influencing bus and rail public transport patronage in New Zealand. Econometric models were developed to forecast public transport demand in three major urban centres (Auckland, Wellington and Christchurch). Using a partial adjustment (PA) model, the relationship between patronage (measured as the total number of bus/train trips per capita) was modelled as a function of:

- service level (expressed in bus/train kilometres per capita)
- real fare (expressed as real revenue per passenger)
- real disposable income per capita
- car ownership (expressed as number of cars per capita)
- real fuel price.

The models were estimated using quarterly data going back to the mid to late 1990s (depending on the city). The author found the effect of different drivers (both in terms of magnitude and statistical significance) varied across the three cities. For example, bus fares were found to affect bus demand in Wellington and Christchurch, but not in Auckland. Income was found to have a positive effect on Auckland rail patronage, but a negative effect on Wellington rail patronage.[5]

Since public transport can be viewed as a substitute for (at least some types of) private car transport, these results suggest that a prudent course of action when modelling passenger car demand would be to investigate regional demand patterns, even if the primary interest is in forecasting travel demand at the national level. Regional differences can be investigated/modelled by including region-specific dummy variables. The estimated income effects also point to a frequently overlooked limitation of econometric models, which is that they are not particularly well suited to dealing with changes in the quality of service.

Victorio (2011) used quarterly national-level traffic volume data to investigate drivers of freight and passenger road travel. Freight demand was modelled as a function of GDP, retail volume (in tonnes) and diesel prices. Light vehicle travel (of which the majority were passenger cars) was modelled as a function of total national spending, total national employment and petrol prices. Quarterly data for the period 2002 to 2010 was used to estimate three models: an OLS model, an autoregressive model, and an error-correction (EC) model.[6] These models were at an early stage of development, and the preliminary findings were not very satisfactory. The main reason, we believe, is that the modelling period used for the analysis was too short, especially considering a number of recent changes that affected road freight transport demand (ie significant increases in diesel price and road user charges).

2.3 Available road transport data

Upon completing this preliminary review of recent New Zealand transport models, we contacted data owners at the Transport Agency and the New Zealand Ministry of Transport (MoT) to investigate data availability. Both the Transport Agency and MoT collect data on road transport demand, although there are

[5] The author postulates that the positive relationship between income and rail patronage in Auckland is due to the opening of a new train station in 2003 near the city's central business district, which attracts a high proportion of high income patronage.

[6] We explain these models in the next chapter.

differences in how transport demand is measured (traffic volume vs vehicle kilometres travelled), data aggregation level (ie whether the measure captures travel on state highways only or travel on all roads) and data frequency (ie whether data is available on a monthly or annual basis).

Tables 2.1 and 2.2 provide a summary of the Transport Agency and MoT data on road freight transport and passenger car transport. The data summarised in these tables can be used to develop econometric demand models. In the next two sections we explore what analytical steps should be considered when developing such models.

Table 2.1 Freight demand data – heavy vehicle travel

Demand measure	Benefits/drawbacks	Data availability
Traffic volume[(a)]	Covers travel on state highways only Some data available from 1989 Data available by region Data available *monthly* The measure does not include any indication of the amount of goods transported Heavy vehicles are defined as vehicle weighing more than 3.5 tonnes and therefore may include buses	Available from the Transport Agency Continuous data available for about 100 telemetric sites from 1989 Continuous data for a large sample of telemetric sites available from 2002 The most comprehensive (in terms of network coverage) continuous dataset available from 2005
Vehicle kilometres travelled (VKT)[(b)]	Covers travel on state highways only Data available by region Data available *annually* The measure does not include any indication of the amount of goods transported Heavy vehicles are defined as vehicles weighing more than 3.5 tonnes and therefore may include buses	Available from the Transport Agency Continuous data available for about 100 telemetric sites from 1989 Continuous data for a large sample of telemetric sites available from 2002 The most comprehensive (in terms of network coverage) continuous dataset available from 2005
Vehicle kilometres travelled (VKT)[(c)]	Covers travel on state highways and local roads Data available by vehicle type and weight Data available *quarterly* Based on road user charge (RUC) licences purchased	Available from the Transport Agency Annual data available from 1987
Vehicle kilometres travelled (VKT)[(d)]	Covers travel on state highways and local roads Data available *annually* The measure does not include any indication of the amount of goods transported Data available for trucks and trailers	Available from the MoT Annual data available from 1992
Tonnes kilometres travelled (TKM)	Covers travel on state highways and local roads Data available *annually* Calculations are based on RUC and require an assumption regarding average truck load Data available for trucks and trailers	Available from the MoT Data available for the period 1992 to 2011 Due to simplification of RUC scheme in 2012, the series is discontinued [(e)]

[(a)] A measure of how many times telemetric counters around the country were driven over.

[(b)] Using traffic volume data from telemetric sites, annual average daily traffic (AADT) is calculated for every road section of the entire state highway network. The VKT for a given road section is calculated by multiplying the AADT for that road section by the road section's length. Summing across road sections and multiplying by the number of days in a year, gives the total kilometres travelled in a given year.

[c] RUC distance licences are purchased in units of 1000km, with charges based on vehicle characteristics.

[d] Derived from the change in certificate of fitness (CoF) odometer readings

[e] Based on communications with Stuart Badger, Principal Scientist at Ministry of Transport (MoT), on 1 August 2013, the Ministry is now using a new approach combining weigh-in-motion (WIM) and CoF data to produce TKM estimates. As of the writing of this report, the new data series is under review and not yet publicly available.

Table 2.2 Passenger vehicle travel data

Demand measure	Benefits/drawbacks	Data availability
Traffic volume[a]	Covers travel on state highways only Some data available from 1989 Data available by region Data available *monthly* Data are for all light vehicles (majority of which are private passenger cars)	Available from the Transport Agency Continuous data available for about 100 telemetric sites from 1989 Continuous data for a large sample of telemetric sites available from 2002 The most comprehensive (in terms of network coverage) continuous dataset available from 2005
Vehicle kilometres travelled (VKT)[b]	Covers travel on state highways only Some data available from 1989 Data available by region Data available *monthly* Data are for all light vehicles (majority of which are private passenger cars)	Available from the Transport Agency Continuous data available for about 100 telemetric sites from 1989 Continuous data for a large sample of telemetric sites available from 2002 The most comprehensive (in terms of network coverage) continuous dataset available from 2005
Vehicle kilometres travelled (VKT)[c]	Covers travel on state highways and local roads Data available *annually* Data are for passenger vehicles	Available from the MoT Annual data available from 2001
Passenger kilometres travelled	Covers travel on state highways and local roads Data available *annually* Derived by multiplying VKT by an estimated average number of passengers per vehicle	Available from the MoT Annual data available from 2001

[a] A measure of how many times telemetric counters around the country were driven over.

[b] Using traffic volume data from telemetric sites, AADT is calculated for every road section of the entire state highway network. The VKT for a given road section is calculated by multiplying the AADT for that road section by the road section's length. Summing across road sections and multiplying by the number of days in a year, gives the total kilometres travelled in a given year.

[c] Derived from the change in warrant of fitness odometer readings.

3 Literature review

Our scoping analysis indicated that, with the currently available data and without substantial efforts in data collection and manipulation, econometric models could be developed and estimated to investigate the historical relationship between economic activity and road transport demand. Such models could be used to forecast future road transport demand.

Given the scope of our study, our literature review is limited to studies that have estimated econometric models. This review is not intended to be comprehensive; our goal is to gain some insight into the type of analytical approaches recently used to investigate the relationship between economic activity and transport demand (freight and passenger), and the analytical issues that researchers have had to address.

In addition to reviewing recent New Zealand transport studies, we have focused on the following streams of literature:

- Studies that investigate the degree of (de)coupling of transport demand (freight in particular) and economic activity. These studies are concerned with whether the historical relationship between economic activity and transport demand has changed over time and the underlying drivers of the change. In these studies particular attention is paid to how economic activity should be defined, and how the magnitude of the estimated income elasticity depends on the chosen definition. This literature is not explicitly focused on developing demand forecasts. However, estimated elasticities from these studies can be used to derive forecasts (eg by multiplying estimated coefficients by the projected values for the demand drivers).
- Studies that investigate the use of different modelling techniques to forecast aggregate transport demand, and explore the benefits and drawbacks of each. These studies use, what are considered to be, state-of-the-art time series techniques. The focus is more on the technical properties of different models and estimators and less on how the economic activity variable included in these models should be defined.

In addition, we also consider relevant recent models developed by equivalent organisations to the Transport Agency in other countries.

Because the objective of the present research was to inform the development of models that would enable the Transport Agency to estimate the relationship between transport demand and economic/income drivers, the literature review findings have been organised in terms of the key analytical steps to be considered when developing the models. In particular:

- selection of the demand variable
- selection of explanatory variables
- data testing and model specification
- model validation.

3.1 How should road transport demand be defined?

Demand for transport is a derived demand, and as such depends on demand for other goods and activities. This poses a challenge in transport modelling as it is not obvious what measure (or definition) of demand should be used (unlike, for example, in modelling demand for soft drinks). While in practice the choice of demand measure may be limited by data availability, the decision should be driven by the policy question one is trying to address, with transport demand defined in a manner that most directly captures the response one is trying to estimate. If one is interested in how changes in the economy affect road freight demand, then using tonne kilometres travelled is a more direct measure than vehicle count, for example, as it better reflects the notion that freight demand is driven by the demand for commodities being transported. If, however, one is interested in assessing the effect of an increase in toll road charges, then vehicle count (or traffic flow) may be a more appropriate measure of demand.

In the reviewed literature, Agnolucci and Bonilla (2009), who investigated whether decoupling between economic activity and freight demand occurred in the UK, define transport demand in terms of freight tonne kilometres. They note that other measures used in the literature on decoupling included tonnes lifted and revenue share of tonne kilometres. Richard Paling Consulting (2008) forecast sectoral freight transport demand in terms of both tonnage and vehicle-tonne kilometres travelled. Other measures of freight demand used in the literature include traffic volume counts and vehicle kilometres (or miles) travelled. Tonne kilometres travelled (TKM) appears to be the most widely used measure of freight transport demand.

With respect to passenger vehicle demand, measures of demand used in the literature we reviewed included traffic volume counts, vehicle miles travelled (VMT), and petrol consumption. Kennedy and Wallis (2007), for example, used both traffic volume counts on state highways and total petrol purchases as measures of passenger car demand when investigating the impact of changes in fuel prices on passenger car travel.

3.2 How should economic activity be defined?

Studies that focus on the long-term relationship between road freight and economic activity, most commonly define economic activity in terms of real GDP and real GDP per capita. McKinnon (2007) and Agnolucci and Bonilla (2009) argue that GDP is not the most appropriate measure to use in this type of investigation as it does not adequately capture changes in the structure of the economy.

At the core of this argument is an understanding that both the volume of economic activity (as measured by aggregate GDP) and the changing structure of the economy (ie the increasing share of the service sector relative to the goods-producing sectors) will have an effect on freight transport demand. As changes in the structure of the economy are likely to follow long-term secular trends, they will underpin/ drive long-term changes in freight transport demand.

Researchers have explored this issue in a number of ways, with their choice of approach being largely dictated by data availability. Agnolucci and Bonilla (2009), who investigated the relationship between road freight transport and economic activity in the UK, defined economic activity in terms of gross value added (GVA) from goods-producing sectors (namely, manufacturing, mining and quarrying, construction, and the electricity, gas and water supply sectors).[7] They also constructed an alternative measure of economic activity by adding the value of imports to the GVA from goods-producing sectors.

Using data for Sweden, Andersson and Elger (2012) formally investigated whether GDP, GVA for goods-producing sectors, or trade variables better explained changes in aggregate freight demand (road, rail, sea and the total across the three modes). They found that trade variables had the largest effect on freight transport demand in the short run, while GDP was the main driver in the long run.

One way to take account of the changes in the structure of the economy is to undertake modelling at the industry or commodity level. Shen et al (2009) used commodity-level data on freight demand (road plus rail) in the UK to estimate eight econometric models – a separate econometric model for each of the seven commodity groups investigated and one model in which freight demand and output values were aggregated across all commodities.[8] Sectoral economic activity was defined in terms of total production output plus imports (in tonnes). They found that the estimates of income elasticity (ie elasticity with respect to economic activity) varied widely across commodity groups, indicating that the structure of the economy has a material effect on estimated industry-wide income elasticities. For example, using the PA model they estimated the income elasticity of transport demand for the food, drink and agricultural products to be 3.3, while for metal and ore products, the income elasticity was estimated to be close to 1. In the aggregate model, the estimated elasticity represents a weighted average elasticity across the seven commodity groups, and was estimated to be 1.5.

Data on transport activity at the industry/commodity level, however, is not commonly available. MoT and the Transport Agency have started to address this data deficiency by commissioning the *National transport demand study* (Paling 2008). The study used a bottom-up approach to forecast freight demand by commodity type. Namely, the study estimated transport intensity by commodity group and then used those estimates, together with projections of output for each modelled commodity group, to forecast transport demand. As the collected data series started in 2008, the data does not yet lend itself to investigating the historical relationship between economic activity and transport demand.[9]

Pickrell et al (2012), tasked with developing an econometric forecasting model for the US Federal Highway Administration, modelled road transport demand by vehicle type (eg light-duty vehicles, single unit truck and combination truck) and by road type (eg urban interstate, urban other, rural intestate, rural arterial, etc). Although not explicitly stated, the logic behind this approach is that different types of trucks are used for different purposes and hence their use may, at least to some extent, proxy freight transport demand by different industries (ie smaller trucks are usually used for shorter distances and time-sensitive deliveries and hence may be more sensitive to changes in retail sales, while larger trucks are used for long-distance heavy freight and hence may be more responsive to changes in goods-producing sector output). The results from Pickrell et al (2012) national-level models indicate that the sensitivity of the demand for road freight transport with respect to GDP is likely to vary widely across different types of trucks. For example, the estimated coefficient on the GDP variable in the single unit truck demand model

[7] It is not clear whether the agricultural sector is excluded intentionally from the definition of goods-producing sectors.

[8] The commodity groups investigated are: food, drink and agricultural products; coal and coke, petroleum and petroleum products; metals and ores; construction; chemicals and fertilisers; and others (including manufactured goods, machines, etc.).

[9] An earlier study by Bolland et al (2005) undertook a similar exercise. However, the methodologies used in the two studies differ and, hence, the results are not directly comparable. For example, in the 2005 study, about 60% of all commodities were not assigned to a specific commodity group (they were aggregated to a group named ‘Other’). In Kennedy and Wallis (2007) only 30% of the total commodity output was assigned to a commodity group ‘Other’.

was 0.28, while the analogous coefficient in the combination truck model was 0.59. Alternative variables of economic activity used were retail sales (value and as a percentage of GDP), value of durable and non-durable goods, and electronic and mail order sales (as a percentage of all retail sales).[10]

In a recent inquiry into infrastructure pricing, the Australian Productivity Commission (PC) estimated long-run price and income elasticities for rail and road freight transport. To capture the diversity of road freight transport, the PC estimated separate models by load type (bulk and non-bulk), area (urban and non-urban), and vehicle type (rigid and articulated[11]). In addition to GDP, each demand model was estimated using its own and substitute-mode's freight price (in cents/TKM), and a trade variable (ie import or export value). The results were reported for only a few selected models; nevertheless, they indicated that long-run income elasticity estimates (ie elasticity with respect to GDP) varied considerably across the models. The PC estimated the long-run income elasticity to be 0.74 for total road freight, 0.56 for non-urban road freight and 1.11 for articulated freight. The range of import elasticities across the three models was somewhat narrower, from 0.62 (total road freight) to 0.94 (non-urban road freight).

West et al (2011) sought to identify a set of regularly generated, well-documented and easily obtainable variables with high statistical significance in explaining the variability of freight demand in the US, with a goal of developing a freight demand forecasting model for infrastructure investors and capital planners. Among the economic drivers investigated were real GDP, housing starts (thought to be a good proxy for economic optimism and consumption), the Industrial Production Index (which measures the relative output of manufacturing, mining and energy producers), import and export values, retail sales, employment (total and in the wholesale industry) and the inventory to sales ratio (thought to provide a counter-cyclical measure of the quantity of goods consumed). They encountered the challenge, which we discuss in the subsection on data issues, that most of the drivers were highly correlated with each other.

Unlike with freight demand modelling, passenger vehicle travel models do not seem to have attracted much discussion about how the economic activity variable should be specified. Frequently used measures are GDP per capita, disposable income per capita and employment figures.

3.3 What other explanatory variables should be considered?

Pickrell et al's (2012) work (ie developing an econometric forecasting model for the US Federal Highway Administration) bears similarities to models the Transport Agency is considering developing. For that reason, in table 3.1, we summarise the explanatory variables used (or tested for inclusion) in their roadway transport demand models.[12] This table is included as a reference for the Transport Agency when considering which drivers to test for inclusion in their models (in addition to an economic/income variable). However, additional drivers not on this list could also be considered, for example a dummy variable for 2011 to pick up the effect of the Christchurch earthquake.

[10] The model is still in development phase and results for different model specifications are not publicly available.

[11] As explained in the PC report, articulated trucks consist of a prime mover plus a semi trailer. This is the most widely used type of vehicle for non-urban freight transport in Australia.

[12] In appendix A we provide summaries of some of the papers we reviewed.

In his introductory econometrics book Wooldridge states that:

> *[s]ometimes it is natural for the partial effect, elasticity, or semi-elasticity of the depended variable with respect to an explanatory variable to depend on the magnitude of yet another explanatory variable.*[13]

This suggests that interaction terms, derived by multiplying two (or more) explanatory variables, should also be tested for inclusion in regression models to capture any variation in the extent to which a particular driver affects transport demand.

For example, in addition to including personal income and gasoline price when modelling passenger vehicle travel, one may also want to include an interaction term between the two variables. The interaction term, if statistically significant, will capture the extent to which the responsiveness of consumers to changes in gasoline price increases or decreases as income changes. Also, in models with regional dummy variables, the income variable could be interacted with the dummy variables to investigate whether and to what extent responsiveness to income changes varies across regions.

When selecting drivers for inclusion in its demand models, the Transport Agency will need to consider whether any of the selected drivers is endogenous. We discuss this issue in section 3.4.

Table 3.1 Examples of explanatory variables tested by Pickrell et al (2012)

Variable type	Light-duty vehicle	Single-unit trucks	Combination trucks
Dependent variable	Total annual vehicle miles travelled (VMT)	Total annual VMT	Total annual VMT
Economic activity/ income measures	Total GDP Disposable personal income Median household income	Total GDP Real value of durable plus nondurable goods Real retail sales Real retail sales (% of GDP) Electronic and mail-order sales (as % of retail sales) Real value of service sector (% of GDP)	Total GDP Value of durable plus nondurable goods Value of durable plus nondurable goods (% of GDP) Imports plus exports of goods (% of GDP) US industrial production
Demographic characteristics	Total population % of population aged 20–65 years Number of households Average persons per household % of households that are families % of families with children <18 years %t of population in urban areas Regional population variables	None	None

[13] For an introduction to models with interaction terms see chapter 6 in Wooldridge (2005).

Variable type	Light-duty vehicle	Single-unit trucks	Combination trucks
Cost of driving	Gasoline price per gallon Fuel economy (MPG) Fuel cost per mile driven	Diesel price per gallon Single unit truck miles per gallon (MPG) Fuel cost per mile Driver wages	Diesel price per gallon Combination truck MPG Fuel cost per mile Driver wages
Vehicle price	New vehicle price index Used vehicle price index Vehicle parts and price index New vehicle price index/consumer price index New vehicle real sales price	Producer price index (transportation equipment) New vehicle price index	Producer price index (transportation equipment) New vehicle price index
Road supply	Total road-miles Road-miles per vehicle	Total highway-miles Total highway-miles per all vehicles Highway-miles in urban areas % of population in urban areas	Total highway miles per all vehicles Total highway-miles Total public road-miles
Employment	Total employment Labour force participation rate (%) Employed persons per household	None	None
Transit service	Vehicle-miles of bus and rail transit service Vehicle-miles of rail transit service Number of cities with rail transit service	None	None

Source: Pickrell et al (2012).

3.3.1 Multicollinearity

One of the challenges faced by researchers is that many of the potential transport demand drivers (eg GDP, population, trade variables, petrol price) tend to be highly correlated.

When two or more explanatory variables are highly correlated (ie when there is multicollinearity), the parameters on the individual explanatory variables cannot be estimated with a high degree of precision. If multicollinearity is present, we may find, for example, that the relationship between transport demand and some of the explanatory variables is not statistically significant (and may even have the wrong sign).[14]

[14] It is good practice to calculate the variance inflation factor (VIF) for the candidate explanatory variables when checking for multicollinearity (rather than just rely on the correlation coefficients). A rule of thumb is that VIF values greater than 10 warrant further examination. We have been informed that the Transport Agency uses the Stata statistical package; the Stata command for calculating VIF values after running a regression is ‘estat vif’.

Greene (1997) suggests that an obvious practical remedy is to exclude the variable suspected of causing the problem, but warns that this course of action may come at a high cost. If the excluded variable does actually belong in the demand model, then the remaining coefficients will be biased, possibly severely so. Another approach is to use principal component analysis (PCA). This is a statistical technique that allows one to derive one or more summary measures (ie principal components) from the original set of explanatory variables. Simply stated, each principal component is a weighted average of two or more variables from the original set of explanatory variables.

One problem with this approach is that there is no clear economic interpretation of the estimated coefficients on the principal components. As Greene (1997) points out '[h]ow do we interpret the price elasticity plus minus twice the income elasticity?'.

Another problem is that there is no guarantee that the linear combination of the explanatory variables (ie the structure of a particular principal component) will stay the same into the future. This is especially problematic if the model is to be used for forecasting purpose.

West et al (2011) explored using PCA as a way to combine the explanatory powers of multiple demand drivers. They grouped drivers into those that measured employment, consumption, production, commodity prices and foreign exchange; and then used the variables within each group to derive principal components. For example, one of the candidate models for predicting road freight demand included a commodity price principle component, a consumption principal component, and a dummy variable indicating the start of the North American Free Trade Agreement. They found that for modelling truck and rail demand, the PCA led to improved statistical results compared with simpler regression models (ie higher R-squared and better forecasting performance).

3.4 Model selection

3.4.1 Static vs dynamic models

Regression models which do not include a lagged dependent variable or lagged explanatory variables are referred to as 'static' models. Implicit in such models is the assumption that the observed demand is in equilibrium. These models have been criticised on the ground that they fail to take into consideration that consumers take time to adjust their consumption fully to a particular change, such as a price or policy change. It is argued that static models are likely to produce intermediate-run elasticities and that dynamic models are more appropriate for capturing the full time frame of responses, from short-run to long-run (Oum et al 2000).

One approach to deal with the issue of multi-period responses is to include a lagged dependent (ie demand) variable as an additional explanatory variable in the demand model. This type of model is known as a partial adjustment (PA) model.

Both static and PA models can be regarded as special cases of the more general autoregressive distributed lag (ARDL) model, which can accommodate different lag structures on both the dependent and the independent variables.

The general form of an ARDL model is presented in equation 3.1:

$$Y_t = \alpha + \sum_{i=1}^{I} \beta_i Y_{t-i} + \sum_{s=1}^{S} \sum_{j=0}^{J} \gamma_{sj} X_{s,t-j} + \varepsilon_t \qquad \text{(Equation 3.1)}$$

where α, β_i , and γ_{si} are coefficients to be estimated and ε_t is a random error term.

A static model is the case where I=0 and J=0, while a PA model is the case with only J=0 and I≠0. If all variables are expressed in the logarithmic form, then the estimated coefficients on the continuous variables can be interpreted as elasticities. In general, the estimated elasticities from a static model are considered to be long-run elasticities, although some critics argue that the estimated coefficients are more representative of intermediate-run elasticities. PA models and ARDL models produce both short-run elasticities and (after some algebra) long-run elasticities.

Let's suppose we have a simple log-log PA model, with truck tonne-kilometres travelled (VKT) expressed as a function of last period's VKT and this period's GDP. The estimated coefficients on the two explanatory variables are β_1 and γ_1, respectively. The coefficient on the GDP variable, γ_1 gives us the short-run elasticity of VKT with respect to GDP. The long-run elasticity can be derived from the expression: $\gamma_1/(1-\beta_1)$.[15]

3.4.2 Non-stationarity

It is well understood by econometricians that using non-stationary time series data[16] in regressions may lead to spurious results – results which erroneously indicate, through misleading values of R-squared and t-statistics, that there is a meaningful relationship among the regression variables. It is therefore a standard procedure to test time series data for stationarity (ie to perform what is known as the 'unit root' test) and, if the variables are found to be non-stationary, to choose an appropriate statistical approach to mitigate against spurious regression findings.

When time series data is found to be non-stationary, it is a standard procedure in econometrics to test whether the related series are co-integrated; that is, whether there exists a stationary long-run linear combination of these non-stationary variables. If, in fact, the variables are co-integrated, the presence of non-stationary variables is not considered to be problematic.

Non-stationary series could be made stationary by using differencing. However, by estimating the model with differenced variables, one would lose information from economic theory concerning the long-run equilibrium properties of the data. A growing literature on the subject has shown that cointegration and error correction (EC) methods are appropriate and useful ways to analyse trending variables (eg Greene 1997). In addition to accounting for the non-stationarity of the variables, these techniques provide a

[15] One drawback of PA and ARDL models is that standard errors for the calculated long-run elasticities are not readily available since the elasticities are derived as ratios of coefficient estimates. This issue, and the various methods for addressing it, are more widely discussed in the time series literature than in the applied transport studies that use these modelling techniques.

[16] Non-stationary data behaves similarly to a random walk with drift. In such a process the current period's value is equal to the last period's value plus a random error. As a result, the mean and the variance of the data are not constant over time.

framework for estimating long-run equilibrium relationships while allowing for short-term dynamics in the modelled relationship.[17, 18]

The general form of an EC model is presented in equation 3.2:

$$\Delta Y_t = \alpha + \sum_{i=1}^{I} \beta_i \Delta Y_{t-i} + \sum_{s=1}^{S}\sum_{j=0}^{J} \gamma_{sj} \Delta X_{s,t-j} + \lambda\left(Y_{t-1} - \sum_{s=1}^{S} \delta_s X_{s,t-1}\right) + \varepsilon_t \quad \text{(Equation 3.2)}$$

where α, β_i, γ_{sj}, δ_s , and λ are coefficients to be estimated and ε_t is a random error term. The estimated coefficients on the differenced variables (denoted with the Δ operator) represent short-run elasticities, while the estimated coefficients on the variables in levels (the parameters in the parenthesis) represent long-run elasticities.

This methodology was recently used by the Productivity Commission (2006) to estimate long-run price and income elasticities for road and rail freight transport. The demand models were estimated using data for the period 1964 to 2000.

Despite its desirable technical properties, EC models require richer datasets compared with simpler regression models. For example, Wang (2011) gave consideration to PA and EC techniques when trying to model public transport demand in New Zealand. Having no more than nine years of data for some cities, Wang (2011) determined that only estimation of a PA model was feasible.

Pickrell et al (2012) found that single-equation error correction specifications produced models that were very similar to the analogous versions of the same equations estimated using OLS regression. These result led them to conclude that standard OLS models produced acceptably accurate estimated representations of the long-term relationships and hence used the OLS models to produce demand forecasts.

3.4.3 Functional form

The functional form used to link transport demand to the explanatory variables determines the shape of the demand curve. If incorrectly specified, it may be the cause of forecasting errors.

In its 2011 literature review, the NZIER found the most commonly used functional form for transport demand estimation is a double-logarithm linear model (also referred to as log-log or log-linear). The appeal of log-log models is not unique to modelling transport demand; it is a default model across most industries. The reason is that: 1) estimation of log-log models is relatively straightforward and 2) the estimated coefficients can be interpreted as elasticities (ie a percentage change in the dependent variable resulting from a percentage change in an explanatory variable).

A study by Li et al (2011), which was limited to freight transport, suggests that the translog model is also widely used. This type of model allows for greater flexibility in the demand specification.

At the onset of the Transport Agency model development, it may be wise to start with a simpler model (ie log-log) to serve as a baseline against which, over time, more ambitious models can be benchmarked.

---

[17] In addition to transport demand, these techniques have recently been used in modelling postal demand, water demand and electricity demand. For example, see Soteri et al (2009), Elgar and Martínez-Espiñeira (2007) and Australian Energy Market Operator (2013).

[18] A very good guide for practitioners is Franses (1998).

3.4.4 Endogeneity

In a well-specified model all explanatory variables should be exogenous (ie the error term in the regression should be uncorrelated with the explanatory variables). If this assumption is violated (ie some of the explanatory variables are endogenous), then the estimated parameters will be biased. Endogeneity may arise as a result of omitting an important driver from the regression (if it cannot be observed and/or measured), simultaneity (ie causality between the dependent and an explanatory variable is bidirectional), and measurement errors in the variables.

A standard way of addressing this issue is to use the instrumental variable (IV) approach. Let's assume we are interested in estimating a demand model, but recognise that one of our explanatory variables, call it X, may be endogenous. The application of the IV method involves estimating (ie predicting) X using a selected set of instruments; and then using the predicted X in the demand model. Because there are two stages in the estimation (ie the estimation of the endogenous variable X and the estimation of the demand variable) this method is referred to as the 2SLS method.[19]

When selecting an IV, the intent is to approximate a randomised trial using the exogenous variation provided by the instrument(s) (Angrist and Krueger 2001). Finding IV(s), however, is never an easy task. We want a variable or variables that are correlated (positively or negatively) with the endogenous variable, but are not correlated with the unexplained component in our outcome variable. Statistical tests should be implemented to test a selected IV for validity.[20]

What transport demand drivers could be endogenous and why? Several recent papers recognise the possibility that the supply of roads and VKT are simultaneously determined (ie there is a feedback loop between road supply and travel demand). To disentangle this simultaneous relationship, they use the IV approach. Namely, Noland and Cowart (2000) used land area and population to instrument for the stock of road measured in lane kilometres of road. Fulton et al (2000) instrumented growth in lane kilometres of highway using the lagged values of the same variable. Cervero and Hansen (2002) used about 20 instruments describing physical geography and political climate. Duranton and Turner (2011) used three historical variables as instruments: planned highway kilometres from a mid-20th century highway plan, railroad route kilometres at the end of the 19th century; and the incidence of major expeditions of explorations in the 19th century.

Duranton and Turner (2011) also postulated that regional population (or population growth) could be endogenous (ie people are attracted to areas with better transport infrastructure). In one of their models, they constructed an IV for metropolitan statistical area (MSA) population growth using historical shares of sectoral employment in the MSAs and interacted them with the national growth rate of each sector during the study period. This possibility of feedback loops between transport demand and some of its key drivers makes demand forecasting complex. Duranton and Turner (2011) estimated a number of different models controlling for possible endogenous variables and provided a range of elasticity estimates for the parameters of interest.

[19] See, for example, chapter 15 in Wooldridge (2005).

[20] In Stata, IV estimation can be implemented with the 'ivreg2' command, followed by 'estat overid', the command for the validity of the selected instrument.

3.4.5 Simultaneous equation models

Among the reviewed papers which used the IV approach to control for endogeneity of road supply, Cervero and Hansen's (2002) approach differed from the rest as they modelled demand for transport and the supply of road as simultaneous equation models. This approach allowed them to simultaneously estimate both the effect of lane miles on VMT (so called induced demand), and the effect of VMT on lane miles (so called induced investment). They argue that treating both the supply of road and VMT as endogenous variables is a more rigorous approach to modelling transport demand. Cervero and Hansen (2002), however, added an additional step in their estimation. They state:

> *To obtain more efficient estimates, a third stage of estimation was introduced that explicitly accounted for the cross-equation correlation of error terms as well as, unlike 2SLS, the presence of a right-hand side endogenous variable.*

They refer to this method as the three-stage least squares (3SLS) method.

This method was also used in a recent paper by Souche (2010), who estimated passenger vehicle travel demand and public transport demand as a system of simultaneous equations. Passenger vehicle travel demand, measured as the number of daily car trips per person, was estimated as a function of the average user cost of a car trip; the average user cost of a public transport trip; the length of the roads per 1000 inhabitants divided by the area of the city; per capita income; and urban density. Public transport demand, measured as the number of daily public transport trips per person, was estimated using the same variables except that the road supply variable was replaced by a public transport supply variable (defined as the number of per capita public transport VKT divided by the area of the city). To estimate the system of equations, Souche combined the 2SLS method with the seemingly unrelated regression (SUR) approach. The SUR approach uses information about correlation among contemporaneous error terms across the equations to improve the efficiency of parameter estimates.[21] Souche notes that this approach – 2SLS plus SUR – has only occasionally been used in the transport literature.

3.5 Model validation and testing

Model validation and testing should be an integral part of any model development; both to select a preferred model from a set of candidate models and, once the preferred model is chosen, to test its forecasting accuracy on an ongoing basis. Broadly speaking, we can distinguish between in-sample and out-of-sample criteria. Simply put, in-sample criteria tell us how well a model predicts the values which were used to estimate the model; out-of-sample criteria tell us how well the model predicts the values which were not included in its estimation.

The most commonly used in-sample criterion is **adjusted** R-squared; two commonly used alternative measures are the Akaike and the Schwartz (or Bayesian) information criteria.[22] Since forecasting is essentially an out-of-sample problem, assessing out-of-sample forecasting performance of candidate models is paramount. Although there are different criteria that could be used, they all involve using a sub-sample (say the first 20 years out of 22 in the sample) to estimate the parameters of the model, then

[21] For more on the SUR method, see Greene (1997, chapter 14).

[22] For more information, see Greene (1997, chapter 8).

using the estimated parameters to predict the values for the two hold-out years for which we have real data, and then calculating the difference between the predicted and the actual values.

A very good example of the use of out-of-sample methodology in model selection is available in Shen et al (2009). The authors claim that only a few studies of freight demand have employed the recent developments in multivariate dynamic econometric time series modelling. Their paper aims to fill this gap in the literature by applying, what they refer to as, state-of-the-art econometric time series models to model aggregate freight demand (road plus rail) in the UK. They compare the forecasting performance of two widely used modelling techniques (the static OLS model and the PA model) with the forecasting performance of four advanced time series models. The latter include: ARDL model, the vector autoregressive (VAR) model, the time varying parameter (TVP) model, and the structural time series model (STSM).[23] Using the mean absolute percentage error, a standard measure for assessing forecasting performance, they found that for relatively longer forecasting horizons (defined in their study as five-years-ahead), the PA model and the ARDL model performed best.

[23] The general form of these models can be found in Shen et al (2009). For brevity, we have not reproduced them in this report.

4 Data exploration

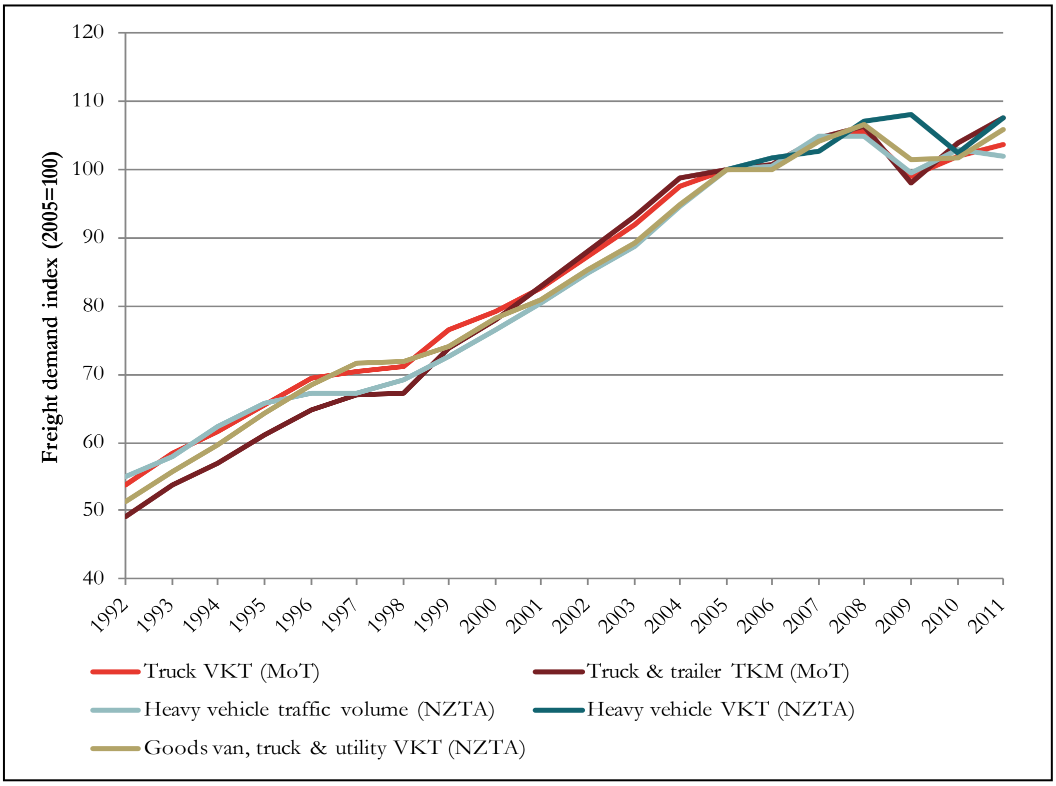
4.1 Considerations for freight demand model construction

As explained in chapter 2, there are several measures of road freight demand currently available from the Transport Agency and the MoT. To investigate whether the overall trends in these different measures of road freight transport are consistent, we graphed the five series for which data is readily available (see figure 4.1).

It can be seen that the two series from the MoT, the total truck VKT (light red line) and the total truck and trailer TKM (dark red line), exhibit very similar trends to the heavy vehicle traffic volume series from the Transport Agency (light blue line), even though the Transport Agency data is for state highways only and is collected from a limited number of telemetric sites around the country. The Transport Agency's VKT data series (based on RUC licence purchases) also exhibits the same pattern (tan line). These series indicate that the rate of growth in freight demand started to decline somewhere between 2003 and 2004, well before the onset of the global financial crises.

Somewhat surprisingly, the Transport Agency's heavy vehicle VKT data series (based on traffic volume and available from 2005), shows a dip in 2010, a year later than the other four series. We suggest this be investigated to establish whether the apparent pattern is due to a data entry error, a lag in reporting the data, or some other reason.

Figure 4.1 Freight demand – heavy vehicle travel



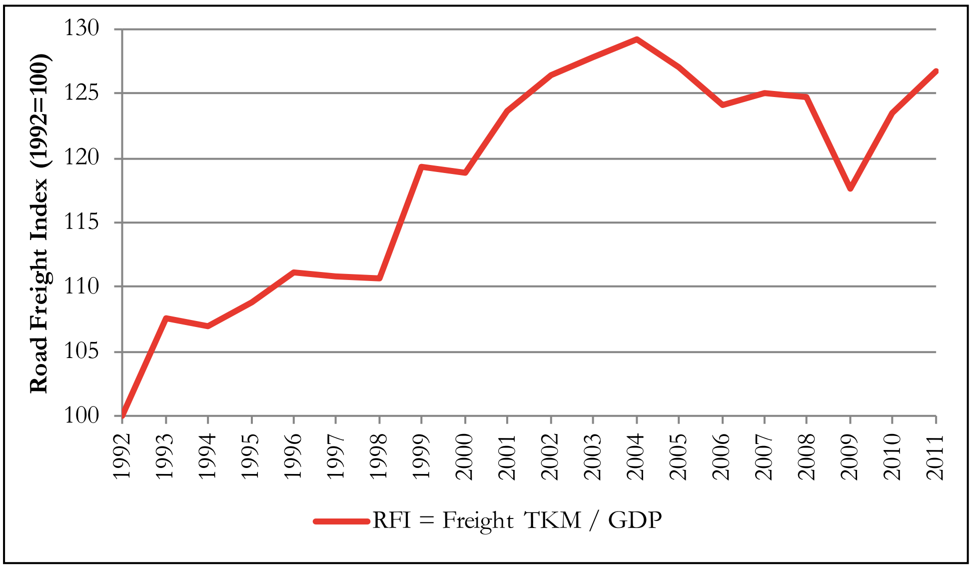
Note: Truck VKT = total truck vehicle kilometres travelled on state highways and local roads. Truck & trailer TKM = total truck and trailer tonnes kilometre travelled on state highways and local roads. Heavy vehicle traffic volume = count of vehicle flow on state highways collected from about 100 telemetric sites around the country. Heavy vehicle VKT = heavy vehicle kilometres travelled on state highways; calculations based on traffic flow data collected from all telemetric sites around the country. Goods van, truck & utility VKT (Transport Agency) = total goods van, truck and utility vehicles kilometres travelled on state highways and local roads; calculation based on RUC license purchases.

Source: Frontier Economics

We next investigate the historical relationship between freight demand and economic activity. We use MoT's TKM series as a measure of road freight demand as it reflects the amount of goods transported.[24]

Figure 4.2 presents a Road Freight Index (RFI), which we constructed as the ratio of road freight TKM to real GDP. The RFI points to a possible structural break around 2004. Before 2004, road freight transport demand grew, on average, faster than GDP (6.5% per year compared with 4.1%). After 2004, road freight transport demand grew, on average, by 1.2% per year compared with 1.3% growth in GDP, causing the RFI to decline slightly.

Figure 4.2 Road Freight Index



Source: Frontier Economics using data from Ministry of Transport and Statistics New Zealand

Because the available freight transport data series are relatively short, a prudent course of action is to first explore the available data using simple analysis tools to gain some insight into possible drivers of demand, before building any econometric models.

4.1.1 Exploratory data analysis

We followed the work of McKinnon (2007) by investigating some of the factors he found to be behind the decoupling of economic and freight transport growth in the UK.

4.1.1.1 Data issues

McKinnon (2007) found that about one third of the apparent decoupling in the UK was actually due to inconsistencies in data coverage. Namely, the official UK statistics on road travel are based on operations by UK-registered operators. Over time, however, the share of foreign road haulage operators increased

[24] As explained in chapter 1, the MoT has discontinued this series and there is no indication at present that it will be re-started. Nevertheless, there is merit in using this series to investigate demand drivers.

significantly. This means that the official statistics on the total tonne kilometres travelled were underestimated/under-reported.

We investigated whether the MoT made any changes in the methodology used to calculate its truck and trailer TKM data series. Our communication with MoT (Stuart Badger, MoT, pers comm) and review of information available on the MoT website indicate that the only change in the methodology was a revision of the truck-load utilisation assumption, which was implemented in 2008.[25]

4.1.1.2 Road's shares of the freight market

We next investigated whether the observed change in road freight transport growth could be explained by changes in the freight modal split.

Although there is no comprehensive data on New Zealand freight transport by mode, two reports provide a snapshot of modal shares in 2002 and 2008 (Bolland et al 2005; Richard Paling Consulting 2008). Data from these reports, summarised in table 4.1, shows that road's share of the freight market increased from about 67% in 2002 to about 70% in 2008. It appears that this increase was due mainly to substitution from rail to road.

Table 4.1 Mode shares of the freight market (by tonne-km)

Mode	2002	2008
Road	66.5%	~ 70%
Rail	17.8%	~ 15%
Sea	15.3%	~ 15%
Air	0.3%	< 1%

Source: The 2002 figures are from Bolland et al (2005) and the 2008 figures are from Richard Paling Consulting (2008).

4.1.1.3 Relative price of freight transport

To investigate how road freight transport competitiveness relative to rail freight transport has changed over time, time series data on surface freight transport costs by mode is needed. Such data, however, is not readily available. Statistics NZ does have a Producer Price Index (PPI) for 'Road Transport' and a PPI for 'Rail, Water, Air and Other Transport'. The latter, however, includes international transport and as such is not a valid measure of road freight transport alternatives.[26]

Comparison of the PPI for road freight transport and the PPI for all industries points to a significant increase in road freight transport costs in 2006 (see figure 4.3). Since 2006, road freight transport costs have been growing in line with the overall PPI.

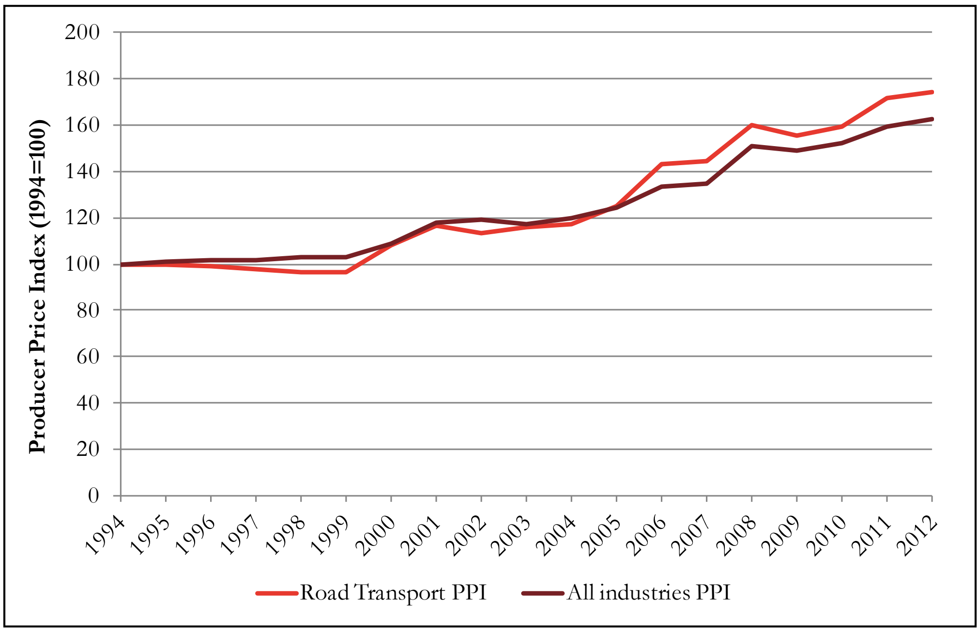
Closer investigation of the major components of road freight transport operating costs shows significant increases in both diesel price and RUC over the past several years. As illustrated in figure 4.4, the average real price of diesel in the period 1989 to 2004 was about 60 cents/litre, compared with close to 90 cents/litre in the post-2004 period.

[25] Information is available at www.transport.govt.nz/ourwork/TMIF/Pages/FT007.aspx

[26] The Transport Agency may want to investigate obtaining a similar index from Statistics NZ which would exclude international travel.

The RUC charges were declining (in real terms) until 2007, after which they started to increase sharply on an annual basis. In figure 4.5 we present historical RUC charges for two types of trucks: two-axle up to 10 tonnes and three-axle up to 20 tonnes. According to the Road Transport Forum NZ (2010), two-axle trucks between 7 and 10 tonnes are the most common type of truck because of the demand for interregional and time-sensitive delivery, while three-axle trucks between 16 and 20 tonnes travel the greatest distances.

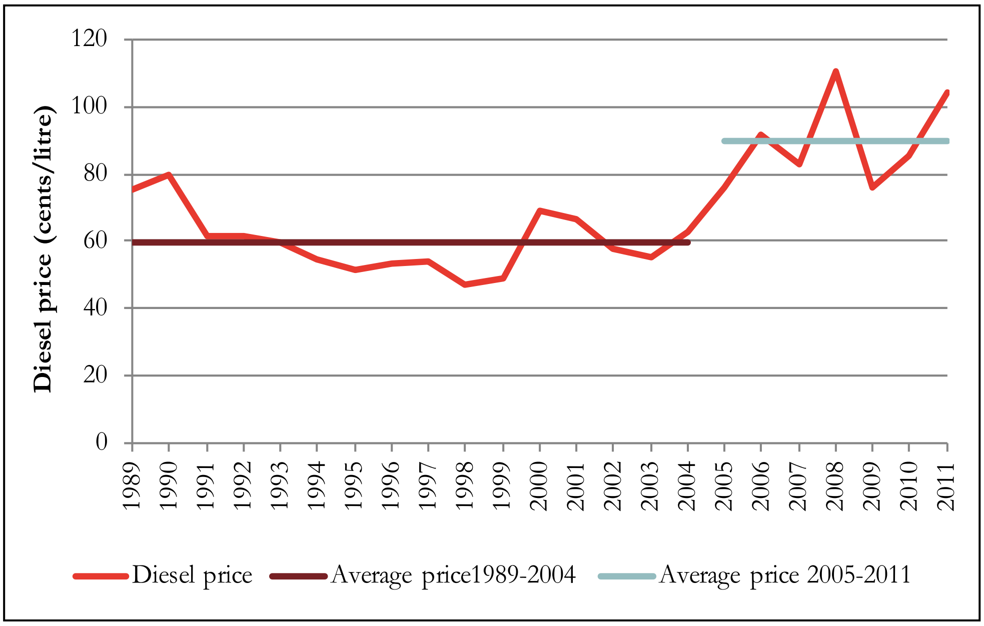
Figure 4.3 Producer Price Index



Note: PPI measures changes in prices relating to the supply (output). Road freight transport PPI is available from 1994. Data is for the second quarter of each year.

Source: Frontier Economics using data from Statistic New Zealand

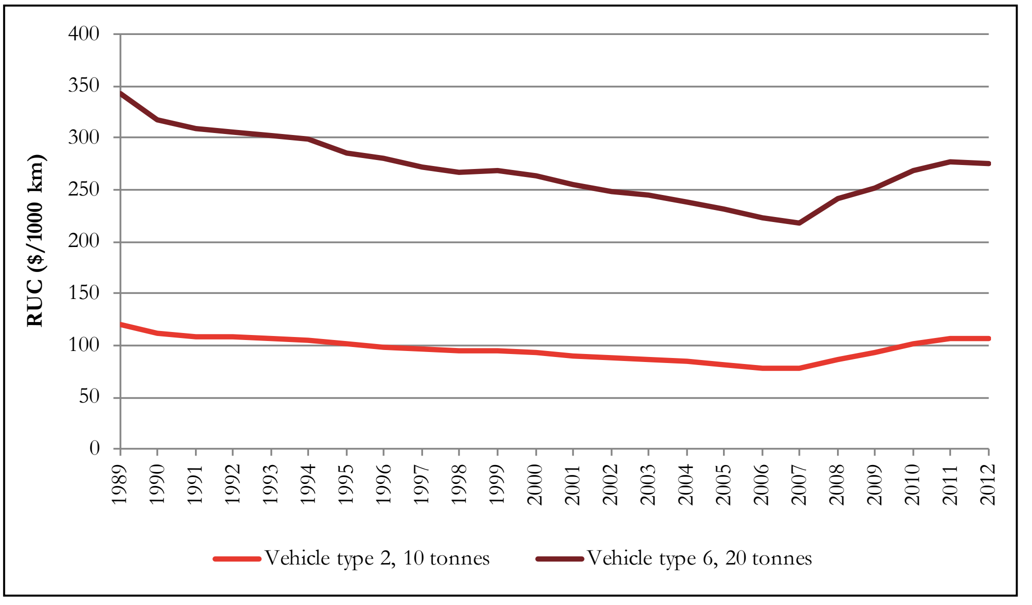
Figure 4.4 Diesel price



Note: Prices converted to 1996 dollars.

Source: Frontier Economics using data from New Zealand Ministry of Business, Innovation and Employment

Figure 4.5 RUC charges



Note: RUC charges converted to 1996 dollars.

Source: Frontier Economics using data from New Zealand Ministry of Transport

4.1.1.4 Productivity changes

Leung and Tantirigama (2011) estimated productivity changes in road and rail transport in New Zealand using 2003 and 2007 national accounts input-output (I-O) tables. They found that multi-factor productivity of road and rail freight transport improved in that period. The magnitude of the estimated total productivity increase over the 2003 to 2007 period varied between 1.35% and 3.78% depending on the methodology used.

A very commonly used approach for accounting for technological/productivity improvements in econometric models is to use a time trend. If this is not possible, due to high correlation with other explanatory variables, then the effect of technological/productivity changes will be confounded with income and price effects. Hence, the estimated elasticity estimates will probably not be 'pure' income/price elasticities, but rather 'efficiency adjusted' income/price elasticities.

4.1.1.5 Transport intensity

McKinnon (2007) states it is widely accepted that the increase in industrial output, rather than GDP growth, is the main driving force behind freight transport. The reason is that goods-producing sectors are generally more transport intensive compared with service-producing sectors.

We investigated this by constructing a proxy measure of transport intensity for goods-producing sectors (agriculture, forestry, mining, manufacturing and construction) and service-producing sectors. This proxy transport intensity measure was constructed as the ratio of the monetary value of freight transport services 'consumed' by these industry to their GVA.

The New Zealand national accounts I-O tables provide information on the use of products and services by industry. The 1996 and 2007 I-O tables contain information on consumption of 'road freight transport services' by industry. For example, in 1996, all goods-producing industries consumed $1263 million of road transport freight services. With the GVA of $27,554 million, their transport intensity in 1996 was 0.05. This can be interpreted as using $0.05 of road freight services per $1 of GVA.

As can be seen from table 4.2, goods-producing sectors are significantly more transport intensive than service-producing sectors. As goods-producing industries' share of GDP declines, we would expect road freight demand growth to slow.

Table 4.2 Transport intensities

	Goods-producing sectors		Service-producing sectors	
	1996	2007	1996	2007
Consumption of road freight transport services (in millions)	$1263	$2055	$1013	$1629
Gross value added (in millions)	$27,554	$34,847	$68,742	$103,741
Transport intensity	0.046	0.059	0.015	0.016

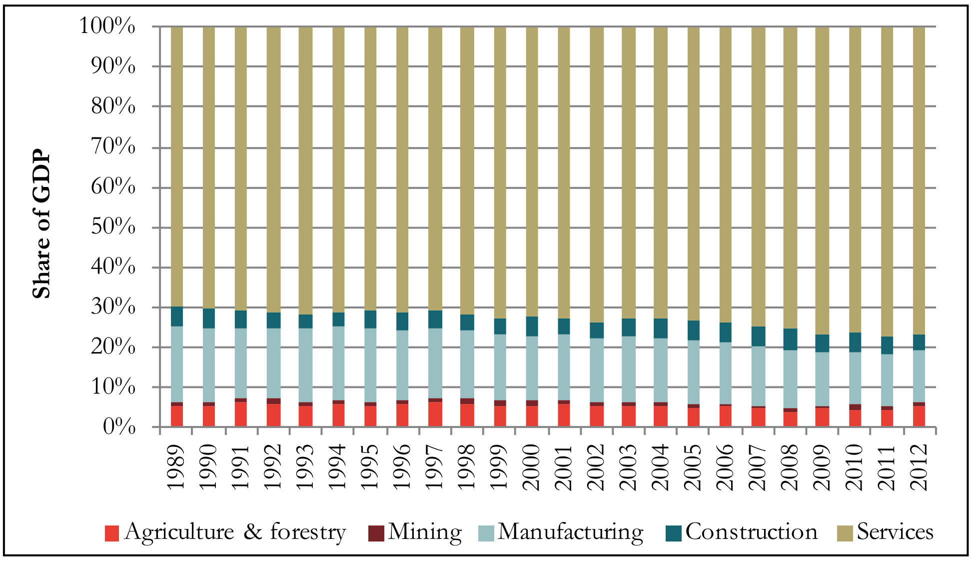
Note: All values are in 1996 dollars.

Source: Frontier Economics using data from Statistics NZ

4.1.1.6 Change in the composition of GDP

New Zealand, like many industrial countries, is undergoing long-term structural change, with goods-producing sectors contracting their share of GDP and service-producing sectors expanding their share of GDP. As can be seen in figure 4.6, goods-producing sectors accounted for 30% of GDP in 1989. Their share, however, dropped to 23.5% by 2012.

Figure 4.6 Composition of GDP



Source: Frontier Economics using data from Statistics NZ

The decline in the goods-producing sectors' share of GDP accelerated in the last decade. This is the result of a significantly lower average annual growth rate compared with the service sector (see table 4.3).

Table 4.3 Average annual growth rates

Sectors	Average annual growth rate 1992 to 2002	Average annual growth rate 2003 to 2012
Goods-producing sector GVA	2.6	0.4
Service sector GVA	3.9	2.5

Source: Frontier Economics using data from Statistics NZ

4.1.1.7 Issues to consider in model development

Our analysis indicates that three factors could have affected the slowdown in road freight transport growth from 2004: an increase in the price of diesel, an increase in road user charges (only since 2007), and a decline in the goods-producing sectors as a share of GDP.

Investigating the effect of diesel price and RUC charges on freight transport demand is relatively straightforward (ie by including them as explanatory variables in an econometric model). It is less clear, however, how to account for the changing structure of the New Zealand economy when modelling freight transport demand.

One approach, feasible with the currently available data, is to develop demand models by truck type. Different types of trucks are used for different purposes and their use hence may, at least to some extent, proxy freight transport demand by different industries (ie smaller trucks are usually used for shorter distances and time-sensitive deliveries and hence may be more sensitive to changes in retail sales, while larger trucks are used for long-distance heavy freight and hence may be more responsive to changes in goods-producing sector output).

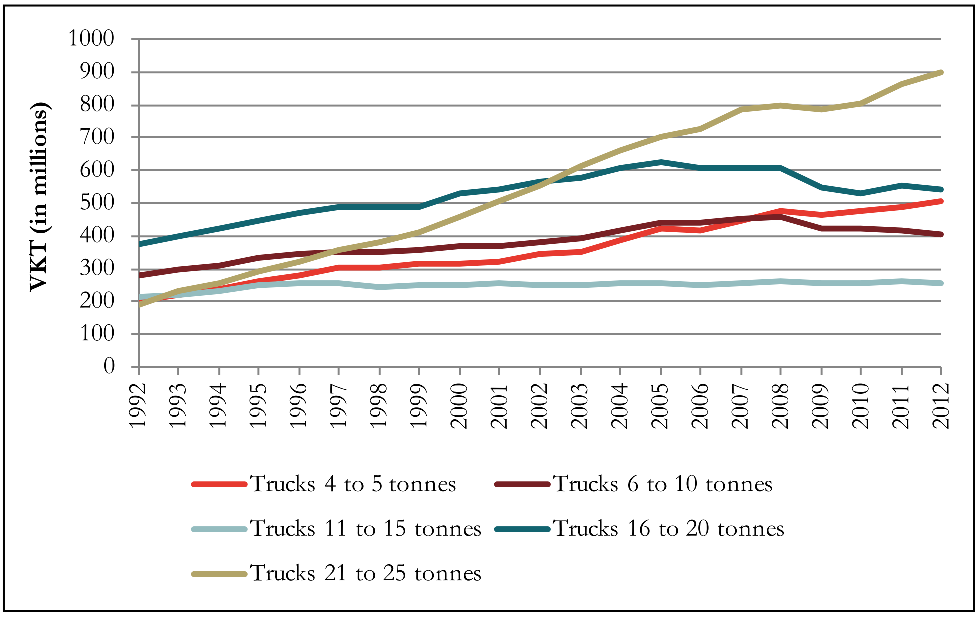
The Transport Agency collects data on RUC licences, which are purchased in units of 1000km, and uses this data to calculate VKT by vehicle type and weight category. In the vehicle category 'goods van, truck and utility vehicle', VKT data is available for vehicles types ranging in weight from 1 tonne to 45 tonnes. This data can be used to develop econometric demand models for freight transport. Vehicles weighing 1 to 3 tonnes are likely to represent goods vans and utility vehicles and hence should be analysed separately.

Although the data is available from 1987, we recommend limiting the sample to the period 1992 onwards. The VKT values for the pre-1992 period appear unreasonably low, casting some doubt on their accuracy.[27]

Figure 4.7 indicates that freight demand (as proxied by purchased VKT) differs significantly over time across different truck types. For illustrative purposes, we grouped trucks weighing 4 tonnes to 25 tonnes into five weight categories. When developing demand models, however, consideration will need to be given to the most appropriate grouping, which at the initial step should be informed by industry knowledge. Let us illustrate this with an example. It can be seen in figure 4.7 that VKT by trucks weighing 21 to 25 tonnes grew faster than VKT by trucks weighing 16 to 20 tonnes (the latter actually seems to have been declining since 2005). If some of this is driven by freight operators switching from smaller to larger trucks due to, for example, increasing driver wages, then it may be sensible to combine VKT for trucks in these two truck categories and model them together.

[27] For example, the data indicates that vehicles weighing between 3 and 4 tonnes purchased, on average, 105 thousand VKT per quarter from 1987Q3 to 1991Q1, but close to 33 million VKT per quarter from 1992Q2 to 1994Q4. This does not appear to be a simple scaling issue (ie reporting values in thousands vs in millions).

Figure 4.7 VKT by truck size



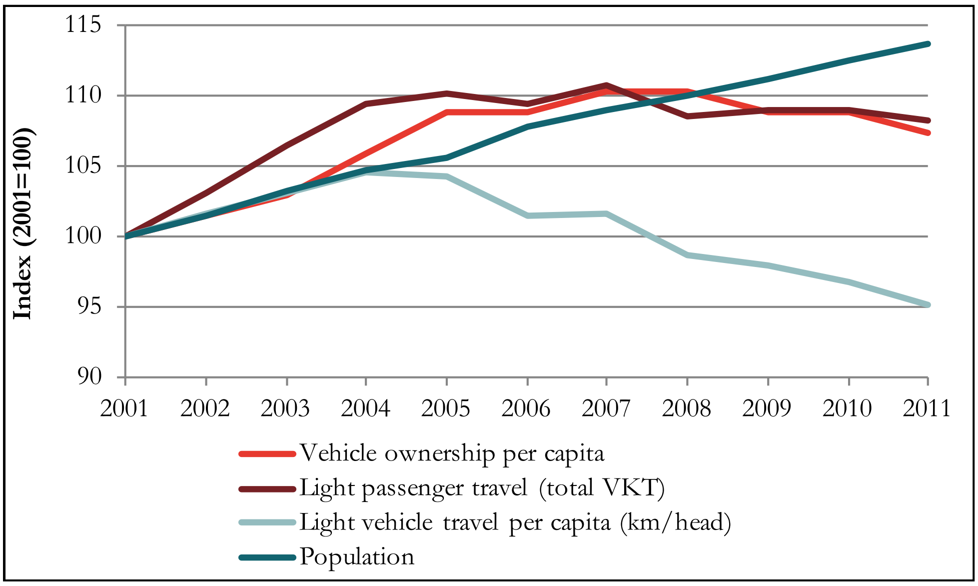
Note: VKT based on RUC licence purchases.

Source: Frontier Economics using data from NZ Transport Agency

4.2 Considerations for passenger demand model construction

Annual data on passenger vehicle ownership and vehicle use (ie VKT on state highways and local roads) is available from the MoT starting from 2001. As can be seen in figure 4.8, total passenger vehicle travel and vehicle ownership per capita began to plateau around 2005. These trends, together with steady population growth, have resulted in a decline in per capita vehicle use from 2005 onwards.

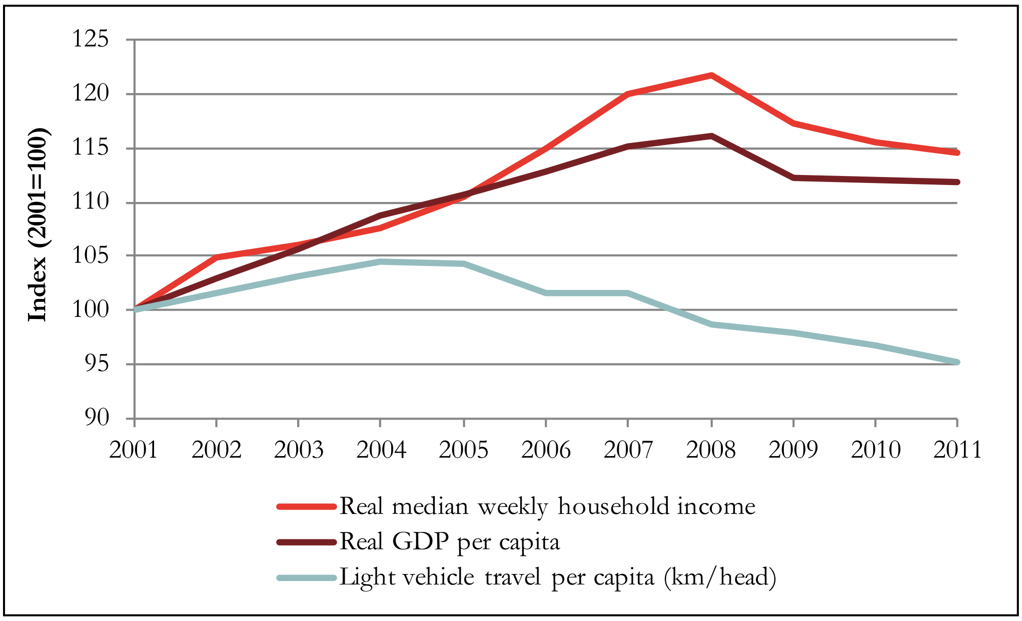
Figure 4.8 Passenger vehicle ownership and use (total and per capita)



Source: Frontier Economics using data from Ministry of Transport

Examination of per capita passenger vehicle use and income trends does not provide a clear picture as to their relationship (see figure 4.9). The decline in vehicle use per capita started well before the onset of the global financial crises in 2009.

Figure 4.9 Passenger vehicle use and income



Source: Frontier Economics using data from Ministry of Transport and Statistics NZ

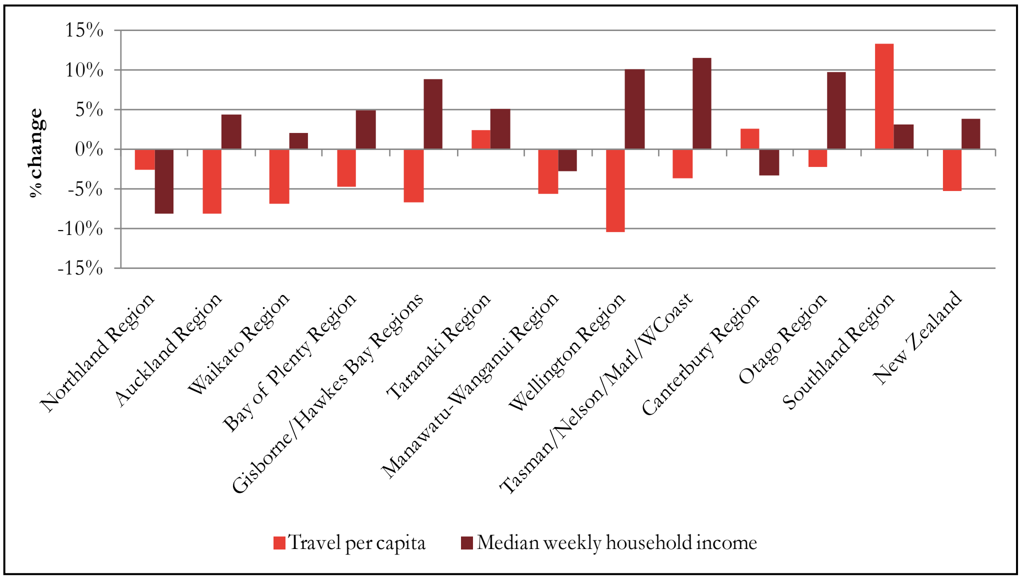
Because the available data on passenger vehicle travel covers a relatively short period of time, it may be insightful to analyse driving patterns at the regional level, namely by regional council area. The only dataset available for this type of analysis is the Transport Agency's light vehicle VKT data calculated from traffic flow information.[28] The VKT data for 14 regional council areas is readily available from 2005. A somewhat less comprehensive dataset (in terms of network coverage) could be constructed going back to 2002. There are two drawbacks to this data as explained in section 2.3: 1) the data covers travel on state highways only; and 2) the data combines travel by passenger cars and light commercial cars.

Statistics NZ does not produce regional GDP series on a regular basis. However, income data by regional council area is available from the New Zealand Income Survey, which is conducted annually by Statistics NZ. Relevant data from the survey includes: average and median weekly household income, average and median weekly earnings, number of households and number of people in paid employment.

A comparison of changes in vehicle use and income are presented in figure 4.10. Most of the regions experienced a decrease in vehicle use between 2005 and 2012, although the rate of decrease varied.

[28] As explained in the notes to table 2.2, VKT is calculated using traffic volume data from telemetric sites. Namely, AADT is first calculated for every road section of the entire state highway network. The VKT for a given road section is then calculated by multiplying the AADT for that road section by the road section's length and multiplying by the number of days in a year.

Figure 4.10 Comparison of vehicle use and income across regions in 2005 and 2012 (% change)



Source: Frontier Economics using data from the NZ Transport Agency

5 Recommendations

5.1 What data to use

5.1.1 Road freight activity

As discussed in chapters 3 and 4, it is widely accepted that the increase in industrial output, rather than GDP growth, is the main driving force behind freight transport (since goods-producing sectors are generally more transport intensive compared with service-producing sectors). Hence, when modelling freight transport demand, it is necessary to account for the changing structure of the economy. This could be accomplished using time series data on tonne kilometres travelled by commodity type; such data, however, is not available in New Zealand.[29]

In the absence of transport data by commodity type, we recommend the Transport Agency use RUC-based VKT data. This data is available by heavy vehicle class and weight category, which would enable the Transport Agency to **model freight demand by vehicle type**. The reason is that different types of trucks are used for different purposes and their use may, at least to some extent, proxy freight transport demand by different industries (ie smaller trucks are usually used for shorter distances and time-sensitive deliveries and hence may be more sensitive to changes in retail sales, while larger trucks are used for long-distance heavy freight and hence may be more responsive to changes in goods-producing sector output).

The Transport Agency's VKT data is not available for trucks alone; however, trucks account for the majority of VKT in the heavy vehicle class 'goods van/truck/utility'. Because there are over 40 weight categories in this vehicle class (ranging from 3 tonnes to over 45 tonnes), it would be impractical to develop separate demand models for each. We therefore recommend that the Transport Agency groups vehicles into a smaller number of weight categories. Consideration will need to be given as to the most appropriate grouping, which at the initial step should be informed by industry knowledge. Although the data series are available from 1987, VKT values in the first few years appear unreasonably low. We therefore recommend limiting the sample to include data from 1992 onwards. Data on economic drivers can be obtained from Statistics NZ.

5.1.2 Passenger vehicle travel

As explained in chapters 2 and 4, data on passenger vehicle travel is available for a relatively short period of time (at best 11 years) which severely limits any analysis of how the effect of income on travel demand has changed over time. Income levels, however, differ across regional council areas in New Zealand, and this difference can provide some insight into the relationship between income and travel demand. We therefore recommend the Transport Agency investigates developing a **panel model** using its regional VKT data.[30], If a panel model is structured with regional dummy variables, then the interaction terms between

[29] As we noted in chapter 2, the aggregate TKM data published by the MoT has been discontinued.

[30] Panel data is data for multiple entities in which each entity is observed at two or more time periods (ie panel data has both cross sectional and time series features). The NZTA's VKT data is a panel data set – annual VKT are available for each regional council area for a period of about 10 years. For an application of panel models in modelling transport demand, see Duranton and Turner (2011).

the income variable and the regional dummy variables could tell us the extent to which income elasticities differ across regions (and across different income levels).

To investigate the impact of income growth on passenger vehicle travel, we recommend the Transport Agency use its VKT data calculated from traffic flow information. Although data is readily available from 2005, a somewhat less comprehensive dataset (in terms of network coverage) could be constructed going back to 2002. We recommend the Transport Agency investigate using data from 2002. The VKT data is available for 14 regional council areas. Historical data on income and employment by regional council areas can be obtained from Statistics NZ. Projected values for these drivers, however, will need to be sourced from experts.

5.2 What econometric models to consider

Table 5.1 summarises the most frequently used econometric techniques for modelling road transport demand. All these models belong to the same broad 'family' of models and hence one cannot recommend *ex ante* which of them should be used. Model selection is an integral part of model development, and should be guided by data properties and diagnostic checks.

Particular attention needs to be paid to the small-sample properties of the estimation methods. Many of the widely used methods for estimating long-run parameters (including the Johanson vector error correction model approach) are based on having large data samples to be effective. We recommend the Transport Agency investigates methods which may not have been used widely in applied transport literature, but which come from the more general field of time series analysis. For example, one well-known method for working with co-integrated variables is the so-called dynamic ordinary least squares estimator, developed by Stock and Watson (2007). [31]

Table 5.1 Properties of most-commonly used econometric models in transport literature

Model properties	Static OLS	PA	ARDL	ECM
Long-run elasticity	Yes	Yes	Yes	Yes
Short-run elasticity	No	Yes	Yes	Yes
Dynamic structure	No	Yes	Yes	Yes
Can be extended to panel analysis	Yes	Yes	Yes	Yes

Note: Static OLS = static ordinary least squares regression; PA = partial adjustment model; ARDL = autoregressive distributed lag; ECM = error correction model.

As explained in chapter 3, recognising that there are possibly feedback loops between transport demand and some of its drivers further complicates demand modelling. Some recent studies have dealt with this issue by using, what is known in the econometrics literature as, the 2SLS approach in which the

[31] Simply stated, the dynamic ordinary least squares method is a single equation approach that includes leads and lags of first differences of the explanatory variables, and that corrects for serially correlated errors by a general least squares procedure.

endogenous explanatory variable is instrumented. To our knowledge, a very few studies have used a more complex modelling framework in which the 2SLS method is combined with the SUR method to attempt to improve the efficiency of the estimated parameters.

5.3 Modelling steps

We recommend the Transport Agency adopt the analytical steps described below, and the order in which they are presented, as a general strategy for the development of its econometric models.

5.3.1 Data preparation

This is an important stage, very often under-appreciated by people who do not regularly interrogate data, or who fail to understand that data preparation and handling procedures can affect the results (possibly materially so). As already explained, when modelling road freight transport, we recommend the Transport Agency investigate grouping vehicles in the 'goods van/truck/utility' class into a few weight categories, giving primary consideration to how vehicles are used (ie time sensitive deliveries vs long-haul transport).

When modelling passenger car use with VKT data derived from traffic flow information, consideration needs to be given to whether observed fluctuations and level shifts are due to factors such as missing telemetric data or additions of new road segments/bypasses, and how best to address such data issues.

Victorio (2011), in an internal Transport Agency study on developing top-down econometric models of transport demand, suggests the Transport Agency could try using monthly data instead of quarterly as a way to increase the number of observations in its sample. We caution against this approach for the following reasons:

- Using monthly data may require including additional explanatory variables in the models (possibly up to 11 monthly dummies) to pick up monthly traffic fluctuations; monthly fluctuations may not be of much interest for long-run demand forecasting
- Monthly data is likely to exhibit autocorrelation; aggregating data to a higher level may ameliorate this issue
- Economic/income variables are likely to affect demand slowly over time
- Many economic variables are available on a quarterly or annual basis only. One can, of course, interpolate data to monthly values, but this will (by construction) introduce additional autocorrelation which will then need to be addressed.

5.3.2 Selection of candidate demand drivers

When modelling passenger vehicle demand, we recommend the Transport Agency use median household income available by regional council area from Statistics NZ. When modelling freight demand by truck type, we recommend the Transport Agency test which of the following measures of economic activity perform best in each model: GDP (plus imports), GVA for goods producing sectors, GVA for service sectors, trade (imports plus exports), and retail sales.

Our literature review presented in chapter 3 provides some guidance to the Transport Agency as to what demand drivers, other than income, should be considered for inclusion in its demand models. We do not recommend using principal component analysis.

When selecting demand drivers, we recommend the Transport Agency consider whether any of the selected drivers are endogenous and what instrumental variables it could use to address the issue. To aid the Transport Agency in model development, we have provided a summary of instrumental variables which have been used in recent studies to instrument road supply.

Since the Transport Agency intends to use its demand models for forecasting, it will also need to consider whether projected values for the selected drivers are available.

5.3.3 Data investigation and model selection

All candidate drivers will need to be tested for stationarity and co-integration, before a preferred model can be selected. As explained in section 3, using non-stationary time series data in regressions may lead to spurious results – results which erroneously indicate, through misleading values of R-squared and t-statistics, that there is a meaningful relationship among the regression variables.

Model selection should be guided by data properties and diagnostic checks (eg checking regression residuals for autocorrelation). When deciding which of the candidate variables should be included in the preferred models, the selection should be based both on statistical significance and economic significance of the tested drivers (economic significance is usually understood to mean the magnitude of a regression coefficient). Furthermore, we caution against using an automated rule that would drop any candidate explanatory variable that is not statistically significant at the 5% significance level; excluding a key driver from a model may cause more harm than leaving it in (ie may lead to omitted variable problem).

5.3.4 Model testing and validation

As explained in chapter 3, model testing and validation should be an integral part of any model development. We therefore recommend that, even after the preferred models have been selected, the Transport Agency continue to check their models forecasting performance on a regular basis (eg as models are re-run with new data). We envisage that modelling will be done in a statistical package where models can be run in a batch mode (such as using .do file in Stata); and hence updating and testing the models should not be too burdensome.

During model development, it may be necessary to repeat some of the above analytical steps. For example, if results in some of the later steps are not satisfactory (eg the regression model does not fit well, the main drivers are not statistically significant, forecasting performance in a hold-out sample is poor), it may be necessary to revisit some of the earlier steps (eg revisit the method used to group trucks into specific weights categories). Multiple iterations will likely improve the understanding of the drivers of transport demand and lead to more precise estimates of the effect of economic activity on transport demand.

We note that this paper, while written to address the specific requirements of the Transport Agency's modelling of road freight transport, may have relevance across a broader range of applications.

5.4 Concluding remarks

Developing econometric models will not be without challenges. We would therefore recommend that, in addition to developing econometric models, the Transport Agency consider alternative modelling approaches based on survey techniques to derive income and/or own- and cross-price elasticities.[32] Once such elasticities are derived they can be used:

- to provide an alternative set of forecasts
- as a check on the elasticity estimates derived from the Transport Agency's econometric models
- to impose constraints on the parameters in the econometric models.

Another approach for the Transport Agency to assessing reasonableness of the parameters estimated with its econometric models would be to compare them to published travel demand elasticities using a meta-analysis. Meta-analysis is a quantitative tool frequently used by researchers to combine and contrasts results from different studies in order to identify patterns among study results. In a recent study, Li et al (2011) used this method to identify sources of systematic variation in freight demand elasticity studies. They explain how analysts can utilise their model outputs to adjust the empirical evidence from specific studies.

[32] For example, Hensher et al (2013) used survey techniques to estimate own-price and cross-price elasticities for 18 different vehicle classes.

6 References

Agnolucci, P and D Bonilla (2009) UK freight demand: elasticities and decoupling. *Journal of Transport Economics and Policy* 43, no.3: 317–344.

Andersson, FNG and T Elger (2012) Swedish freight demand. *Journal of Transport Economics and Policy 46*, no.1.

Angrist, JD and AB Krueger (2001) Instrumental variables and the search for identification: from supply and demand to natural experimentation. *Journal of Economic Perspective 15*, no.4.

Australian Energy Market Operator (2013) Forecasting methodology information paper. Accessed 1 July 2013. www.aemo.com.au/Electricity/Planning/Forecasting

Bolland, J, D Weir and M Vincent (2005) Development of a New Zealand national freight matrix. *NZ Transport Agency research report 283.*

Bureau of Transport and Regional Economics (2006) *Freight Measurement and Modelling in Australia.* Report 112, BTRE, Canberra ACT.

Cervero, R and M Hansen (2002) Induced travel demand and induced road investment. *Journal of Transport Economics and Policy 36*, no.3: 469–490.

Conder, T (2009) Development and application of a New Zealand car ownership and traffic forecasting model. *NZ Transport Agency research report 394*.

de Jong, G, J Fox, M Pieters, AJ Daly and R Smith (2004) A comparison of car ownership models. *Transport Reviews 24*, no.4: 379-408.

Duranton, G and MA Turner (2011) The fundamental law of road congestion: evidence from US cities. *American Economic Review 101*, no.6: 2616–2652.

Elgar, E and R Martínez-Espiñeira (2007) An estimation of residential water demand using co-integration and error correction techniques. *Journal of Applied Economics X,* no.1: 161–184.

Franses, PH (1998) *Time series models for business and economic forecasting.* Cambridge University Press.

Fulton, LM, RB Noland, DJ Meszler and JV Thomas (2000) A statistical analysis of induced travel effects in the US mid-Atlantic region. *Journal of Transportation and Statistics 3*, no.1: 1–30.

Greene, W (1997) *Econometric analysis.* Chapter 4, 3rd edition. New Jersey: Prentice Hall.

Hensher, DA, AT Collins, JM Rose and NC Smith (2013) Direct and cross elasticities for freight distribution access charges. *Transport Research part E* (forthcoming).

Kennedy, I and D Kennedy (2007) Impacts of fuel price changes on New Zealand transport. *NZ Transport Agency research report 331*.

Leung, J and T Tantirigama (2011) Contribution of transport to economic growth and productivity in New Zealand. *Proceedings of Australasian Transport Research Forum 2011*, 28–30 September 2011, Adelaide, Australia.

Li, Z, DA Hensher and JM Rose (2011) Identifying sources of systematic variation in direct price elasticities from revealed preference studies of inter-city freight demand. *Transport Policy 18*: 727–734.

McKinnon, AC (2007) Decoupling of road freight transport and economic growth trends in the UK: an exploratory analysis. *Transport Reviews: A Transnational Transdisciplinary Journal 27*, no.1: 37–64.

Noland, RB and WA Cowart (2000) Analysis of metropolitan highway capacity and the growth in vehicle modes of travel. *Transportation 27*, no.4: 363–390.

Oum, TH, WG Waters and X Fu (2000) Transport demand elasticities. In *Handbook of transport modelling*, DA Hensher and KJ Button (Eds), Amsterdam: Pergamon.

Oxera and Arup (2010) *How has the preferred econometric model been derived*? Accessed 14 March 2013. www.gov.uk/government/uploads/system/uploads/attachment_data/file/4233/econometric-approach.pdf

Pickrell, D, D Pace, R West and G Hagemann (2012) Developing a multi-level vehicle miles of travel forecasting model. *91st Annual Meeting of the Transportation Research Board*. Washington DC. 22–26 January 2012.

Productivity Commission (PC) (2006) *Road and rail freight infrastructure pricing*. Report no.41. Canberra. Accessed 14 March 2013. www.pc.gov.au/__data/assets/pdf_file/0003/47532/freight.pdf

Richard Paling Consulting (2008) *National freight demand study*. Report prepared for the Ministry of Transport, NZ Transport Agency and Ministry of Economic Development.

Road Transport Forum NZ (2010) *Road freight facts*. Accessed 14 March 2013. www.rtfnz.co.nz/cms_show_download.php?id=267

Shen, S, T Fowkes, T Whiteing and D Johnson (2009) Econometric modelling and forecasting freight transport demand in Great Britain. *Proceedings of the European Transport Conference*, 2009.

Soteri, SF, F Fève, JP Florens and F Rodriguez (2009) Internet advertising and direct mail: trends and analysis for the UK. Pp 209–222 in MA Crew and PR Kleindorfer (Eds). *Progress in the competitive agenda in the postal and delivery sector.* Edward Elgar Publishing.

Souche, S (2010) Measuring the structural determinants of urban transport demand. *Transport Policy 17*, no.3: 127–134.

Statistics NZ (2012) *Using national accounts input-output tables*. Accessed 10 April 2013. www.stats.govt.nz/browse_for_stats/economic_indicators/NationalAccounts/input-output%20tables.aspx

Stock, JH and MW Watson (2007*) Introduction to econometrics*. Chapter 16, 2nd edition. Boston: Pearson Education.

Truong, TP and D Hensher (2012) Linking discrete choice to continuous demand within the framework of a computable general equilibrium model. *Transportation Research part B 46*: 1177–1201.

Victorio, A (2011) New Zealand transport demand: some evidence and forecasts. Unpublished report for NZ Transport Agency.

Wang, J (2011) Appraisal of factors influencing public transport patronage. *NZ Transport Agency research report 434*.

West, R, D Rubin and JC Villa (2011) *Identification and evaluation of freight demand factors.* Report prepared for the US National Transportation Research Board.

Wooldridge, JM (2005) *Introductory econometrics: a modern approach.* 3rd edition. South-Western College Publishing.

Appendix A: Literature review summary

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
Agnolucci and Bonilla (2009) *UK freight demand: elasticities and decoupling*	Estimation of road freight demand in the UK over the period 1957–2003 and to assess the decoupling of freight demand and economic activity between 1997 and 2004.	Road freight. Total including heavy goods vehicles and vans.	Times series econometric estimation using general-to-specific methodology. Annual data. Application of STSM that incorporates stochastic trend.	Road freight tonne-km.	GVA from production all sectors plus construction (equal to GDP plus imports less GVA from services sector and imports); price of road and pipeline transport services, trend variable. .
BTRE (2006) *Freight measurement and modelling in Australia* (section 3.2 Urban road freight demand model)	Forecast individual capital city and aggregated road freight transport tasks.	Urban road freight.	Panel model for capital city analysis. Static OLS for aggregated model. Annual data.	Freight tonne-km (in logs).	Log (city population * real per capita national GDP); log real freight rate for short haul transport; city specific dummies for panel model; time-specific dummies (city specific for panel model).
Conder, T (2009) *Development and application of a New Zealand car ownership and traffic forecasting model*	Developing and showing the application of improved methods for forecasting future car ownership in NZ, taking into account data constraints.	Cars	Time series using OLS. Predictive equation with reducing time trend. Annual data	Car ownership per capita.	GDP per capita; car price index; time trend.

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
Li and Hensher (2009) *Road freight demand in Australia: Key drivers and forecasts*	Identify the key drivers of freight demand in Australia and evaluate a number of times series models for forecasting accuracy.	Road freight (total road freight task).	Various time series models: trend model, autoregressive integrated moving average model, vector autoregressive model, vector autoregressive model with exogenous regressors, and Bayesian vector autoregressive model. Annual data. Data is for period 1971 to 2003.	Total road freight task.	Real GDP, real road freight variables (not specified further).
McKinnon (2007) *Decoupling of road freight transport and economic growth: trends in the UK: An exploratory analysis*	Examines possible causes of the decoupling of freight movement demand from GDP growth in the UK between 1997 and 2004.	Road freight. HGVs only.	Eyeballing time series data of possible factors contributing to a levelling off of freight movements. Annual data.	Road tonne-km (for trucks with gross weight greater than 3.5 tonnes = HGVs).	Statistical accounting changes; dematerialisation (reduction in material resources required per unit of GDP); change in composition of GDP, change in road freight's share of freight market, increase in penetration of foreign operators, displacement of freight from trucks to vans, reduction in the number of links in the supply chain; diminishing rate of special concentration; improvements in efficiency of vehicle routing; domestic supply chains becoming fully extended; erosion of industrial activity to other countries; real costs of road freight transport.

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
Oxera & Arup (2010) *Econometric approach report – How has the preferred econometric model been derived?*	Derivation of a preferred econometric approach for forecasting rail passenger demand in the UK.	Rail passenger.	Panel data models including: ARDL, EC, panel effects (fixed and random). Of the former two dynamic models the Arellano & Bond and Blundell & Bond estimation techniques were tried. The almost ideal demand systems and structural time series models were eliminated for practical or theoretical reasons prior to testing. Annual and route-specific data.	Journeys (in logs)	Preferred model included the following variables: Fare (log); total jobs at destination (log); working age population at origin (log); proportion of households without a car; car journey time; fuel price (log); generalised journey time (log); service quality index; lag of dependent variable; Hatfield derailment dummy.
Pickrell et al (Volpe National Transportation Systems Center, US Department of Transportation) (2011) *Developing a multi-level vehicle miles of travel forecasting model*	Provides overview of a model developed to provide forecast of future changes in passengers and freight vehicles in the US for various vehicle classes at national level, by state and for road categories. Only the national model is reviewed here.	Passenger and freight vehicle road travel.	Partial adjustment model. Annual data. Also developed fleet model (not reviewed here).	Light vehicle VMT per capita Single unit trucks VMT Combination trucks VMT	Real personal disposable income per capita (actual and squared); fuel costs per mile; lagged dependent variable; new vehicle price. Real GDP; fuel cost per mile; % urban road miles; lagged dependent variable. Real GDP; fuel cost per mile; interstate miles, lagged dependent variable; interaction of dummy for road freight deregulation in 1979 and GDP; interaction of deregulation dummy and fuel costs.

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
Productivity Commission (Australia 2006) *Road and rail freight infrastructure pricing*	Estimates price and income elasticities for rail and road freight transport. To capture the diversity of road freight transport, the estimates separate models by load type (bulk and non-bulk), area (urban and non-urban), and vehicle type (rigid and articulated).	Rail and road freight transport.	EC, with all variables expressed in logs	Freight in TKM	Each demand model is estimated using GDP, its own and substitute-mode's freight price (in cents/TKM), and a trade variable (ie import or export value).
Shen et al (2009) *Econometric modelling and forecasting of freight transport demand in Great Britain*	Estimates road and rail freight using six econometric models at an aggregate and disaggregate level for period 1974 to 2006.	Road and rail freight.	Alternative times series econometric models: OLS; PA model; ReADLM; VAR; TVP; STSM. Annual data. Cointegration and EC approaches rejected on basis that some data series were not suitable for these models.	Road and rail freight transport demand in tonne-km (in logs) by commodity level and total.	Indexes of industrial production (in logs) by commodity group and total. Dummy variables for various shocks including strikes, oil price hikes and major shift in 'Other' commodity demand since 1988. Also tried real GDP, total production output plus imports in tonnes and aggregate industrial production as measures of economic activity, as well as a price variable comprising an index of real road operating costs.

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
Victorio (2011) *New Zealand transport demand: Some evidence and forecasts* NZ Transport Agency internal study (not published)	To estimate the drivers of total, freight and passenger vehicle demand in New Zealand between 2002Q1 and 2010Q4 and to provide forecasts until 2015Q2.	Freight and passenger road vehicles.	OLS, auto regression and EC models. Quarterly data.	Total vehicle use (vehicle counts)	Average of petrol and diesel prices, real GDP; population. Included lag dependent and independent variables and change variables as in some model specifications.
				Freight vehicle use (vehicle counts)	Diesel prices; volume of retail goods in tonnes; real GDP. Included lag dependent and independent variables and change variables as in some model specifications.
				Passenger vehicle use (vehicle counts)	Petrol prices; national employment; total consumption spending by households and government. Included lag dependent and independent variables and change variables as in some model specifications.
Wang, J (2011) *Appraisal of factors influencing public transport patronage*	Development of a model to estimate drivers of public transport demand in three major New Zealand cities.	Public transport (rail and bus)	Partial adjustment model. Quarterly data, imputed from annual in some instances (other than Wellington for which annual data was used).	Log of patronage (trips per capita)	Logs of service levels (bus/train km per capita); real fare (revenue per capita), real income (disposable income per capita); cars per capita, real fuel price. Dummies for seasonal effects by quarter for Auckland and Christchurch. Stepwise regression procedure led to different models being estimated for each city.

Author(s) and title	Purpose	Sector	Estimation methodology	Dependent variable(s)	Independent variables
West, R et al (2011) *Identification and evaluation of freight demand factors*	Identify a set of regularly generated, well-documented, and easily obtainable variables with high statistical significance in explaining freight demand in the US to develop a freight demand forecasting model for infrastructure investors and capital planners.	Road, rail and waterborne freight.	OLS	Road: tonne-miles, vehicle-miles Rail: tonnage, tonne-miles, train-miles, car-miles, revenue tonne-miles. Water: waterway tonnes, waterway tonne-miles	Explore various drivers, including: real GDP, housing starts (thought to be a good proxy for economic optimism and consumption), the Industrial Production Index (which measures the relative output of manufacturing, mining and energy producers), import and export values, retail sales, employment (total and in the wholesale industry), and the inventory to sales ratio (thought to provide a counter-cyclical measure of the quantity of goods consumed). Group drivers into those that measure employment, consumption, production, commodity prices and foreign exchange; and then use the variables within each group to derive principal components.

Source: Frontier Economics

Appendix B: Glossary

2SLS	two-stage least squares
3SLS	three-stage least squares
AADT	annual average daily traffic
ARDL	autoregressive distributed lag (model)
CoF	certificate of fitness
EC	error correction (model)
Frontier	Frontier Economics
GDP	gross domestic product
GVA	gross value added
I-O	input-output
MPG	miles per gallon
MoT	Ministry of Transport
MSA	metropolitan statistical area
NLTDM	National Land Transport Demand Model
NZIER	New Zealand Institute for Economic Research
OLS	ordinary least squares
PA	partial adjustment (model)
PC	Australian Productivity Commission
PCA	principal component analysis
PPI	producer price index
RFI	road freight index
STSM	structural time series model
SUR	seemingly unrelated regression
TKM	tonnes kilometres travelled
TVP	time varying parameter (model)
VAR	vector autoregressive (model)
VIF	variance inflation factor
VKT	vehicle kilometres travelled
VMT	vehicle miles travelled