CHAPTER-III

Research Design
DESIGN OF THE STUDY

3.1 INTRODUCTION

Design of the study plays a vital role in conducting any study. A good design of the study ensures smooth execution of study. It not only gives direction to the entire study, but also creates a platform for the study. It is defined as the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure. The objective of this chapter is to present a blueprint of the study. This chapter includes the following sections.

- Need for the study
- Research Statement
- Scope of the Study
- Research Objectives
- Sampling Design
- Research Instrument
- Data Collection Method
- Framework of the Model
- Tools and Techniques of Analysis.
3.2 NEED FOR THE STUDY

More and more women are becoming financial decision makers. It is estimated that by 2020, half of the U.S. wealth would be in the hands of women and two-thirds by 2030. The trend is similar in other parts of the world (US Securities Exchange Commission Report 2010). Almost one third of women are financially independent and 23% of them earn more than men (UN statistics 2010).

According to the analysts, money is not solely meant wealth for women. It is security, independence, freedom and the ability to leave a legacy that they wanted most from money. It was earlier perceived that women needed funding to purchase jewellery or get discounts at the beauty salons (Barletta Marti 2001). Later market researchers (Hargreaves Lansdown 2004), (Sinha Pradip, Banerjee Prashant 2004) had proved that, in financial services, a women look for a complete financial plan that would take care of her special investments as well as other financial needs. However, this financial plan depends on the four distinct stages during her life-span, such as a young single woman, married woman, family-centered priorities and single status (Learned Andrea & Johnson Lisa 2001). It is estimated that women usually have less working hours and a longer life-expectancy, which has created a distinct drive for the financial services among women when compared with that of men (US Labor Daprtment Statistics).
According to marketing researchers (Eckel, C., and P. Grossman 2002), women want to keep a larger part of their current earnings as savings and investment, put every penny at its best use and invest conservatively without exposing too much of capital risk.

Women earn only about 80 cents for every dollar earned by men. Because they earn less, women often are unable to invest as much as men. However, in order to make up for other discrepancies in retirement benefits, women may actually need to invest more. Women often leave work to bring up children or care for elderly relatives; they have fewer total working years. On an average, they spend seven years out of the workforce to care for family members (US Labour Department Statistics)

This may mean that women qualify for lower pension benefits. Fewer years in the workforce, fewer years with a single employer, and lower pay are all factors that may contribute to a lower average pension for female retirees. At the same time women on average live longer than men (UN Statistics of women 2011).

The above mentioned facts signify the need for women to invest for their future. A considerable number of studies took place in the recent past comparing men and women investment behavior and many are under impression that equity investment, day-trading and playing the stock market has always been a very male affair. Yet, over the past decade more and more women have joined in. In the last
ten years the percentage women investors trading equity online doubled from 17% to 34% of their total investors, reported by U.K. online broker Selftrade. This trend is not restricted to the one website in U.K. Stockopedia, a site specialising in broker research and stock market analysis, says between 30% and 40% of its current user base is female. "Gone are the days when online investing was completely dominated by men," says Edward Croft, chief executive of Stockopedia.

This trend of women investing on equity oriented securities is not confined only to advanced nations, in a developing nation like India, National Council of Applied Economic Research (NCAER) found in its survey, the number of working women investing on equity oriented securities has doubled in just two years span from 2009 to 2011 and it further stated that the proportion of investment on equity oriented securities as a percentage of total investment on all financial assets.

Given the above backdrop, it is felt that, there is a greater need to study on “investment behavior of working women with a special focus of equity oriented securities.

3.3. STATEMENT OF RESEARCH PROBLEM

The global economy went through one its worst ever recession in the recent past. Financial markets across the world have suffered with lack of liquidity to reduce the ill effects of recession. Coincidentally, it was also the time when
proportion of working women also started growing rapidly and so as their investment capacity (UN Statistics of women 2010).

Hence, it is felt appropriate to do research on the investment behavior of working women. The research process began with a question “is investment behavior of women same as men?” in a general way. After the initial review of literature, it is found that women behave differently than men while investment. This prompted the researcher to conduct an initial survey of working women in the state of Andhra Pradesh in India and also brainstorming session with personal financial planners which revealed that there are numerous internal and external factors affecting working women’s investment decision making. This entire process led this researcher to formulate the final research problem as “to what extent various internal and external factors are affecting women investment behavior?”

3.4 SCOPE OF THE STUDY

Every academic research should serve a purpose of uplifting the society. This study is aimed at motivating the equity investment culture among working women so that, economy of a nation will benefit apart from upliftment of socio-economic status of women. In this study working women with in the age group of 18 years to 60 years are taken as respondents as they form the majority of working women.
As this study is based on equity investment culture and the respondents are drawn from India, only those respondents with minimum annual income of Rupees One Hundred Twenty Thousand are considered, thinking that they are financially capable of making investment.

The scope of the study is extended to the following segments of society.

1. **Working Women**: working women form the nuclease of this study. as their financial needs are increasing every day, they need to focus more than ever on their personal financial planning in general and investment planning in particular. As the equity investment world is becoming more complex with the introduction of a range of equity oriented securities and working women are finding it difficult to spare time on their, investment planning, this study is expected to help them in identifying various factor affecting equity investment so that they are able to invest in more organized way.

2. **Financial Institutions**

Financial institutions play vital role in economy by mobilizing funds from the investors in favor of the companies that are needed to start or expand their business which in turn develop the economies. Certain financial intermediaries like Mutual Fund Companies, Insurance Companies, Pension Funds, Stock Broking agencies, etc., are expected to find this
useful in terms of understanding the needs of working women in a better manner and they may reach to them in a better way.

3. **Professional Advisors**

As the literature suggests, more number of working women are taking the help of professional advisors while making equity investment decision. Profession advisors must know what working women are looking from their investment and how they are behaving and which are the factors influencing their investment behavior. Once the professional advisors are aware of these issues, they can perform their role with much ease.

4. **Government**

Government of any nation has to take care of socio-economic status of its citizens. In the older days, women were suppressed in majority of the nations. With the changes in socio-economic status of women in the recent past, Governments across the globe have to formulate policies to ensure that the emergence of working women will last longer. This will happen when working women are able to decide their future on their own and they are having enough opportunities to invest their funds.

Apart from the above mentioned segments of society, the study also extends to media and academicians.
3.5 **RESEARCH OBJECTIVES**

Based on the research problem and research gap, the following objectives are framed to guide the research process.

1. To gain insight about investment behavior among women with special focus on equity oriented securities

2. To assess the existing investment practices towards equity oriented securities among women.

3. To identify various internal and external variables affecting investment decision of women on equity oriented securities.

4. To classify all the identified variables into different factors and name them in accordance with their underlying nature and dimensions.

5. To ascertain the extent to which these factors affect the equity investment decision of working

3.6 **NATURE OF RESEARCH**

Descriptive research is chosen to attain the research objectives as descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group. Descriptive research also concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation (Kothari C.R. 2002). Most of the social
research comes under this category. Some important features of descriptive research can be presented in the following table.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall design</td>
<td>Rigid design (design must make enough provision for protection against bias and must maximize reliability)</td>
</tr>
<tr>
<td>Sampling design</td>
<td>Probability sampling design (random sampling)</td>
</tr>
<tr>
<td>Statistical design</td>
<td>Pre-planned design for analysis</td>
</tr>
<tr>
<td>Observational design</td>
<td>Structured or well thought out instruments for collection of data</td>
</tr>
<tr>
<td>Operational design</td>
<td>Advanced decisions about operational procedures</td>
</tr>
</tbody>
</table>

### 3.7 METHOD OF DATA COLLECTION

#### 3.7.1. Key Variables and their importance

All descriptive research studies are based around variables. A variable is the characteristic or attribute of an individual, group, educational system, or the environment that is of interest in a research study. Variables can be straightforward and easy to measure, such as gender, age, or course of study. Other variables are more complex, such as socioeconomic status, academic achievement, or attitude toward an object. Characteristics of the environment may also be variables. Therefore, once the general research topic has been
identified, the researcher should identify the key variables of interest. Identifying the key variables is important for the following reasons:

- The key variables provide focus when writing the Introduction section.
- The key variables provide focus to the Methods section.
- The Instrument will measure the key variables. These key variables must be directly measured or manipulated for the research study to be valid.

Hence relevant literature is reviewed and variables influencing women investment on equity oriented securities were identified.

3.7.2. Variables Measuring Instrument

A well constructed structured questionnaire is designed and it is divided into two parts namely part-A and part-B. In part-A, a set of three 5 point dichotomous statements are used to measure each of the variable identified, as it is not possible to measure them directly. Summated scale values of each variable are taken in turn for further data analysis. Interval scale is used in part-A with an objective of collecting metric data that helps the researcher to perform various parametric tests in later stages of the research. Part-B of the questionnaire consists of questions related to the demographic and other information for the purpose of preparing profile of the respondents.

3.7.3. Unidimensionality of Summated Scale
An underlying assumption and essential requirement for creating a summated scale is that the items are unidimensional, meaning that they are strongly associated with each other and represent a single concept. As Factor analysis plays a vital role in the assessment of dimensionality of a set of items by determining the number of factors and loadings on each factor. The test of unidimensionality is that each summated scale should consist of item loading highly on a single factor. Hence Factor analysis is used to check unidimensionality of each variable while constructing summed scale (Hattie, J 1985).

3.8. **SAMPLE DESIGN**

3.8.1. **Type of Universe and Sampling Unit.**

The first step in developing any sample design is to clearly define the set of objects, technically called the Universe, to be studied. The universe can be finite or infinite. In finite universe, the number of items is certain. But in case of an infinite universe the number of items is infinite. In this research, the type of universe consists of those working women who are aged between 18 years to 60 years. As it is not possible to identify the exact number of such working women, the universe is infinite.
A decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual. In this research, the sampling unit is a geographical region i.e. India.

India is selected as the sampling unit due to the following reasons

1. India's **gross domestic product** in purchasing power parity (PPP) terms stood at $4.46 trillion in 2011, marginally higher than Japan's $4.44 trillion, making it the third-biggest economy after the United States and China. (IMF)

2. India is second most populous nation in the world (UN Statistics).

3. Human sex ratio of India in the 15 to 64 age group (which is considered to be economically active group) is 1.07 against 1.05 of world average (World Fact Book 2011).

4. Growth rate of working women is on par with growth rate of world (ILO statistics 2011).

5. Women investors in equity oriented securities category in India account for 38% of total investors (NCEAR-SEBI survey 2011)
3.8.2. Sample Size Determination

The Respondents Population and Sample

The premeditated objectives of the empirical investigation required a reasonably large and representative sample of the targeted respondent population. It is fairly well known from the available facets of the residents of India, that only certain segments of the population are of direct interest from the viewpoint of the ‘objectives of the present study’. As such the focus for collection of data has been on relevant segments of the population. The ‘targeted relevant population segment’ comprises those respondent investors who:

(i) possess reasonable level of income i.e. those women who are earning a minimum of Rupees One hundred twenty thousand per annum, and well employed or engaged in economic activities such as professionals, entrepreneurs, etc.

(ii) are well aware of their social/economic needs and of the investment scenario; and

(iii) have availability of investible funds and are presently engaged in investment activity.

After considering the above points, the sample size is decided as 500 respondents.
3.9. METHOD OF DATA COLLECTION

The Internet is increasingly being used as a medium for social science research. To assess the validity of such efforts, an electronic version self-monitoring questionnaire was placed at a site on the World Wide Web by Tom Buchanan and John L. Smith in 2010. In all, 963 responses were obtained through the Internet and these were compared with those from a group of 224 undergraduates who completed a paper-and-pencil version. Comparison of model fit indices obtained through confirmatory factor analyses indicated that the Internet-mediated version had similar psychometric properties to its conventional equivalent and compared favourably as a measure of self-monitoring.

Internet survey has many advantages also. Some of them are

1. It involves no cost while collecting data
2. It consumes very less time and has high response rate than conventional method Dejana Braithwaite at all, (2003)
3. Unless all the questions are answered, the web system will not process it further. Hence, there will not be any problem of missing data.
4. As the responses are automatically saved in a spread sheet, there will not be any data entry mistakes.

Due to the above mentioned advantages of Internet survey, this researcher created a questionnaire on “google documents”.
After preparing the online questionnaire, the biggest challenge was to get a source list of respondents. For this purpose, the two companies; National Securities Depository Services Limited (NSDL) and Central Depository Services Limited (CDSL), which are authorized to hold the electronic accounts called as Demat accounts of investors in India were contacted to get the email IDs of women investors. Unfortunately, both the companies did not respond for long and one of them CDSL replied later saying that they cannot share the data even for academic research. Thankfully, one stock broking company in India i.e. Reliance Securities Limited has agreed to share the email IDs.

To know why women are not investing on equity oriented securities, some of the non equity investors are also needed as respondents. Hence, some banks were approached to get the email IDs of their women customers. Except two private sector banks in India; ICICI Bank and Axis Bank have agreed to provide access with a rider. They both replied to the request by saying that they cannot give the email IDs as it may create problem to them. But, they said if the questionnaire is sent to their authorities, they will forward that to their women customers chosen randomly. They concluded that under Corporate Social Responsibility, this is the maximum that they can do.

Immediately, the online questionnaire is sent to the women customers of Reliance Securities Limited and also to the authorities of ICICI bank and Axis bank. But, to the dissatisfaction of this researcher, only 78 responses have come
back during two months period of October & November, 2011. Though the number of responses is not even 20% of total expected sample size, they served an important purpose in identifying certain mistakes and later the questionnaire was redesigned after eliminating those mistakes.

After corresponding with the authorities of both the Banks ICICI & Axis and also with the authorities of Reliance Securities on the poor response of the women, it is understood that majority of the women customers of those institutions might not be interested in responding to such a lengthy questionnaire which requires at least an hour to answer them. So it is understood that only those women who are willing to spend time and have genuine interest in responding the questionnaire have to be contacted.

At this juncture, this researcher decided to paste the questionnaire on the wall of social networking website “Facebook” requesting the working women to understand the importance of the research on the topic of equity investment by women and respond. It was well informed in the request that it will be a time consuming exercise. The mobile telephony number and email ID of this researcher is also given for clarification of any doubts.

To the surprise of this researcher in approximately two months duration between
December 2011 to end of January 2012, all the targeted responses have been received on the email of this researcher. It also gave happiness to receive a few phone calls and emails appreciating the efforts of this researcher on such a topic and in fact, some of the respondents have stated that the questionnaire is an eye opener to them to act seriously on their financial and investment planning. So once the desired 500 responses are received, the post is withdrawn from the wall after thanking all the respondents.

3.10. DATA ANALYSIS

As per the research objectives listed in the previous section, the first activity to be performed is to formulate different groups of variables based on their underlying dimensions. For performing this activity Factor Analysis is chosen.

3.10.1. Factor Analysis

Factor analysis is an interdependence technique, whose primary purpose is to define the understanding structure among the variables in the analysis. If anyone employs multivariate techniques, by their very nature, the number of variables increases. Univariate techniques are limited to a single variable, but multivariate techniques can have tens, hundreds, or even thousands of variables. As the researcher adds more and more variables, more and more overlapping (i.e. correlation) may likely to result among the variables. In some instances, such as when one is using multiple measures to overcome measurement error by multivariable measurement, the researcher even strives for correlation among the
variables. As the variables become correlated, the researcher now needs ways in which to manage these variables – grouping highly correlated variables together, labeling or naming the groups, and perhaps even creating a new composite measure that can represent each group of variables.

Factor analysis provides the tools for analyzing the structure of the interrelationships (correlations) among a large number of variables (e.g. test scores, test items, questionnaire responses) by defining sets of variables that are highly interrelated, known as factors. These groups of variables (factors), that are by definition highly intercorrelated, are assumed to represent dimensions within the data. If the researcher is only concerned with reducing the number of variables, then the dimensions can guide in creating new composite measures. On the other hand, if he has a conceptual basis for understanding the relationships between variables, then the dimensions may actually have meaning for what they collectively represent. In the latter case, these dimensions may correspond to concepts that cannot be adequately described by a single measure.

3.10.1. a) Prerequisite of Factor Analysis

Required sample size.

Generally, factor analysis is not performed for a sample of less than 50 observations, and preferably the sample size should be 100 or larger. As a general rule, the minimum is to have at least 5 times as many observations as the number of variables to be analyzed, and the more acceptable sample size would have 10:1
ratio. Some researchers even propose a stringent ratio of 20:1 (Hair et al 2010). In this research, there are 20 variables to be studied and subjectively decided sample size of 500 meets even this stringent required sample size.

**Statistical tests**

In order to justify the usage of factor analysis the researcher must ensure that the data matrix has sufficient correlations. To verify this, the following tests were performed.

Bartlett test of sphericity: It is a statistical test for the presence of correlations among the variables. It provides the statistical significance that the correlation matrix has significant correlations among the variables. The obtained test value must be positive definite.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy: it is used to quantify the degree of intercorrelations among the variables and the appropriateness of Factor analysis. This index ranges from 0 to 1, reaching 1 when each variable is perfectly predicted without errors by the other variables. The measure can be interpreted with the following guidelines.

0.8 or above, meritorious; 0.7 or above middling; 0.6 or above mediocre; 0.5 or above, miserable; and below 0.5, unacceptable (Kaiser H.F1970).
### 3.10.1.b. Factor extraction method

As the primary objective of using factor analysis is to identify the latent dimensions or constructs represented in the original variables and reducing all the variables into a smaller number of factors, “principle component analysis” method of factor extraction is used in this research as it considers the total variance and derives factors that contain small proportion of unique variance and, in some instances, error variance. However, the first few factors do not contain enough unique or error variance to distort the overall factor structure. Specifically, with principle component analysis, unities (values of 1.0) are inserted in the diagonal of the correlation matrix, so that the full variance is brought into the factor matrix (Mulaik S.A. 1990).

#### Criteria for number of factor to extract

In this research, “Latent Root Criterion” is applied to extract the required number of factors. The rationality behind applying this widely used technique is that any individual factor should account for the variance of at least a single variable if it is to be retained for interpretation. With principle component analysis each variable contributes a value of 1 to the total eigenvalue. Thus, only the factors having latent roots or eigenvalues greater than 1 are considered significant. (Eigenvalue or latent root is defined as column sum of squared loadings)
for a factor. It represents the amount of variance accounted for by a factor) (Stewart, D.W. 1981)

3.10.1.c Factor rotation

The un-rotated factor matrix simply helps the researcher in data reduction. But, in order to achieve the research objective of identifying a structure with the help of underlying dimensions of all variables, “Varimax” factor rotation method is used. This criterion centers on simplifying the columns of the factor matrix. With the Varimax rotational approach, the maximum possible simplification is reached if there are 1s and 0s in the column. That is, the Varimax method maximizes the sum of variance of required loadings of the factor matrix.

3.10.1.d. Judging Significance of Factor Loadings

In interpreting factors, a decision must be made regarding the factor loadings for the purpose of ensuring practical significance. Using practical significance as criteria, one can assess the loadings as follows

- Factor loadings in the range of 0.3 to 0.4 are considered to meet the minimum level for interpretation of structure.
- Loadings that are 0.5 and above are practically considered as significant
- Loadings exceeding 0.7 are considered indicative of well defined structure.
These guidelines are applicable when the sample size is 100 or larger. Based on the sample size of the research and its practical significance, only those variables whose factor loadings are 0.5 and above are retained in this study.

3.10.1. e. Reliability of the Structure

It is an assessment of the degree of internal consistency between multiple measurements of a variable or factor. “Cronbach’s Alpha”, the widely used reliability coefficient value is calculated for all the factors extracted. An alpha value of 0.6 is desirable in social science research (Cronbach, L.J. 1951)

3.11. FRAMEWORK OF THE MODEL

3.11.1. Relationship between Behaviour and Intention

Intentions are assumed to capture the motivational factors that influence a behavior; they are indications of how hard people are willing to try, of how much of an effort they are planning to exert, in order to perform the behavior. As a general rule, the stronger the intention to engage in a behavior, the more likely should be its performance (Icek Ajzen 1991).

3.11.2. Theory of planned behavior

In psychology, the theory of planned behavior is a theory about the link between attitudes and behavior. The concept was proposed by Icek Ajzen to improve on the predictive power of the theory of reasoned action by including perceived behavioural control. It is one of the most predictive persuasion
theories. It has been applied to studies of the relations among beliefs, attitudes, behavioral intentions and behaviors in various fields such as advertising, public relations, advertising campaigns and healthcare. The theory states that attitude toward behavior, subjective norms, and perceived behavioral control, together shape an individual's behavioral intentions and behaviors.

**Concepts of key variables**

Behavioral beliefs and attitude toward behavior

- Behavioral belief: an individual's belief about consequences of particular behavior. The concept is based on the subjective probability that the behavior will produce a given outcome.

- Attitude toward behavior: an individual's positive or negative evaluation of self-performance of the particular behavior. The concept is the degree to which performance of the behavior is positively or negatively valued. It is determined by the total set of accessible behavioral beliefs linking the behavior to various outcomes and other attributes.

Normative beliefs and subjective norms

- Normative belief: an individual's perception about the particular behavior, which is influenced by the judgment of significant others (Amjad, N. & Wood, A.M. 2009)
Subjective norm: an individual's perception of social normative pressures, or relevant others' beliefs that he or she should or should not perform such behavior.

Control beliefs and perceived behavioral control

- Perceived behavioral control: an individual's perceived ease or difficulty of performing the particular behavior. It is assumed that perceived behavioral control is determined by the total set of accessible control beliefs.
- Control beliefs: an individual's beliefs about the presence of factors that may facilitate or impede performance of the behavior. The concept of perceived behavioral control is conceptually related to self-efficacy.

Behavioral intention and behavior

- Behavioral intention: an indication of an individual's readiness to perform a given behavior. It is assumed to be an immediate antecedent of behavior.
- It is based on attitude toward the behavior, subjective norm, and perceived behavioral control, with each predictor weighted for its importance in relation to the behavior and population of interest.
- Behavior: an individual's observable response in a given situation with respect to a given target. Behavior is a function of compatible intentions and perceptions of behavioral control in that perceived behavioral control is expected to moderate the effect of intention on behavior, such that a
favorable intention produces the behavior only when perceived behavioral control is strong.

3.11.3. Application of theory

So far, the theory of planned behavior has more than 1200 research bibliographies in academic databases such as Communication & Mass Media Complete, Academic Search Premier, PsycARTICLES, Business Source Premier, PsycINFO, and PsycCRITIQUES.

Another application of the theory of planned behavior is in the field of environmental psychology. Generally speaking, actions that are environmentally friendly carry a positive normative belief. That is to say, sustainable behaviors are widely promoted as positive behaviors. However, although there may be a behavioral intention to practice such behaviors, perceived behavioral control can be hindered by constraints such as a belief that one's behavior will not have any impact (Stern, P.C. 2005).

Philmore Alleyne & Tracey Broome (2010) are the first to use TPB in factors influencing investment decision. The study used the dimensions of Ajzen’s (1991) theory of planned behaviour (attitudes, subjective norms, and perceived behavioural control), and Sitkin and Weingart’s (1995) risk propensity as predictors of investments intentions. The study used a self-administered questionnaire of a sample of business students in an undergraduate institution. It
was found that attitudes, subjective norms, perceived behavioural control, and risk propensity were significant predictors of investment intentions. They also found that risk propensity did not moderate the relationship between the predictors (attitudes, subjective norms, and perceived behavioural control) and the dependent variable, intentions to invest. These findings are consistent with prior research and do show that education in business finance can help to influence investment decisions.

3.11.4. Proposed Model

Based on the concept of theory of planned behavior and the study of Philmore Alleyne & Tracey Broome (2010), the following model is designed for conducting the research.

Figure-3.1- Proposed Model
The proposed model says that actual behavior of an individual is preceded by Investment Intention which is caused by the above mentioned four factors.

3.12. MULTIPLE REGRESSION ANALYSIS

As per the research objectives stated earlier, this research is intended to know to what extent each of the various factors affecting investment decision of working women is to be ascertained. Multiple regression is chosen to attain this objective due to the following reasons.

- There is only one dependent variable called intention to invest on equity oriented securities. This is measured by using a 3 item scale.
- There are four independent variables namely Personality, Influence of Society, Behavioural Control & Demography and Financial Liberty
- All the variables including dependent and independent are in the form of metric data.
- Multiple Regression establishes casual relationship.

Multiple regression provides means of objectively assessing the degree and character of relationship between dependent and independent variables by forming the variate of independent variables and then examining the magnitude, sign and statistical significance of the regression coefficient for each independent variable.

In this manner, the independent variables, in addition to their collective prediction of the dependent variable, may also be considered for their individual
contribution to the variate and its predictions. Interpretation of the variate may rely on any of the three perspectives: the importance of independent variables, the types of relationships found or the interrelationships among the independent variables. The following are the main advantages of performing multiple regression.

- The most direct interpretation of the regression variate is the determination of relative importance of each independent variable in the prediction of the dependent measure.
- In addition to assessing the importance of each variable, multiple regression also affords the researcher a means of assessing the nature of relationship between independent and dependent variables.
- Finally multiple regression provides insight into the relationships among independent variables in their predictions of the dependent measure.

3.12.1. Selection of dependent and independent variables

The dependent variable “intention to invest” is selected after careful study of earlier research in the area of investment behavior (Philmore Alleyne & Tracey Broome 2010) and theory of planned behavior (Icek Ajzen 1991). The set of four independent variables have been attained after performing Exploratory Factor Analysis in the previous stage.
3.12.2. Required sample size for Multiple Regression

The sample size used in the multiple regression analysis is the single most influential element under the control of the researcher in designing the analysis. The effects of sample size are seen most directly in the generalisability of the result. There is no mathematical formula available to determine the ratio of optimal number of variables to the sample size for carrying on the multiple regression analysis. However, it has been observed empirically by various researchers that by changing the sample sizes, the ratio of 1:50 results in the stabilization of R square and adjusted R square in the case of multiple regression analysis (Arya and Pal, 2001). The subjectively set sample size of 500 is meeting this requirement very comfortably.

3.12.3. Assumptions in Multiple Regression Analysis

Multivariate techniques and their univariate counterparts are all based on a fundamental set of assumptions representing the requirements of the underlying statistical theory. The need to test statistical assumptions is increased in multivariate applications because of two characteristics of multivariate analysis.

First, the complexity of relationships, owing to the typical use of large number of variables, makes the potential distortions and biases more potent when the assumptions are violated, particularly when the violations compound to become even more detrimental than if considered separately.
Second, the complexity of the analysis and results may mask the indications of assumption violations apparent in the univariate analysis. In almost all instances, multivariate procedures will estimate multivariate model and produce results even the assumptions are severely violated. Thus, the researcher must be aware of any assumption violations and implications they may have for the estimation process or the interpretation of the results.

Although many assumptions or requirements come into picture in one or more of the multivariate techniques, four of them potentially affect every univariate and multivariate statistical technique. They are,

1. Normality
2. Linearity
3. Homoscedasticity
4. Multicollinearity

3.12.3. a. Normality

The most fundamental assumption in multivariate analysis is normality, referring to the shape of the data distribution for an individual metric variable and its correspondence to the normal distribution, the benchmark for statistical methods. If a variation from normal distribution is sufficiently large, all resulting statistical tests are invalid, because normality is required to perform “F and t
The severity of nonnormality is based on two dimensions: the shape of offending distribution and the sample size.

**Graphical analysis of Normality**

**Histogram**

The simplest diagnostic test for normality is a visual check of the histogram that compares the observed data values with a distribution approximating the normal distribution. For all independent variables and dependent variables histograms are drawn by superimposing the normal distribution curve on the histogram.

Although appealing because of its simplicity, this method is problematic for small sample sizes, where construction of histogram is difficult. In this research, as the sample size is considerably large at 500, one can trust the histogram. But, as mentioned earlier, normality is the most crucial assumption, all other possible and available tests also have been performed.

**Normal Probability Plot (P-P Plot)**

There is another useful graph that one can inspect to see if a distribution is normal called a p-p (probability-probability) plot. This graph plots the cumulative probability of a variable against cumulative probability of a particular distribution. In this case it is fitted against normal distribution. It means that the data are ranked and sorted. Then for each rank the corresponding z-score is calculated. This is
expected value that the score should have in a normal distribution. Next the score itself is converted to a z-score. The actual z score is plotted against the expected z-score. If the data are normally distributed then the actual z-score will be same as the expected z-score and the researcher gets an approximately a straight diagonal line. If the values fall on the diagonal of the plot then the variable is normally distributed, but, deviations from the diagonal show deviations from normality. In this research, p-p plots for normality for all independent and dependent variables have been drawn to ensure normality.

**Stem-and-Leaf Plots**

Stem-and-leaf plot was introduced by the statistician John Tukey (1977) in *Exploratory Data Analysis*. Stem-and-leaf plots provide information about the frequency of a quantitative variable’s values by incorporating the actual values of the distribution. These plots are composed of three main components.

On the far left side of plot one can find the frequency with which a particular value (the one shown for that row) occurred. In the center of the figure is the “stem” and the far right portion is the “leaf.” The stem is the base value that which we combine with the leaf portion to derive the full value.

Stem-and-leaf plots ordinarily combine a range of individual values under a single stem. Depending on how tightly the scores are grouped, one can have either a finer or more wider picture of the distribution. By observing the distribution of
“leaves,” researchers can quickly assess the general shape of the distribution; that is, they can form an impression as to whether it is normal, positively skewed (scores are more concentrated toward the low end of the distribution), or negatively skewed (scores are more concentrated toward the high end of the distribution). It is also possible to generally see whether its kurtosis is more positive (a peaked distribution among the middle values) or more negative (a relatively flat distribution) than a normal curve. Stem and leaf picture are drawn for all variables included in multiple regression analysis of this research to confirm normality of the data.

**Statistical tests of Normality**

Though, the above mentioned graphical tests and observations are giving satisfactory results, but still it is better to perform statistical tests for ensuring normality. A simple test is a rule-of-thumb based on skewness and kurtosis values.

**Skewness**

Skew is the tilt (or lack of it) in a distribution. The more common type is right skew, where the tail points to the right. Less common is left skew, where the tail is points left. A common rule-of-thumb test for normality is to run descriptive statistics to get skewness and kurtosis, and then divide these by the standard errors. Skew should be within the +2 to -2 range when the data are normally distributed. Some authors use +1 to -1 as a more stringent criterion when normality is critical.
Kurtosis

Kurtosis is the peakedness of a distribution. A common rule-of-thumb test for normality is to run descriptive statistics to get skewness and kurtosis, then use the criterion that kurtosis should be within the +2 to -2 range when the data are normally distributed (a few authors use the more lenient +3 to -3, while other authors use +1 to -1 as a more stringent criterion when normality is critical). Negative kurtosis indicates too many cases in the tails of the distribution. Positive kurtosis indicates too few cases in the tails. Note that the origin in computing kurtosis for a normal distribution is 3 and a few statistical packages center on 3, but the foregoing discussion assumes that 3 has been subtracted to center on 0, as is done in SPSS and LISREL. Some authors are under opinion that the acceptable range of both Skeness and Kurtosis is that their values should fall under maximum of 3 times of their respective standard error values. Skewness, Kurtosis and their respective standard error values are calculated for all variables.

3.12.3. b. Linearity

An implicit assumption of all multivariate techniques are based on correlational measures of associations, including multiple regression is linearity. Because correlations represent only the linear association between variables, nonlinear effects will not be represented in the correlation values. This omission results in an under estimation of actual strength of the relationship. It is always
prudent to examine all relationships to identify any departures from linearity that may affect the correlation.

**Graphical Method**

The use of bivariate scatterplots is the most typical way of assessing linearity between two variables. Variables that are both normally distributed and linearly related to each other will produce scatterplots that are oval shaped or elliptical. If one of the variables is not normally distributed, linearity will not be achieved. We can recognize this situation because the resulting scatterplot will be nonelliptical (Tabachnick & Fidell, 2001b).

However, there is a downside to running a plethora of bivariate scatterplots, as Tabachnick and Fidell (2001b) aptly note: “Assessing linearity through bivariate scatterplots is reminiscent of reading tea leaves, especially with small samples. And there are many cups of tea if there are several variables and all possible pairs are examined” As the sample size in this research is 500, bivariate scatterplots for independent variables against the dependent variable have been prepared and linearity for all of them are established.

**3.12.3. c. Homoscedasticity**

It refers to the assumption that dependent variables exhibit equal level of variance across all independent variables. Homoscedasticity is desirable because the variance of dependent variable being explained in the dependence relationship
should not be concentrated in only a limited range of the independent values. In most situations researchers may find many different values of the dependent variable for each value of the independent variable. For this relationship to be fully captured, the variance of the dependence values must be relatively equal at each value of the predictor variable.

**Levene's test of homogeneity of variances**

Levene's test of homogeneity of variance, which is the most common test, tests the assumption that each group (category) of one or more categorical independent variables has the same variance on an interval dependent. If the Levene statistic is significant at the .05 level or greater, the researcher rejects the null hypothesis that the groups have equal variances. The Levene test is robust in the face of departures from normality. To verify the assumption of homoscedasticity, both graphical tests as well as Levene’s tests have been performed for the dependent variable across all four independent variables.

**3.12.3. d. Multicollinearity**

It is assumed in regression that each predicted value is independent, which means that the predicted value is not related to any other prediction; that is, they are not sequenced by any variable. Multicollinearity occurs when any single variable is highly correlated with a set of other independent variables. This leads to duplication of prediction and hence the results of multiple regression analysis cannot be accepted. This is a problem of data but not of model specification. The
ideal situation for a researcher would be to have high correlation between set of independent variables and dependent variable and at the same time very little correlations among the set of independent variables. Hence the task of researcher would be as follows

- Assess the degree of multicollinearity
- Determine the impact of results
- Apply the necessary remedied if needed

**Identifying multicollinearity**

The simplest and most obvious means of identifying collinearity is an examination of correlation matrix for the independent variables. The presence of high correlations (generally 0.9 or greater) is the first indication of substantial collinearity. Lack of high correlations values however, does not ensure lack of collinearity. Collinearity may be due to the compound effect of two or more independent variables which is termed as multicollinearity.

To assess multicollinearity, the researcher needs a measure expressing the degree to which each individual independent variable is explained by the set of other independent variable. In simple terms, each independent variable becomes dependent variable and it is regressed against the remaining independent variables. The two most common measure for assessing both pairwise and multiple variable collinearity are tolerance and its reverse Variance Inflation Factor (VIF).
Tolerance

It is a direct measure of multicollinearity which is defined as the amount of variability of the selected independent variable which is not explained by the other independent variables. Mathematically it is defined as 1 - R square. Tolerance value should be high, which means small degree of multicollinearity i.e. the other independent variables do not collectively have any substantial amount of shared variance.

Variance Inflation Factor (VIF)

It is calculated simply as the reverse of tolerance value. VIF gets its name by the fact that the square root of VIF is the degree to which the standard error has been increased due to multicollinearity. The rules of thumb for the VIF are as follows:

- VIF < 3: not a problem
- VIF > 3; potential problem
- VIF > 5; very likely problem
- VIF > 10; definitely problem

To verify the absence of multicollinearity Tolerance and VIF values of all independent variables are calculated.

3.12.4. Selection of estimation technique

In most instances of multiple regression, the researcher has a number of possible independent variables from which to choose for inclusion in the regression equation. Sometimes the set of independent variables is exactly
specified and the regression model is essentially used in confirmatory approach. In other instances, the researcher may use the estimation technique to pick and choose among set of independent variables with either sequential search methods or combinational process. One such sequential method is stepwise estimation method.

**Stepwise estimation method**

This is the most popular sequential method. This approach enables the researcher to examine the contribution of each of the independent variable to the regression model. Each variable is considered for inclusion prior to the developing the equation. The independent variable with the greatest contribution is added first. Independent variables are then selected for inclusion based on their incremental contribution over the variables already in the equation. In this research, this selection method is adopted to select independent variables and enough care is taken on the incremental contribution to the predicting capacity of the overall equation.

3.12.5. Examining the model fit

Finally the model analyzed by looking at the R square and adjusted R square values and R square change values. Descriptive statistics are used to know the predicting capacity of each of the independent variable for this purpose the standardized beta coefficients of respective are used. The same are also used in formulating regression equation.
3.13. CHAPTERISATION

The research is divided into six chapters as follows

- **Chapter-I**: This is the introductory chapter to the research. This chapter introduces the concept of investment. Various investment alternatives available to investors are discussed in this chapter. The role of investment in general and equity investment in specific in a nation’s economic development are presented. Further, the role working women in economy and their investment practices are also discussed.

- **Chapter-II**: This chapter is completely dedicated to review of literature with an objective of identifying various internal and external factors affecting equity investment decision of working women. At the end of this chapter, all such factors are listed along with the names of contributor authors.

- **Chapter-III**: This chapter consists of comprehensive research design that includes need for the research and the objectives of research. This chapter also explains the entire research sampling procedure and methodology. The Research Model and Techniques of analysis are explained in this chapter.

- **Chapter-IV**: This chapter comprehensively describes the profile of respondents i.e. working women in India. All the important demographic, sociographic and psychographic variable are thoroughly analyzed. The
investment behavior of working women is completely examined and important inferences are drawn.

- **Chapter-V**: This chapter deals with data analysis. Exploratory factor analysis is used in this chapter to determine the underlying structure of all the variables and this has set precedence for further analysis. Casual relationship between set of independent variables and dependent variable is established in this chapter with the help of multiple regression analysis.

- **Chapter-VI**: This is the conclusive chapter of the research in which all important findings related to investment behavior of working women are incorporated. Along with findings of the research, suggestions for motivating working women to invest on equity oriented securities are also mentioned in this chapter.