Modelling and Forecasting Demand for Electricity in New Zealand: A Comparison of Two Approaches

K. Fatai, Les Oxley and F. Scrimgeour

Department of Economics, University of Waikato, Private Bag, Hamilton, New Zealand
(scrim@waikato.ac.nz, loxley@waikato.ac.nz, kolf@waikato.ac.nz)
Department of Economics, University of Canterbury, Private Bag, Christchurch, New Zealand (loxley@econ.canterbury.ac.nz)

Abstract: Modeling energy demand in New Zealand is typically based on either a partial general equilibrium model or models constructed from spreadsheet packages. The results show that electricity is forecast to be the fastest growing energy demand by households and the industrial sector for the next two decades. In this paper we attempt to model and forecast electricity demand using two econometric approaches: Engle-Granger's error correction model (ECM), and the AutoRegressive Distributed Lag regression (ARDL) approach. We investigate which model has the lowest forecasting error using a series of forecasting measures. The ARDL approach has better forecasting performance than the Engle and Granger ECM in this exercise.

Keywords: Electricity; Forecasting

1. INTRODUCTION

One of the first applications of the Engle and Granger [1987] cointegration method was to forecast monthly electricity sales (Engle et.al, 1989). Lately, however, a number of other methods have been applied to this task. Abeysinghe and Boon [1999] tested six methods for estimating elasticity and found that they give conflicting results. They concluded that Engle and Granger's error correction method (ECM) outperformed the other alternatives for estimating elasticity. This paper discusses two methods, namely, Engle and Granger's [1987] ECM and the ARDL approach due to Pesaran et.al. [1996], and apply these two methods to modelling electricity demand in New Zealand.

2. THE EXISTING LITERATURE

Dahl [1994] surveyed studies on electricity demand and found that most of the studies used static or partial adjustment models. Most of the papers, however, did not test the forecasting ability of the models. However, Jones [1993] tested the forecasting ability of the static and partial adjustments models comparing them with the general to specific (GTS) or dynamic regression model. Jones [1993] concluded that “the general to simple approach appear to offer a satisfying new methodology for generating superior forecast models of petroleum consumption and other energy use patterns” (p.698).

However, Chan and Lee [1997] disputed the Jones results, arguing that although the GTS approach had the advantage of reducing potential misspecification errors, it may overlook that most of the time series data are non-stationary. To overcome this problem, an alternative solution is the use of an ECM. Chan and Lee [1997] favoured the Engle and Granger ECM approach. In this paper we compare the forecasting ability of the Engle and Granger ECM and the ARDL approach using data from the electricity sector in New Zealand.

3. DATA AND METHODOLOGY

Most data were taken from the energy database of the International Energy Agency (IEA). The series are annual for New Zealand, 1960-1999, and cover the industrial, commercial and residential sectors. All energy data are in million of tons oil equivalent (Mtoe) units. For the ECM and long-run models additional data on fuel prices, the CPI to proxy prices of other goods, real GDP (PPP adjusted) were utilised. Data on average annual temperature (1960-1999), were taken from the New Zealand National Institute of Water and Atmosphere (NIWA). Although we would have preferred to use
quarterly data our analysis was constrained by
the annual energy data from the IEA. All data
were transformed to natural logarithm.

In the household sector Beenstock et al [1999]
use both a nested and a non-nested demand
function. In the nested household demand
function, the model is of the form of a
behavioural equation:

\[ H = H(C, P; T) + u. \]  (1)

where \( H \) denotes household sector, \( P \) is the price
of electricity and \( T \) measures meteorological
influences. It is also possible to formulate a
non-nested model of the form:

\[ H = H(I(C), P; T) + v. \]  (2)

In the non-nested model the variable \( C \) is
replaced by Household Income, which in turn,
determines how much of the household income
is to be spent on consumption as implied by
\( I(C) \).

For the industrial sector they use a
behavioural model of the form:

\[ I = I(Q, P_e; T) + e. \]  (3)

where \( Q \) denotes industrial production, \( P_r \) is the
relative price of electricity in the industrial
sector, and \( e \) is an assumed iid disturbance term.
Because of data restrictions Beenstock et al.
[1999] were restricted to using a nested demand
function in both the household and industrial
sectors.

In this paper, we follow the theoretical
model in Beenstock et al. [1999] using annual
data. The model for the household sector is
non-nested of the form:

\[ H = (C, P_e, P_e; T) + e. \]  (4)

where \( H \) is household electricity consumption,
\( C \) is consumer spending proxied by real GDP, \( P_e \)
is the relative price of electricity, \( P_e \) refers to the
prices of other goods and is proxied by the
consumer price index (CPI). \( T \) measures
meteorological influences proxied by
temperature. For the industrial sector the model
is of the form:

\[ I = (Q, P_r, P_e; T) + v. \]  (5)

where \( Q \) is industrial production proxied by
real GDP, \( P_r \) is the relative price of electricity in the
industrial sector and \( P_e \) is relative price of
other goods proxied by the CPI. Finally, for the
final energy consumption, we use a model of the
form:

\[ TFC = (G, P_r, P_e; T) + u. \]  (6)

where \( G \) measures total production proxied by real
GDP, \( P_r \) is relative price of fuel, \( P_e \) is the CPI and \( T \)
is the same temperature variable used in the models
for the household and industrial sector.

The above models were estimated using the Engle
and Granger's ECM and AutoRegressive
Distributed Lag (ARDL) approaches. The merit of
the ARDL approach is that it can be applied
irrespective of the order of integration of the
variables. The merits of the Engle and Granger
ECM are that the short and long run effects can be
consistently modelled via the disequilibrium
adjustment to a long-run equilibrium.

3.1 Integration properties of the data

The first step in the estimation and testing of the
models is to determine the order of integration of
the data. Hendry and Juselius [2000a,b] regard
testing for integration properties as an essential first
tests are used for this purpose with the results
presented as Table 1 below.

3.2 Engle and Granger's Error Correction
Approach

The first model estimated uses the Engle and
Granger approach and can only be applied if the
variables are I(1) and cointegrated. In this case the
variables can be first-differenced and the error-
correction (ECM), term from the cointegrating
regression added as in equations (7).

\[ \Delta X_t = \alpha + \sum_{i=1}^{k} \beta_i \Delta X_{t-i} + \sum_{j=1}^{1} \phi_j \Delta X_{t-j} + \varepsilon_{t} \]  (7)

The lag lengths \( k \) and \( l \) are particularly important as
the lag length chosen in the VAR can significantly
alter the result. We used the Akaike Information
Criteria (AIC) to determine the optimal lag length.

3.3 Autoregressive Distributed Lag
Regression (ARDL) approach of
Pesaran et al [1996].

The main advantage of this approach to testing and
estimation is that it can be applied whether the
regressors are I(0) or I(1) and avoids the pre-test
problems associated with standard cointegration
analysis.

The first stage of the process involves establishing
the existence of a long-run relationship between the
variables and is tested by considering the joint
significance of the coefficients of the lagged
levels variables $Y_{-1}$ and $X_{-1}$ in an equation like
(8) below:

$$\Delta X_t = \alpha + \sum_{i=1}^{p} \alpha_i \Delta X_{-i} + \sum_{j=1}^{q} \beta_j \Delta Y_{-j} + \gamma \Delta X_{-1} + \epsilon_t$$

(8)

using tables presented in Pesaran et al. [1996].
If the null hypothesis of no long-run relationship
is rejected, the ARDL model can be established
and either a long-run or ECM version of the
model constructed. The model can be used for
dynamic forecasts on either the levels or first-
difference version.

4. **EMPIRICAL RESULTS**

Table 1 presents the results of testing for the
order of integration of the data. Both the ADF
and PP tests imply that the price of both
residential and industrial electricity appear to be
I(0), as well as temperature, while the rest of
the variables are I(1). Given most of the
variables are I(1), we then test for cointegration
between the I(1) variables.

**Table 1. Testing for unit roots (industrial and
commercial), log-levels and first differenced,
1960-1999: Augmented Dickey Fuller.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levels</th>
<th>Differented</th>
<th>CV</th>
<th>Lags</th>
</tr>
</thead>
<tbody>
<tr>
<td>industrial</td>
<td>-2.24</td>
<td>-3.94</td>
<td>3.53</td>
<td>2</td>
</tr>
<tr>
<td>commercial</td>
<td>-1.75</td>
<td>-4.70</td>
<td>3.53</td>
<td>1</td>
</tr>
<tr>
<td>indust/com**</td>
<td>-1.64</td>
<td>-4.02</td>
<td>3.53</td>
<td>2</td>
</tr>
<tr>
<td>tfc*</td>
<td>-1.16</td>
<td>-4.09</td>
<td>3.53</td>
<td>1</td>
</tr>
<tr>
<td>residential</td>
<td>-2.59</td>
<td>-4.21</td>
<td>3.53</td>
<td>1</td>
</tr>
<tr>
<td>gdp</td>
<td>-2.9</td>
<td>-3.85</td>
<td>3.53</td>
<td>2</td>
</tr>
<tr>
<td>pind</td>
<td>-4.74</td>
<td>-4.8</td>
<td>3.53</td>
<td>2</td>
</tr>
<tr>
<td>pres</td>
<td>-3.53</td>
<td>-4.65</td>
<td>3.54</td>
<td>1</td>
</tr>
<tr>
<td>temp</td>
<td>-4.03</td>
<td>-6.05</td>
<td>3.53</td>
<td>1</td>
</tr>
</tbody>
</table>

*Total final energy consumption. ** = industrial/commercial
electricity consumption. pind = industrial electricity price.
pres = residential electricity price. temp = annual average
temperature.

Cointegration is investigated between four
measures of electricity consumption - industrial
electricity consumption, commercial electricity
consumption, industrial and commercial
electricity consumption combined, residential
electricity consumption, and total final
electricity consumption - and the other
variables (real income (gdp), temperature and
the relative price of both industrial electricity
and residential electricity). The results are
presented in Tables 2 – 6.

For industrial and commercial electricity
consumption, there is no cointegrating relationship
with the other variables, either bivariate or
multivariate.

**Table 2. Testing for bivariate cointegration
between variables and indus/comm electricity
consumption, 1960-1999.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>eigen</th>
<th>Trace</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lgdp</td>
<td>10.05</td>
<td>14.9</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4.85</td>
<td>4.85</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
<tr>
<td>Ltemp</td>
<td>14.21</td>
<td>19.33</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5.12</td>
<td>5.12</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>13.96</td>
<td>18.97</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5.01</td>
<td>5.01</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
</tbody>
</table>

Proceeding to test for cointegration between
commercial electricity consumption and its
determinants, tests for bivariate cointegration found
none as shown in Table 3.

**Table 3. Testing for bivariate cointegration
between variables and commercial electricity
consumption, 1960-1999.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>eigen</th>
<th>Trace</th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lgdp</td>
<td>12.36</td>
<td>18.39</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6.03</td>
<td>6.03</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
<tr>
<td>Ltemp</td>
<td>14</td>
<td>23.23</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>9.23</td>
<td>9.23</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>17.01</td>
<td>26.84</td>
<td>$r = 0$</td>
<td>$r = 1$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>9.82</td>
<td>9.82</td>
<td>$r &lt;= 2$</td>
<td>$r = 2$</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 considers the existence of cointegration
between residential electricity consumption and
income, temperature and residential electricity price
but found no bivariate cointegration relationships.

Given no cointegrating relationships were found
between either industrial, commercial or residential
electricity consumption and the other variables, we
then tested whether a cointegrating relationship
could be found between total final end-user
electricity demand consumption and the other
variables.

We found that there is a cointegrating relationship
between total final end-user electricity
consumption, real income, the price of electricity
and fuel and the price of other goods, as shown in
Table 5. Because the other measures of electricity
were not cointegrated with income, temperature or
the relative price of electricity, we decided to use
the ECM and GTS methods to model only for final
electricity consumption. The results are presented
below as section 4.1 and 4.2.
plot the forecast values of both approaches. It seems that the ARDL approach is favoured over the sample period studied. Further studies on the forecasting ability of the ARDL approach is needed.

6. ACKNOWLEDGEMENTS

All three authors wish to thank the New Zealand Public Good Science Fund, grant number UOWX:0010, *Energy Resources and Energy Resource Economics*, for supporting this research.

7. REFERENCES


5. CONCLUSIONS

In this paper we use two approaches to analyse the pattern of electricity consumption in New Zealand. We start by examining the integration and cointegration properties of the energy disaggregates (coal, oil, gas, electricity) and found that they do not cointegrate with energy determinants. However, we found that total final electricity demand cointegrates with some of it's determinants. This enables us to model total final electricity as an ECM using the Engle and Granger's approach. However, the ARDL approach enables us to model the ECM without pretesting whether the variables are I(1) or I(0).

We conducted various stability tests and found no evidence of instability in the variables used in both approaches. We compared the forecasting errors using four measures and also