CHAPTER 3

RESEARCH METHODOLOGY

"An expert is one who knows more and more about less and less until he knows absolutely everything about nothing."

– Nicholas Murray Butter

This chapter explains the event study methodology, univariate and multivariate techniques used in analyzing the primary and secondary data.

3.1 INTRODUCTION

The study is divided into two parts. The first part of the study is causal in nature and caters to objectives one and two. Inclusion of a stock in an index is an information free event as discussed in chapter one. However evidence from the past shows that there are reactions to index inclusions. Hence the first two objectives study the impact of announcement and change on the price and volume traded of shares included in an index.

The second part of the study is exploratory in nature and pertains to the next three objectives which aim at studying the psychological biases that influence investment decisions. Eight most important psychological biases were considered and their presence were measured in the sample respondents along with an analysis of these biases among the different groups based on their demographic profile. Next, the behavioral factors that emerged out of these biases were extracted and those factors that best discriminate the investors with varying trading frequency were identified for a better understanding of investor behavior.
3.2 DATA FOR INDEX INCLUSIONS

Data used is secondary in nature. Since index inclusions were considered, two major indices were taken up for the study. One was the NSE Nifty index which comprises of 50 shares and the other was the S&P CNX 500 index. Both these indices are computed and maintained by IISL (India Index Services and Products Ltd.).

3.2.1 Stock Market Indices

A stock market index is a measure of the relative value of a group of stocks in numerical terms. As the stocks within an index change, the index value changes. An index is important to measure the performance of investments. Stock market index is created by selecting a group of stocks that are representative of the whole market or a specified sector or segment of the market. An index is calculated with reference to a base period and a base index value. Stock market indices are meant to capture the overall behavior of the equity market.

NIFTY

The NIFTY is a well diversified fifty stock index accounting for 23 sectors of the Indian economy. It is a popular stock index and is used for a variety of purposes such as benchmarking fund portfolios, index based derivatives and index funds (source NSE website).

CNX 500

The CNX 500 index represents about 95% of the free float market capitalization of the stocks listed on NSE and is India’s first broad based benchmark of the Indian capital market. The CNX 500 companies are
disaggregated into 73 industry indices. The industry weightages in this index represent the industry weightage in the market (source NSE website).

Both these indices are maintained by India Index Services and Products Limited (IISL), a subsidiary of NSE Strategic Investment Corporation Limited. The main objective of IISL is to develop, construct and maintain indices on Indian equities that serve as useful benchmarks and to provide data on trading activity in the Indian stock markets. IISL has constituted an Index Maintenance Sub-Committee that takes all decisions on replacements to an index.

The reason for choosing these two indices is that there are marked differences in the composition and maintenance of these indices during the period of study which are mentioned as follows:

1. Nifty is a smaller index with only 50 companies whereas CNX 500 is a larger index with 500 companies.

2. During the period of study, the changes to the composition of Nifty index was done systematically, that is, there was a gap of about six weeks between the date of announcement and the date of effective change. However with respect to CNX 500, changes were sudden and irregular. The time gap between the date of change and announcement was very less.

3. Nifty index is more popular among the general public as well as analysts whereas CNX 500 is mostly followed by analyst and fund managers to understand the direction of the market.

For the above reasons it was of interest to take the two distinctly different indices and study the price and volume reaction for companies which were included in these indices.
NSE website provides the complete details of the names and effective change dates of all the stocks that were included in the Nifty and S&PCNX 500. The announcement dates were inferred from the date of the circular intimating the change. Since circulars are released in the evening after the close of the market, the day following the circular date is considered as the announcement day.

Prowess database was used to download the daily returns and volume traded for the companies taken up for the study and NSE website provided details regarding closing values of the indices and volume traded on these indices.

3.3 EVENT STUDY METHODOLOGY FOR INDEX INCLUSIONS

The analysis was carried out in the event study framework, which is a proven methodology for testing the impact of a particular event on stock prices. An event study measures the impact of any event, economic in nature on the share prices for a selected period of time, which is normally a few days before and after the event. Empirical literature on finance used event studies for quite sometime to examine market efficiency, price effects of stock splits, dividend announcements, earnings announcements etc. Briefly in an event study the first step is to choose an event of interest and the impact of the same is studied on price and volume. The event taken up for the study was inclusion of shares in an index. The reason why this was taken up is because a survey of past literature revealed that stock prices jump when shares are included in the index even in developed markets in spite of the fact that it is an information free event. There are two significant dates when it comes to index inclusion, one is the ‘date of announcement’ and the other is ‘the effective date of change’.
In this study the price and volume effects of stocks that were added to the Nifty and S&P CNX 500 were analysed around the announcement date and the effective date of change. Three event windows namely ‘estimation’, the ‘announcement’ and ‘effective date of change’ windows were formed. The estimation window comprises of 100 trading days immediately before the announcement window. The announcement and change windows had 10 trading days prior to and 10 trading days after the event date which is either the announcement date or the date of effective change respectively. Though windows as short as three days and as long as 120 can be constructed for event studies, based on earlier studies on index composition changes, a 21 day window was considered apt for the study. Recent research articles by SSS Kumar (2007), Josipura and Janakiraman (2015) revealed that it was prudent to consider event window of 21 days. Thus the stock price returns and volumes traded over a short period of time were analysed to study the impact of announcement and the impact of change on both the indices.

3.3.1 Price Effect

There are several models used in research to estimate abnormal returns. The Market Model approach, which was considered to be well specified under a variety of situations when daily returns are used (Brown and Warner, 1985) was taken here to estimate the alpha and beta coefficients. For every security under the market model, the daily returns in the estimation window are expressed as follows:

\[ R_{jt} = \alpha_j + \beta_j R_{mt} + \epsilon_{jt} \]

Where \( R_{jt} \) denotes returns to stock \( j \) and \( R_{mt} \) denotes returns to the market portfolio on day \( t \) respectively.
The Abnormal Returns (AR) are calculated as under:

\[ AR_{jt} = R_{jt} - \alpha - \beta_j R_{mt} \]

where,

- \( AR_{jt} \) = the abnormal return of the particular stock \( j \) on the day \( t \);
- \( R_{jt} \) = the return of the particular stock \( j \) on the day \( t \);
- \( \alpha \) = the average returns of the firm compared to the market average;
- \( \beta_j \) = the market risk of this stock; and
- \( R_{mt} \) = the returns on a market index for day \( t \).

In order to make conclusions the abnormal returns are aggregated along time and across securities.

Mean Abnormal Returns (MAR\( t \)) ie the average of abnormal returns across \( N \) firms on day \( t \) is calculated as follows:

\[ MAR_t = \frac{1}{N} \sum_{j=1}^{N} AR_{jt} \]

Then Cumulative Abnormal Returns (CAR\( t \)) which is defined as the sum of all the excess returns over the window of interest is calculated as:

\[ CAR_t = \sum_{i=1}^{t} AR_{jt} \]

The Mean Cumulative Abnormal Returns (MCAR\( t \)) across the \( N \) firms over the event period is defined as:

\[ MCAR_t = \frac{1}{N} \sum_{j=1}^{N} CAR_{jt} \]
3.3.2 Volume Effect

To find out the change in trading activity around these event dates, volume study was undertaken. Though there are several models, the mean and market adjusted volume measure, similar to those of Harris and Gurel (1986) to examine abnormal volumes around event days is considered. Thus volume ratio for each day in the announcement and effective window was calculated as follows:

$$VR_j = \frac{V_j/V}{V_m/V_m}$$

Where $V_j$ and $V_m$ are daily volume traded, of stock j and the market respectively. $V_j$ and $V_m$ are the mean trading volume of stock j and the market during the trading days covered in the estimation window. The volume ratio is expected to have a value of 1 under normal conditions and will be away from 1 to denote impact of any event under study.

3.3.3 Testing for Statistical Significance

The cross sectional t statistic for daily abnormal returns was calculated using cross sectional variance under the assumption that abnormal returns are cross sectional independent and normally distributed.

$$t \text{ statistic} = \frac{\text{MAR}}{\text{Standard Deviation}/\sqrt{N}}$$

The same was used for calculating t statistic with respect to volume effect as well.
3.4 DATA FOR PSYCHOLOGICAL BIASES

Stock market investments require individuals to be informed to take optimal decisions. However very often it is found that investors trade for no fundamental reason and are greatly influenced by recent market trends. Shiller’s (1991) survey on individual and institutional investors and stock brokers behavior during the October 1987 market crash revealed that many trading decisions were influenced by emotions and psychological biases rather than news about fundamentals. Hence psychological biases in investment decisions were proposed to be studied by collecting primary data from the target population.

3.4.1 Target Population

Investors who invest in the stock market were the target population for the study. In this study the term ‘investor’ and ‘trader’ was used synonymously and refers to a person who holds a demat account, buys and sells securities in the stock markets.

3.4.2 Sampling Technique

Primary data was collected through questionnaires during the period 1st February 2013 to 30th November 2013. Data was collected on a stratified basis. Two stage stratified random sampling technique was employed for sample selection. In the first stage, out of the metropolitan cities in India, two cities namely Mumbai and Chennai were selected. Mumbai was selected because it is the trading capital of India and Chennai was selected because it is the place of residence of the researcher. The second stage of selection was made based on the frequency of trading as recommended by experts in the field, thereby giving representation to different categories of traders. This was done in order to understand investor behavior based on frequency of trading.
3.4.3 Sample Size

Initially it was proposed to collect approximately 400 responses for which 500 questionnaires were distributed. 368 completed questionnaires were collected out of which 24 questionnaires were rejected for incomplete and improper information. Hence the final sample settled at 344 which represents a response rate of 68.8% which was considered as reasonable for this type of study.

3.4.4 Questionnaire

Survey of relevant literature revealed the biases that influence investor behavior. Statements to measure these biases were then drawn from different sources. These statements along with questions to measure demographic variables were included in the questionnaire. Expert opinion was sought from academicians and practitioners and certain statements were modified and used for the purpose of the study.

The questionnaire consisted of two sections with 50 questions. Section A was aimed at studying the personal profile of the respondents consisting of gender, age, educational qualification, occupation, annual income, value of investment in the stock market, number of years of experience in trading in the stock market and trading frequency.

Section B was aimed at studying the psychological biases that influence investment decisions. This section consisted of statements that capture the different psychological biases. Survey of past literature helped in identifying eight psychological biases that contribute to individual irrationality thereby leading to sub optimal decisions. The statements were constructed in such a way that they do not straightaway reveal the psychological bias that they intend to measure. Most of the statements were situation based. A conscious effort was made to include wherever possible
statements that relate to investment decisions as well as statements that relate to general life situations to understand the biases at all times. The statements that study each of these biases were jumbled and given so that respondents do not get an idea as to what psychological bias they are answering for. This approach helped in making the individuals respond only to the situation on hand. Likert’s five point scale was used in studying the responses which ranges as follows: 5-strongly agree, 4-agree, 3-neutral, 2-disagree, 1- strongly disagree. This scaling allowed for the standardisation of results, as well as made it easy for the respondents to answer the questionnaire.

3.5 ANALYSIS OF PSYCHOLOGICAL BIASES

Questionnaires were circulated among investors who held demat accounts and traded in shares. First a pilot study was done and necessary changes were made to the questionnaire.

3.5.1 Reliability and Validity

Reliability of the questionnaire was tested through Cronbach alpha score which was 0.776. This was more than the acceptable lower limit of 0.6 (Hair et.al, 2006). For content validity the questionnaire was given to academicians and practitioners.

SPSS was used for the analysis of responses collected through questionnaires.

3.5.2 Univariate Analysis

In order to test the third objective of identifying and measuring psychological biases in individual investors, the weighted mean score and standard deviation of responses to each statement was found. The overall mean score for each of the biases were then found using the mean score of statements. This helped in determining the existence of biases for the sample
respondents. Percentages were used to provide summary statistics of the survey respondents.

To have a better understanding of these psychological biases for the different categories of respondents, the weighted mean scores for each bias with respect to different groups in the demographic profile was found. One way ANOVA, to find whether there are differences among the different groups of respondents with respect to psychological biases was calculated.

3.5.3 Factor Analysis

In order to identify behavioral factors that emerge out of psychological biases measured, exploratory factor analysis was used. Out of the 42 statements used in the study, 28 specific statements which relate to investment biases alone were chosen for factor analysis, since the other statements related to general life situations.

Factor analysis is an interdependence technique with its primary purpose being to define the underlying structure among the variables in the analysis (Hair et al 2006). Thus factor analysis is a multivariate technique that provides for analyzing the structure of interrelationships among large number of variables. It is a useful tool in reducing the number of variables and creating new dimensions within the data.

Factor analysis cannot be done for a sample fewer than 50 observations. However as a general rule, the minimum is to have at least five times as many observations as the number of variables to be analysed and the more acceptable sample size would have a 10:1 ratio.
Assumptions in Factor Analysis

1. A basic assumption of factor analysis is that some underlying structure does exist in the set of selected variables.

2. Another desirable situation is the assumption of some degree of multi collinearity among the variables because the intention here is to identify interrelated sets of variables.

Selecting the Factor Extraction Method

There are basically two methods of extraction namely (1) Common factor analysis (2) Principal Component Analysis.

The selection of one of these methods is based on two criteria (1) the objective of factor analysis and (2) the amount of prior knowledge about the variance in the variables.

Principal Component Analysis (PCA) is used when the objective is to summarise most of the original information into a minimum number of factors for prediction purposes. This method is employed in the present study.

Criteria for the Number of Factors to be Extracted

Three basic criteria are employed to decide the number of factors to be extracted or in other words when to stop factoring.

**Latent Root Criterion:** This is the most commonly used technique. Here only factors having latent roots or eigen values greater than 1 are considered significant. All factors with latent roots less than 1 are considered insignificant and are disregarded.
Percentage of Variance Criterion: Here the purpose is to ensure practical significance for the derived factors by making sure that they explain at least a specified amount of variance. In social sciences, where the information is less precise, it is common to consider a solution that accounts for 60% of the total variance.

Scree Test Criterion: The Scree test is derived by plotting the latent roots against the number of factors in their order of extraction and the shape of the resulting curve is used to evaluate the cut off point. The point at which the curve first begins to straighten out is considered to indicate the maximum number of factors to be extracted.

In the present study all the three criteria are used to retain the number of behavioral factors that emerge out of the psychological biases measured.

Factor Rotation

Unrotated factor solutions achieve the objective of data reduction but it may not provide factor structure that offers the most adequate interpretation of the variables under examination.

Factor rotation is a process where the reference axes of the factors are buried about the origin until some other position is reached. The ultimate effect of rotating the factor matrix is to redistribute the variance from earlier factors to later ones to achieve a simpler, theoretically more meaningful factor pattern.

The commonest case of rotation is an orthogonal factor rotation, in which the axis is maintained at 90 degrees.

Orthogonal Rotation Methods: The rotational methods namely Quartimax, Varimax and Equimax approaches have been developed. Among
the three the most popular approach and the one that is used in the present study is explained in the following paragraph.

The Varimax criterion centers on simplifying the columns of the factor matrix. This method maximises the sum of variances of required loading of the factor matrix. VARIMAX has proved to give a clearer separation of the factors. The VARIMAX rotation has proved successful as an analytical approach to obtaining an orthogonal rotation of factors.

**Significance of Factor Loadings in Case of Interpretation**

1. Factor loading in the range of ± 0.030 to ± 0.40 are considered to meet minimal level for interpretation of the structure.

2. Loadings ± 0.50 or greater are considered practically significant.

3. Loadings ± 0.70 are considered indicative of a well-defined structure.

4. Extremely high loading of ± 0.80 and above are not typical and practically achievable (Hair et.al, 2006)

**3.5.4 Multiple Discriminant Analysis**

In order to find out the behavioral factors which distinguish traders from non-traders multiple discriminant analysis was used. For this purpose investors were divided into three categories based on the frequency of trading. Investors who trade daily and weekly form category I, investors who trade monthly and few times in a year form category II and investors who trade once in a year form category III.
Discriminant analysis is the appropriate statistical technique when the dependent variable is categorical and the independent variables are metric in nature.

**Use of Centroids:** Centroids are group means. Group means are arrived at by averaging the discriminant scores for all the individuals within a particular group.

When the analysis involves two groups there will be two centroids. The centroids indicate the most typical location of any member from a particular group and a comparison of the group centroids shows how far apart the groups are, in terms of the discriminant function.

**Number of Discriminant Functions:** NG-1 functions will be calculated, where NG represents the number of groups. Thus if the dependent variable consists of more than two groups, discriminate analysis will calculate one discriminant function. Each discriminant function will calculate a separate discriminant Z score.

**Steps Involved In Discriminant Analysis**

**Selection of Variables:** Discriminant Analysis requires a single metric dependent measure and one and more metric independent measures to provide differentiation between groups.

**Sample Size:** As a thumb rule, a minimum of five observations for each independent variable is required.

**Division of Sample:** In order to validate results the sample has to be divided into sub-samples. It is essential that each sub-sample be of adequate size to support conditions from results.
Estimating Discriminant Model and Assessing Overall fit: Overall model fit can be assessed using discriminant Z scores. Comparison of the group means (centroids) on the X scores provides one measure of discrimination between groups.

Interpretation of Results: If the discriminant function is statistically significant substantive interpretations can be made. The relative importance of each independent variable in discriminating between the groups can be determined with the following methods.

a. Standardised discriminant weights

b. Discriminant loadings

c. Partial F values

When two or more significant discriminant functions are arrived at they can be rotated to redistribute the variance.

3.6 CONCLUSION

Thus the event study methodology was used for analysis of the secondary data related to market reactions whereas percentages, weighted means, ANOVA, factor analysis and discriminant analysis were used to analyze the primary data relating to investor psychology. In the chapter that follows, the findings and interpretations of the event study methodology are discussed in detail.