CHAPTER – IV

METHODOLOGY

According to Collis and Hussey (2003)\(^4\), Methodology refers to overall approach to research process which includes underpinning of theory, collection of data and analyzing it. The research process adopted depends to a great extent on the approach taken by the researcher. Research design is the general plan of how to go about answering the research question. It gives the logic behind every interpretation. Due to nature of research carried out, focus has been on gathering the secondary data which is relevant to the analysis being carried out. This chapter elaborates the methods adopted for collecting data, analyzing for the research study. The current research would aim at understanding the flow of FIIs in Indian stock market. Data that is collected and tools used to understand the impact of FIIs in Indian stock market are explained in this section.

4.1 Sources of Data

The main source of obtaining necessary data for the study is secondary in nature. The study is empirical in nature hence secondary data is used primarily to conduct the research. The data is collected from the Internet and hand books by exploring the secondary data available on websites like bse.com, rbi.com,

nse.com, etc. The data collected are analysed statistically using Microsoft Excel and SPSS 20.0 (Statistical Package for Social Science 20th Version).

Secondary Data: The secondary data constitutes FII flows data which is collected from Money Control and Equity Master, the daily returns of SENSEX and NIFTY from BSE and NSE websites respectively. The data to understand the trend in FII flows is collected from RBI website and information about FII from SEBI website. The data is also collected from the data bases and by exploring the secondary sources available on websites. The data collected is analyzed to understand FIIs behaviour and impact on Indian stock market, as the most dominant investor group in the domestic market.

The data gathered from various sources are primarily studied and necessary data is sorted out sequentially keeping in mind the procedure of the study. The impact of FII inflow or outflow on the Indian stock market and the influence of the stock market on the FIIs investment pattern depended on various factors. To examine the trend, a few indicators and measurements of the stock market are taken in to account. To measure the volatility of stock market and impact of economic indicators on the FIIs investment and vice versa, the information collected includes:

1. BSE Sensitive Index (SENSEX)
2. NSE Market Index (S&P CNX NIFTY)
3. Total Turnover of SENSEX and NIFTY
4. Market Indices - SENSEX and NIFTY

FII flows are chosen in terms of two variables viz., FII inflow (gross purchases) and FII outflows (gross sales). Data on FII equity purchases (inflows) and sales (outflows) on daily and monthly basis are taken from the website of SEBI. Apart from the above stock market indicators, market prices of 15 stocks are collected. From the 50 companies listed in NSE, the first 15 companies are selected based on the performance. It include: ACC Ltd., Ambuja Cements Ltd., Asian Paints Ltd., Axis Bank Ltd., Bajaj Auto Ltd., Bank of Baroda, Bharat Heavy Electricals Ltd., Bharat Petroleum Corporation Ltd., Bharti Airtel Ltd., Cairn India Ltd., Cipla Ltd., Coal India Ltd, DLF Ltd., Dr. Reddy's Laboratories Ltd., GAIL (India) Ltd.

4.2 Study Period

The trend in FII investment is studied for the complete period from the registration of FII in India i.e., from 1992 to 2012. The trend for FII purchases, sales, investment in equity and debt, and net investment is studied from 2000 to 2012. The trend for flow of FII in BSE SENSEX and CNX NIFTY is studied from 2000 to 2012. The relationship between FII flows and economic factors is studied by taking monthly data from 2000 to 2012. To study the impact of FIIs Net Investment pattern and Market Indices data are considered from 2000 to 2012.
4.3 Tools for Data Analysis

As the foreign investors emerged as one of the reasons for the movement of SENSEX and NIFTY in the Indian stock market, the present work is framed to analyze the influence of FII flows to understand the performance of stock market. The tools used for analysis include: Trend Analysis, Correlation analysis, Regression Analysis, Augmented Dickey Fuller test, and Granger causality test.

Yearly trend with increase and decrease in the investment made by FIIs is specified. The coefficient of trend variable explains the average growth of investment made by FIIs for the years. The percentage increase or decrease in the trend value explains the year on year growth in the investments of FIIs. The number of registered FIIs during the study period 1992 to 2012 is also analysed with increase or decrease in trend percentage. Percentage share of FIIs investment in different sectors of companies in NSE are given in percentage to the total investment. FIIs investment in equity and debt is given for the study period with their share to total investment on yearly basis and monthly basis. Volatility in stock market is studied for the 15 selected companies with their daily range, three months, weekly one year, two years, and monthly averages.

The net foreign investment and market indices are studied and their relationship is analyzed with correlation coefficient. Correlation analysis is a statistical measure of how two variables move in relation to each other. Pearson
correlation is ‘+1’ in the case of a perfect positive (increasing) linear relationship (correlation), ‘−1’ in the case of a perfect decreasing (negative) linear relationship and some value between ‘−1’ and ‘+1’ in all other cases, indicating the degree of linear dependence between the variables. As it approaches zero there is less of a relationship (closer to uncorrelated). The closer the coefficient is to either ‘−1’ or ‘+1’, the stronger the correlation between the variables. If the variables are independent, Pearson's correlation coefficient is ‘0’. If we have a series of n measurements of X and Y written as xi and yi where i = 1, 2, ..., n, then the sample correlation coefficient can be used to estimate the pearson correlation ‘r’ between X and Y. The sample correlation coefficient is written as:

\[ r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}, \]

where \( x \) and \( y \) are the sample means of X and Y, and \( S_x \) and \( S_y \) are the sample standard deviations of X and Y. This can also be written as:

\[ r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n - 1)s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}. \]

If \( x \) and \( y \) are results of measurements that contain measurement error, the realistic limits on the correlation coefficient are not ‘−1’ to ‘+1’ but a smaller range. The extent of existence of relation between the market indices and the FII
investment is also studied with regression analysis. For this purpose, r-square
value is used to understand the extent of influence of market indices on the net
foreign investment.

Regression analysis is a statistical process for estimating the relationships
among variables. It includes many techniques for modeling and analyzing several
variables, when the focus is on the relationship between a dependent variable and
one or more independent variables. More specifically, regression analysis helps
one to understand how the typical value of the dependent variable (or 'criterion
variable') changes when any one of the independent variables is varied, while the
other independent variables are held fixed. Most commonly, regression analysis
estimates the conditional expectation of the dependent variable given the
independent variables – that is, the average value of the dependent variable when
the independent variables are fixed. Less commonly, the focus is on the location
parameter of the conditional distribution of the dependent variable given the
independent variables. In all cases, the estimation target is a function of the
independent variables called the regression function.

Regression analysis is widely used for prediction and forecasting. Many
techniques for carrying out regression analysis have been developed. Familiar
methods such as linear regression and ordinary least squares regression are
parametric, in that the regression function is defined in terms of a finite number of
unknown parameters that are estimated from the data. The statistical equation developed for the Regression Analysis is:

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \]

The coefficients ‘\( \beta_1 \)’ and ‘\( \beta_2 \)’ explain the extent of variation caused by the independent variables in predicting dependent variable. The analysis of variation is a method which separates the variation to one set causes from the variation ascribable in the other set. Analysis of variance is referred to as ANOVA. This technique was developed by R.A. Fisher. ANOVA helps to examine the significant mean differences among more than two groups on an interval or ratio-scaled dependent variable. The results of ANOVA show whether or not the means of the various groups are significantly different from one another, and indicated by the F statistic. ANOVA table is used to understand the relationship between dependent and independent variables.

Correlation analysis is also performed between FII fund flow into India and various parameters such as rupee, stock market returns, and macro economic performance to understand the linkage between FII flows and the movement of the currency and stock prices. The determinants of FII are also analysed with correlation analysis.

The Granger causality test is a statistical hypothesis test for determining whether one time series is useful in forecasting another. Ordinarily, regressions
reflect ‘mere’ correlations, but Clive Granger argued that causality in economics could be reflected by measuring the ability of predicting the future values of a time series using past values of another time series. Since the question of ‘true causality’ is deeply philosophical, econometricians assert that the Granger test finds only ‘predictive causality’. A time series $X$ is said to Granger-cause $Y$ if it can be shown, usually through a series of $t$-tests and $F$-tests on lagged values of $X$ (and with lagged values of $Y$ also included), that those $X$ values provide statistically significant information about future values of $Y$. Granger defined the causality relationship based on two principles:

1. The cause happens prior to its effect.
2. The cause has unique information about the future values of its effect.

The Granger-Causality test is used in time series data. For any time series data analysis, all data series must be stationary. A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap or lag between the two time periods and not the actual time at which the covariance is computed. After checking the stationarity of the data series, the dynamic linkage is established between the two stationary series using Grangers Casuality test. For this analysis, index returns is taken by using the following formula.
\[ Rt = \ln P_t - \ln P_{t-1} \]

In the above equation, \( Rt \) represents return at time \( t \)

\( \ln \) is the natural logarithm

\( P_t \) represents Closing sensex Value at time, \( t \).

\( P_{t-1} \) represents previous day’s (t-1) sensex closing value.

Before making the data analysis, the time series properties of various data are identified. More specifically both the dependent and independent variables are examined. For verifying the time series properties of data, Augmented Dickey-Fuller test (ADF) is used for checking the unit root of the major variables. It is an augmented version of the Dickey–Fuller test for a larger and more complicated set of time series models. The augmented Dickey–Fuller (ADF) statistic, used in the test, is a negative number. The more negative it is, the stronger is the rejection of the hypothesis that it is a unit root at some level of confidence.

The testing procedure for the ADF test is the same as for the Dickey–Fuller test but it is applied to the model:

\[ \Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t, \]

where \( \alpha \) is a constant, \( \beta \) the coefficient on a time trend and \( p \) the lag order of the autoregressive process.
The main objective of this study is to analyze the role of FIIs in the Indian stock market and their contribution towards the performance of sensex, the benchmark index of the Indian stock market. The study encompasses the monthly net FIIs investment, collected from the “Database on Indian economy” of RBI, and the monthly index price of sensex, collected from BSE website. This study includes only the net monthly investment of FIIs and it does not include the cumulative investment. The study is carried out in threefold. First the FII flows into the Indian capital market are analyzed after 1993, during which India has opened up its economy thereby facilitating the flow of funds across the borders. As the foreign investors has emerged as the one of the routine reasons offered by market pundits for the movement of Sensex and Nifty in the Indian stock market, the present work is framed in order to analyze the influence of FII flows in the performance of stock market.

Vector Auto regression Model (VAR) is used to study the dynamic relationship between the two variables. The VAR estimation procedure requires selection of variables to be included in the model. The variables should be selected according to the relevant economic model and should fairly account for and explain the link between real and financial sides of the economy. The important steps in VAR estimation are as follows,

- Checking the stationarity of the variables
- Selection of appropriate lag length
• Ordering of variables.

For any time series analysis, all the data series must be stationary. Stationary condition has been tested using Augmented Dickey Fuller (ADF). It is a test for a unit root in a time series sample. The ADF, used in the test is a negative number. The more negative it is, the stronger the rejection of null hypothesis that there is a unit roots at some level of confidence. Therefore, the hypothesis of a stationary series can be evaluated by testing whether the $a_1$ value is strictly less than one. In time series literature, unit root tests like ADF tests are used to check whether a variable or series included in the model is stationary or not. For the VAR estimation all the variables included in the model should be stationary.

An unrestricted VAR assumes that the variables are related to both their own lagged values as well as lagged values of other variables. In this study estimated VAR are of reduced form since they only use lagged values of variables on the right hand side indicating non - existence of Simultaneity in the system. Numbers of variables included in the system depend upon theoretical considerations and decision about lag length is based upon statistical tests. For an unrestricted VAR it is necessary that the same numbers of lags of all of the variables are used in all equations. The optimal lag length is for minimum value of multivariate information criteria based on AIC, SC and HQ. It is essential that all variables included in the VAR should be stationary so as to conduct joint significance test on the lags of identified variables. The significance of all lags of each variable within the VAR framework is examined jointly by F tests which will establish the joint significance of all lags of the individual variables.