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

K. Fatai,[a] **Les Oxley**[a,b] **and F. Scrimgeour**[a]

[a]*Department of Economics, University of Waikato, Private Bag, Hamilton, New Zealand (scrim@waikato.ac.nz, loxley@waikato.ac.nz, koli@waikato.ac.nz)*
[b] *Department of Economics, University of Canterbury, Private Bag, Christchurch, New Zealand(l.oxley@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 demanded 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:

$$H = H(C, P; T) + u. \quad (1)$$

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:

$$H = H(I(C), P; T) + v. \quad (2)$$

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:

$$I = I(Q, P_{I},; T) + e. \quad (3)$$

where Q denotes industrial production, P_i 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:

$$H = (C, P_e, P_o; T) + e. \quad (4)$$

where H is household electricity consumption, C is consumer spending proxied by real GDP, P_e is the relative price of electricity, P_o 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:

$$I = (Q, P_{i,} P_{o;} T) + v. \quad (5)$$

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

$$TFC = (G, P_{i,} P_{o;} T) + u. \quad (6)$$

where G measures total production proxied by real GDP, P_i is relative price of fuel, P_o 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) approachs. 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 step. The ADF and Phillips and Perron [1988] (PP) 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).

$$\Delta X_t = \alpha + \sum_{i=1}^{k} \zeta_i \Delta X_{t-i} + \sum_{j=1}^{l} \varphi_j \Delta Y_{t-j} + \xi ECM_{-1} + \varepsilon_t \quad (7)$$

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:

$$\Delta X_t = \alpha + \sum_{i=1}^{k} \zeta_i \Delta X_{t-i} + \sum_{j=1}^{1} \varphi_j \Delta Y_{t-j} + \beta Y_{t-1} + \gamma X_{t-1} + \varepsilon_t \quad (8)$$

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.

Variables	Levels	First Differenced	CV	Lags
industrial	-2.24	-3.94	3.53	2
commercial	-1.75	-4.70	3.53	1
indust/com**	-1.64	-4.02	3.53	2
tfc*	-1.16	-4.09	3.53	1
residential	-2.59	-4.21	3.53	1
gdp	-2.9	-3.85	3.53	2
pind	-4.74	-4.8	3.53	2
pres	-3.53	-4.65	3.54	1
temp	-4.03	-6.05	3.53	1

*total final energy consumption. ** = industrial/commercial energy 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.

Variable	Max eigen.	Trace	H_0	H_1	VAR
Lgdp	10.05	14.9	r = 0	r = 1	1
	4.85	4.85	r <= 2	r = 2	
Ltemp	14.21	19.33	r = 0	r = 1	1
	5.12	5.12	r <= 2	r = 2	
Price	13.96	18.97	r = 0	r = 1	1
	5.01	5.01	r <= 2	r = 2	

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.

Variable	Max. eigen.	Trace	H_o	H_1	VAR
Lgdp	12.36	18.39	r = 0	r = 1	1
	6.03	6.03	r <= 2	r = 2	
ltemp	14	23.23	r = 0	r = 1	1
	9.23	9.23	r <= 2	r = 2	
price	17.01	26.84	r = 0	r = 1	1
	9.82	9.82	r <= 2	r = 2	

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.

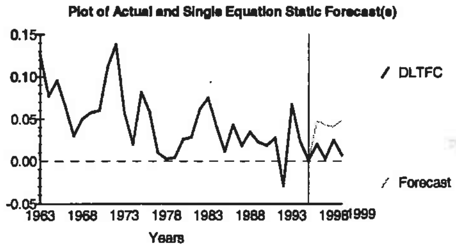

Figure 3. Actual and forecasted values using Engle and Granger's ECM.

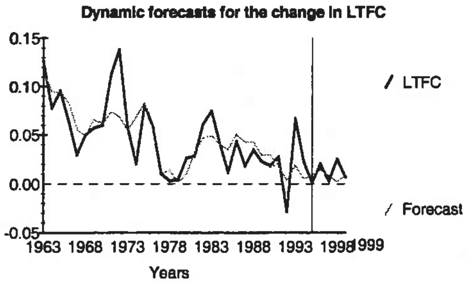

Figure 4. Actual and forecasted values using the ARDL ECM.

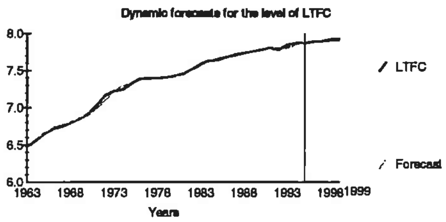

Figure 5. Plot of the actual and forecasted values using the ARDL approach.

The trend seems to show that although the level of TFC is increasing, it is increasing at a decreasing rate.

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 Grangers 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 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

Abeysinghe, T. and T.K. Boon , Small Sample Estimation of a Cointegrating Vector: An Empirical Evaluation of Six Estimation Techniques. *Applied Economics Letters.* 6, 645-48, 1999.

Beenstock, M., E. Goldin and D. Nabot, The demand for electricity in Israel, *Energy Economics*, 21, 168-183, 1999.

Chan, H.L. and S.K. Lee, Modelling and forecasting demand for coal in China, *Energy Economics* 19, 271-287, 1997.

Dahl, C., A survey of energy demand elasticities for the developing world, *Journal of Energy Development.* 1-47, 1994.

Davidson, J.H., D.H., Hendry, F., Srba, and S. Yeo, Econometric modeling of the aggregate time series relationship between consumers' expenditure and income in the United Kingdom, *The Economic Journal,* 88, 661-692, 1978.

Engle, R. F. and C.W.J. Granger, Cointegration and error-correction: representation, estimation, and testing. *Econometrica,* 55, 251-276, 1987.

Engle, R.F. C. Mustafa and J. Rice. Modelling Peak Electricity Demand. University of California, San Diego Department of Economics Working Paper: 35-89. July 1989.

Hendry, D.F. and K. Juselius, Explaining cointegration analysis: part 1, *The Energy Journal*, 21, 1-42, 2000a.

Hendry, D.F. and K. Juselius, Explaining cointegration analysis: part 2, *The Energy Journal*, 22, 75-120, 2000b.

Jones, C.T. (1993) A single-equation study of U.S. Petroleum consumption: the role of model specification. *Southern Economic Journal*, 59, 687-700, 1993.

Pesaran, M.H., Y. Shin, and R.J. Smith, Testing for the existence of a long-run relationship, DAE Working Paper No. 9622, Department of Applied Economics, University of Cambridge.

Phillips, P.C.B. and P. Perron, Testing for a unit root in time series regression, *Biometrika,* 75, 335-346, 1988.