Chapter 4

Research Methodology
4.1 Research Process (Framework for carrying Research)
4.2 Problem Statement
4.3 Research Objectives
4.4 Hypotheses Formulation
4.5 Scope of Study
4.6 Research Variables
4.7 Sources of Data
4.8 Sample Size
4.9 Test of Reliability
4.10 Tools for Data Analysis
4.11 Limitations of the Study
Chapter Overview

This chapter presents issues relating to Research Methodology adopted for the present study. The chapter starts with Research Process Section. The second section deals with Problem statement. The third section listed out Research Objectives. The fourth section noted down Hypotheses formulation. The fifth sections deals with scope of study. The sixth section discuss about the Research variables. The seventh section list out the sources of the data. The eight section discuss about the Sample Size and Selection. The ninth section tested the Reliability. The tenth section described the Tools for Data Analysis and it’s justification. The chapter ends with the limitation section.

4.1 Research Process (Framework for carrying Research)

The following section introduces meaning of research, research methodology and flow of research methodology.

Research is undertaken within most professions. More than a set of skills, it is a way of thinking: examining critically the various aspects of one’s professional work. It is a habit of questioning what one does, and a systematic examination of the observed information to find answers with a view to instituting appropriate changes for a more effective professional service. Research in common parlance refers to a search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. The Oxford Dictionary lays down the meaning of research as an “endeavour to discover new or collate old facts etc. by scientific study of a subject, course of critical investigation”. More precisely research is a systematized effort to gain new knowledge. Some people consider research as a movement, a movement from the unknown to the known. It is actually a voyage of discovery. The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, one may think of research objectives as falling into a number of following broad groupings:
1. To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as exploratory or formulative research studies);

2. To portray accurately the characteristics of a particular individual, situation or a group (studies with this object in view are known as descriptive research studies);

**Figure 4.1 Flow Chart of Research Methodology**

- Problem Discovery and Definition
- Define Research Objectives
- Selection of Research Techniques
- Exploratory/Descriptive
- Selection of Sampling Design
- Randomly selected 50 PSEs out of 61 listed in BSE
- Probability Sampling
- Data Gathering
- Collection of Data (Field Work)
- Data Processing and Analysis
- Editing and Coding of Data
- Tools used - Descriptive statistics, Correlation, Regression, SEM, Panel data analysis, GFA and MPL
- Data Processing and Analysis
- Drawing Conclusion and Preparation of Report
- Interpretation and Findings
- Secondary Data Study
- Audited Annual Reports
- Report
To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as diagnostic research studies);

4. To test a hypothesis of a causal relationship between variables (such studies are known as hypothesis-testing research studies).

The researcher has opted for combined descriptive and exploratory research studies. VAIC is an established model (Descriptive) and which he has tested other variables (Exploratory) also to enhance the model. A detail flow chart is attached below.

Research methodology is a way to systematically solve the research problem. It may be described as a science of studying how research is done scientifically. In the following paragraphs, the researcher explains the various steps that are adopted by him in studying his research problem along with the logic behind them. Research methodology is a way to systematically solve the research problem (Kothari, 2008).

4.2 Problem Statement

The following section listed the research questions.

So far Indian Public Sector was taking their financial measures as their performance yardstick. There was no reference of performance with respect to knowledge capital or intellectual capital which is one influential factor of Corporate Governance mechanism (Idrianita Anis, 2013) leading to valuable contribution to corporate performance. Therefore, with this purpose in mind a set of research questions were identified which this research endeavour seeks to answer. These are enlisted below:

Research Question 1 – What is the relationships between different elements of Knowledge Capital.

Research Question 2 – What are the important elements of Knowledge Capital for Public Sector Enterprises in India.

Research Question 3 – How do the different elements of Knowledge Capital relate to performance of the Public Sector Enterprises in India.

Research Question 4 – How can the different elements of Knowledge Capital be measured.
Research Question 5 - How can the different elements of Knowledge Capital be reported.

4.3 Research Objectives
The following section lays down objective of this research study.

The primary objective of this study is to evaluate the performance of Knowledge Capital of the Public Sector Enterprises in India. The study is based on the argument that the shareholders earning is an outcome of Knowledge Capital efficiency. It views that Knowledge Capital is responsible to bring more earnings per share (EPS) for shareholders after attaining maximum value creation efficiency. The objectives are listed below:

- Evaluate performance of Knowledge Capital (KC) of Public Sector Enterprises (PSEs) in India.
- Establish relationship with Value Added Intellectual Coefficient (VAIC) and Earning Per Share (EPS).
  - Establish relationship with Human Capital Efficiency (HCE) and Earning Per Share (EPS).
  - Establish relationship with Structural Capital Efficiency (SCE) and Earning Per Share (EPS).
  - Establish relationship with Capital Employed Efficiency (CEE) and Earning Per Share (EPS).
- Establish relationship with other variables Size of the Assets (ASSET), Frequency of Board meeting (MEETING), Number of Executives (NOEXE), Remuneration of CEO and Directors (CEOEXDIR), Ratio of Non-Executive Director to Total Number of Directors (NONR) with Earning Per Share (EPS).

4.4 Hypotheses Formulation
The following section registering list of hypotheses to be tested.

When a prediction or a hypothesised relationship is to be tested by scientific methods, it is termed as research hypothesis. The research hypothesis is a predictive statement that relates an independent variable to a dependent variable. Usually a research
hypothesis must contain, at least, one independent and one dependent variable. Predictive statements which are not to be objectively verified or the relationships that are assumed but not to be tested are not termed research hypotheses (Kothari, 2008). With these objectives in mind the researcher formulated ten hypotheses and ten alternative hypotheses in all. They have been listed below:

H0 1: There is no significant successful business performance for Public Sector Enterprise (PSE).

H1 1: There exists significant successful business performance for Public Sector Enterprise (PSE).

H0 2: There exists no significant relationship of Human Capital Efficiency (HCE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 2: There exists significant relationship of Human Capital Efficiency (HCE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 3: There exists no significant relationship of Structural Capital Efficiency (SCE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 3: There exists significant relationship of Structural Capital Efficiency (SCE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 4: There exists no significant relationship of Capital Employed Efficiency (CEE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 4: There exists significant relationship of Capital Employed Efficiency (CEE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 5: There exists no significant relationship of Size of Assets (ASSET) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 5: There exists significant relationship of Size of Assets (ASSET) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 6: There exists no significant relationship of Frequency of Board Meeting (MEETING) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 6: There exists significant relationship of Frequency of Board Meeting (MEETING) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 7: There exists no significant relationship of Number of Executive (NOEXE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).
H1 7: There exists significant relationship of Number of Executive (NOEXE) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 8: There exists no significant relationship of Remuneration of CEO and Directors (CEOEXDIR) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 9: There exists significant relationship of Remuneration of CEO and Directors (CEOEXDIR) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 9: There exists no significant relationship of Number of Independent Board member versus Total Number of Directors (NONR) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H1 9: There exists significant relationship of Number of Independent Board member versus Total Number of Directors (NONR) with Earnings Per Share (EPS) for Public Sector Enterprise (PSE).

H0 10: The panel data has no Fixed effect

H1 10: The panel data has Fixed effect

4.5 Scope of Study
The following section giving outline of the scope of study.

Scope has been defined in terms of time frame for the study and in terms of selection of units and also in terms of sectors. The study covers Public Sector Enterprises (PSEs) in India only which are listed in Bombay Stock Exchange (BSE) (www.bsepsu.com). The sectors covered are Finance (Banks and Financial Institutions), Agriculture (Fertiliser companies as it is termed as by BSE), Capital Goods, Metal and Mining, Miscellaneous (diversified), Oil and Gas, Power and Transport. The period of study is for the Financial Year 2001-02 to 2010-11 (ten years).

4.6 Research Variables
The following section listed the research variables and model adopted.

The variables for study have been described below.

The Value Added Intellectual Coefficient (VAIC) used in this study as a basic methodology to measure the Intellectual Capital or Knowledge Capital. The core
concept of VAIC is that human capital is mainly responsible for overall value creation performance of the organization. VAIC is the universal indicator which shows value creation ability of a company in quantitative terms and represents as measure of business efficiency in knowledge based economy. VAIC has three components which are—

1) Human Capital Efficiency (HCE)
2) Structural Capital Efficiency (SCE)
3) Capital Employed Efficiency (CEE)

The researcher has taken all the three as independent research variables. Earnings Per Share (EPS) is the earning per share held by the shareholders which has been taken as dependent variable as this is best indicator for return to shareholders.

In addition to this, the researcher has taken five additional variables to enhance the VAIC model—

Frequency of Board meeting (MEETING) – If the Board (top level decision making body) meets frequently then more meaningful knowledge of company policy will be inducted to the organization which will create more values (Vafeas, 1999).

Size of the total asset of the company (ASSET) – The size of the asset is helping to increase value of the organization (Ho and Williams, 2003).

Salaries of CEO, directors (top level executives) (CEOEXDIR) – If the good executive directors are in the board who can be inducted with higher remuneration, they would dedicate to the creation of value to the organization (Merhebi, 2006).

Number of Executives in the company (NOEXE) – the number of executives the company has who are knowledge workers also contributes to the value creation for enhancing corporate performance (Weill, 2005) and

Ratio of Non-executive directors to total number of directors in the board (NONR) – The function of the Board of Directors as an internal control mechanism is enhanced
by the inclusion of outside directors. This has been a one of the emphasis of corporate governance guidelines and laws (Cadbury, 1992; Sarbanes-Oxley Act, 2002; Indian Companies 1956, clause 49 of the agreement). This factor has influence on value creation as more and more non-executive director will bring out-side appropriate knowledge to the organization to create value (S.S.Devi, 2008; Z.Z.Abidin et al, 2009).

The research framework is described below.

Figure 4.2 Conceptual diagram of Research Framework

Literature is in agreement with three dimensions of knowledge capital, i.e. human capital, structural or organizational capital and relational or customer capital. This
research is based on Ante Pulic's VAIC model, which includes capital employed rather than relational capital and relational capital is considered as a part of structural capital. Human capital is treated as the most important asset of an organization which not only increases the operational efficiency in using tangible assets but also creates intangible assets. From the literature it is confirmed that successful companies always invest in HC to develop their overall working capabilities and environment. Investment in employee capabilities has direct impact on financial performance of a firm. As HC is the most influential factor of knowledge capital, therefore; the study checks empirically whether HC has direct relationship with shareholders earning or not. It includes financial capital/capital employed as part of IC efficiency considering better utilization of physical and financial assets is possible only due to efficient HC and SC. The study tests the relationship among Human Capital Efficiency (HCE), Capital Employed Efficiency (CEE) and Structural Capital Efficiency (SCE). The researcher also proposes five additional variables to examine the relationship with shareholders earning and these variables are : Frequency of Board meeting (MEETING), Size of the total asset of the company (ASSET), Salaries of CEO, directors (top level executives) (CEOEXDIR) and Number of Executives in the company (NOEXE) and Ration of Non-executive directors to Total number of directors in the board (NONR).

4.7 Sources of Data

The following section discussed about the sources of data.

The study is exploratory in nature. It is based on 10 years data from secondary sources. The period of study was Financial Year 2001-02 to 2010-11. Data were gathered from the audited annual reports of BSE-PSU companies. These annual reports were gathered through companies’ websites, BSE website, direct contact, www.reportjunction.com, Prowess database of CMIE. Companies in the sample cover eight industrial sectors (1-Agriculture, 2-Capital Goods, 3-Finance & Banking, 4-Metal & Mining, 5-Oil & Gas, 6-Power, 7- Miscellaneous and 8-Transport) making the sample representative.
4.8 Sample Size

The following section gives details about Sample Size and its selection.

Out of a total of 280 numbers of Public Sector Enterprises, 247 constitutes Central Public Sector Enterprises (CPSEs), 27 constitutes Public Sector Banks (PSBs), 6 constitute State Level Public sector Enterprises (SLPEs). There are 61 listed companies in BSE (http://www.bsepsu.com/bsepsu_index.asp). Out of these 61 listed companies, 3 companies (State Bank of Mysore-SBM, State Bank of Travancore-SBT and State Bank of Bikaner and Jaipur-SBBJ) are subsidiaries of State Bank of India (SBI), one company Chennai Petroleum Corporation(CPC) is subsidiary of Indian Oil Corporation(IOC) and one company Mangalore refinery and Petroleum Corporation(MRPL) is a subsidiary of Oil and Natural Gas Corporation(ONGC). The financial data of subsidiaries company's annual report (here input data) are consolidated in the holding company's annual report. Researcher is not taking the subsidiaries data as he is taking holding companies data. So, effectively only 56 independent CPSE as are listed in BSE (N=56). For population size(N) is known (Krejcie, Morgan 1970) the sample size, \( s = X^2NP(1-P) + d^2(N-1) + X^2P(1-P) \) where \( s \) = required sample size.

\( X \)= the table value of chi-square for 1 degree of freedom at the desired confidence level (3.841).

\( N \)= the population size.

\( P \)=the population proportion (assumed to be 0.50 since this would provide the maximum sample size).

\( d \)= the degree of accuracy expressed as a proportion (0.05).

The amount of \( s \) is calculated as 48.99. The researcher has taken as 50. The researcher has randomly selected 50 companies (enclosed in Annexure – 1) for the purpose of the study. The annual reports for the period from financial year 2001-02 to 2010-11 only of these companies are reviewed for the study.

4.9 Test of Reliability

The following section examined the reliability of data by Cronbach alpha method.
There are several different reliability coefficients. One of the most commonly used is by Cronbach's Alpha which is based on the average correlation of items within a test if the items are standardized; it is based on the average covariance among the items. Cronbach's Alpha is a measure of internal consistency, that is, how closely related a set of items are as a group. A "high" value of Alpha is often used (along with substantive arguments and possibly other statistical measures) as evidence that the items measure an underlying (or latent) construct.

Cronbach's Alpha can be written as a function of the number of test items and the average inter-correlation among the items. Below, for conceptual purposes, the formula is shown for the standardized Cronbach's Alpha:

$$\alpha = \frac{N \times \beta}{\nu + (N - 1) \times \beta}$$

Here N is equal to the number of items, \(\beta\) is the average inter-item covariance among the items and \(\nu\) equals the average variance. Because Cronbach's Alpha can be interpreted as a correlation coefficient, it ranges from 0 to 1.

**Table 4.1 Reliability Analysis**

<table>
<thead>
<tr>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.660</td>
<td>0.767</td>
</tr>
</tbody>
</table>

Here the Cronbach's Alpha is 0.767 which is suggesting that the items have relatively high internal consistency given for the number of observation being 500 nos. Reliability coefficient of 0.70 or higher is considered "acceptable" in most social science research situations (Hinton et al., 2012).

**4.10 Tools for Data Analysis**

The following section illustrated different Tools used and their importance in the study.

The major tools for analysis are –

i) Descriptive Statistics (SPSS)

ii) Correlation and Regression (SPSS and Minitab)

iii) Structural Equation Model (AMOS)
iv) Panel data Analysis (Gretl, Gnu Regression, Econometrics, and Time-Series Library)
v) Grey Relational Analysis (MS-Excel)
vi) Malmquist Productivity Index (Efficiency Management System developed by Holger Scheel) (http://www.holger-scheel.de/ems/)

Descriptive statistics
Descriptive statistics are very important because if raw data is simply presented it would be hard to visualize what the data was showing, especially if there was a lot of it. Descriptive statistics allows a researcher to carefully describe and understand behaviour of the variables. Descriptive statistics therefore enables to present the data in a more meaningful way, which allows simpler interpretation of the data, shows or summarizes data in a meaningful way such that, for example, patterns might emerge from the data. This is the basic thing required to handle data.

Correlation and Regression
Correlation analysis is to find correlation coefficient which is a measure of linear association between two variables. Regression analysis involves identifying the relationship between a dependent variable and one or more independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables. They can indicate only how or to what extent variables are associated with each other. Researcher wants to study the relationship among all the variables as well as relationship of dependent variable with other variables. This is also the fundamental analysis to find relationship among variables.

Structural Equation Model (SEM)
SEM is a largely confirmatory, rather than exploratory, technique. That is why, one is more likely to use SEM to determine whether a certain model is valid, rather than using SEM to find a suitable model, although SEM analyses often involve a certain exploratory element. The researcher has used this analysis for the confirmatory purpose only.

Panel Data Analysis
If a cross-sectional sample are measured two or more times, the resulting observations are described as forming a panel or longitudinal data set. The data for the current
study is an example of panel data as the variables are taken for 50 companies for 10 years. As it has a time component, researcher has felt the need to include panel data analysis also as the earlier analysis does to include time effect (Wasim-ul-Rehman, 2013).

**Grey Relational Analysis and Malmquist Productivity Index**

One of the objectives of this study is Performance evaluation of Indian Public Sector Enterprises in respect to Knowledge Capital of Intellectual Capital. To evaluation performance of the companies, one thing haunts the researcher about adequacy of data. Grey Relational Analysis does the estimation of efficiency while Malmquist Productivity Index does the estimation of productivity. Efficiency relates to the quality of work, which is creating output with less waste, using fewer resources or spending less money while productivity is measured by output during comparable time periods. This is the reason for taking these two methods in calculation for performance evaluation (Wang and Chang, 2011). Here researcher has taken EPS as output and HCE, SCE and CEE are the three inputs which are already part of VAIC and established. He has not considered other variables as these are yet to be considered part of VAIC.

**i) Descriptive Statistics**

- For better numerical understanding of the data, basic descriptive statistics has been applied.

- Descriptive statistics gives a way of accurately describing large datasets quickly and easily. The most common description statistics used are the measures of central tendency (mean, median and mode) and measures of dispersion (standard deviation and variance). Mean - the mean is the numerical average of all values in the sample. 

> Researcher has analysed mean of all the variables for all sectors of Public Sector Enterprises.

**ii) Correlation and Regression analysis**

These are used for quantitative measure of the strength of the linear relationship between independent and dependent variables. Correlation analysis is the statistical tool that is used to describe the degree to which one variable is linearly related to another. The regression line gives an estimation of the linear relationship between a
dependent variable and one or more independent variable or covariates. To understand the dynamics of IC and its impact on EPS, it is advantageous to apply regression model to the dependent and independent variables. Correlation and regression analysis are related in the sense that both deal with relationships among variables.

General syntax for multiple regression equation for population is:

\[ Y = a + b_1 X_{1i} + b_2 X_{2i} + b_3 X_{3i} + \ldots + b_k X_{ki} + \epsilon_i \]

Left side \((Y)\) of the equation specifies the outcome variable while right side \(b_1\) specifies the coefficient of the first independent variable \((X_1)\) and \(b_2\) is the coefficient of second independent variable \((X_2)\), \(b_n\) specifies the coefficient of nth independent variable \((X_n)\). \(a\) is the constant and \(\epsilon_i\) is the difference of predicted and observed value of \(Y\) for the ith participant.

Multiple regression results links IC efficiency with earnings performance. Regression equation considers earning per share (EPS) as dependent variable. The three independent variables are capital employed efficiency (CEE), human capital efficiency (HCE) and structural capital efficiency (SCE). While the five additional variables are size of the total assets of the company (ASSET), frequency of board meetings (MEETING), Salaries of CEO, directors and other executives (CEOEXDIR), number of executives in the company (NOEXE) and Ratio of Non-Executive to Total number of Directors in the Board (NONR). All independent and additional variables were taken together in the regression equation. Considering that the earning per share depends on intellectual capital efficiencies, the following equation is developed:

\[ Y_{EPS} = a + b_1(HCE) + b_2(SCE) + b_3(CEE) + b_4(ASSET) + b_5(MEETING) + b_6(CEOEXDIR) + b_7(NOEXE) + b_8(NONR) + \epsilon_i \]

Using residual analysis the assumptions of the regression model is tested. In order to use residual analysis to test the assumptions of the regression model, the researcher has to test the following assumptions:

a) Linearity of the regression model

b) Constant error variance (homoscedasticity)
c) Independence of error

d) Normality of error

Apart from this the overall model is tested by the ANOVA table generated by SPSS. A researcher may face problems because of the collinearity of independent (explanatory) variables while performing multiple regressions. This situation occurs when two or more independent variables are highly correlated with each other. In a multiple regression analysis when two independent variables are correlated, it is referred to as collinearity and when three or more variables are correlated, it is referred to as multicollinearity. Collinearity is measured by the Variance Inflationary Factor (VIF) for each explanatory variable. If explanatory variables are uncorrelated then Variance Inflationary Factor (VIF) if equal to 1. Variance Inflationary Factor (VIF) being greater than 10 is an indication of muticollinearity problems.

First researcher has developed a regression model that explains most of the variation in the dependent variable by all the explanatory variables. Second, the regression model has been developed which is parsimonious, easy to interpret and implement for a manager. In Enter method of regression, all the variables are taken into account.

In order to strengthen the competitiveness of public sector enterprises and to measure the performance of intellectual capital management the researcher has used Grey Relational Analysis (GRA) to measure operational performance and Malmquist Productivity Index (MPI) to judge productivity evaluation (R. Costa, 2008).

iii) Structural equation model (SEM)

It is a very general, very powerful multivariate technique. Researcher has done this modeling with the help of AMOS. SEM is an extension of the general linear model (GLM) that enables researcher to test a set of regression equations simultaneously. SEM software can test traditional models, but it also permits examination of more complex relationships and models, such as confirmatory factor analysis and time series analyses.
The model consists of a set of relationships among the measured variables. These relationships are then expressed as restrictions on the total set of possible relationships.

SEM users represent relationships among observed and/or unobserved variables using path diagrams.

The researcher has a database with eight continuous predictor variables: HCE, SCE, CEE, ASSET, MEETING, NOEXE, CEOEXIDR and NONR measured in 2001-02 to 2010-11. There is one continuous dependent variable, EPS measured in 2001-02 to 2010-11. These data are simulated based on the results reported in a similar study by Chen and Lee (2012).

The researcher has done the analysis for the eight sectors. The model is accepted or not is decided by the following table.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Good fit</th>
<th>Acceptable fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>(0 \leq \chi^2 \leq 2)</td>
<td>(2 \leq \chi^2 \leq 3)</td>
</tr>
<tr>
<td>CFI(Comparative Fit Index)</td>
<td>(0.97 \leq \text{CFI} \leq 1.00)</td>
<td>(0.95 \leq \text{CFI} \leq 0.97)</td>
</tr>
<tr>
<td>NFI(Normed Fit Index)</td>
<td>(0.95 \leq \text{NFI} \leq 1.00)</td>
<td>(0.90 \leq \text{NFI} &lt; 0.95)</td>
</tr>
<tr>
<td>C.R(Critical Ratio)</td>
<td>(&gt; 1.96 \text{ under 5% level of significance})</td>
<td>(&gt; 1.96 \text{ under 5% level of significance})</td>
</tr>
</tbody>
</table>

Source: Rao (2013)

iv) Panel Data Analysis

There are several reasons for the increasing interest in panel data sets. An important aspect is that their use offers a solution to the problem of omitted variable bias caused by unobserved heterogeneity, a common problem in the fitting of models with cross-sectional data sets. Data sets that combine time series and cross sections are called longitudinal or panel data sets. Panel data sets are more orientated towards cross section analyses – they are wide but typically short (in terms of observations over time). Heterogeneity across units is central to the issue of analysing panel data. To collect panel data one collects data on the same units for two or more time periods.
In panel data analysis there are broadly two type of approach Fixed effects approach and Random effects approach.

**Fixed Effects (LSDV) versus Random Effects model(ECM)**

Panel data models examine group (individual specific) effects, time effects or both. These effects are either fixed effect or random effect. A fixed effect model examines if intercepts vary across groups or time periods whereas a random effect model explores differences in error variances. A one-way model includes only one set of dummy variables(e.g firm) while a two way model considers two sets of dummy variable(firm and year). If it is assumes the error component and X's are uncorrelated, ECM may be appropriate. Whereas if they are correlated FEM may be appropriate. Keeping this fundamental difference in the two approaches in mind, the choice between FEM and ECM may be done by:

- If T(the number of time series data) is large and N(the number of cross-sectional units) is small there is likely to be little difference in the values of the parameters estimated by FEM and ECM. Hence the choice here is on computational convenience. On this score FEM may be preferable.
- When N is large and T is small, the estimates obtained by the two methods can differ significantly. Recall that in ECM $\beta_{it} = \beta_i + e_{it}$ but in FEM it is treated $\beta_i$ as fixed and non-random.
- If the individual error components $e_i$ and one or more repressor are correlated then the ECM estimators are biased whereas those obtained from FEM are unbiased.
- If N is large and T is small and if the assumption underlying ECM hold, ECM estimators are more efficient than FEM estimators.

Hausman Test is used to choose between FEM and ECM. The null hypothesis underlying the Hausman test is that FEM and ECM estimators do not differ substantially. The test statistic developed by Hausman has an asymptotic chi-square distribution. The Hausman test statistic has an approximate chi-square distribution with k degrees of freedom, where k is the number of slope parameters in the model. If the null hypothesis is rejected, the conclusion is that ECM is not appropriate and that
may better off using FEM in which case statistical inferences will be conditional on the error components in the sample.

The data being a panel data, the researcher has analysed different types of panel data analysis.

Which type of analysis is suitable for these data is determined by the following table (table 4.3).

**Table 4.3 Selection of approaches (OLS,FEM or ECM)**

<table>
<thead>
<tr>
<th>No.</th>
<th>Fixed effect (F test)</th>
<th>Random effect (B-P LM test)</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H0 is not rejected (No fixed effect)</td>
<td>H0 is not rejected (No random effect)</td>
<td>Pooled OLS</td>
</tr>
<tr>
<td>2</td>
<td>H0 is rejected (fixed effect)</td>
<td>H0 is not rejected (No random effect)</td>
<td>Fixed effect model</td>
</tr>
<tr>
<td>3</td>
<td>H0 is not rejected (No fixed effect)</td>
<td>H0 is rejected (random effect)</td>
<td>Random effect model</td>
</tr>
<tr>
<td>4</td>
<td>H0 is rejected (fixed effect)</td>
<td>H0 is rejected (random effect)</td>
<td>Fixed effect model is chosen if the null hypothesis of a Hausman test is rejected; otherwise, random effect model is fit.</td>
</tr>
</tbody>
</table>

Source: Park (2011)

The objective in this study is to use a sample of data on Indian Public Sector Enterprises to obtain an unbiased estimate of the effect of Earning Per Share. The researcher believe that the most important variables that affect the Earning per Share are HCE, SCE, CEE etc. and one unobserved variable say governmental interference. HCE,SCE, CEE are observable confounding variables. Because the researcher can obtain data for HCE, he can control it by including it as an explanatory variable in his model. Governmental interference is an unobservable confounding variable. Because the researcher cannot observe governmental interference and collect data for it, he cannot control for it by including it as an explanatory variable. However, the researcher believe that governmental interference differs across Indian Public Sector Enterprises, but is constant over time. Therefore, if he can collect panel data on
HCE, SCE, CEE etc. and he can specify a fixed-effects model and statistically control for innate ability.

v) Grey Relational Analysis (GRA)
Data Envelopment Analysis (DEA) is a relatively new "data oriented" approach for evaluating the performance of a set of peer entities called Decision Making Units (DMUs) which convert multiple inputs into multiple outputs. For the research under study each PSE is a DMU. The advantage of the Grey Relational Analysis is that it is designed to study uncertainty. It is shown that this method is superior to other methods in theoretical analysis of systems with uncertain information and incomplete data samples (Chiang, Hsieh, 2009). Especially, it can be used if the large samples are not available or if the user is uncertain whether the data is representative. It can also be used in early effective factor assessment (Chiu, 2009). Therefore, this study adopts Grey Relational Analysis approach to propose a feasible and effective method for performance evaluation of companies.

The generation of Grey relation for performance evaluation of organization is shown in Fig. 4.3.

**Figure 4.3 Flowchart to determine Grey relation degree**

- Setting up eigenvalue matrix, input original data
- Standardized data transformation, formulas:
  I) the bigger the better or
  II) the smaller the better or
  III) nominal-the best
- Calculation of Grey relational degree:
  - getting absolute difference of compared series and referential series using formula
  - find out minimum and maximum
  - choose the constant $p$ (set to 0.5)
  - calculation of relational coefficient and relational degree
- Set up the ranking of organizations based on influence factors

Source: Wang, Chang et al. (2011)
This study adopted the above mentioned research steps to develop an influence factors evaluation model based on GRA, and apply to influence factors evaluation and selection. The Grey relational analysis uses information from the Grey system to dynamically compare each factor quantitatively.

vi) Malmquist Productivity Index

This study also uses DEA’s malmquist model by using listed Indian Public Sector Enterprises information to analyse efficiency change for all the relevant companies and to measure technical efficiency scores during two particular periods. Secondly, the study analyzes technical change and measures the condition of efficiency frontier-shift between two particular periods. Finally, the study also carries out a comparison between the period efficiency and productivity change, in order to understand the situation of every annual growth and decline of efficiency and productivity.

Productivity is generally defined in terms of the efficiency improvement and technical change with which inputs are transformed into outputs in the production process (Costa, Campisi., 1998). There are two types of production efficiency: technical efficiency (TE), which evaluates a firm’s ability to obtain the maximum possible output from a given set of inputs, and allocative efficiency (AE) which measures a firm’s ability to maximize its profits by comparing marginal revenue product with marginal costs of inputs. However, this econometric approach requires the specifications of production function technology. Recently, mathematical programming approaches, such as Data Envelopment Analysis (DEA) are developed to measure TE by combining the firm’s production to the best production frontier. Specifically, the productivity can be measured by using narrow measures like partial productivity indices or a more comprehensive Total Factor Productivity (TFP). Partial Productivity Indices refer to ratios of output to each of categories of input for which separate data exist. Total factor productivity (TFP) is an overall indicator of how well an organization uses all of its resources to create its products and services. Moreover, TFP is a broader measure of economic and technical efficiency reflecting a diversity of factors including managerial efficiency, economies of scale, R & D, market structure and human capital utilization. TFP can be split up into two major
components viz: technological progress and improvement in technical efficiency. It is important at the outset to distinguish between technological progress and improvement in technical efficiency. Technological progress may be attributed to the introduction of new technology, which leads to an expansion of the best production frontier and hence gives higher output even with given input of resource. The other component is improvement in technical efficiency which yields higher output being the result of improved management practices, better industrial relationships, and diffusion of new technological knowledge as well as short run adjustments to shocks, external to the enterprise as technical efficiency change.

As researcher has panel data, Malmquist allows the decomposition of productivity changes into two components (technical efficiency change, and technical change or changes in the best practice). Following Costa et al. (2008) the Malmquist TFP index calculates the change in productivity between two points by estimating the ratio of the distances of each point relative to a common technology. The Malmquist input oriented TFP change index between the base period t & the following period t+1 is defined as:

\[ M(y_t, x_t, y_{t+1}, x_{t+1}) = \left[ \frac{d_t(Y_{t+1}, X_{t+1})}{d_t(Y_t, X_t)} \right] \times \left[ \frac{d_{t+1}(Y_{t+1}, X_{t+1})}{d_{t+1}(Y_t, X_t)} \right]^{1/2} \]

A value of M greater than unity implies a positive TFP growth from period t to period t+1. Otherwise, a value of M less than one indicates a TFP decline. The above equation is geometric mean of two TFP indices. The first index is calculated with respect to period t technology, while the second index is evaluated with respect to period t+1 technology.

This study uses companies which are Indian Public Sector Enterprises as DMUs. A total of 50 companies with data from the year 2001-02 to 2010-11 are chosen to be DMU as empirical sample. There are three inputs, HCE, SCE and CEE and one output EPS. The steps are as follows: when preceding the part of localization grey relational analysis, the first step must set up referential sequence and comparative sequence. This study factors belong to the small identity, then select the minimum and the large identity, then select the maximum to setup referential sequence. So those 50
companies are comparative sequence. When proceeding, the original data into the
grey relational generation, it mainly deals with data processing of the original data
that are yet to be true according to actual situation and promotion of data's
visualiziability. This study adopts Hsia’s method (Hsia and Wu, 1998) and proceeds
the original data of the HCE, SCE, CEE and EPS (all larger the better). Then calculate
the grey relational coefficient and grey relational grade. Followed by the value of the
grey relational grade, calculate the grey relational rank ordinal.

4.11 Limitations of the Study

The following section discussed about the limitations of this study

Due to the non availability of non listed and proprietary sector data, this research is
limited to BSE-PSU index companies. At present different countries apply different
national accounting standards, disclosure and listing requirements prevailing in that
country. These differences can also affect the results of this model in other countries
as they may be following different accounting framework. The results of the study
cannot, therefore, be generalized to apply to all the Public sector or whole stock
market or the non listed sector. Human and structural capital efficiency is not
comparable among different sectors within the BSE listed companies since different
industries are composed of different IC-related factors.