Forecasting Automobile Petrol Demand in Australia: An Evaluation of Empirical Models

Zheng Li
John M. Rose
David A. Hensher
Institute of Transport and Logistics Studies
Faculty of Economics and Business
The University of Sydney
NSW 2006 Australia
Zhengl@itls.usyd.edu.au
Johnr@itls.usyd.edu.au
Davidh@itls.usyd.edu.au

July 9 2008

Abstract

Transport fuel consumption and its determinants have received a great deal of attention since the early 1970s. In the literature, different types of modelling methods have been used to estimate petrol demand, each having methodological strengths and weaknesses. This paper is motivated by an ongoing need to review the effectiveness of empirical fuel demand forecasting models, with a focus on theoretical as well as practical considerations in the model-building processes of different model forms. We consider a linear trend model, a quadratic trend model, an exponential trend model, a single exponential smoothing model, a Holt's linear model, a Holt-Winters' model, a partial adjustment model (PAM) and an autoregressive integrated moving average (ARIMA**)** model. More importantly, the study herein identifies the difference between forecasts and actual observations of petrol demand in order to identify forecasting accuracy. Given the identified best-forecasting model, Australia's automobile petrol demand from 2007 through to 2020 is presented under the "business-as-usual" scenario.

Keywords: petrol demand forecasting, automobiles, time series data, the partial adjustment model, the autoregressive integrated moving average model, elasticities, trend-fitting approaches, exponential smoothing, forecasting effectiveness

1 Introduction

The influences underlying the consumption of fuel for transport activity have received a great deal of attention since the first oil *crisis* in the early 1970s (Espey, 1996). In addition to attempts to determine the key influences on petrol consumption, many studies examining fuel demand have been undertaken to predict future demand (Banaszak *et al.* 1999; Murat and Ceylan, 2006; Ediger and Akar, 2007). More recently, environmental concerns such as climate change have become increasingly important in the desire to understand global fuel demand, with a particular emphasis on the transport sector, given that 14 percent of global greenhouse gas emissions are estimated to be produced from this sector (Hensher, 2008). In looking at transport fuel demand, many studies have concentrated on automobile petrol demand, given that cars represent one of the major consumers and petrol is the dominant fuel source for the current passenger car fleet (for studies that do not look solely at automobile petrol demand, see e.g., Birol and Guerer, 1993; Samimi, 1995).

The prediction of fuel consumption has become an increasingly important tool for energy planning, with the primary purposes often cited as: to i) help policy makers develop appropriate pricing and taxation systems, ii) help decide future investments and decisions on oil reserves to improve energy security, iii) aid in addressing emission and pollution issues in advance, and iv) allow for planning of future energy needs, as well as to identify national infrastructure and research and development requirements. Understanding the determinants of transport fuel demand represents a key to the development of transport and environmental policies. Moreover, it is critical for decision makers to recognize the nature of fuel demand so that they can implement corresponding policies and regulations to ensure sustainable development.

In the transport and energy literature, petrol demand forecasting is an important topic with hundreds of studies having been undertaken to date (see e.g., Banaszak *et al.* 1999; Murat and Ceylan, 2006; Ediger and Akar, 2007). Researchers have used various types of modelling methods to estimate petrol demand., Whilst some studies have examined different approaches, these have typically only explored theoretical differences, usually without undertaking an empirical comparison of the practical usefulness in forecasting fuel demand (see e.g., Hunt *et al.*, 2003).

The purpose of this paper is to address this gap in the literature by empirically comparing the effectiveness of different forecasting models on fuel demand forecasts. In doing so, we describe not only the theoretical elements of the various models, but also the set of practical considerations that define the appeal of specific models. We test the accuracy of each of the forecast models by measuring forecast errors from a hold out sample of data. In total, eight models are built, namely a linear trend model, a quadratic trend model, an exponential trend model, a single exponential smoothing model, a Holt's linear model, a Holt-Winters' model, a partial adjustment model (PAM), and an autoregressive integrated moving average (ARIMA**)** model. The empirical data used to test alternative model specifications is drawn from Australia.

The organisation of this paper is as follows. In the following section, an overview of Australian automobile petrol demand is provided. This is followed by a brief literature review and data description, and the different forecasting models are presented. Model results are then provided, including an evaluation of the forecasting performance of each model, as well as the presentation of long term demand forecast predictions from the best-forecasting models. Conclusions are then drawn along with a discussion of the major findings as well as some key recommendations.

2 Australian Automobile Petrol Demand

In 2005, the total number of road transport vehicles in Australia was estimated to be over 13.9 million. Of these, approximately 80 percent were classified as passenger cars (ABS, 2006a). A significant characteristic of automobile usage in Australia is the high reliance on petrol, with 94 percent of automobiles using petrol as the primary source of combustion in Australia (ABS, 2007a). Over the period 1 November 2004 to 31 October 2005, approximately 28,967 million litres of road transport fuel was consumed. 64.6 percent (i.e., 18,712.7 million litres) was petrol, 30.0 percent diesel fuel (i.e., 8,690.1 million litres) with the remaining consumption representing other fuels (ABS 2006b). During the period, passenger automobiles represented the major end users of petrol. In total, automobiles consumed 15,856 million litres of petrol or 85 percent of total road petrol consumption in Australia (ABS, 2006b).

These figures can be further broken down into different vehicle types. In Australia, approximately 85 percent of total petrol attributable to road travel was consumed by passenger cars over the period 1 November 2004 to 31 October 2005, and articulated and rigid trucks (two main types of freight vehicles) used 65 percent of road diesel during the same period (ABS, 2006b). If light commercial vehicles and non-freight carrying trucks are also considered, the diesel share of goods vehicles would be much higher than 65 percent.

Coupled with high petrol consumption, Australia has also exhibited strong growth in car ownership. From 2001 to 2005, the number of automobiles increased by 12.5 percent, with fuel consumption by road motor vehicles increasing by 3,019 million litres. Road traffic for the corresponding period increased from 206,383 to 31,972 million tonne-kilometres (ABS, 2006b). Whilst there exist many possible causes for this, the two key drivers of these increases are thought to be i) continuing growth in household incomes and ii) increases in population, given that Australian population increased by 1.2 million over the period 2001-05, and meanwhile the average individual income jumped by 24 percent according to the Australian Bureau of Statistics.

3 Literature Review

The study of automobile fuel consumption is not new. Over the past four decades, many econometric studies have been examined the demand for fuel. Whilst the purposes of these studies are diverse, a significant concern has been to analyse the effects on petrol consumption resulting from the threat of fuel energy scarcity (Espey, 1996). More recently, environmental concerns have been cited as a key reason behind the desire to understand and model global fuel demand.

It is expected that an increase in price would lead to a decrease in the quantity demanded (a negative price elasticity), and a rising income is expected to stimulate petrol usage (a positive income elasticity). Also it can be expected that petrol demand would decline when income decreases, as less transport activities would be required in a less affluent economy. An interesting question is whether a decrease in price would lead to rising demand for fuel.

Breunig and Murphy (2007) have found that Australia's petrol demand response to price decreases is not significantly different from zero in the short run. That is, petrol demand would remain at the same level given a decrease in price in the short run. In another world, a lower price would not have an impact on petrol consumption in a short term. Breunig and Murphy also showed that the long-run demand responsiveness to petrol price increases and decreases is symmetrical. The estimated elasticities represent the percentage change in petrol demand resulting from one percent change of price or income.

One of the most common models for forecasting transport fuel demand and estimating elasticities is the partial adjustment model (PAM). The rational for a PAM is that it always takes time to fully respond to a change, and hence the instantaneous reaction is part of the ultimate level (Sterner and Dahl, 1992). For example, car drivers may respond to an increase in petrol price through short term actions such as switching to public transport, or medium to longer term actions such as replacing current cars with more fuel-efficient vehicles, or changing residential location. As a result of inflexibility in the stock of durables (e.g., car stock, home location, available public transport) and in conjunction with other factors such as personal habits or a lack of information, consumer reactions tend not to be instantaneous with significant lag effects evident in behaviour. For this reason, it is common to include a lagged dependent variable into the right-hand side of the equation (e.g., as with the partial adjustment model). Once responses to a change in an explanatory factor are complete, this situation is defined as the long run. Short run solutions in these models typically suggest that many responses are not fully complete due to some fixed factors.

Sterner and Dahl (1992) employed different models for estimating price and income elasticities for petrol demand, and apply them to the same OECD data from 1960 to 1985. They also reported that the short-run elasticities estimated from the partial adjustment model with time series data ranges -0.1 to -0.3 for price, and 0.15 to 0.55 for income. Birol and Guerer (1993) also used the PAM to estimate transport sector fuel demand for six developing countries, with annual historic data from 1971 to 1990. The study examined the demand for petrol and diesel fuels separately. The study focused on generating price and income elasticities for transport petrol and diesel demand under three separate scenarios: (i) low oil price and high GDP growth, (ii) base oil price and base GDP growth, and (iii) high oil price and low GDP growth. Under all three economic scenarios, the forecasted results predicted that each country's transport sector would consume significantly more fuel into the future.

Al-faris (1997) estimated the demand for petrol and other oil products with respect to price and income for six Gulf Cooperation Council (GCC) countries, by using a PAM with annual time series data between 1970 and 1991. Results of this study showed a significant variation in the estimates among fuel types (e.g., petrol, LPG, jet fuel, etc.) as well as across countries. Moreover, Al-faris found both price and income effects were inelastic in the short run, with an average short-run elasticity for the six GCC countries being -0.14 for fuel price and 0.13 for income.

Banaszak *et al.* (1999) also employed a PAM to examine automobile fuel demand for Taiwan and Korea using yearly data covering the period 1973–92. Banaszak *et al.* (1999) also estimated both short- and long-run income and price elasticities for fuel demand as well as providing a detailed comparison with similar studies conducted earlier. Banaszak *et al.* (1999) concluded that fuel demand is price inelastic in the short run, with the income effect being more significant than the price effect.

As well as the use of PAM, another well-developed modelling framework for analysing fuel consumption is cointegration[1] with error correction (EC)[2]. The approach is commonly associated with the development of unit root tests involving there steps; i) examine the stationarity of the dependent and independent variables through the augmented dicky-fuller (ADF) test; ii) estimate cointegrating relationships by checking the stationarity of the residuals; and finally iii) build an error correction model to estimate the short-run elasticities.

Samimi (1995) estimated the road transport fuel demand elasticities with respect to price and income for Australia. In that study, Samimi used the cointegration framework with an error correction model to quantify the short- and long-run elasticities. The short-run income elasticity is 0.25, while the short-run price elasticity is insignificant. The long-run income and price elasticities are 0.52 and -0.12 respectively.

Eltony and Al-Mutairi (1995) employed cointegration with an EC model to examine petrol demand in Kuwait. In this study, time series data related to petrol consumption per capita, petrol price and income per capita for the years 1970 to 1989 were used. Similar to Banaszak *et al.* (1999), Eltony and Al-Mutairi (1995) found that the income effects on petrol demand are much more significant than price effects. They report short- and long-run price elasticities of -0.37 and -0.46 respectively, and income elasticities of 0.47 in the short run and 0.92 in the long-run. Further, they found that 52 percent of the total adjustment towards the long-run level occurred within the first year, suggesting that petrol demand is price inelastic in comparison to income effects.

Ramanathan (1999) used the same framework to investigate the relationship between national income, petrol price and petrol demand in India. The data used for estimation by Ramanathan (1999) included yearly petrol consumption, GDP and petrol price for the period of 1972 to 1993. The short-run income elasticity was estimated as 1.18 compared to a long-run elasticity of 2.68. Similar to both Banaszak *et al.* (1999) and Eltony and Al-Mutairi (1995), Ramanathan (1999) also found that income effects were greater than price effects, with price elasticities estimated to have lower magnitudes (short-run: -0.21 and long-run: -0.32). Further, the study found that petrol demand tends to be price inelastic due to the relatively low gasoline consumption and steady economic growth in India. Road transport represents the main consumer of petrol, with the passenger car fleet increasing dramatically in India over the period of the study and continuing to do so. Nevertheless, it was found that during the period examined, India's petrol demand increased faster than economic growth.

---

[1] Cointegration in this context means that a long-run relationship exists, and hence the long-run elasticities may be calculated through a cointegration regression.

[2] The EC model examines the causal relationships between independent and dependent variables and generates short-run elasticities.

Most of the transport fuel demand studies discussed above have a common theme that is the examination of price and income effects. Table 1 presents the corresponding price and income elasticities from some reviewed studies in this paper, where the reviewed short-run price elasticities vary from -0.1 to -0.385, and short-run income elasticities are between 0.13 and 1.178. For the long-run estimates, the price elasticities vary from -0.12 to -0.866, and income elasticities are between 0.52 and 2.682. It has been found that each of those studies produced larger long-run estimates than the short-run, and income effects (both the short-run and long-run) are more significant than the corresponding price effects (with the exception of Al-faris' findings for six GCC countries). These similar results can also be found in some transport fuel demand reviews (see e.g., Graham and Glaister, 2002).

Table 1: Transport Fuel Demand Elasticities (S: Short-run elasticity; L: Long-run elasticity)

Author(s)	**Energy type**	**Data**	**Periodicity**	**Methodology**	**Price Elasticities**	**Income Elasticities**
Al-faris (1997)	Petrol and other fuels	Time series 1970-91	Yearly	Patrial adjustment	S=-0.14 (GCC countries)	S=0.13 (GCC countries)
Sterner and Dahl (1992)	Petrol	Time series 1960-85	Yearly	Patrial adjustment	S=-0.1 to -0.3 (OECD)	S=0.15-0.55 (OECD)
Banaszak *et al*. (1999)	Petrol and diesel	Time series 1973-92	Yearly	Patrial adjustment	S=-0.124 L=-0.519 (Taiwan) S=-0.385 L=-0.866 (Korea)	S=0.233 L=0.977 (Taiwan) S=0.439 L=0.989 (Korea)
Eltony and Al-Mutairi (1995)	Petrol	Time series 1970-89	Yearly	Cointegration	S=-0.37 L=-0.46 (Kuwait)	S=0.47 L=0.92 (Kuwait)
Sammi (1995)	Petrol and diesel	Time series 1980-93	Quarterly	Cointegration	S: insignificant L=-0.12 (Australia)	S=0.25 L=0.52 (Australia)
Ramanathan (1999)	Petrol	Time series 1972-93	Yearly	Cointegration	S=-0.209 L=-0.319 (India)	S=1.178 L=2.682 (India)

Besides general elasticity approaches (e.g., the PAM, cointegration modelling, etc.) which are designed to estimate the short- and long-run relationships between dependent and explanatory variables, the autoregressive integrated moving average (ARIMA) model represents an alternative statistical technique for modelling energy demand. Instead of considering the effects of independent variables on fuel consumption, ARIMA models analyse the stochastic properties of economic time series data on there own (Gujarati, 1995). An ARIMA model combines several time series techniques such as differencing, autoregressive (AR) models, and moving average (MA) models (Kumar *et al*., 2004). Compared with the elasticity analysis methods discussed above, ARIMA models focus on lagged observations from time series data.

Ediger and Akar (2007) predicted Turkey's primary energy demand from 2005 to 2020 using an ARIMA model. The data used in this study included annual consumption figures for different types of energy between 1950 and 2004. The forecasts show that Turkey's energy demand will continuously grow from 2005 to 2020. At the time of writing, Turkey's average growth rate for natural gas was 6.8 percent and oil products 1.6 percent. Another important result of the study was that fossil fuels would play an increasingly significant role in terms of Turkey's economic growth, and as a share of total energy is predicted to increase from 87.6 percent in 2005 to 91.6 percent in 2020.

Based on a review on over one hundred empirical road transport fuel consumption studies, Goodwin *et al.* (2004) drew the conclusion that passenger cars are more sensitive to price changes than freight vehicles. One explanation of the difference in sensitivity to fuel price is that direct fuel costs for freight vehicles account for a smaller proportion of total operating costs than for passenger cars. Therefore, the price elasticity for passenger cars' fuel demand is more significant than for goods vehicles' fuel demand.

The studies cited above all used time series data – observations on a single event over multiple time periods. Similarly, we use time series data analysis in estimating automobile fuel demand. In the next section we describe the data.

4 Data

Quarterly time series data is used for forecasting in this study, including total road petrol consumption (TPC), real gross domestic product (GDP), and real petrol price (RPP) for Australia over the period 1977q1 (the first quarter of 1977) to 2006q4 (the fourth quarter of 2006). The consumer price index (CPI) and petrol price index (PPI) are adjusted to 1998q1 as the base. Total road petrol consumption is measured in megalitres (millions of litres), obtained from three sources: Department of Primary Industries, Bureau of Transport and Communications Economics (BTCE), and Department of Industry, Tourism and Resources. Australian Bureau of Statistics (ABS) and BTCE are the main sources for gross domestic product data. Real GDP is seasonally adjusted to 1998 as the base, measured in million dollars. Nominal retail petrol price (NRPP) in cents is obtained from BTCE and ABS. Real petrol price (RPP) can be calculated by Equation (1).

$$RPP_t = NRPP_t * (\frac{PPI_{base}}{PPI_t}), \tag{1}$$

where RPP_t and $NRPP_t$ are the real and nominal retail petrol price at time period *t*, and PPI_{base} and PPI_t are petrol price indices at base time and period *t* respectively.

The data is divided into two parts. The first is from 1977q1 to 2005q1 which is used for modelling and estimation. The remaining data, from 2005q2 to 2006q4 is used as a hold out sample to examine the forecasting effectiveness of different forecasting approaches estimated on the first segment of data. The time series of each key data item is presented in Figures 1-3.. Petrol consumption tended to increase from 1977 to 2006, with increasing seasonal fluctuations over time. Overall, GDP had shown a steady increase over the same period. Real petrol price fluctuated between 60 and 80 cents before 1991, and then kept constant around 70 cents until 2006, except that it rapidly jumped to 78 cents during the second quarter of 2006.

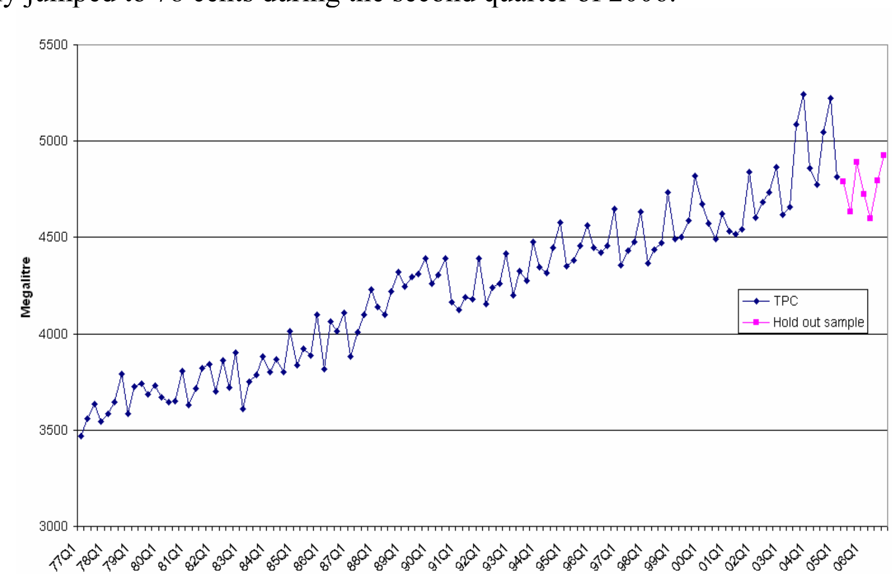

Figure 1: Total Road Petrol Consumption in Australia from 1977q1 to 2006q4

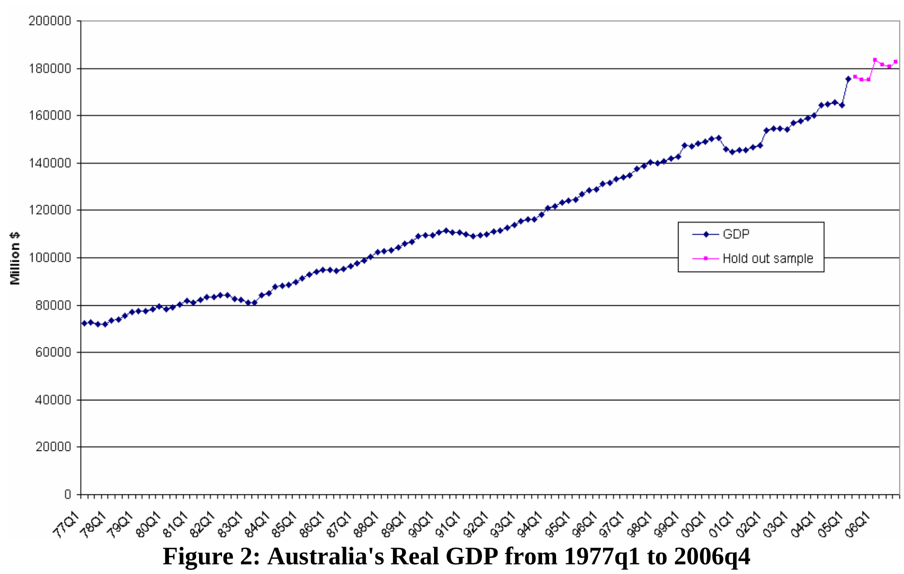

Figure 2: Australia's Real GDP from 1977q1 to 2006q4

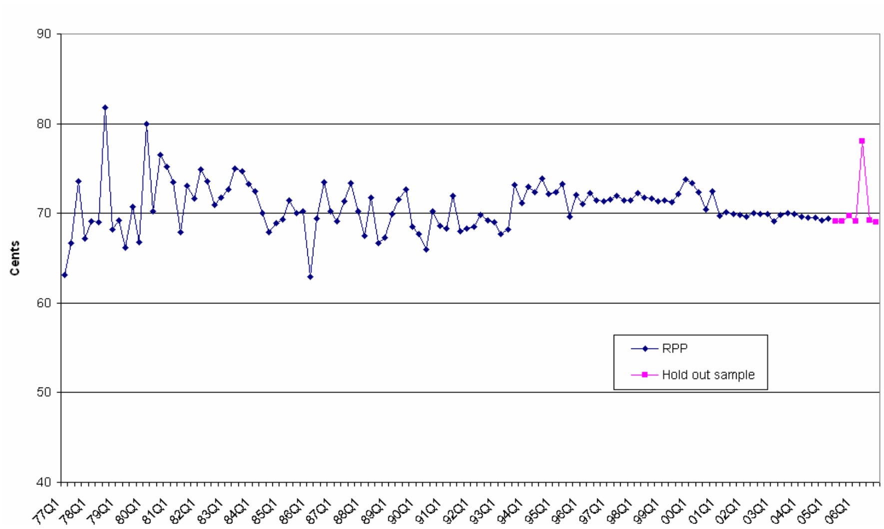

Figure 3: Real Petrol Price from 1977q1 to 2006q4

5 Modelling Methods

Approaches for econometric modelling and forecasting may be divided into four categories: (i) models used to estimate relationships between explanatory and dependent variables over periods of time, incorporating underlying economic processes; (ii) models that depict relationships between the past and current values, and forecast future events on the basis of historical outcomes only; (iii) cross sectional methods that analyse relationships between various variables at a point in time for different units; and (iv) approaches that consider relationships between dependent and independent variables for different units over time (Verbeek, 2004). The four types of econometric models require different data types, with the first two methods requiring time series data (i.e., observations on a single event over multiple time periods) with the third method requiring one off cross-sectional data, and the last requiring panel data incorporating the two dimensions of time series and cross-sectional data simultaneously.

In this study, time series data are collected to predict petrol demand in Australia. As such, we limit ourselves to only the first two types of econometric models. Among the first category of econometric models, PAM and cointegration with an ECM are two well-developed approaches. The autoregressive integrated moving average (ARIMA**)** model is an example of a model belonging to the second category. Both the PAM and cointegration with an ECM can clearly identify short- and long-run elasticities. However, PAM is capable of specifying all variables at the same level, while the ECM mechanism always uses differenced variables (e.g., a first-differencing transformation) to ensure stationarity before estimating. Thus, PAM is able to retain more information on a time series, compared with ECM. Sterner *et al.* (1992) also concluded that PAM is the most appropriate dynamic model for short- and long-run estimates at the individual country level. Given the focus herein on petrol demand for a single country and other advantages of a PAM, PAM is used for the empirical analysis.

5.1 The Partial Adjustment Model

Transport fuel demand is typically presented as a function of a series of economic variables (Samimi, 1995), with many studies identifying income and fuel price as the major parameters to determine energy consumption (Dahl and Sterner, 1991; Sammi, 1995; Wohlgemuth, 1997; Banaszak *et al*., 1999, Ramanathan and Subramanian, 2003, De Vita *et al*., 2006). Most studies have found that income (which often refers to GDP) is the most significant macroeconomic factor in determining transportation related fuel demand (e.g., Wohlgemuth, 1997) given that higher GDP usually indicates more business and trade activity, which in turn drives more transport requirements as well as increase fuel consumption. Typically petrol demand is represented as

$$TPC_t = f(GDP_t, RPP_t), \qquad (2)$$

where TPC_t is road petrol consumption at time period *t*, GDP_t is real gross domestic product at time period *t*, and RPP_t is the real petrol price at time period *t*.

Before developing an empirical econometric model for petrol demand, we discuss the concept of partial adjustment which is often used in conjunction with regression based models. Partial adjustment is used to account for the fact that, for time series data, adaptation may take time after factors such as price and income are changed, with consumers unable or unwilling to adjust as a result of inflexibility in the stock of consumer durables (Sterner and Dahl, 1992). Instead they adapt partially to the situation, given as "s" in the equation (3). The double-log functional form is a convenient form to generate elasticities.

$$\ln TPC_t - \ln TPC_{t-1} = s(\ln TPC_t^* - \ln TPC_{t-1}) + \mu_{t1}, \qquad (3)$$

and

$$\ln TPC_t^* = c_0 + c_1 \ln GDP_t + c_2 \ln RPP_t + \mu_{t2}, \qquad (4)$$

$\ln TPC_t$ is the natural log of instant road petrol consumption at time period *t*, and $\ln TPC_t^*$ is the natural log of ultimate level of desired consumption due to changes in price and/or GDP. By combining equations (3) and (4), total petrol demand can be expressed as equation (5).

$$\ln TPC_t = sc_0 + sc_1 \ln GDP_t + sc_2 \ln RPP_t + (1-s)\ln TPC_{t-1} + \mu_t, \qquad (5)$$

where c_0 is a constant, c_1, c_2 and c_3 are the unknown coefficients for corresponding independent variables, s lies between 0 and 1 and μ_t (random error term) = $\mu_{t1} + \mu_{t2}$.

When analysing time series data, seasonal effects may also play a significant role. This is particularly so for most types of energy demand, where seasonal variations are fairly obvious discernable patterns within the data. Given that quarterly time series data are employed in this study, seasonal dummy variables are introduced to examine the seasonal fluctuations, given in Equation (6).

$$S(D) = a_2 D_{2t} + a_3 D_{3t} + a_4 D_{4t}, \quad (6)$$

where D_{2t} = 1 for the second quarter and D_{2t} = 0 otherwise, D_{3t} = 1 for the third quarter and D_{3t} = 0 otherwise, D_{4t} = 1 for the fourth quarter and D_{4t} = 0 otherwise, D_{2t}, D_{3t} and D_{4t} are all equal to 0, for the first quarter and a_2, a_3 and a_4 are the unknown coefficients. Given (6), the petrol demand model is revised as Equation (7).

$$\ln TPC_t = sc_0 + sc_1 \ln GDP_t + sc_2 \ln RPP_t + (1-s)\ln TPC_{t-1} + a_2 D_{2t} + a_3 D_{3t} + a_4 D_{4t} + \mu_t. \quad (7)$$

Both short- and long-run elasticities can be estimated based on Equation (7). The corresponding short-run elasticities are the coefficients' (sc_1 and sc_2) values, and the long-run elasticities are the coefficients divided by *s*, which are c_1 and c_2.

5.2 The Autoregressive Integrated Moving Average Model

The Autoregressive integrated moving average (ARIMA) model combines several time series techniques such as differencing, autoregressive (AR) models, and moving average (MA) models (Kumar *et al.*, 2004). The ARIMA model allows for an analysis of the stochastic properties of economic time series on their own (Gujarati, 1995). Compared with regression models, ARIMA models focus on past or lagged periods of a time series.

A stationary time series always has the same mean, variance and autocovariance at any period of time (Gujarati, 1995). If a time series is stationary, it can be modelled by an AR process or a MA process or both. According to Bowerman *et al.* (2005), AR considers lagged effects of the time series itself (y_t), with a *p*-order autoregressive or AR (p) is presented as equation (8).

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \alpha_t, \quad (8)$$

δ is a constant, y_t is the dependent variable at time period *t*, ϕ_p are the parameters of y_{t-p} (p = 1, 2, 3 … p) and α_t is an uncorrelated random error term with zero mean and constant variance.

MA considers the effects of the current and past error terms (or randomness), a q-order moving average, or MA (q) is listed as

$$y_t = \mu_t - \theta_1 \mu_{t-1} - \theta_2 \mu_{t-2} - ... - \theta_q \mu_{t-q}, \quad (9)$$

where μ_t is an uncorrelated random error term with zero mean and constant variance, and θ_q is the parameters of μ_{t-q} (q = 1, 2, 3 … q).

Often a stationary time series has the characteristics of both AR and MA. Such data can therefore be modelled as a combination of both past values and past errors. This form of model, known as a ARMA (p, q) model, can be presented as equation (10).

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \mu_t - \theta_1 \mu_{t-1} - \theta_2 \mu_{t-2} - ... - \theta_q \mu_{t-q}. \quad (10)$$

The above models are based on stationary time series. However, time series data do not necessarily have to be stationary. To identify whether a time series data is stationary or not, it is common to plot the observations against time. If the values fluctuate with a constant variation around a constant mean, then they are said to be stationary. Stationary time series data can be modelled by an ARMA (p, q) model, as per equation (9). A non-stationary time series data can be estimated by adding an additional term to the ARMA model giving an ARIMA (p, d, q) model, where d indicates that the time series is differenced *d* times before it becomes stationary (Kumar *et al*., 2004).

A standard approach to ARIMA modelling is the Box-Jenkins methodology, consisting of four stages: identification, estimation, diagnostic checking and forecasting. Based on the characteristics of time series data, Bowerman *et al.* (2005) also divided the Box-Jenkins method into two categories for the purposes of modelling data exhibiting non-seasonal and seasonal trends. In this case, Bowerman *et al.* (2005) defined seasonal data as time series data with the presence of seasonal fluctuations. Seasonality is identified when similar patterns are observed at particular times of the year. Seasonality may occur when weekly, monthly or quarterly data is involved, which is a significant component of time series data.

5.3 Simple Methods for Forecasting Petrol Demand

In addition to the above econometric models, other simpler methods are available for use in demand forecasting. These simpler statistical approaches typically provide a straightforward means of directly calculating forecasts and include such models as the linear trend model, the quadratic trend model, and the exponential trend model. These models analyse trends of time series data and make forecasts based on the observed trends. The independent variable (x) is the time period code, and the first observation in the time series is assigned a code value of $x = 0$, then followed by the time period codes: 1, 2, 3 … n. The observed data value is the dependent variable (y). The method of least-squares is used to compute the values of coefficients. The forecasts of the dependent variable are achieved by substituting the corresponding time period code values into different forms of trend equations. For example, a linear trend equation is given as Equation (11).

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (11)$$

Where x_t is the time period code, which is set at 0 initially, β_0 is a constant, β_1 is the unknown coefficient and ε_t is a random error term. A quadratic trend model is given as Equation (12).

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \varepsilon_t \quad (12)$$

where β_0 is a constant, β_1 and β_2 are the unknown coefficients and ε_t is a random error term. An exponential trend model is given as Equation (13).

$$y_t = \beta_0 \beta_1^{x_t} \varepsilon_t \quad (13)$$

This equation can be transformed into a log-linear form:

$$\log y_t = \log \beta_0 + x_t \log \beta_1 + \log \varepsilon_t \quad (14)$$

where β_0 is a constant, β_1 is the unknown coefficients and ε_t is a random error term.

Besides those trend-fitting models, other simple methods are also possible, including exponential smoothing, Holt's linear method and Holt-Winters' method. In addition to the effect of smoothing time series data, exponential smoothing also can be used to estimate forecasts. Exponential smoothing allocates more weight to recent values in forecasting than the older observations, with the weight assigned to observations determined by smoothing parameters (Makridakis *et al.,* 1998). Holt's linear method and Holt-Winters' method are developed based on exponential smoothing. Detailed information on these simpler forecasting methods can be found in many sources (e.g., Levine *et al.*, 2005; Makridakis *et al.*, 1998). Interestingly, these simpler methods are often forsaken within the literature with a tendency towards the use of more econometrically advanced methods.

The research question addressed in this paper is: which method(s) provides the best forecast, with minimal forecasting error? To address this issue, we define and measure forecasting error using the mean absolute deviation (MAD) technique as defined in Equation (15).

$$MAD = \frac{\sum_{t=1}^{n} \left| y_t - f_t \right|}{n}, \quad (15)$$

where y_t is the actual observation in time period *t* and f_t is the forecast in time period *t*. According to Levine *et al*. (2005), the mean absolute deviation (MAD) is an effective measure of the average of the absolute differences between the actual observations and the predicted values of a time series. If a forecasting model fits the actual data in a given time series accurately, the MAD will be small. Thus, the method that produces the minimum MAD value in the hold out sample is selected as the best-forecasting model amongst the eight models examined herein.

6 Results, Analysis and Evaluation

All models reported here were estimated using Statistical Package for the Social Sciences (SPSS 15.0) and Statistical Analysis System (SAS 9.1). We now outline the results of the eight models.

6.1 The Partial Adjustment Model Results

Firstly, we estimate a PAM model (Equation 7) using the lag of the dependent variable (lagged LnTPC) as the partial adjustment mechanism. As quarterly data are used in this study, a lag structure up to the fourth lag is considered for each of the three independent variables (LnGDP, and lagged LnRPP and LnTPC). Dummy variables are also estimated to account for possible seasonal effects. The results are given Table 2.

The *F* statistic of the model (425.92) is large enough to indicate a strong linear relationship between the dependent and independent variables of the model. The adjusted R^2 of the model is 0.961 suggesting a very strong model fit. All the independent variables are significant at the 95 percent level of confidence level. The significant dummy variables indicate that seasonal variations for corresponding quarters do exist.

Table 2: The PAM Results

	Lagged Period	Par.	(*t-ratio*)	VIF
(Constant)	-	3.922	(9.131)	
D2	-	.022	(3.450)	2.305
D3	-	.028	(4.908)	1.938
D4	-	.054	(9.764)	1.770
LnGDP	0	.267	(8.046)	21.672
LnTPC1	1	.265	(2.994)	22.510
LnRPP1	1	-.216	(-4.498)	1.042

Variance inflation factors (VIF)[3] were also estimated as part of the model. VIF values estimated larger than 10 (see Table 1) suggest the presence of multicollinearity, with the presence of multicollinearity possibly influencing the stability of the regression coefficients. The presence of high VIF values is likely to be a result of endogeneity, given that, as an example, GDP may be considered endogenous as changes in petrol consumption may trigger changes in GDP. Thus, the high VIF values, which indicate highly correlated explanatory variables, may also represent other effects within the data also.

The use of a double log model suggests that the parameter estimates themselves represent short run elasticities. Although up to the fourth lag is applied to all variables (except seasonal dummy variables), only those statistically significant ones are shown in the final results (see Table 1). As such, the short-run price elasticity for real petrol prices (LnRPP1) is -0.216, where the first lag price is significant. This suggests that the price response only occurs one quarter after a change in price is realised, as quarterly data is used in modelling. Also, the short-run income elasticity (LnGDP) is estimated as 0.267, which is statistically significant for the current period of

[3] The variance inflation factor (VIF) test is a method to measure collinearity for each explanatory variable. It is given as $VIF_j = \frac{1}{1-R_j^2}$ where R_j^2 is the coefficient of multiple determination of explanatory variable x_j with respect to all other explanatory variables.

estimation. The coefficient for the first lag of LnTPC is also significant and equal to 0.265. The partial adjustment coefficient (*s*), calculated using Equation (5), is 0.735.

Long-run elasticities can be also calculated by dividing the estimated short-run elasticities by *s*. The long-run elasticity with respect to price is -0.294 and 0.363 with respect to income. These figures support the common finding by other researches that income has a larger impact on petrol demand than petrol prices (Dahl, 1994; Eltony and Al-Mutairi, 1995; Wohlgemuth, 1997; Ramanathan, 1999; Kayser, 2000; De Vita et al., 2006).

Table 3 compares price and income elasticities from Samimi (1995) and this study, both using Australian data. The short run income elasticity for this study is similar to that found by Samimi (1995), however, different long-run estimates are produced; 0.364 compared with 0.52. Explanations for the differences in the long-run estimates are provided by Espey (1998), who concludes that short-run income elasticities tend to be constant over time, whereas long-run income elasticities tend to decrease. Given that the time series data used by Samimi covers the period between 1980 and 1993, whilst this study employs data from 1977 to 2005, this finding is not surprising. Samimi also found that the short-run price elasticity was insignificant, whereas we have found that the single period lagged price elasticity is significant. Further, we have identified a larger in magnitude long-run price elasticity (-0.294) in comparison to that found by Samimi's (-0.12), again in line with one of Espey's conclusions that long-run income elasticities increase over time.

Table 3: The Comparison between the results from this Study and Samimi (1995)

	Samimi (1995)	**This study**
Energy type	Petrol and diesel	Petrol
Country	Australia	Australia
Data	Time series	Time series
Periodicity	Quarterly	Quarterly
Time frame	1980 to 1993	1977 to 2005
Methodology	Cointegration with error correction model	Partial adjustment model
Elasticities		
Short-run income	0.25	0.267
Long-run income	0.52	0.363
Short-run price	insignificant	-0.216
Long-run price	-0.12	-0.294

In addition to different time periods being examined between the two studies, another significant difference is that Samimi considered both diesel and petrol, whilst we concentrate solely on automobile petrol consumption. Although both passenger cars and freight vehicles can consume petrol and diesel, petrol is the main fuel for automobiles and diesel is the major fuel for goods vehicles. Moreover, different modelling methods are used. Samimi (1995) employs a cointegration with error correction model whereas we report the findings using a partial adjustment model. When using the cointegration with error correction model, the relationship between short- and long-run elasticities for price and income is not restricted to be the same as is the case with the partial adjustment model. These differences may also explain different long-run elasticities estimated from two studies.

The estimated elasticities are used to predict petrol demand during the sample-testing period from 2005q2 to 2006q4. Actual changes in real petrol price and income (GDP) over the sample-testing period can be calculated (seven quarters in total). The estimation of transport fuel elasticities is usually classified in a short run and long ru, where the long run may be over five years (Gallini, 1983; Eltony, 1993) or even over 15 years (Litman, 2007). The sample-testing period is less than two years. Thus, short-run price and income elasticities are used to generate forecasts for seven quarters' forecasting horizon.

Given our earlier discussion on short-run demand responsiveness to increases and decreases in price and income (i.e., symmetrical demand response to income increases and decreases, lower demand resulting from higher petrol price and no impact on demand if price decreases), petrol demand can be predicted based on actual percentage changes in price and income between 2005q2 and 2006q4. The results are summarised in Table 4.

Table 4: Forecasts Estimated by using the Short-run Elasticities

Year/Quarter	Real values ($2x4$ MA)	Forecasts	Errors
05Q2	4790.63	4978.62	-187.99
05Q3	4633.43	4969.71	-336.28
05Q4	4891.58	4958.98	-67.40
06Q1	4726	5021.88	-295.88
06Q2	4596.84	4882.88	-286.04
06Q3	4791.9	4875.47	-83.57
06Q4	4923.3	4890.75	32.55
		MAD=	**184.24**

Applying Samimi's short-run elasticities, the MAD value is 233.47, which is larger than the MAD value obtained for this study. Lower MAD values suggest better forecasts, as lower forecast errors are achieved. As such, the elasticities estimated from this study may be judged to be more accurate than those found by Samimi (1995). The evaluation of long-run forecasts has not been considered, as the long run has variously been described as being at least over five years (Eltony, 1993) and up to over 15 years (Litman, 2007). Evaluating long-run forecasts therefore requires a longer time period for the hold out sample than we have[4].

Time series data also often exhibit trend-cycles, seasonal and random effects. The trend-cycle component of time series data refers to changes in the level of a time series, seasonality refers to variations due to seasonal factors, and randomness to stochastic fluctuations present within the data (Makridakis *et al.*, 1998). Often, it is worthwhile to investigate each of these characteristics present to determine the likely

[4] Taking a longer hold out sample period will result in a significant loss in the degrees of freedom, which is expected to negatively impact upon the extrapolation of patterns observed within the time series data.

impacts on forecasts. In the current context, the use of quarterly data suggests the likely presence of seasonal effects. As such, a centred four moving average (2×4 MA) technique is used to average out the seasonal variation and randomness of quarterly time series data, so as to allow for a clear examination of trend effects. In doing so, we are able to not only explore the forecasting abilities of the eight forecasting models, but to compare their forecasting performance, accounting for both seasonally and non-seasonally adjusted data[5].

Figure 4 shows the non-seasonal TPC scatterplot after averaging out seasonality and randomness, by applying a centred four MA process. Comparing Figure 4 to Figure 1, it is clear that this process results in much smoother and cleaner data.

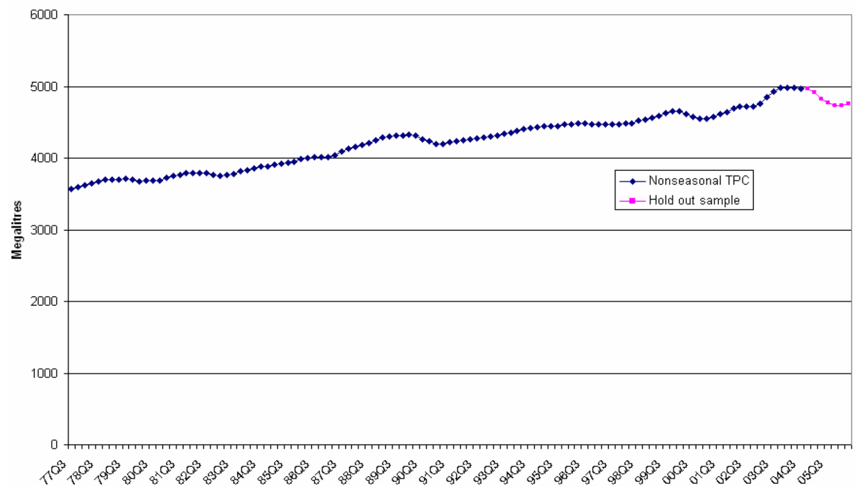

Figure 4: Non-seasonal TPC-Trend and Cycle Component

A centred four MA removes the beginning two and end two quarters from the original data. As the original data set for modelling and estimation ranges from 1977q1 to 2005q1, the non-seasonal time series after a 2×4 MA transformation starts at 1977q3 and ends at 2004q3. Because seasonality is averaged out, seasonal dummy variables should also be taken out of the PAM with seasonal data. As such, the PAM for non-seasonal data simplifies too equation (16).

$$\ln TPC_t = sc_0 + sc_1 \ln GDP_t + sc_2 \ln RPP_t + (1-s)\ln TPC_{t-1}. \qquad (16)$$

The modelling results based on non-seasonal TPC are given in Table 5 (only with those statistically significant lags).

Table 5: The PAM Results Using Non-seasonal TPC

	Par.	(*t-ratio*)	VIF
(Constant)	2.947	(6.900)	-
LNGDP	.231	(7.421)	33.956
LNRPP	-.076	(-2.160)	1.061
LNTPC4	.365	(4.245)	34.286

[5] In this study, the original time series of petrol consumption is denoted as "seasonal TPC" and the one after a centred four moving average transformation is denoted as "non-seasonal TPC".

The value of adjusted R^2 of the new model is 0.978, superior to the previous model with 0.961. Further, there is no obvious pattern in the regression residual scatterplots, which suggests that residuals are random and uncorrelated from one period to the next. All independent variables are statistically significant at 95 percent of confidence level. The coefficient for LNRPP (short-run price elasticity) is statistically significant at the current period, with a value of -0.076. The coefficient for LNGDP shows that the short-run income elasticity is 0.231. The coefficient for the lagged TPC is 0.365. Therefore, the partial adjustment mechanism (s) is $1-0.365=0.635$. The long-run elasticity estimated by using non-seasonal data is -0.12 for price and 0.364 for income. The forecasts for the period between 2004q4 and 2006q2 (the testing period for non-seasonal data) are given in Table 6.

Table 6: Forecasts Estimated by Using the Short-run Elasticities with Non-seasonal Data

Year/Quarter	Real values (2×4 MA)	Forecasts	Errors
04Q4	4966.86	4963.07	3.79
05Q1	4917.22	5037.99	-120.77
05Q2	4824.19	5045.50	-221.31
05Q3	4771.62	5037.60	-265.97
05Q4	4736.19	5033.34	-297.16
06Q1	4731.77	5088.57	-356.80
06Q2	4755.54	5031.59	-276.04
		MAD=	**220.26**

Interestingly, the fourth lag partial adjustment mechanism (LNTPC4) is significant whilst other lags for the variable are not (see Table 5). In comparison to the results from the PAM using seasonal data (see Table 2) where the first lag of the dependent variable was significant, the new results are rather unrealistic as they suggest that the impact of changes in the explanatory variables (i.e., price and income) on petrol demand behave as a step function. That is, the effect of a price change after one quarter is the same as after two quarters and three quarters but then increases after one year. As such, the elasticities generated from the model may be more biased than the original model. This partly explains the larger MAD value observed for this model.

Given the above, our preferred estimates are for the data are therefore -0.216 for the short-run price elasticity, 0.267 for the short-run income elasticity, -0.294 for the long-run price elasticity and 0.363 for the long run income elasticity.

6.2 ARIMA Modelling Results

Using the Box-Jenkins method, estimation of an ARIMA model is based on the sample autocorrelation (SAC) and the sample partial autocorrelation (SPAC) functions. The ARIMA model has different processes for seasonal and non-seasonal data, compared to all the other models, denoted as seasonal ARIMA modelling and non-seasonal ARIMA modelling. The modelling processes are summarised in Appendix A. After estimating the appropriate seasonal and non-seasonal ARIMA models, the forecasts are given in Table 7.

Table 7: Forecasts Estimated from Seasonal and Non-seasonal ARIMA Modelling

Year/Quarter	Seasonal ARIMA Forecasts	Year/Quarter	Non-seasonal ARIMA Forecasts
05Q2	4809.25	04Q4	4964.14
05Q3	5037.24	05Q1	4960.28
05Q4	5219.94	05Q2	4985.36
06Q1	4880.98	05Q3	5003.81
06Q2	4869.63	05Q4	5022.96
06Q3	5098.52	06Q1	5039.20
06Q4	5282.66	06Q2	5051.95

6.3 Forecast Results

MAD values are estimated to determine the forecasting ability of the different models for both seasonal and non-seasonal data. The results of these are given in Tables 8 and Table 9 respectively. The examination of the results produces some interesting conclusions. Firstly, all models provide reasonably accurate forecasts, with the mean absolute deviation in percentages varying from 3.63 percent to 6.09 percent. The short-run forecasting accuracy is between 93.91 percent and 96.37 percent. Secondly, the quadratic trend model produces the best forecasts for the seasonal data, with a MAD value of 170.14 and a forecasting error of 3.63 percent. For the non-seasonal data, the same model also produces the best forecasts, with a MAD value of 122.39 and an average forecasting error of 2.56 percent. Further, with the exception of Holt-Winters' method, simpler models tend to outperform more sophisticated models such as the ARIMA model.

Thirdly, with the exception of the PAM, the MAD values based on the non-seasonal data are smaller than the MAD values based on the seasonal data. Smaller MAD values suggest less-biased forecasts and as such, the forecasting performance of each model improves after replacing seasonal data with non-seasonal data (again, with the exception of the PAM).

Fourthly, the results show that forecasting accuracy tends to decrease, as the forecasting horizon increases. Fifth, most estimated forecasts are larger than the real value of the observations as most forecast errors are negative. Finally, the demand for road petrol in Australia shows an increasing trend over time, which the forecasts show are likely to continue into the future.

Table 8: MADs of Different Forecasting Models Using Seasonal Data

Year/Quarter	**Real values**	**Linear Trend**			**Quadratic Trend**			**Exponential Trend**			**Single Exponential Smoothing**		
		Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %
05Q2	4790.63	4901.86	-111.23	-2.32%	4899.60	-109.23	-2.27%	4902.30	-111.67	-2.33%	4958.81	-168.18	-3.51%
05Q3	4633.43	4913.77	-280.34	-6.05%	**4911.39**	-278.22	-6.00%	4915.87	-282.43	-6.10%	4958.81	-325.38	-7.02%
05Q4	4891.58	4925.68	-34.10	-0.70%	**4923.18**	-31.86	-0.65%	4929.47	-37.89	-0.77%	4958.81	-67.23	-1.37%
06Q1	4726.00	4937.58	-211.59	-4.48%	4934.96	-209.24	-4.42%	4943.11	-217.11	-4.59%	4958.81	-232.82	-4.93%
06Q2	4596.84	4949.49	-352.64	-7.67%	4946.74	-350.18	-7.61%	4956.78	-359.94	-7.83%	4958.81	-361.97	-7.87%
06Q3	4791.90	4961.39	-169.49	-3.54%	4958.52	-166.91	-3.48%	4970.50	-178.60	-3.73%	4958.81	-166.91	-3.48%
06Q4	4923.30	4973.30	-50.00	-1.02%	4970.30	-47.29	-0.95%	4984.25	-60.95	-1.24%	4958.81	-35.51	-0.72%
		MAD =	172.77	3.68%	**MAD =**	**170.14**	**3.63%**	MAD =	178.37	3.80%	MAD =	194.00	4.13%

Year/Quarter	**Real values**	**Holt's Linear Method**			**Holt-Winters' Method**			**ARIMA**			**Partial Adjustment Model**		
		Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error%	Forecasts	Errors	Error %
05Q2	4790.63	4978.34	-187.71	-3.92%	4900.09	-109.46	-2.28%	**4809.25**	-18.62	-0.39%	4978.62	-187.99	-3.92%
05Q3	4633.43	4991.47	-358.03	-7.73%	5042.89	-409.46	-8.84%	5037.24	-403.81	-8.72%	4969.71	-336.28	-7.26%
05Q4	4891.58	5004.59	-113.01	-2.31%	5209.74	-318.16	-6.50%	5219.94	-328.36	-6.71%	4958.98	-67.40	-1.3%
06Q1	4726.00	5017.71	-291.72	-6.17%	4918.55	-192.56	-4.07%	**4880.98**	-154.99	-3.28%	5021.88	-295.88	-6.26%
06Q2	4596.84	5030.83	-433.99	-9.44%	4950.29	-353.45	-7.69%	**4869.63**	-272.79	-5.93%	4882.88	-286.04	-6.22%
06Q3	4791.90	5043.95	-252.05	-5.26%	5094.43	-302.53	-6.31%	5098.52	-306.62	-6.40%	**4875.47**	-83.57	-1.71%
06Q4	4923.30	5057.08	-133.78	-2.72%	5262.84	-339.54	-6.90%	5282.66	-359.36	-7.30%	**4890.75**	32.55	0.066%
		MAD=	252.90	5.36%	MAD=	289.31	6.09%	MAD=	263.51	5.53%	MAD=	184.24	3.92%

Table 9: MADs of Different Forecasting Models Using Non-seasonal Data

Year/Quarter	**Real values (2X4MA)**	**Linear Trend**			**Quadratic Trend**			**Exponential Trend**			**Single Exponential Smoothing**		
		Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %
04Q4	4966.86	4868.93	97.93	1.97 %	4859.51	107.35	2.16%	4905.58	61.27	1.23 %	**4970.05**	-3.19	-0.06 %
05Q1	4917.22	4880.69	36.52	0.74 %	4870.76	46.46	0.94 %	**4919.34**	-2.13	-0.04 %	4970.05	-52.84	-1.07 %
05Q2	4824.19	4892.46	-68.27	-1.42 %	**4882.00**	57.81	-1.20 %	4933.15	-108.95	-2.26 %	4970.05	-145.86	-3.02 %
05Q3	4771.62	4904.22	-132.60	-2.78 %	**4893.23**	121.61	-2.55 %	4946.99	-175.36	-3.68 %	4970.05	-198.43	-4.16 %
05Q4	4736.19	4915.99	-179.80	-3.80 %	**4904.46**	168.27	-3.55 %	4960.86	-224.68	-4.74 %	4970.05	-233.87	-4.94 %
06Q1	4731.77	4927.75	-195.98	-4.14 %	**4915.67**	183.90	-3.89 %	4974.78	-243.01	-5.14 %	4970.05	-238.28	-5.04 %
06Q2	4755.54	4939.51	-183.97	-3.87 %	4926.87	171.33	-3.60 %	4988.74	-233.19	-4.90 %	4970.05	-214.51	-4.51 %
		MAD=	127.87	2.67%	**MAD=**	**122.39**	**2.56%**	MAD=	149.80	3.14%	MAD=	155.28	3.26%

Year/Quarter	**Real values (2X4MA)**	**Holt's Linear Method**			**Holt-Winters' Method**			**ARIMA**			**Partial Adjustment Model**		
		Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %	Forecasts	Errors	Error %
04Q4	4966.86	4962.24	4.62	0.09%	4984.13	-17.27	-0.35%	4964.14	2.71	0.05%	4963.07	3.79	0.076%
05Q1	4917.22	4954.43	-37.21	-0.76 %	4998.21	-80.99	-1.65 %	4960.28	-43.07	-0.88 %	5037.99	-120.77	-2.46%
05Q2	4824.19	4946.62	-122.43	-2.54 %	5012.29	-188.09	-3.90 %	4985.36	-161.17	-3.34 %	5045.50	-221.31	-4.59%
05Q3	4771.62	4938.81	-167.18	-3.50 %	5026.36	-254.74	-5.34 %	5003.81	-232.18	-4.87 %	5037.60	-265.97	-5.57%
05Q4	4736.19	4931.00	-194.81	-4.11 %	5040.44	-304.25	-6.42 %	5022.96	-286.78	-6.06 %	5033.34	-297.16	-6.27&
06Q1	4731.77	4923.19	-191.42	-4.05 %	5054.52	-322.75	-6.82 %	5039.20	-307.43	-6.50 %	5088.57	-356.80	-7.54%
06Q2	4755.54	**4915.38**	-159.83	-3.36 %	5068.60	-313.05	-6.58 %	5051.95	-296.41	-6.23 %	5031.59	-276.04	-5.80%
		MAD=	125.36	2.63%	MAD=	211.59	4.44%	MAD=	189.96	3.99%	MAD=	222.71	4.62%

6.4 Optimal Forecasts

For the seasonal data, the quadratic trend model is identified as the best-forecasting model. Thus, it is used to estimate quarterly petrol demand from 2007 to 2020. The seasonal forecasts for several time spans are given in Table 10.

Table 10: Quarterly Petrol Demand Forecasts (Megalitres) 2007-20 Using Seasonal Data

	2007	2010	2015	2020
Q1	4982.08	5123.23	5357.84	5591.62
Q2	4993.85	5134.98	5369.54	5603.29
Q3	5005.62	5146.73	5381.25	5614.96
Q4	5017.39	5158.47	5392.96	5626.62

For the non-seasonal data, the quadratic trend model has the best forecasting ability, and Table 11 provides the estimated quarterly forecasts between 2007 and 2020.

Table 11: Quarterly Petrol Demand Forecasts (Megalitres) 2007-20 Using Non-seasonal Data

	2007	2010	2015	2020
Q1	4960.43	5093.82	5313.18	5528.84
Q2	4971.60	5104.88	5324.05	5539.53
Q3	4982.75	5115.93	5334.92	5550.20
Q4	4993.90	5126.96	5345.77	5560.87

Although quarterly seasonal and non-seasonal data are different for the same quarter, theoretically, there should be no difference after aggregating the quarterly data into yearly data, assuming that the seasonal component of the data is constant from year to year. A centred four MA transformation is a combination of two simple four moving averaging processes, given as the following equations.

$$T_{2.5} = \frac{(Y_1 + Y_2 + Y_3 + Y_4)}{4} \quad \text{(Simple four MA)} \qquad (17)$$

$$T_{3.5} = \frac{(Y_2 + Y_3 + Y_4 + Y_5)}{4} \quad \text{(Simple four MA)} \qquad (18)$$

$$T_3 = \frac{T_{2.5} + T_{3.5}}{2} = \frac{(0.5Y_1 + Y_2 + Y_3 + Y_4 + 0.5Y_5)}{4} \quad \text{(Centred four MA)} \qquad (19)$$

Y_1 and Y_5 represent the same quarters, but in two different years. Thus, the seasonal and random components are not removed, but averaged out by giving equal weight (0.25) to each quarter in Equations (17) to (19). Therefore, the two sets of yearly data should correspond when aggregating from quarterly into yearly data. By adding up quarterly figures into yearly figures from 1978 to 2005, the mean absolute difference in percentages between seasonal and non-seasonal data (yearly) is only 0.22 percent over the same period. This slight difference can be explained by the inconstant seasonal component in different years. Given the slight difference between yearly seasonal and non-seasonal TPC, and the more accurate forecasts when using non-seasonal data, the quadratic trend model with non-seasonal TPC is identified as the best forecasting model to estimate annual forecasts until 2020. Table 12 presents the forecasts using this model.

Table 12: Annual Petrol Demand Forecasts (Megalitres) 2007-20 Using Non-seasonal Data

	2007	2010	2015	2020
Annual	19908.68	20441.59	21317.92	22179.44

The estimated forecasts show that annual road petrol demand is expected to reach 22,179.44 megalitres by 2020 which is 20.8 percent more than consumption in 2000 (i.e., 18,360.6 megalitres)[6]. The passenger car sector has been the major road petrol consumer in Australia and its petrol share remained unchanged for almost a decade, according to a series of ABS Survey of Motor Vehicle Use. Approximately 85 percent of Australian road petrol was consumed by automobiles in 2006 (ABS, 2007b). Assuming that this share figure keeps constant at 85 percent until 2020, 18,852.52 megalitres of petrol would be consumed by automobiles by 2020, which is approximately 16.7 percent higher than automobile petrol consumed in 2000, given that 88 percent of total road petrol was consumed by automobiles in 2000.

The Bureau of Infrastructure, Transport and Regional Economics (BITRE) estimated that total greenhouse gas emissions from automobiles in 2020 in Australia will be approximately 23 percent higher than in 2000 (BITRE, 2002). Given that emissions are correlated with petrol consumption, the BITRE's prediction supports the predictions found in this current study. That is, automobile petrol consumption would reach 18,852.52 megalitres by 2020, increasing by 16.7 percent over the 2000 to 2020 period.

7 Discussion and Conclusions

This paper provides an analysis of future petrol demand in Australia. Eight models (classified into simple and complicated methods) were employed. Different models were estimated allowing for a comparison of forecasts to establish which model is most likely to produce the more accurate forecasts. To evaluate forecast performance, a hold out sample was employed to test the effectiveness of corresponding forecasts.

For the seasonal data, the best-forecasting model is the quadratic trend model, with a MAD value of 170.14 and a mean absolute error of 3.63 percent in a short run. For the nonseasonal data, the same model produced the most accurate short-term forecast, with the lowest MAD value of 122.39 and a mean absolute error of 2.56 percent in a short run. After aggregating quarterly data into yearly data, the quadratic trend model with the non-seasonal data is the best model for yearly forecasts. The modelling results indicate that the annual demand for road transport petrol in Australia is expected to increase by over 20.8 percent between 2000 and 2020, with automobile petrol demand forecast to increase by approximately 16.7 percent over the same period, and reach 18,852.52 megalitres in 2020.

[6] Our data series goes up to 2006. The big spike in petrol prices occurred in 2008, with average prices increasing from $Aud1.20 to $Aud1.60 in the first half of 2008. This is likely to deflate the forecasts, although we suspect that this will be a small adjustment given the relatively small amount of aggregate disposable income spent on car use (in contrast to car ownership). We also expect that manufacturers will respond by delivering more fuel efficient cars to the market, like they did in the 1970's after the price increases. The big influence on demand growth will remain population growth.

The evaluation of the forecasting ability of the models shows that simple methods (e.g., the quadratic trend model) have produced better forecasting results than sophisticated methods (e.g., ARIMA). Compared with simple models, sophisticated models may be more suitable in delivering improved goodness of fit of the history; however they do not always deliver more accuracy in predicting the future. The results from this study show that simple methods better extrapolate the patterns of a time series so as to estimate more accurate forecasts. Compared with the PAM model in this study, Breunig and Gisz (2008) developed a far more complicated regression model (where the time trend is added in the equation as a independent variable, residuals (μ_t) is considered as a q^{th}-order moving average, etc) with quarterly seasonal Australian data from 1966 to 2006 to estimate the short-run price elasticity being -0.13 and 0.16 for income. Applying Breunig and Gisz' estimates, the MAD value over the same testing period is 187.3, which is larger than the MAD estimated from the PAM with seasonal data in this study (i.e., 184.24). Therefore, the responsiveness of petrol demand to price and income changes were better captured in the relatively simpler model in this study than the Breunig and Gisz model. This evidence supports our finding that simple methods may produce more accurate estimates.

The key conclusion of this paper is that more econometrically sophisticated methods do not always produce better forecast results than simpler models. Our findings suggest that simpler statistical techniques may improve forecasting performance significantly, alongside data adjustment. For example, our findings suggest that forecasts became more accurate after using a centred four moving average to average out seasonal components of time series data. It may be time to re-think the current trend in the literature of increasing model complexity. .

References

ABS (2000) *Survey of Motor Vehicle Use, Australia, Oct. 1998*, Australian Government.

ABS (2001) *Survey of Motor Vehicle Use, Australia, Oct. 2000*, Australian Government.

ABS (2006a) *Australia's Environment: Issues and Trends*, Australian Government.

ABS (2006b) *Survey of Motor Vehicle Use, Australia, 01 Nov 2004 to 31 Oct 2005*, Australian Government.

ABS (2007a) *Motor Vehicle Census, Australia*, Australian Government.

ABS (2007b) *Survey of Motor Vehicle Use, Australia, 12 months ended 31 October 2006*, Australian Government.

Al-faris, A.F. (1997) Demand for oil products in the GCC countries**,** *Energy Policy*, 25(1), 55-61.

Banaszak, S., Chakravorty, U. and Leung, P. S. (1999) Demand for ground transportation fuel and pricing policy in Asian Tigers: A comparative study of Korea and Taiwan, *The Energy Journal*, 20(2), 145-166.

Birol, F. and Guerer, N. (1993) Modelling the transport sector fuel demand for developing economies, *Energy Policy***,** 21(12), 1163-1173.

Bowerman, B. L., O'Connell, R. T. and Koehler, A. B. (2005) *Forecasting Time Series and Regression: an Applied Approach, 4th Edition*, Thomson Brooks/Cole, California, USA.

Breunig, R. V. and Murphy, C. (2007) *Single-equation estimates of Australian petrol demand: methodological issues and implications of different modeling strategies*, the Australian National University.

Breunig, R. V. and Gisz, C. (2008) An exploration of Australian petrol demand: unobservable habits, irreversibility, and some updated estimates, Resubmitted to *the Economic Record*.

BTRE (2002) *Greenhouse Gas Emissions from Transport*, Australian Government.

Dahl, C. A. and Sterner, T. (1991) Analysing gasoline demand elasticities: a survey, *Energy Economics*, 13(3), 203-210.

De Vita, G., Endresen, K. and Hunt, L. C. (2006) An empirical analysis of energy demand in Namibia, *Energy Policy*, 34(18), 3447–3463.

Ediger, V. S. and Akar, S. (2007) ARIMA forecasting of primary energy demand by fuel in Turkey, *Energy Policy*, 35(3), 1701-1708.

Eltony, M. (1993) Transport gasoline demand in Canada, *Journal of Transport Economics and Policy,* 27, 249–253.

Eltony, M. N. and Al-Mutairi, N. H. (1995) Demand for gasoline in Kuwait: An empirical analysis using cointegration techniques, *Energy Economics,* 17(3), 249-253.

Espey, M. (1996) Explaining the variation in elasticity estimates of gasoline demand in the Unite States: A meta-analysis, *Energy Economics*, 17(3), 315-323.

Gallini, N. (1983) Demand for gasoline in Canada, *Canadian Journal of Economics,* 16(2), 299–234.

Goodwin, P., Dargay, J. and Hanly, M. (2004) Elasticities of Road Traffic and Fuel Consumption with Respect to Price and Income: A Review, *Transport Reviews*, 21(3), 275–292.

Graham, D. J. and Glaister, S. (2002) The demand for automobile fuel: a survey of elasticities, *Journal of Transport Economics and Policy*, 36(1), 1-26.

Gujarati, D. N. (1995) *Basic Econometrics*, *3rd Edition*, McGraw-Hill, New York.

Hensher, D.A. (2008) Climate change, enhanced greenhouse gas emissions and passenger transport- what can we do to make a difference? *Transportation Research D* 13(2), 95-111.

Hunt, L.C., Judge, G. and Ninomiya, Y. (2003) Underlying trends and seasonality in UK energy demand: a sectoral analysis, *Energy Economics*, 25(1), 93–118.

Kumar, K., Yadav, A. K., Singh, M. P. and Jain, V. K. (2004) Forecasting daily maximum surface ozone concentrations in Brunei Darussalam-An ARIMA modelling approach, *Journal of the Air & Waste Management Association*, 54(7), 809-815.

Levine, D. M., Stephan, D., Krehbiel, T. C. and Berenson, M. L. (2005) *Statistics for Managers, Using Microsoft Excel, 4th Edition*, Pearson Prentice Hall.

Litman, T. (2007) *Transportation Elasticities: How Prices and Other Factors Affect Travel Behavior*, Victoria Transport Policy Institute.

Makridakis, S., Wheelwright, S. C. and Hyndman, R. J. (1998) *Forecasting: Method and Applications, 3rd Edition*, John Wiley & Sons, NJ.

Murat Y. S. and Ceylan, H. (2006) Use of artificial neural networks for transport energy demand modelling, *Energy Policy*, 34(17), 3165–3172.

Ramanathan, R. (1999) Short- and long-run elasticities of gasoline demand in India: An empirical analysis using cointegration techniques, *Energy Economics*, 21(4), 321-330.
Ramanathan, R. and Subramanian, G. (2003) Elasticities of gasoline demand in the Sultanate of Oman, *Pacific and Asian Journal of Energy*, 13(2), 105-114.
Samimi, R. (1995) Road Transport Energy Demand in Australia: A Cointegration Approach, *Energy Economics*, 17(4), 329-339.
Sterner, T. and Dahl, C. A. (1992) Modelling transport fuel demand, in T. Sterner (ed), *International Energy Economics*: 65-81, Chapman and Hall, London.
Sterner, T., Dahl, C. and Franzén, M. (1992) Gasoline tax policy: carbon emissions and the global environment, *Journal of Transport Economics and Policy*, 26, 109–119.
Verbeek (2004) *A guide to modern econometrics, 2nd Edition*, John Wiley & Sons, NJ.
Wohlgemuth, N. (1997) World transport energy demand modelling, *Energy Policy,* 25(14-l5): 1109-1119.

Appendix A

Seasonal ARIMA modelling results

The results of seasonal ARIMA modelling are listed as follows, including pre-differencing transformation, identification, estimation, diagnostic checking and forecasting.

Pre-differencing transformation. The initial step is to choose the right pre-differencing transformation. The comparison of the scatterplots from three pre-differencing transformations (square root, quartic root and a natural logarithm) shows that only the natural logarithm transformation produces relatively constant seasonal variances over time. Therefore, the appropriate pre-differencing (y^*) should be the natural log of y ($\ln y$).

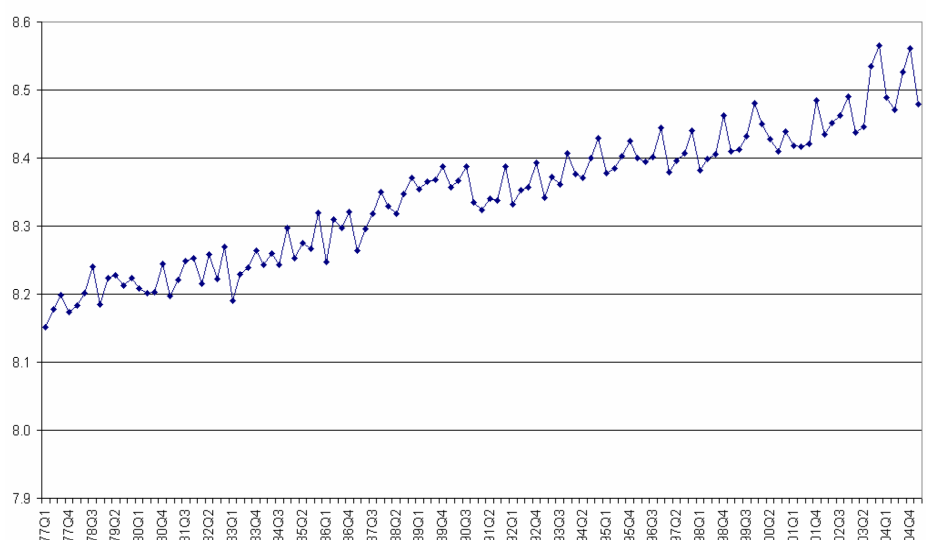

Figure 5: Natural Logarithm Pre-differencing Transformation of Seasonal TPC

Identification. The identification of seasonal ARIMA modelling is based on evaluating the SAC and SPAC functions of four types of stationary transformations. Only the third stationary transformation ($z_t = y^*_t - y^*_{t-4}$) has produced the SAC which dies down quickly at the non-seasonal and seasonal levels. This means that the time series of seasonal TPC has been transformed into stationary. Moreover the SAC spikes at lag four (seasonal level) only, and the SPAC spikes at lag one (non-seasonal level). Therefore, it can be identified that MA term at lag four is significant, and AR terms at lag one are significant.

```
Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
  1       0.28733   |                .   |******              |
  2       0.11731   |                .   |** .                |
  3       0.00151   |                .   |   .                |
  4      -0.35254   |             *******|   .                |
  5       0.10345   |                .   |** .                |
  6       0.02084   |                .   |   .                |
  7      -0.07318   |                .  *|   .                |
  8      -0.15659   |                .***|   .                |
  9       0.13210   |                .   |***.                |
 10      -0.07343   |                .  *|   .                |
 11      -0.11312   |                . **|   .                |
 12      -0.19577   |                ****|   .                |
 13      -0.02302   |                .   |   .                |
 14       0.18868   |                .   |****                |
 15       0.02193   |                .   |   .                |
 16      -0.12661   |                .***|   .                |
 17       0.04341   |                .   |*  .                |
 18      -0.13169   |                .***|   .                |
 19       0.01166   |                .   |   .                |
 20       0.00069   |                .   |   .                |
 21      -0.03467   |                .  *|   .                |
 22      -0.01318   |                .   |   .                |
 23       0.00231   |                .   |   .                |
 24      -0.14516   |                .***|   .                |
```

Figure 6: Sample Autocorrelation Function of Seasonal TPC

```
Lag   Covariance   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1    Std Error
  1   0.00016758       0.28733   |               .    |******              |     0.095783
  2   0.00011092       0.19019   |               .    |****                |     0.103388
  3   0.00004859       0.08331   |               .    |** .                |     0.106550
  4   -0.0001605       -.27519   |              ******|    .               |     0.107146
  5   -0.0000343       -.05877   |               .   *|    .               |     0.113445
  6   -0.0000454       -.07779   |               .  **|    .               |     0.113724
  7   -0.0000679       -.11635   |               .  **|    .               |     0.114211
  8   -0.0000417       -.07150   |               .   *|    .               |     0.115294
  9     2.8737E-6      0.00493   |               .    |    .               |     0.115700
 10   -0.0000388       -.06647   |               .   *|    .               |     0.115702
 11   -0.0000382       -.06546   |               .   *|    .               |     0.116051
 12   -0.0000816       -.13991   |               . ***|    .               |     0.116390
 13   -0.0001092       -.18729   |               .****|    .               |     0.117922
 14   0.00003149       0.05399   |               .    |*   .               |     0.120620
 15   0.00003649       0.06256   |               .    |*   .               |     0.120842
 16   0.00004391       0.07529   |               .    |**  .               |     0.121139
 17   0.00008637       0.14808   |               .    |*** .               |     0.121567
 18   -0.0000670       -.11489   |               .  **|    .               |     0.123211
 19   -0.0000177       -.03043   |               .   *|    .               |     0.124190
 20    5.49996E-6      0.00943   |               .    |    .               |     0.124258
 21   -0.0000606       -.10389   |               .  **|    .               |     0.124265
 22   0.00001449       0.02485   |               .    |    .               |     0.125059
 23    1.50465E-7      0.00026   |               .    |    .               |     0.125105
 24   -0.0000507       -.08691   |               .  **|    .               |     0.125105
```

Figure 7: Sample Partial Autocorrelation Function of Seasonal TPC

Estimation. Table 13 presents the estimated coefficients of AR and MA.

Table 13: Estimated Coefficients for Seasonal ARIMA Modelling

	Par.	(*t-ratio*)	Lag
MU	.01185	(8.47)	0
MA1,1	.59698	(7.19)	4
AR1,1	.38852	(4.31)	1

Diagnostic checking. According to the criteria of diagnostic checking, the estimated model satisfies the following two key points to an appropriate ARIMA model. First, the p-values for K = 6, 12, 18 and 24 are all greater than 0.05, and secondly there are no spikes in the Sample Autocorrelation function of the residuals (RSAC) or the Sample Partial Autocorrelation function of the residuals (RSPAC). Therefore it can be proven that the above model is adequate and can be used for forecasting.

Table 14: Seasonal ARIMA Diagnostic Checking

To Lag	Chi-Square	DF	P	Autocorrelations					
6	5.77	4	.2169	-.081	.165	.069	.097	.039	.038
12	8.34	10	.5953	-.062	-.061	.011	-.012	-.009	-.114
18	16.07	16	.4481	-.145	.029	.030	-.049	.084	-.165
24	18.96	22	.6480	.037	.005	-.091	-.026	.013	-.101

```
Lag  Covariance   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1   Std Error
  1  -0.0000364     -.08096    |                . **|   .                |   0.095783
  2  0.00007405     0.16454    |                .   |***.                |   0.096408
  3  0.00003100     0.06889    |                .   |*  .                |   0.098951
  4  0.00004379     0.09730    |                .   |** .                |   0.099390
  5  0.00001742     0.03871    |                .   |*  .                |   0.100260
  6  0.00001708     0.03796    |                .   |*  .                |   0.100397
  7  -0.0000280     -.06221    |                .  *|   .                |   0.100529
  8  -0.0000273     -.06057    |                .  *|   .                |   0.100881
  9  4.97124E-6     0.01105    |                .   |   .                |   0.101214
 10  -5.3874E-6     -.01197    |                .   |   .                |   0.101225
 11  -3.8699E-6     -.00860    |                .   |   .                |   0.101238
 12  -0.0000515     -.11440    |                . **|   .                |   0.101245
 13  -0.0000651     -.14466    |                .***|   .                |   0.102424
 14  0.00001315     0.02923    |                .   |*  .                |   0.104282
 15  0.00001330     0.02956    |                .   |*  .                |   0.104357
 16  -0.0000222     -.04944    |                .  *|   .                |   0.104434
 17  0.00003774     0.08386    |                .   |** .                |   0.104648
 18  -0.0000743     -.16510    |                .***|   .                |   0.105263
 19  0.00001647     0.03659    |                .   |*  .                |   0.107613
 20  2.20109E-6     0.00489    |                .   |   .                |   0.107727
 21  -0.0000410     -.09105    |                . **|   .                |   0.107729
 22  -0.0000116     -.02588    |                .  *|   .                |   0.108432
 23  6.03752E-6     0.01342    |                .   |   .                |   0.108489
 24  -0.0000453     -.10077    |                . **|   .                |   0.108504
```

Figure 8: Sample Autocorrelation Function of the Residuals (RSAC) for Seasonal ARIMA

```
Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 1     -0.08096     |                 . **|  .                 |
 2      0.15902     |                 .   |***.                |
 3      0.09600     |                 .   |** .                |
 4      0.08711     |                 .   |** .                |
 5      0.02865     |                 .   |*  .                |
 6      0.00833     |                 .   |   .                |
 7     -0.08846     |                 . **|   .                |
 8     -0.10144     |                 . **|   .                |
 9      0.00804     |                 .   |   .                |
10      0.02325     |                 .   |   .                |
11      0.01794     |                 .   |   .                |
12     -0.10102     |                 . **|   .                |
13     -0.17105     |                 .***|   .                |
14      0.02805     |                 .   |*  .                |
15      0.10227     |                 .   |** .                |
16      0.00253     |                 .   |   .                |
17      0.10518     |                 .   |** .                |
18     -0.15644     |                 .***|   .                |
19     -0.05612     |                 .  *|   .                |
20     -0.00475     |                 .   |   .                |
21     -0.10066     |                 . **|   .                |
22      0.01256     |                 .   |   .                |
23      0.07086     |                 .   |*  .                |
24     -0.09009     |                 . **|   .                |
```

Figure 9: Sample Partial Autocorrelation Function of the Residuals (RSPAC).for Seasonal ARIMA

Forecasting. The forecasts from the estimated ARIMA model are listed in Table 15

Table 15 Forecasts Estimated from Seasonal ARIMA Modelling

Year/Quarter	Seasonal ARIMA Forecasts
05Q2	4809.25
05Q3	5037.24
05Q4	5219.94
06Q1	4880.98
06Q2	4869.63
06Q3	5098.52
06Q4	5282.66

Non-seasonal ARIMA modelling results

The results of non-seasonal ARIMA modelling are listed as follows, including identification, estimation, diagnostic checking and forecasting.

Identification. After the second differencing transformation ($z_t = y_t - 2y_{t-1} + y_{t-2}$), the SAC of non-seasonal TPC dies down quickly, given in Table 15. Therefore, the time series becomes stationary at I(2). Also both SAC and SPAC die down quickly. Therefore, it is an autoregressive integrated moving average model, ARIMA (p, 2, q).

Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1	Std Error
1	46.644989	0.16898	***	0.096674
2	-58.725528	-.21274	****	0.099396
3	-24.508629	-.08878	**	0.103564
4	-107.847	-.39068	********	0.104273
5	-5.336546	-.01933		0.117157
6	30.139626	0.10918	**	0.117187
7	-34.237547	-.12403	**	0.118134
8	8.465515	0.03067	*	0.119344
9	39.070937	0.14154	***	0.119418
10	10.881073	0.03942	*	0.120976
11	0.155136	0.00056		0.121096
12	-56.834332	-.20589	****	0.121096
13	-65.963638	-.23896	*****	0.124324
14	28.950365	0.10487	**	0.128545
15	29.222736	0.10586	**	0.129342
16	34.080330	0.12346	**	0.130149
17	39.542049	0.14324	***	0.131239
18	-44.437363	-.16098	***	0.132692
19	-6.699900	-.02427		0.134505
20	22.437924	0.08128	**	0.134546
21	-28.528283	-.10335	**	0.135004
22	17.500956	0.06340	*	0.135742
23	12.344386	0.04472	*	0.136018
24	-32.924434	-.11927	**	0.136155

Figure 10: Autocorrelation Function of Non-seasonal TPC

Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	0.16898	***
2	-0.24838	*****
3	-0.00213	
4	-0.46559	*********
5	0.18110	****
6	-0.22773	*****
7	-0.08219	**
8	-0.16894	***
9	0.19692	****
10	-0.12665	***
11	-0.01072	
12	-0.36372	*******
13	0.01850	
14	-0.06902	*
15	-0.05085	*
16	-0.03802	*
17	0.01811	
18	-0.16388	***
19	0.06691	*
20	-0.08836	**
21	0.03671	*
22	0.01962	
23	0.02680	*
24	-0.16098	***

Figure 11: Partial Autocorrelation Function of Non-seasonal TPC

Estimation and diagnostic checking. Based on spikes in SAC, SPAC, and RSAC, the ARIMA model (p=1,4, d=2, q=2) is the most appropriate, as it has the lowest standard error (12.523) compared with other selected models. Also according to Figure 12 and Figure 13, no significant spikes in RSAC or in RSPAC have occurred. The estimated coefficients for AR(1), AR(4) and MA(2) are 0.37727, -0.25933 and 0.95754 respectively (see Table 16. From the results of diagnostic checking in Table 17 *p*-values are all greater than 0.05 (expect for K=6 the *p*-value is 0.0424).

Table 16Estimated Coefficients for Non-seasonal ARIMA Modelling

	Par.	(*t-ratio)*	Lag
MU	.07251	(-.59)	0
MA1,1	.95754	(22.94)	2
AR1,1	.37727	(4.16)	1
AR1,2	.25933	(-2.67)	4

Table 17 Non-seasonal ARIMA Diagnostic Checking

To Lag	Chi-Square	DF	P	Autocorrelations					
6	8.18	3	.0424	-.022	.149	.104	-.074	.180	.035
12	10.70	9	.2968	-.060	-.083	.048	-.010	-.021	-.088
18	20.17	15	.1654	-.167	.096	.053	.035	.127	-.132
24	23.23	21	.3316	.029	.066	-.094	.049	-.008	-.077

```
Lag   Covariance   Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1   Std Error
 1    -3.400357     -.02168     |                .   |   .                |   0.096674
 2    23.344764     0.14885     |                .   |***.                |   0.096719
 3    16.336042     0.10416     |                .   |** .                |   0.098837
 4   -11.671465     -.07442     |                .  *|   .                |   0.099857
 5    28.169650     0.17961     |                .   |****                |   0.100374
 6     5.426741     0.03460     |                .   |*  .                |   0.103335
 7    -9.368660     -.05974     |                .  *|   .                |   0.103443
 8   -12.966519     -.08268     |                . **|   .                |   0.103765
 9     7.604430     0.04849     |                .   |*  .                |   0.104379
10    -1.622773     -.01035     |                .   |   .                |   0.104589
11    -3.338357     -.02129     |                .   |   .                |   0.104598
12   -13.779486     -.08786     |                . **|   .                |   0.104639
13   -26.113105     -.16650     |                .***|   .                |   0.105326
14    15.015057     0.09574     |                .   |** .                |   0.107758
15     8.351608     0.05325     |                .   |*  .                |   0.108550
16     5.551199     0.03540     |                .   |*  .                |   0.108794
17    19.910579     0.12695     |                .   |***.                |   0.108901
18   -20.659885     -.13173     |                .***|   .                |   0.110276
19     4.553639     0.02903     |                .   |*  .                |   0.111737
20    10.316896     0.06578     |                .   |*  .                |   0.111807
21   -14.713829     -.09382     |                . **|   .                |   0.112168
22     7.642842     0.04873     |               .    |*    .              |   0.112899
23    -1.189779     -.00759     |               .    |     .              |   0.113096
24   -12.141805     -.07742     |               .  **|     .              |   0.113100
```

Figure 12: Sample Autocorrelation Function of the Residuals (RSAC) for Non-seasonal ARIMA

```
Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
 1     -0.02168     |                .   |   .                |
 2      0.14845     |                .   |***.                |
 3      0.11266     |                .   |** .                |
 4     -0.09400     |                . **|   .                |
 5      0.14878     |                .   |***.                |
 6      0.05921     |                .   |*  .                |
 7     -0.09802     |                . **|   .                |
 8     -0.14785     |                .***|   .                |
 9      0.09792     |                .   |** .                |
10      0.02600     |                .   |*  .                |
11     -0.06484     |                .  *|   .                |
12     -0.10932     |                . **|   .                |
13     -0.10500     |                . **|   .                |
14      0.12774     |                .   |***.                |
15      0.10226     |                .   |** .                |
16      0.01651     |                .   |   .                |
17      0.12154     |                .   |** .                |
18     -0.09009     |                . **|   .                |
19     -0.06779     |                .  *|   .                |
20      0.01466     |                .   |   .                |
21     -0.07282     |                .  *|   .                |
22      0.02481     |                .   |   .                |
23      0.05157     |                .   |*  .                |
24     -0.07262     |                .  *|   .                |
```

Figure 3 Sample Partial Autocorrelation Function of the Residuals (RSPAC) for Non-seasonal ARIMA

Forecasting. The forecasts estimated from the non-seasonal ARIMA model are given in Table 18

Table 18: Forecasts Estimated from Non-seasonal ARIMA Modelling

Year/Quarter	Non-seasonal ARIMA Forecasts
04Q4	4964.14
05Q1	4960.28
05Q2	4985.36
05Q3	5003.81
05Q4	5022.96
06Q1	5039.20
06Q2	5051.95