CHAPTER -4

RESEARCH METHODOLOGY

4.1. Introduction

This chapters of the study deals with the details of research design, data sources and research methods applied in order to achieve the stipulated objectives.

4.2. Research Design and Time Frame

The present study is empirical in nature wherein the causality analysis is done between electricity consumption and economic growth in India. Further, the study extends to evaluate the technical performance of the power sectors in India. For this, data envelopment analysis (DEA) is used. This study also assesses the financial performance of the power sectors in India by applying descriptive analysis and trend analysis. As India emerged as one of the most energy demanding countries due to rapid growth after economic reforms especially in the beginning years of the new millennium, the time frame, chosen for the study, is 2001-02 to 2015-16.

4.3. Data Description and Sources

The study is based on secondary data. There is no comprehensive source of data used in the study. The data used in this study, therefore, taken from multiple sources. For causality analysis, which is the examination of causal relationship between electricity consumption and economic growth in India, GDP at constant price (Base Year 2004-05) has been taken as the proxy variable for the economic growth and electricity consumption is taken in kWh. The time series data for these two variables are taken from the data book of *Planning Commission of India*. To evaluate the technical performance of the power sectors, data of installed electricity generation capacities of thermal and hydro power units across India is taken from *Indiastat.com*. For the assessment of financial performance of power sector, the data of cost incurred and revenue generated by power sector was taken from various annual reports of *Ministry of Power, GOI*.

4.4. Modeling Criterion

4.4.1. Stationarity

A time series is required to be stationary for the sake of feasibility and forecasting. It leads to a reliable prediction of its future behavior. A time series have characteristics to return to its mean and fluctuations occurs around its mean.

A time series Y_t is said to be stationary if:

i) $E(Y_t)$ = constant for all t;

ii) $Var(Y_t)$ = constant for all t; and

iii) $Cov(Y_t, Y_{t+k})$ = constant for all t and all $k \neq 0$

Unit root test is a test that examines the null hypothesis of unit root against the alternative hypothesis of stationarity. This includes Augmented Dickey-Fuller (ADF) test and Phillips-Perron (PP) test. This study has used the Augmented Dickey-Fuller (ADF) to check the stationarity of the variables. The equation for the ADF test is constructed as follows:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha i \, \Delta Y_{t-1} + \varepsilon_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.1)$$

Where,

ε_t is pure white noise error term[1].

The number of lagged difference term to include is often determined empirically, the idea being to include enough terms so that the error term in equation (4.1) is serially uncorrelated, so that we can obtain unbiased estimate of δ that is the coefficient of lagged Y_{t-1} can be obtained (Gujrati, 2004).

[1]A stochastic process is purely white noise process if it has zero mean, constant variance, and is serially uncorrelated (Gujrati,2004).

4.4.2. Co-integration Test

Co-integration test refers to check whether there is long run relationship over time among the variables. In the regression, a non-stationary time series variables should be differentiated to avoid the spurious regression. It helps to get more feasible inferences in the study. The existence of long run relationship or co-integration among variables leads to go for the regression model in the level without leading to a spurious regression. To check the co-integration among the variables this study has applied Johansen co-integration test. It is so because this test has an advantage over other co-integration tests as it takes into consideration the possibility of multiple co-integrating vectors. In this test co-integrations are usually examined on the basis of trace statistics and max statistics, developed by Johansen. The statistics are formulated as follows:

$\lambda_{trace}(r) = -T\sum_{i=r+1}^{n}(1 - ri)$ ……………………………...(4.2)

$\lambda_{max}(r, r_{+1}) = -T\ (1\text{-}r\text{+}1)$ ……………………………(4.3)

Where,

R = Number of co-integrating vectors

λ = Estimated value of r^{th} characteristic root (eigen value)

T= Number of observations

When the appropriate values of ‘r’ are clear these statistics are referred to the λ trace and λ max. The first statistics tests the null hypothesis that the number of distinct co-integration vector is less than or equal to r against an alternative hypothesis. From the discussion it is clear that λ and λ max are equal to zero when all the λ = 0. Further the estimated characteristic roots are from the zero, the more negative is (1- λ_i) and larger is λ trace statistics. The second statistics tests the null hypothesis that the number of co-integrating vector is r against the alternative hypothesis of (r+1) co-integrating vector. If the estimated value of the characteristic root is close to zero, λ max will be small. It shows that if there is one co-integrating equation for the given series, the results continues the presence of one co-integrating relationship among the variables.

4.4.3. Granger Causality Test

Apart from co-integration test, the present study examines the causal relationship as well as the directions of the relationship between variables electricity consumption and GDP. For the examination of the causality, this study has used the Granger causality test. According to Granger (1969,1988), a time series x granger causes another time series y if series y, can be predicted with better accuracy by the help of past values of x , rather than by not doing so, other information being identical. The causality can be estimated by the help of following equation:

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta Yt\text{-}_1 + \alpha_2 \Delta Yt\text{-}_2 + \ldots\ldots\ldots + \alpha_p \Delta Yt\text{-}_p + u_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4.4)$$

Where,

α_0 is vector constant term

α_1 , α_2,……..,α_p are parameter to be estimated

u_t is error term.

4.4.4. LM Test for Auto-correlation

Auto correlation is defined as the correlation between the members of series of observations ordered in time. In time series analyses the detection of auto correlation is crucial. For this purpose this study has adopted the Lagrange Multiplier (LM) test, developed by Breusch (1978) and Godfrey (1978). It became a standard tool in applied econometrics. The test is performed through an auxiliary regression of the residuals on their lags and the independent variables (Doornik, 1996). In this test two forms are computed:

i) TR_2, where T is the sample size and the R_2 is the co-efficient of multiplier correlation in the auxiliary regression. This statistic has an asymptotic chi-square distribution.

ii) The F- test on the lagged residuals in the auxiliary regression.

Here the null hypothesis is there is no autocorrelation. This null hypothesis can be rejected if the probability value is less than 5%.

4.4.5. The Jarque -Bera (JB) Test for Normality

In the empirical study the detection of normality of residual terms is much important for the problem of spurious relations. This study has applied the JB test for the purpose of diagnosis the presence of normal distribution of residuals. It was developed to test normality, hetero scedasticity and serial correlation or autocorrelation of regression residuals (Jarque & Bera, 1980). The statistics in this test is computed from skewness and kurtosis. It follows the chi-squared distribution with two degrees of freedom. Here, the null hypothesis is residuals are normally distributed which can be rejected if the probability value is less than 5%.

4.4.6. Data Envelopment Analysis (DEA)

DEA is a non-parametric mathematical programming approach for frontier estimation. It involves the use of linear programming methods to examine the relative efficiency of decision making units (DMUs). This non- parametric technique was originally proposed by (Charnes, W.W, & E, Measuring the Efficiency of Decision Making Units, 1978) and referred as a CCR model (Liu & Zhang, 2011). This technique has several variety of models. Among them following three models are utilized in various studies:

i) Standard CRS, VRS and DEA Model.

ii) Cost and Allocative Efficiencies Analysis Model.

iii) Malmquist DEA Method.

The standard CRS, VRS and DEA model consist the calculation of technical and scale efficiencies. This model was extended and further developed as the technique for the analysis of cost and allocative efficiencies. The CRS, VRS and DEA model Cost and Allocative efficiencies analysis models are outlined by Fare, Grosskopf and Lovell (1994). While the Malmquist DEA method was discussed in Fare, Grosskopf, Norris and Zhang (1994). This method mainly used for panel data to calculate the indices of total factor productivity (TFP) change, technological change, technical efficiency change and scale efficiency change.

4.4.7. Malmquist DEA Method

If a researcher is dealing with a panel data one may utilize the DEA like programme and a (in-put or out-put) Malmquist total factor productivity (TFP) index are used to measure productivity change, and to decompose this productivity changes into technical change and technical efficiency change. Fare et. al. (1994) constructed the DEA based Malmquist Productivity Index which is the geometric mean of two Malmquist Productivity Indices of Caves et al (1982). Among these two indices one measures the change in efficiency and other measures the changes in frontier technology. It is estimated by using DEA for a set of DMUs (Fare, Grosskop, Mary, & Zhang, 1994); (Caves, Christensen, & Diewert, 1982).

There are n DMUs under the examination of their performance. Let x_{ij} and y_{rj} denote the value of the i^{th} in-put (i= 1,2,3,……, m) and the r^{th} out-put (r=1,2,3,….., s), of DMU_j (j=1,2,3,…..,n), respectively. The slack variables for the i^{th} in-put and r^{th} out-put are represented by s_i^- and s_r^+ respectively which indicate the input excess and the output short fall respectively. The variable λ_j denotes the weight of DMU_j at the time of examining the performance θ_0 of object DMU_0.

According to Fare et. al. (1982), out-put based Malmquist Productivity Change Index can be obtained from the formula given as follows:

$$M_0^{t+1}(y_{t+1,}\, x_{t+1} y_{t,}\, x_t) = \left[\frac{d(t)\,\{x(t+1),\ y(t+1)\}}{d(t)\,\{x(t),\ y(t)\}}\,\frac{d(t+1)\,\{x(t+1),\ y(t+1)\}}{d(t+1)\,\{x(t), y(t)\}}\right]^{1/2} \quad \text{…………………(4.5)}$$

Where,

M_0 = Represents the productivity of production point (x_{t+1}, y_{t+1}) relative to the production point (x_t, y_t)

d (t) = Relative efficiency of particular DMU in period t against the performance of those DMUs in period (t+1).

x(t) = The In-put of particular DMUs in time period (t) against the performance of those DMUs in time period (t+1).

y (t) = The relative out-put of particular DMUs in time period (t) against the performance of those DMUs in time period (t+1).

When, $M_0^{t+1} > 1$, signifies a productivity gain,

$M_0^{t+1} < 1$, signifies a productivity loss,

$M_0^{t+1} = 1$, signifies there is no change in productivity (Coelli, 1996), (Liu & Wang, 2007).

One of the objectives of this study is to examine the efficiency of two power units in India which are thermal power unit and Hydro power unit. These are called as DMUs. Thermal power denoted as DMU1 while hydro power sector is denoted as DMU2. One In-put and one output are taken in this study and the time period is fifteen years. The installed capacity in Mega Watt (MW) and gross generation in (MU units) is taken as the In-put and out-put for thermal power unit. For hydro power sector installed generating capacity (in MU) and gross generation of power in (MU units) are taken as in-put and out-put respectively.

Assuming constant returns to scale (CRS), Malmquist DEA method has been applied to check the efficiency of these two power sectors. Productivity at different time period has been calculated on the basis of the equation (4.5) which reveals the efficiency of these DMUs.