CHAPTER - III
RESEARCH METHODOLOGY

3.1. Introduction

This chapter exposes the research design of the study as well as the methods used to collect and analyze data. It starts with discussion, choice of research design by comparing with other types. In connection with respondents’ selection, using sampling technique and the aim to represents the sample in order to generalize for the whole population. In addition, data collection methods are also discussed, followed by the explanation of questionnaire design and the measurements. Especially, this chapter shows how the analysis is carried out once findings are obtained by using AMOS software and statistical techniques including Descriptive Statistics, Chi-Square test, ANOVA (Analysis of Variance), Multiple Regression Analysis, Confirmative Factor Analysis, Cronbach’s Alpha test, and Structural Equation Modeling (SEM).

3.2. Research Design

Research design provides the framework for data collection and analysis (Ghauri & Gronhaug, 2010, p.54; Bryman & Bell, 2007, p.40). Research design is needed because it facilitates the smooth sailing of the various research operations, thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money. Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in the analysis, Research design, in fact, has a greater bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire structure of the research work.

In order to understand the common behaviours of individual investors, Case Study or Experimental or Longitudinal design are not suitable but Descriptive Research Design is most
suitable. More specifically, experimental design is often used for examining the relationship between variables. Experiments tend to be used in order to explore and explain a specific issue. In experimental research, two groups should be established, one is experimental group and one is control group to compare the difference between these two groups (Saunders et al., 2009, p.142). Case study infers the analysis of one single case (Collis & Hussey, 2009, p.82); and longitudinal design is employed to examine the changes and provide the casual influences over time (Collis & Hussey, 2009, p.78), whereas this research needs to study a relative large sample size at one single time. Thus, Descriptive Research Design is preferred within this study. Most of the social research comes under this category the Descriptive Research Study concerned with describing the characteristics of a particular individual, or of a group. This study concerned with specific predication, with narration of facts and characteristics concerning individual and group. The research design must make enough provision for protection against bias and maximise reliability.

Descriptive Research Design feature is relevant to this study, the first because it fits the nature of this study to describe a common trend of investors’ behaviours rather than one specific case, and the second because data in this study has not been collected in stages but carried out in a single time period.

3.3. Method of Data Collection

Self-completion questionnaire seems to be one of the most common methods of quantitative researches. With a self-completion questionnaire, respondents answer questions by completing the questionnaire themselves. This method is chosen for some reasons. The first reason is that as the research questions are defined clearly, questionnaire is the best choice to have standardized data, which is easily to process, and analyze. Moreover, it is cheaper than other methods. Questionnaire also is more convenient for respondents in case they need to provide some sensitive information, in other words; they tend to be more honest
than in an interview (Bryman & Bell, 2007, p.242). Each question is contributing to each research objectives. After designing the questionnaire, opinions from experts (academicians) were gathered and necessary corrections were carried out. Before collecting data, a pilot study was conducted in order to counter check the questionnaire. During the pilot study, problems faced by the respondents in filling up questionnaire were identified and necessary corrections were incorporated and final draft of the questionnaire was framed.

The respondents were met directly over different places of Cuddalore District. The respondents were asked to tick the appropriate choices on the questionnaire instantly. These collected data were analyzed and its quality was determined by using the Cronbach’s alpha. This questionnaire comprises of 16 part of questions includes demographic variables..

3.4. Sampling Area

The Cuddalore district has an area of 3,564 km². It is bounded on the north by Viluppuram District, on the east by the Bay of Bengal, on the south by Nagapattinam District, and on the west by Perambalur District. The district is drained by Gadilam and Pennaiyar rivers in the north, Vellar and Kollidam River (Coleroon) in the south.

According to 2011 census, Cuddalore district had a population of 2,605,914 with a sex-ratio of 987 females for every 1,000 males, much above the national average of 929. A total of 279,950 were under the age of six, constituting 147,644 males and 132,306 females. The average literacy of the district was 69.66%, compared to the national average of 72.99%. The district had a total of 635,578 households. There were a total of 1,169,880 workers, comprising 136,035 cultivators, 325,599 main agricultural labourers, 19,151 in household industries, 356,486 other workers, 332,609 marginal workers, 29,135 marginal cultivators, 213,813 marginal agricultural labourers, 12,876 marginal workers in household industries and 76,785 other marginal workers. The district has a population density of 702
inhabitants per square kilometre (1,820/sq mi). In Cuddalore District, there are 8 taluks, 13 Blocks, 5 Municipalities and 18 Town Panchayats.
3.5. Sample Size Determination

The formula for computing ‘n’, the sample size required to do the study

\[ n = \left( \frac{Z \times SD}{E} \right)^2 \]

Z : The ‘Z’ value represents the Z score from the standard normal distribution for the confidence level desired, in the study 95% confidence level would indicate (from a standard normal distribution for a two-sided probability value of 0.95) a Z score of 1.96

SD: Population Standard Division (SD) of the concerned variable, a very small sample taken (Pilot Sample) as a test the standard division occurs as follows

E: It is called tolerable error in estimating the variable in questions. The tolerable error is expressed in the same units as the variable being measured or estimated by the study. Only the researcher can decide the error. In this study decided the error, based on the scale value 1 to 5, tolerate in the estimates of the study is 0.3. As for as concern the study calculate the sample size in avenue wise.

Bank & Post Office Investment Avenue

\[ n = \left( \frac{1.96 \times 1.30}{0.3} \right)^2 = 72.13 = 75 \text{ (approx)} \]

Insurance Investment Avenue

\[ n = \left( \frac{1.96 \times 1.27}{0.3} \right)^2 = 68.84 = 75 \text{ (approx)} \]

Shares Investment Avenue

\[ n = \left( \frac{1.96 \times 1.31}{0.3} \right)^2 = 73.17 = 75 \text{ (approx)} \]

Mutual Funds Investment Avenue

\[ n = \left( \frac{1.96 \times 1.32}{0.3} \right)^2 = 74.37 = 75 \text{ (approx)} \]
Bonds Investment Avenue

\[ n = \left( \frac{Z \cdot S_D}{E} \right)^2 = \left[ \frac{1.96 \cdot 1.29}{0.3} \right]^2 = 71.03 = 75 \text{ (approx)} \]

Gold & Silver Investment Avenue

\[ n = \left( \frac{Z \cdot S_D}{E} \right)^2 = \left[ \frac{1.96 \cdot 1.26}{0.3} \right]^2 = 67.76 = 75 \text{ (approx)} \]

Real Estates Investment Avenue

\[ n = \left( \frac{Z \cdot S_D}{E} \right)^2 = \left[ \frac{1.96 \cdot 1.31}{0.3} \right]^2 = 73.24 = 75 \text{ (approx)} \]

Each of which investment avenue has derived the sample size is 75; all put together 525 samples are adequate the study.

### 3.6. Sampling Technique

As the research aims at exploring the Investors behavioural factors at the Cuddalore district area, a relative large sample size is recommended. The larger sample size is more representative and more reliable result (Saunders et al., 2009, p.219). The sample size consists of 525 respondents from Cuddalore district area, regarding this study the Multistage sampling method is chosen, as it is the best technique to get the highest rate of response it would help to save time and money.

Multistage sampling refers to sampling plans where the sampling is carried out in stages using smaller and smaller sampling units at each stage. In a two-stage sampling design, a sample of primary units is selected and then a sample of secondary units is selected within each primary unit. This handout outlines the development of estimators under the general
setting of two-stage sampling, considers the allocation question under the setting of equal
sized primary and secondary units.

Multi-stage sampling represents a more complicated form of cluster sampling in
which larger clusters are further subdivided into smaller, more targeted groupings for the
purposes of surveying. Despite its name, multi-stage sampling can in fact be easier to
implement and can create a more representative sample of the population than a single
sampling technique. Particularly in cases where a general sampling frame requires
preliminary construction, multi-stage sampling can help reduce costs of large-scale survey
research and limit the aspects of a population which needs to be included within the frame for
sampling.

According Iyoke et al. (2006) Researchers used a multi-stage sampling design to
survey teachers in Enugu, Nigeria, in order to examine whether socio-demographic
characteristics determine teachers’ attitudes towards adolescent sexuality education. First-
stage sampling included a simple random sample to select 20 secondary schools in the region.
The second stage of sampling selected 13 teachers from each of these schools, who were then
administered questionnaires.

The study target is to find out the Investors Investment behaviours in Cuddalore
district. In this study adapted multistage sampling method. In Cuddalore district at the first
level 3 taluks randomly selected Cuddalore, Chidambaram, and Virudhachalam each taluks
has 200 samples. In each taluks 4-town panchayat were randomly selected. In each town
panchayat 50 respondents who were willing to participate were selected based on that ‘n’ was
found to be 525 samples. For safety, 600 questionnaires were distributed and the completed
questionnaires were 537.
3.7. Design of Measurements and Questionnaire

The questionnaire is divided into three parts: the first part represents respondents’ personal information (Age, Gender, Marital status, Education, Occupation, Family Dependent, Income and Residential Location). The second part represents respondents’ behavioural factors (Preferred reasons, Information source, Investment objectives, Preference satisfaction, Risk taking ability and Level of knowledge) the third part represents that investors’ investment decision behaviour.

<table>
<thead>
<tr>
<th>Research Segments</th>
<th>Content of Research</th>
<th>Question No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investors’ Personal information (Demography)</td>
<td>Age, Gender, Marital status, Education, Occupation, Family Dependent, Income and Location</td>
<td>Question No-2 to 9</td>
</tr>
<tr>
<td>Investors’ Investment behavioural factors</td>
<td>Preferred reasons, Information source, Investment objectives, Preference satisfaction, Risk taking ability, Level of knowledge</td>
<td>Question No-10 to 12 &amp; 14 to 16</td>
</tr>
<tr>
<td>Investors’ investment decision behaviour</td>
<td>Decision Making Scale</td>
<td>Question No-13 in the subdivision of (13) -3,7,11,14,19 and 24 (13) -1,6,21,25 and 31 (13) -4,10,15,23 and 26 (13) -9,12,17,27 and 30 (13) -5,8,18,22 and 28 (13) -2,13,16,20 and 29</td>
</tr>
</tbody>
</table>

This study is based on the theories of Flinters’ Decision Making Questionnaires (DMQ). Cited many authors’ literature review is synthesizing a set of questions related to Decision behavioural factors influencing investment decisions and investment Preference. In these parts, the 5-point Likert scales, which are rating scales widely used for asking respondents’ opinions and attitudes (Fisher, 2010, p.214), are utilized to ask the individual investors to evaluate the degrees of their agreement with the impacts of decision behavioural
factors on their investment decision as well as with the statements of investment Preference. The five points in the scale are respectively from 5 to 1: Strongly Agree to Strongly Disagree

3.8. Hypothesis of the Study

1. There is no significant association between Investors’ Demography profile and Investment Avenues

2. There is no significant association between Investors’ information sources and various Investment avenues

3. There is no significance difference between the Investors’ Demography profile and collecting information source

4. There is no significant relationship between the decision-making behaviour variables and respective dimensions

5. There is no significant relationship between the decision-making behaviour dimensions and Investors Investment preference

3.9. Data process and analysis

The collected data are processed and analyzed by SPSS and AMOS soft-ware. At first, conduct pilot study conformation of the sample size and reliable and validity of constructs. Then, statistical techniques are used for the data to achieve the research objectives, include Descriptive Statistics, Confirmative Factor Analysis, Cronbach’s Alpha test, and Structural Equation Modeling (SEM).
**Descriptive Statistics**: Descriptive Statistics (Mean, Variance, Critical ratio (C.R), Standard Error (S.E)) are used to describe respondents’ personal information. Descriptive statistics are also used to describe the influence level of behavioural variables on the investment decision making in various Investments.

**Confirmative Factor Analysis**: is a common name of multivariable statistical methods, which aim at defining the core structure in a matrix of data. It helps to analyze the structure of correlations among many variables by identifying a set of core dimensions, called factors (Ghauri & Gronhaug, 2010, p.189). In factor analysis, variables (or items) of the questionnaire are included in homogeneous domains which represent the similar characteristics (O’Brien, 2007, p.143). There are two main types of factor analysis: EFA (Exploratory Factor Analysis) and CFA (Confirmatory Factor Analysis). EFA is the more popular form of factor analysis that attempts to explore the underlying structure of a fairly large number of variables. Whereas, CFA plays a role in confirming the compatibility between the numbers of factors and it’s extracted by the analysis process and those formed by pre established theories. (Liu & Salvend, 2009, p.506).

**Cronbach’s Alpha Test**: is used to test the internal consistency reliability of measurements, which are in formats of continuous variables (for example, 5-point Likert measurements). It includes a statistical summary that describes the consistency of a specific sample of respondents across a set of questions or variables. In the other words, it can help to estimate the reliability of participants’ responses to the measurements (Helms, Henze, Sass & Mifsud, 2006, p.633). Cronbach’s alpha is usually used in social and behavioural researches as an indicator of reliability (Liu, Wu & Zumbo, 2010, p.5). As such, the Cronbach’s alpha is totally suitable for this research because the questionnaire consists of 5-point Likert measurements and the research is in behavioural finance. This research uses Cronbach’s
alpha to test the reliability of the measurements included in the factors that are formed after
the factor analysis.

The Overall reliability for Investors investment decision behaviour variables’
Cronbach’s Alpha is 0.773, and Investors’ Investment behavioural factors Cronbach’s Alpha
is 0.860. Nunnally (1978, p.245) suggests that Cronbach’s alpha should be at least 0.7 to
make sure that the measurements are reliable. However, many statisticians believe that it can
be acceptable if the Cronbach’s alpha is over 0.6 (Shelby, 2011, p.143).

**Structural Equation Modeling (SEM):** is described as the combination of CFA
(Confirmatory Factor Analysis) and multiple regressions. SEM explores the likelihood of
Correlations among the latent variables and includes two parts: (1) a measurement model
(Fundamental of the CFA) and (2) a structural model (the multi-regression model) (Schreiber
et al., 2006, p.325). In this research, SEM is used to confirm which Decision behavioural
factors (formed by the earlier steps of CFA) have the impacts on investment preference of
individual investors as well as estimate the regression weights among them.

SEM is done by the support of AMOS software. The overall model fit of SEM is
determined by some indexes. Asberg, Bowers, Renk and McKinney (2008, p. 491) suggest
that the model is satisfactory with squared error of approximation (RMSEA) less than or
equal to .10, the comparative fit index (CFI) greater than or equal to 0.90, and the
parsimonious fit index (PFI) greater than or equal to 0.60. Schreiber et al. (2006, p.330)
indicate a full set of criteria for an accepted SEM, which is mentioned in the following table
<table>
<thead>
<tr>
<th>Sl. No</th>
<th>Measures of fit</th>
<th>Acceptable level for good fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Chi-square (χ²) at p 0.05</td>
<td>Significant &lt; 0.05</td>
</tr>
<tr>
<td>2</td>
<td>Degree of freedom (d.f)</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>Comparative fit index (CFI)</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td>4</td>
<td>Bentler – Bonett Index or Normed Fit Index (NFI)</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td>5</td>
<td>Root mean squared error of approximation (RMSEA)</td>
<td>&lt; 0.08</td>
</tr>
<tr>
<td>6</td>
<td>RMR</td>
<td>Smaller, the better</td>
</tr>
<tr>
<td>7</td>
<td>GFI</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td>8</td>
<td>AGFI</td>
<td>&gt; 0.90</td>
</tr>
<tr>
<td>9</td>
<td>Non Centrality Parameter (NCP)</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>Non Centrality Parameter, Lower boundary (NCPL90)</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>Non Centrality Parameter, Upper boundary (NCPHI90)</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Parsimony adjusted NFI (PNFI)</td>
<td>&gt; 0.60</td>
</tr>
<tr>
<td>13</td>
<td>Parsimony adjusted CFI (PCFI)</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>Minimum value of Discrepancy (FMIN)</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>Lower Limit of FMIN (LO90)</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>Upper limit of FMIN (HI90)</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>Browne- Cudeck Criterion (BCC)</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>ECVI</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>LO 90</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>HI 90</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>MECVI</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>HOELTER .05</td>
<td>&lt;= 75 Poor fit</td>
</tr>
<tr>
<td>23</td>
<td>HOELTER .01</td>
<td>At least 200</td>
</tr>
</tbody>
</table>