CHAPTER 3

RESEARCH METHODOLOGY OF THE STUDY

3.1 Introduction

Research methodology is a way to solve the research problem scientifically. It can be understood as a science to study that how a research can be done. In this section of study a researcher studies various steps that have been adopted in the research. The logic behind adaptation of certain method or techniques is also mentioned in this section.

This study is based on secondary data compiled from various sources including the budget Expenditure on Education, Department of secondary and higher education, Government of India. The data on variables used in the present study are Government expenditure on education and Gross Domestic Product at constant prices it is so because GDP has been used as a proxy of Economic Growth in this study. Education expenditure consists expenditure at different levels of education such as elementary education, secondary education and university education. In spite of these expenditure there is aggregate education expenditure known as total education expenditure which is the sum of expenditures mentioned above.

3.2 Theoretical Frame Work

Over a long period of time the value of education has been accepted as a significant factor of economic growth. Education decides the performance of development process in an economy. Apart from several contributions in social, cultural and political fields, education also contributes to train and qualify labours. It leads to enable to develop human capital in an economy which results as a rise in productivity, wage increment and thus economic growth. For the human capital development effective education system is essential and for improvement in education, expenditure on education is very crucial (Mercan, 2013, Idrees and Siddiqi, 2013). So education expenditure is known as an important factor of economic growth. In this section Lucas (1988) model on human capital and economic growth has been presented.

3.2.1 Lucas Model

Usually it is stated that quality of education helps to increase the marginal productivity of labour. In the form of education an individual gets training to enhance one's productivity. The enhanced productivity results as an increment of wage. In this model Lucas states that higher the quality of training, the higher will be the marginal product of labour and thus in future higher will be the wage rate. This increment in wage rate appears in terms of higher growth rate in an economy.

In the basic form Lucas model can be presented in the following way:

$$Y = AK^{\alpha}\,(uhK)^{1-\alpha} \qquad \text{……………………………………..} \qquad (3.1)$$

Where,

Y= Output, A = Level of technology, K = Capital stock, L = Labour, u = Total labour time spent on working, h = Stock of human capital.

The production function can be re-written in per capita form as following:

$$y = Ak^{\alpha}(uh)^{1-\alpha} \qquad \text{……………………………………..} \qquad (3.2)$$

Where, k = per capita stock of physical capital. Eqn. (3.2) is the constant returns to scale production function in K and uh. Capital accumulation process via the usual differential equation.

$k = y - c\ (\xi + \delta)k$ ……………………………………… (3.3)

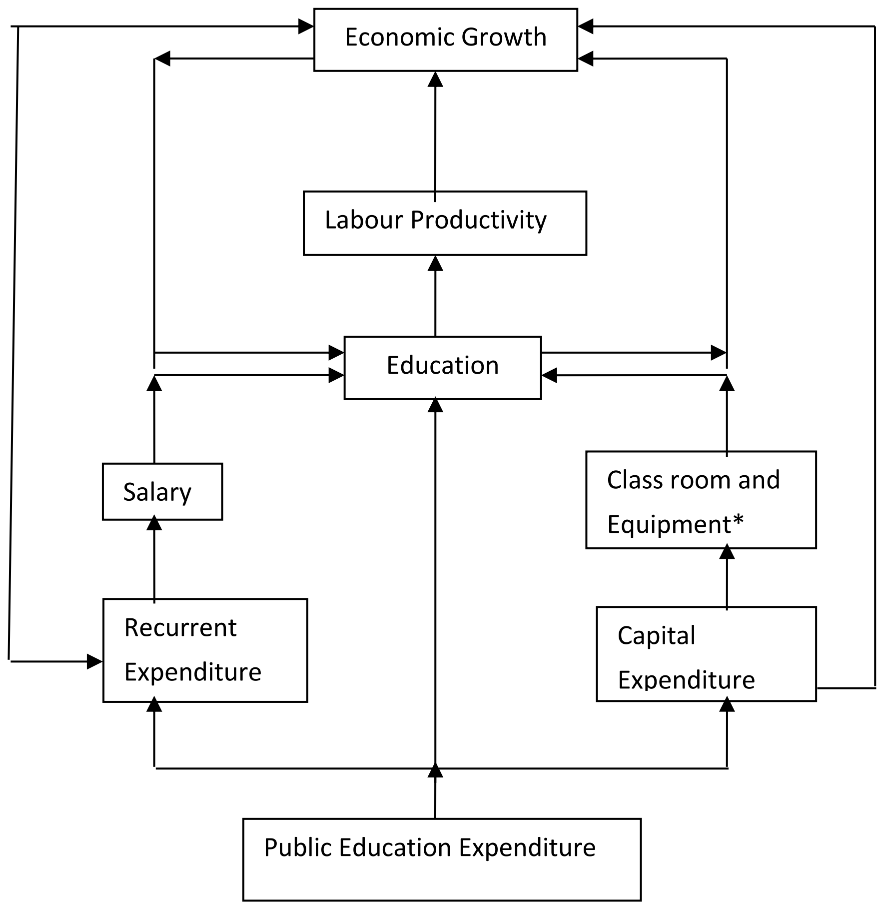
$\frac{dh}{dt} = \phi h(1 - u)$, u > 0 ……………………………………... (3.4)

However equation (3.4) states that the growth rate of human capital is determined by time spent in education or training and equation (3.2) states the way by which human capital affects the current production function.

This model says that human capital is related to increase education. As education increases human capital increases which appears as a result of increase in output. However Lucas model states that time spent to education enables an individual to enhance human capital which helps in capital accumulation due to increase in productivity and wage rate. In this way economic growth of a country increases.

The above theoretical frame work outlines an indirect relationship between education expenditure and economic growth. This frame work also presents the Lucas model of endogenous growth model. According to the model of Barro and Sala-i-Martin public education expenditure directly affects economic growth by Government expenditure multiplier. In summary it can be said that public education expenditure affects by both way directly and indirectly. This relationship between education expenditure and economic growth has been drawn in the following figure:

Figure 3.1: Education Expenditure and Economic Growth: A Frame Work



Source: ESE S. Urhie[1]
*Equipment includes laboratories, library services and teaching aids.

1 ESE S. Urhie Public Education Expenditure and Economic Growth in Nigeria: 1970-2010, Department of Economics and Development Studies, College of Development Studies Covenant University.

3.3 Hypothesis

The following Hypotheses have been formulated in this study:

i. Elementary education expenditure does not cause economic growth in India.
ii. Secondary education expenditure does not cause economic growth in India.
iii. University education expenditure does not cause economic growth in India.
iv. Education expenditure (as a whole) does not cause GDP

3.4 Methodology

In this study various literatures have been reviewed. In these literatures theoretical as well as empirical both are included. After reviewing these literatures it is underlined that the role of education expenditure in models of economic growth (Lucas 1988, Ram 1986, Barro and Sala- i- Martins 1992), the endogenous growth model is specific. The relationship between education expenditure and economic growth has also been identified in the form of structured equation.

In this study I have considered structural equation model with the dependent variable GDP denoted as Y and three independent variables denoted by El, Sec and Univ.

Y = f (Elem, Sec, Univ) ………………………………………… (3.5)

Where,

Y: Gross Domestic Product (GDP).

Elem: Elementary education expenditure.

Sec: Secondary education expenditure.

Univ: University education expenditure.

In econometric form the above equation has been written as follows:

$$Y_t = \beta_0 + \beta_1 El_t + \beta_2 Sec_t + + \beta_3 Univ_t + u_t \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3.6)$$

Where,

β_0 is intercept , β_1, β_2 and β_3 are parameters while u_t is the disturbances or error term.

The set of independent variables includes the measures of expenditure on education in absolute value. These expenditures are divided into elementary education expenditure, Secondary education expenditure and University education expenditure.

3.4.1 Unit Root Test

The test of stationarity (or non stationarity) that has become popular for the past several years in the name of unit root test (Gujrati, D. N.). The stationarity of data has been characterized by a time variant mean and variance. If mean and variance of a data are constant then the data is called stationary. To avoid the problems of spurious regression, it is necessary to confirm whether a stationary co-integration relationship among the variables. The results will be meaning less if it is estimated the relationship without knowing the stationarity of data. Unit root tests are performed to examine the stationarity of series. In this study Augmented Dickey Fuller (ADF) test has been applied to check the unit root of yearly time series. These tests are carried out at the level as well as at their first difference for intercept, trend and intercept and without trend and intercept model.

Here I have started unit root process from Random Walk Model (RWM[2]) without drift.

$Y_t = \rho Y_{t-1} + u_t$ $\quad -1 \le \rho \le 1$ ………………………… (3.7)

Where, u_t is white noise error term[3].

It is known that if $\rho = 1$ that is the case of unit root and y_t is called non stationary series.

$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t$ (subtracting Y_{t-1} from both sides in eqn. 3.7)

$= (\rho - 1) Y_{t-1} + u_t$

$\Delta Y_t = \boldsymbol{\delta} Y_{t-1} + u_t$ ………………………………… (3.8)

Where, $\boldsymbol{\delta} = (\rho - 1)$ and Δ is usually the first difference operator.

At the time of testing null hypothesis that is $\boldsymbol{\delta} = 0$, alternate hypothesis is $\boldsymbol{\delta} < 0$. If $\boldsymbol{\delta} = 0$,

Then, $\rho = 1$ that is the unit root. It means time series is non-stationary.

3.4.2 Augmented Dickey Fuller (ADF) Test

In the case of stationary, the error term (u_t) is assumed to be uncorrelated. But in the case when error term (u_t) is correlated another test is developed which is known as Augmented Dickey Fuller Test (ADF). This test is developed by Dickey and Fuller. In the test of ADF Y_t has been taken as random walk with drift around a deterministic trend which can be written in equation form as follows:

$\Delta Y_t = \beta_1 + \beta_2 t + \boldsymbol{\delta} Y_{t-1} + u_t$ ………………………….. (3.9)

In ADF test the following equation has been estimated

$\Delta Y_t = \beta_1 + \beta_2 t + \boldsymbol{\delta} Y_{t-1} + \sum_{i=1}^{m} \alpha i \, \Delta Y_{t-i} + \varepsilon_t.$ ………………. (3.10)

---

[2] A series is said to be random walk if $Y_t = Y_{t-1} + _u t$. Here u_t is a white noise error term with mean 0 and variance $\boldsymbol{\delta}^2$.

[3] White noise process is a special type of stochastic process or time series which have zero mean, constant varianceσ^2 and is serially uncorrelated.

Where,

ε_t is pure white noise error term[4], $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$, etc.

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 (3.10) is serially uncorrelated, so that we can obtain unbiased estimate of $\boldsymbol{\delta}$ that is the coefficient of lagged Y_{t-1} (Gujrati, D. N.).

3.4.3 Co-integration Analysis

This study has used Johansen co-integration method to test the co- integration relationship among the variables taken in this study. By using this method, I found that how the variables are co- integrated with each other. As Granger notes a test for co-integration that can be thought of as a pre-test to avoid spurious regression problem. So at the first stage study checks the integration order of the series. After that it implies the Johansen co-integration method to investigate the relationship between education expenditure and economic growth. When there are two variables model there can be only one co-integration vector. But when there are more than one variables in a model, the number of co-integrating vectors are more than one. For ‘n’ number of variables there are ‘n-1’ co-integrating vectors. This problem cannot be resolved by Engel-Granger single equation approach. Since there are four variables (more than one) in this study therefore, Johansen approach for multiple equation is applied. Considering ‘n’ variables consisting all of which can be endogenous, a Vector Auto Regressive (VAR) model with higher order can be written as follows:

---

[4] Pure white noise process: A stochastic process is purely white noise process if it has zero mean, constant varianceσ^2, and is serially uncorrelated.

$X_t = A_1 X_{t-1}+A_2X_{t-2}+\ldots\ldots\ldots\ldots+A_p X_{t-p} +e_t$ ……………......... (3.11)

Where,

X_t= (n×1) vector (X_{1t}, X_{2t},…….X_{nt})

e_t = An independent and identically distributed n dimensional vectors with zero mean variance matrix $\sum_e$.

In the present study the number of co-integrations are examined with the help of trace and max statistics which is developed by Johansen. The statistics are formulated as follows:

$\lambda_{trace}(r) = -T\sum_{i=r+1}^{n}(1 - r_i)$ ………………………….. (3.12)

$\lambda_{max}(r, r+1) = -T\ (1-r_{+1})$ ………………………….. (3.13)

Where,

r is the number of co-integrating vectors

λ is the estimated value of r^{th} characteristic root (eigen value)

T is the 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 a general alternative hypothesis. From the discussion it is clear that λ_{trace} and λ_{max} equal to zero when all the λ = 0. Further the estimated characteristic roots are from the zero, the more negative is $(1-\lambda_i)$ 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.

3.4.4 Vector Auto Regression (VAR) Model

Generally, this technology is used to forecast interrelated time series data. It is applicable to analyse the vital impact of random disturbances of the system of variables. In this model all variables are treated on an equal footing. There is no distinction between endogenous and exogenous variables in this model (Gujrati, N.). This model is incapable to incorporate the residual from the past observation in to the regression model for the current observation. This method is easy to understand and it is easily extendable to nonlinear specifications and models that may contain endogenous right hand side variables (Ndem, 2007). The model can be represented as follows:

$$Z_t = \theta_0 + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \theta_3 Z_{t-3} \ldots + \theta_p Z_{t-p} + u_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.14)$$

Where,

Z_t is 4×1 vectors containing GDP, Elem, Sec and Univ.

θ is 4×1 vector of constant terms;

$\theta_1, \theta_2, \theta_3 \ldots \theta_p$ are parameters to be estimated;

P is the optimal lag order;

u_t error term.

3.4.5 Vector Error Correction Model

The vector error correction (VEC) model is just a special case of the VAR for variables that are stationary in their differences (i.e., I (1)). The VEC can also take into account any co-integrating relationships among the variables.

After identification of the co-integration of variables it is needed to investigate the dynamics of the variables. These include short run as well as long run dynamics. All variables are converted into its natural logarithms and they are represented in the following form of Vector Error Correction Model (VECM):

$$\Delta Y_t = \phi_0 + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \ldots + \phi_p \Delta Y_{t-p} + u_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3.15)$$

Where,

ϕ_0 is vector of constant term;

$\phi_1, \phi_2, \phi_3 \ldots \phi_p$ are the parameters to be estimated;

u_t is the error term.

Since the prime task of this study is to determine the causal relationship between the variables. So it proceed to estimate the following vector error correction model and for four variables case the vector error correction model have been specified as follows:

$$GDP_t = \alpha_1 + \sum_{j=1}^{j=p} \beta_1 GDP_{t-j} + \sum_{j=1}^{j=p} \gamma_1 Elem_{t-j} + \sum_{j=1}^{j=p} \delta_1 Sec_{t-j} + \sum_{j=1}^{j=p} \Theta_1 Univ_{t-j} + u_{1t} \quad \ldots\ldots\ldots\ldots \quad (3.16)$$

$$Elem_t = \alpha_2 + \sum_{j=1}^{j=p} \beta_2 GDP_{t-j} + \sum_{j=2}^{j=p} \gamma_2 Elem_{t-j} + \sum_{j=2}^{j=p} \delta_2 Sec_{t-j} + \sum_{j=1}^{j=p} \Theta_2 Univ_{t-j} + u_{2t} \ldots\ldots\ldots\ldots \quad (3.17)$$

$$Sec_t = \alpha_3 + \sum_{j=1}^{j=p} \beta_3 GDP_{t-j} + \sum_{j=2}^{j=p} \gamma_3 Elem_{t-j} + \sum_{j=2}^{j=p} \delta_3 Sec_{t-j} + \sum_{j=1}^{j=p} \Theta_3 Univ_{t-j} + u_{3t} \quad \ldots\ldots\ldots \quad (3.18)$$

$$Univ_t = \alpha_4 + \sum_{j=1}^{j=p} \beta_4 GDP_{t-j} + \sum_{j=2}^{j=p} \gamma_4 Elem_{t-j} + \sum_{j=2}^{j=p} \delta_4 Sec_{t-j} + \sum_{j=1}^{j=p} \Theta_4 Univ_{t-j} + u_{4t} \quad \ldots\ldots\ldots \quad (3.19)$$

Where,

α_1, α_2, α_3 and α_4 are intercept, β_1, β_2, β_3 and β_4, γ_1, γ_2, γ_3 and γ_4, δ_1, δ_2, δ_3 and δ_4, δ_1, δ_2, δ_3 and δ_4, Θ_1, Θ_2, Θ_3 and Θ_4, $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, $\boldsymbol{\Phi}_3$ and $\boldsymbol{\Phi}_4$, and $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, $\boldsymbol{\lambda}_3$ and $\boldsymbol{\lambda}_4$, are parameters to be estimated, u_{1t}, u_{2t}, u_{3t} and u_{4t} are white noise error terms.

3.4.6 Graphical Method

For the purpose of trend analysis or education expenditure this has applied the simple graph method. In this method a separate graph for each variable has been plotted. These separate graphs include GDP, elementary education expenditure, secondary education expenditure and university education expenditure. In these graphs there is also a trend line which indicate about the trend of concerned variables.

3.5 Variables, Data and Sources

The empirical examination has been carried out for the period 1990-91 to 2012-13. In this study variables are Gross Domestic Product (GDP) and sector wise education expenditure. In these sectors of education expenditure only vital sectors have been taken. These sectors are Elementary education expenditure, secondary education expenditure and university education expenditure. The present study uses the secondary data, collected from various departments Government of India. Data of GDP has been collected from planning commission, Government of India and data for sector wise education expenditures have been collected from Ministry of Human Resource

Development (MHRD), Government of India. This study has used the expenditure on education which are financed by central government. These data are in absolute amount. More over the data for GDP is in constant price 2004-05 and their unit is crore.