J. Natn. Sci. Foundation Sri Lanka 1999 **27**(1): 55-64

MODELLING SECTOR-WISE DEMAND FOR ELECTRICITY IN SRI LANKA USING A MULTIVARIATE REGRESSION APPROACH

K. D. R. RUWANTHI and W. N. WICKREMASINGHE*
Department of Statistics and Computer Science, University of Colombo, Colombo 3.

(Received: 01 June 1998; accepted: 08 January 1999)

Abstract : The demand for electricity in Sri Lanka depends mainly on the activities of domestic, industrial and commercial sectors and the three components are highly correlated. Although such correlation does not affect univariate estimation procedures, it may lead to incorrect inferences of influential factors on the demand for electricity. As a result, separate univariate approaches for each sector may not be an accurate method of identifying such factors. Therefore, this study aims to identify such factors using multivariate regression which considers the correlation among different sectors (or dependent variables) and estimates a multivariate demand model for the purpose of forecasting. The overall significance of the fitted demand model and the significant influential factors are assessed by multivariate tests such as Bartlett's using the statistical package SAS. Theoretically, demand is a function of its own price, the income level of consumers, and the price of substitutes. Gross Domestic Product (GDP) at constant (1980) factor prices is used as a proxy for income level of consumers and kerosene is taken as a close substitute for electricity. The analysis uses quarterly data for two periods 1970-1977 and 1978-1994 to assess the effect of the liberalized economy introduced in late 1977. During the period after 1977, the effect of the income level has increased substantially due to the liberalized economy. The substitution between electricity and kerosene is marginal in the post-liberalized period, as electricity is more efficient and convenient than kerosene. Due to such differences between the two periods, the demand for electricity may be explained better by two models rather than a single model estimated for the entire period. The multivariate demand model based on the post-liberalized period is found to adequately forecast the demand for electricity.

Key words: Electricity, energy demand, energy sources, mathematical modelling.

INTRODUCTION

In Sri Lanka, electricity is generated from hydro and thermal power. While hydro power is the cheapest and environmentally beneficial source of energy, thermal power is more expensive and the demand for which varies with the changes in the weather pattern and is mostly used for meeting the shortfall in the supply of hydro electricity in the drought periods. The level of energy consumption depends mainly on the activities of domestic, industrial, and commercial sectors, and the pattern of consumption is associated with characteristics exclusive to different sectors. Therefore it is important to examine the sector-wise consumption of electricity.

---

* Corresponding author
Part of this work was presented at the 54th annual sessions of the Sri Lanka Association for the Advancement of Science (SLAAS).

The domestic sector is the largest energy consumer in Sri Lanka and the household income is an important determinant of the use of electricity for domestic purposes.[2] The industrial sector obtains about 25% of its energy requirements from electricity at present.[6] The consumption pattern of different types of energy sources in the industrial sector indicates a shift from petroleum products and firewood to electricity, as electricity is a more efficient source of energy with minimum wastage in consumption.[6] The commercial sector includes activities of shops, supermarkets, banks and hotels. As a result of expansion in service and commercial oriented activities following the economic liberalization introduced in November, 1977, the commercial demand for electricity has considerably increased. Today, electricity has become the main source of energy for commercial activities, accounting for more than 60% of the total energy consumption by this sector. [6] There has been an increase in the installed capacity of electricity in Sri Lanka from 265 MW in 1970 to 402 MW in 1977 and 1409 MW in 1995. [3,4]

Continued expansion in power generation and the growing demand for electricity, provide a clear case for the need for an efficient forecasting system for the demand. However, there is no recent information available on such forecasting systems. Satisfying the ever increasing demand for electricity is a major problem as long spells of droughts prevailing in catchment areas are more frequent now than in the past, which prevent the hydro power plants from generating to full capacity. Thermal power generation is a substitute in such situations. However, this is expensive and therefore, a sensible pricing policy should be formulated to suppress the unwarranted consumption and to recover the increasing energy cost. In formulating such policies, it is required to identify the economical variables affecting the demand for electricity. However, no recent information is available on the influence of such factors on the demand for different sectors. The earlier studies [10,15] are based on the total demand for electricity and they used univariate regression techniques. However, if the demand in different sectors are inter-related, univariate approach may not be an accurate method of identifying the factors influencing the demand for electricity. Therefore, this paper aims to identify and examine such factors using a more appropriate technique, the multivariate regression. Apart from these, the structural stability of the formulated demand function due to the implementation of liberalized economic policies in November 1977, is also studied. Finally, an attempt is made to forecast demand for electricity for different sectors, at least for the purpose of short-term forecasting based on multivariate regression model.

METHODS AND MATERIALS

The consumption of electricity has been taken as its demand. The demand for electricity is grouped into three major sectors namely, (i) domestic, (ii) industrial, and (iii) commercial. According to the economic theory, the demand for any good or service is a function of its own price, income level of consumers and the price of substitutes. Apart from these factors, demand of a service may also be affected by the changes of economic activities of the country.[13] The most important economic change in Sri Lanka is the implementation of liberalized economic policies in November 1977.

The analysis uses quarterly data for the period 1970 to 1994. The figures for 1995 and 1996 were also available, but due to the frequent power cuts in 1996, the data on consumption of electricity for 1996 cannot be used for this study. The figures for 1995 were not used in the analysis as these figures were used for calculating forecasting errors. Data on consumption of electricity were obtained from Central Bank of Sri Lanka and from Ceylon Electricity Board.[1,3] The tariff structure used by Ceylon Electricity Board is somewhat complicated. In order to overcome problems arising from this, an average price of one unit (1 kWh) of electricity of a particular quarter in Rs/kWh was computed by dividing the total revenue from electricity of that quarter in Rs. Mn. by the number of units consumed within that period in Mn. kWh.

The data on kerosene sales and revenue were collected from Ceylon Petroleum Corporation and from Central Bank of Sri Lanka.[5,1] The price of a metric ton of kerosene of a particular quarter in Rs/M.T. was calculated by dividing the total revenue of that quarter in Rs. Million by the total sales within that period in 000' tons and multiplying by 1000. For commercial and industrial sectors, diesel would perhaps be a better substitute than kerosene. However, as the quarterly figures of sales and revenue of diesel were not readily available, checking of the effect of diesel price was not possible.

Gross Domestic Product (GDP) at constant (1980) factor prices was used as a proxy for income level of the consumers. These are available in Central Bank of Sri Lanka on an annual basis. However, the actual quarterly figures of GDP are not available. Therefore, the quarterly estimates of GDP figures [12] were used in this study for the period 1970-1981. For the period 1982-1995, corresponding figures were estimated by an additive time series model with trend and seasonal variations. Trend was adequately estimated by multiple regression model as it explained 98% of the total variation. The estimated values of GDP were then adjusted for the annual data published by the Central Bank of Sri Lanka. The differences of annual estimates and actual values were small and random.
The multivariate regression model can be stated as:

$$Y = XB + E,$$

where :

Y is an nxp matrix of 'p' dependent variables, $Y_1, Y_2, \ldots, Y_p$,
X is an nx(q+1) matrix of 'q' predictor variables of the form $(1, X_1, X_2, \ldots, X_q)$,
B is a (q+1)xp matrix of unknown parameters, and
E is an nxp matrix of residuals assumed to be from a multivariate normal distribution with zero mean and dispersion matrix Σ.[8]

In this study, there were three dependent variables, Y_1, Y_2, and Y_3,

where :

Y_1 = log (Domestic demand),
Y_2 = log (Industrial demand), and
Y_3 = log (Commercial demand)

and three predictor variables,

X_1 = log (Price of electricity),
X_2 = log (Price of kerosene), and
X_3 = log (Income level of consumers).

The estimators $\hat{B}=(X'X)^{-1}X'Y$ and $\hat{\Sigma}=\frac{1}{n}(\hat{E}'\hat{E})$ of B and Σ can be obtained from the method of maximum likelihood using the statistical package SAS. Multivariate regression considers the correlation among dependent variables. However, the multivariate estimation is achieved by p independent univariate estimations, so the inter-correlation among the Y_i ($i = 1,2,3$) do not influence the estimation of B. On the other hand, the univariate estimation does not consider the off-diagonal elements of Σ which play an important role in testing hypotheses about elements of B. Hence, if correlations among the Y_i ($i = 1,2,3$) are ignored, and three separate univariate regressions are done, one on each Y_i, then the correct parameter estimates will be obtained but incorrect inferences may be drawn about these parameters.[8] The assumption of multivariate normality is assessed by examining whether the scatter diagrams for pairs of first few principal components of E appear to be elliptical.[7] Further, the separate columns of E are assumed to be serially independent[8] and it is tested by the Durbin-Watson (DW) procedure.[14]

The overall significance of the multivariate regression model was tested by the Bartlett's Statistic of the form $-[n-q-1-(p-q+1)/2]\ln \Lambda$ where p, q and n were as defined earlier and Λ is called Wilks' Lambda [8] which may be an equivalent to the F value given by the univariate regression. The value of Λ was given by the MTEST statement of PROC REG in SAS. The Bartlett's statistic has an approximate χ^2 distribution with pxq degrees of freedom under the null hypothesis of non-significant model.[8]

Selection of the most effective variables in the multivariate regression model was obtained by the Bartlett's statistic of the form; $-[n-q-1-(p-q_1+1)/2]\ln \Lambda$, where q_1 is the number of variables that provide no additional predictive information when adjusted for the remaining explanatory variables. It also has a χ^2 distribution with pxq_1 degrees of freedom under the null hypothesis that q_1 number of variables have no effect, when adjusted for the remaining predictor variables in the model.[8]

The effect of the liberalized economic policies introduced in late 1977 on the stability of the demand for electricity was tested by the Chow test [13] where the test statistic is given by :

$$[\{SSR - (SSR_1 + SSR_2)\} / r] / \{ (SSR_1 + SSR_2) / (n - 2r) \},$$

where :

n = sample size,
r = number of parameters being estimated, and
SSR, SSR_1, and SSR_2 are the residual sums of squares of the estimated models for the entire period, the pre-liberalized period, and the post-liberalized period, respectively.

This is an univariate test which is done for each Y_i. The above Chow test statistic has an F distribution with (r, n-2r) degrees of freedom under the null hypothesis of parameter stability. The demand for electricity had experienced a structural change even before the introduction of liberalized economic policies as a result of a major political change in Sri Lanka and the expectations about new economic policies.[10] Therefore at the time these policies were introduced, the patterns of the demand for electricity had already adjusted to the new situation. Thus, we could use 1970-1977 as pre-liberalized period and 1978-1994 as post-liberalized period.

RESULTS AND DISCUSSION

The coefficients of correlation between Y_1 and Y_2, Y_1 and Y_3, and Y_2 and Y_3 are 0.94, 0.96 and 0.93 respectively. This indicates high correlation between dependent variables.

The values of the Bartlett's Statistic of multivariate regression used to test the predictive power of each explanatory variable, adjusted for the effect of remaining variables are tabulated in Table 1 for the three periods. The table also contains the t-statistics of univariate regressions for the comparison of two regressions.

According to Table 1, the effect of variables on the demand are different for the two periods, before and after 1977. Based on the results of multivariate regression, in the post-liberalized period, income level of the consumers is the most influential on the demand ($p < 0.01$), while it is the least effective before 1977, which implied that the effect of the income level has increased substantially due to the liberalized economic policies and the price of electricity has a more significant effect on the demand during the pre-liberalized period. However, the univariate regressions lead to different inferences about price of electricity and income level of consumers. The plots of the demand against X_2 and X_3 over the whole period indicate clear differences before and after 1977. Although the price of kerosene is an important factor in determining the demand before 1977, it became insignificant during the period after 1977. As the use of electricity is more efficient and convenient, the substitution between electricity and kerosene was marginal during the post-liberalized period, and such substitution may be impracticable by the requisite expenses and radical modification of the equipment required, mainly in the industrial and commercial sectors.

Table 1: Effect of explanatory variables on the demand for electricity

Explanatory variable	Bartlett's Statistic of Multivariate Regression and *t- Statistic of Univariate Regression* p=3, q=3 and q_1=1		
	Entire period Sample size (n) =100	Before 1997 Sample size (n) =32	After 1977 Sample size (n) = 68
X_1	25.426 (0.0001)** *1.607 (0.1114)*	15.569 (0.0014)** *1.627 (0.1150)*	10.838 (0.0045)** *3.150 (0.0025)***
X_2	17.517 (0.0004)** *3.978 (0.0001)***	15.809 (0.0013)* *4.395 (0.0001)***	1.27 (0.7306) *-1.421 (0.1602)*
X_3	125.107 (0.0001)** *16.848 (0.0001)***	5.758 (0.1242) *2.635 (0.0136)**	124.786 (0.0001)** *21.345 (0.0001)***

X_1 = log (Price of electricity); X_2 = log (Price of kerosene); and X_3 = log (Income level)

Note : Figures in parentheses are the corresponding p-values of Bartlett's Statistic and t-statistic.
** indicates significance at 1 % level ($p < 0.01$)
* indicates significance at 5 % level ($p < 0.05$)

Due to the above differences between the two periods, the demand for electricity may be explained better by two models rather than a single model estimated for the entire period. Therefore, it is important to test whether the statistical stability of the demand has changed significantly due to the new economic policies. This was identified by the Chow test using the estimates of Residual Sum of Squares (RSS) from the univariate regression on Y_1, Y_2, and Y_3. If estimated parameters in the demand model change with new policies, this implies that the stability of the relationship has changed. The Chow test results appearing in Table 2, show that the parameter stability of the demand model had changed significantly due to the open economic policies implying that the estimated model is not the same for the two periods. Therefore, for the purpose of forecasting, a multivariate demand model estimated for the post-liberalized period is more appropriate than that for the entire period.

Table 2: Chow test results for parameter stability.

Dependent variable	SSR (Entire period)	SSR_1 (Before 1977)	SSR_2 (After 1977)	Chow test statistic
Y_1	0.6103	0.3276	0.1821	6.1843 (0.0002) **
Y_2	0.1633	0.0479	0.1010	3.0302 (0.0214) *
Y_3	0.3321	0.1415	0.1336	6.4922 (0.0001) **

Y_1 = log(Domestic demand); Y_2 = log(Industrial demand); Y_3 = log(Commercial demand)
SSRs = the residual sums of squares of each dependent variable estimated for the corresponding periods.
Note : r = 4
Figures in parentheses are the corresponding p-values of Chow Statistic
* indicates significance at 5 % level ($p< 0.05$)
** indicates significance at 1 % level ($p< 0.01$)

In the post-liberalized period, the price of kerosene had no influence on the demand when its effect was adjusted for price of electricity, income level and also for the inter-correlation among the three dependent variables. Therefore, a multivariate model of the form $\hat{Y}=X\hat{B}$ excluding the price of kerosene, was estimated for that period.

The model is,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = [1 \ X_1 \ X_3] \begin{bmatrix} -6.3479 & -2.1902 & -5.9654 \\ 0.2121 & -0.0153 & 0.0621 \\ 1.9286 & 1.0537 & 1.8465 \end{bmatrix}$$

which gave rise to three univariate models :

(1) $Y_1 = -6.3479 + 0.212 X_1 + 1.9286 X_3$;
(2) $Y_2 = -2.1902 - 0.0153 X_1 + 1.0537 X_3$; and
(3) $Y_3 = -5.9654 + 0.0621 X_1 + 1.8465 X_3$

for the domestic, industrial and the commercial sectors, respectively.

The value of the Bartlett's Statistic for overall significance of this model is 248.789. As the value is greater than the corresponding table value (12.592) at 5% level, it can be concluded that the above regression is highly significant. Therefore, the variables, X_1 and X_3 adequately improve the prediction of the dependent variables, Y_1, Y_2, and Y_3.

The residuals (E) of the estimated model did not indicate any substantial departure from multivariate normal distribution with mean zero. The values of Durbin-Watson (DW) statistic for the serial independence of the residuals of the three models, (1), (2), and (3) were 1.92, 1.40 and 1.74, respectively. This implies that the model (2) contains positive first-order serial correlation as its value of DW statistic is less than the lower bound of the DW critical value, $d_L = 1.74$ and models, (1) and (3) do not indicate such correlation. As the residual plot of model (2) did not show any strong pattern and most of its residuals were within the acceptable range (-0.12,0.12), the amount of serial dependence present in model (2) may have little influence on its estimates.

The forecasting performance of the estimated models are presented in Table 3, where the quarterly predictions of the demand in 1995 are compared with the corresponding actual values.

Table 3 : Forecasting performance of the estimated models for different sectors in 1995.

Model	Demand	Q_1	Q_2	Q_3	Q_4
Y_1 - Domestic sector	Predicted	245.6	250.2	252.8	259.8
	Actual	239.1	253.9	256.8	265.0
	% Forecasting error	-2.7	1.5	1.6	2.0
Y_2 - Industrial sector	Predicted	314.8	315.5	317.1	321.5
	Actual	318.6	320.0	322.5	325.2
	% Forecasting error	1.2	1.4	1.7	1.2
Y_3 - Commercial sector	Predicted	203.5	205.4	207.2	212.5
	Actual	198.5	201.1	204.3	210.2
	% Forecasting error	-2.5	-2.1	-1.4	-1.1

Q_1 = 1st quarter Q_2 = 2nd quarter
Q_3 = 3rd quarter Q_4 = 4th quarter

$Y_1 = -6.3479 + 0.2121 X_1 + 1.9286 X_3$, ($R^2 = 95.61\%$).
$Y_2 = -2.1902 - 0.0153 X_1 + 1.0537 X_3$, ($R^2 = 84.93\%$).
$Y_3 = -5.9654 + 0.0621 X_1 + 1.8465 X_3$, ($R^2 = 94.67\%$).

As the percentage forecasting errors are relatively small, and the tests described indicate no evidence of failures of the model assumptions, we expect the model to be satisfactory. Since future estimates of electricity prices and GDP values can be obtained from Ceylon Electricity Board and Central Bank of Sri Lanka, we believe that the modelling exercise described in this paper is worthwhile. The suggested model can therefore be used to forecast the demand for electricity in any given period, using the price of electricity and the GDP value pertaining to that period. To be on the safe side one can always forecast the next four quarters only, using this model. Once the actual values of these four quarters are available, these values can be used to update the model which could then be used to forecast the next four quarters, and so on.

Acknowledgment

The authors appreciate the valuable assistance received from Mrs. R. Dheerasinghe of the Central Bank of Sri Lanka.

References

1. Central Bank of Sri Lanka, *Bulletin* (1970-1995) Colombo.

2. Central Bank of Sri Lanka (1986/87). *Report on the consumer finance and socio economic survey.* Colombo.

3. Ceylon Electricity Board. *Annual reports of data on sales and generation.* Colombo.

4. Ceylon Electricity Board. *Historical data book :* 1969-1993. Load Forecasting and Tariffs Branch, Statistical Unit, Colombo.

5. Ceylon Petroleum Corporation. *Annual Reports* (1970-1995) Colombo.

6. Dheerasinghe R. (1993). *Sources of energy and sectoral composition of energy consumption in Sri Lanka.* Sri Lanka Economic Association, Colombo.

7. Johnson R. & Wichern D. (1988). *Applied multivariate statistical methods.* Second Edition, Prentice Hall.

8. Krzanowski W.J. (1988). *Principles of multivariate analysis : a user's perspective.* Oxford University press.

9. Mirer T.W. (1988). *Economic statistics and econometrics.* Macmillan.

10. Pathirage R.R. (1983). Factors influencing the demand for commercial energy in Sri Lanka. *Staff Studies* **12**(2), Central Bank of Sri Lanka, Colombo.

11. Perera K.K.Y.W. (1993). *An evaluation of the trends in the energy sector and potential for developing renewable energy.* Sri Lanka Economic Association, Colombo.

12. Savundranayagam T. (1984). *Quarterly estimates of gross domestic product* of Sri Lanka : 1970-1981, *Staff Studies* **14**(122), Central Bank of Sri Lanka, Colombo

13. Thomas R.L. (1993). *Introductory econometrics : theory and applications*. Second Edition, Longman.

14. Weisberg S. (1985). *Applied linear regression*. Second Edition, John Wiley & Sons, Inc.

15. Wijemanne E.L., Pathirage R.R., Jayasekera, D.M. & Gaminiratne, K.H.W. (1986). *Equity, efficiency issues in electricity pricing in Sri Lanka, energy pricing strategies*, **Vol. III**. International Labour Organization and Asian Employment Programme, New Delhi.