Chapter 4

Database and Methodology

This chapter presents detailed idea about how the research is carried out. This chapter describes the research design used, questionnaire development, universe of study, sampling process, data collection and distribution of sample. The chapter ends with an explanation of the statistical techniques used for data analysis.

4.1 Research Design

“A research design is the arrangement of conditions for collection and analysis of data” (Nargundkar 2003; Malhotra and Dash 2010). In the words of Malhotra 2007, “It is a framework or blueprint for conducting the marketing research and lays down the foundation for conducting the research project as it specifies the details of the procedures necessary for obtaining the information needed and/or solve marketing research plans”. An appropriate research design ensures that study (1) will be relevant to the problem and (2) will use economical procedures (Mouton and Marais 1996; Churchill Jr. 2001). According to De Vaus (2001) the purpose of research design is to make certain that the support obtained enables to answer the initial questions as unequivocally as possible.

According Malhotra and Dash 2010; Chawla and Sondhi 2011, “Research designs can be classified into two categories: exploratory and conclusive”. Exploratory research design is used to elucidate thoughts and opinions about the research problem or the respondent population, or to provide insights on how to do more causal research (Malhotra and Dash 2010). In exploratory research design emphasizes on gaining new ideas, background information and insights which are helpful in describing and clarifying the problem more precisely (Churchill Jr. 2001). It is used to generate the hypotheses (Wrenn, Stevens and Louden 2002). “Conclusive research design is used to assist the decision maker in determining, evaluating and selecting the best course of action to take in a given situation” (Malhotra 2007). It is further categorized as descriptive research and
causal research. “Descriptive research is concerned with determining the frequency with which something occurs or the relationship between two variables and is guided by an initial hypothesis” (Churchill Jr. 2001). This type of research is effective in identifying variation between the variable (Wrenn, Stevens and Louden 2002). Descriptive studies are either cross-sectional or longitudinal. “A cross-sectional design involves the collection of information from any given sample of population at a single point of time where as in longitudinal designs; a fixed sample of population elements is measured repeatedly on the same variables” (Malhotra 2007). Causal research is concerned with determining cause-and-effect relationships (Churchill Jr. 2001).

The present study is a cross-sectional study used to investigate relationship among determinants of customer citizenship behaviour. A causal research design is appropriate to study the effect of customer-based corporate reputation, commitment, affective loyalty, intentional loyalty, perceived risk (social and temporal) and customer participation in service delivery on development of positive customer behaviour termed as customer citizenship behaviour and extra-role behaviour.

4.2 Sampling Methodology

4.2.1 Sample Population

The population for the study is individuals visiting supermarkets for shopping and for entertainment purpose. Supermarkets have been taken because this is one of the most important formats that provide customers variety of household and consumer goods at low price under one roof. This retail format is used by people of all categories to satisfy their daily needs. Further, all respondents should have been visiting the supermarkets for more than 6 months. This experience is satisfactory to establish reliable and sufficient perception regarding the extra-role behaviour and citizenship behaviour towards the other fellow customers, employees and company.

4.2.2 Sampling Unit

The sampling unit selected for this research was the customers who visit supermarkets for their daily needs. It is newly established market phenomenon in our economy which came into
existence after liberalization. In Indian retail industry, supermarkets is one of the flourishing format of organized retail sector. It provides variety of goods and services under one roof and cater to need of all categories of all by focusing on customer needs. It is one of the fastest growing market and ranked in the top five retail markets in the world (Mckinsey & Company 2005; CCI 2012).

4.2.3 Sampling Technique

“The selection of a sampling method is based on the various factors like time and money, desired accuracy level, nature of the research question and the data gathering method” (De Vaus 2002). In the study, we first used judgmental sampling in selecting the target population of customers visiting supermarkets. Then convenience sampling was used for selection of customers from supermarkets from neighbouring cities (Chandigarh, Ludhiana, Jalandhar and Amritsar) and the region of the national capital (Delhi, Gurgaon, Noida and Ghaziabad). Selection of customers from these cities ensured a varied sample unit as smaller cities like Amritsar and metros like Delhi were both included so to encompass all segment of customers and bring uniformity in decision making. Finally we used quota sampling to select appropriate quotas of males and females both so as to average out the discrepancy arising due to gender type.

Convenience samples are the least costly and less time consuming of all sampling techniques. Convenience sampling has been commonly used in services marketing research (Erramilli and Rao 1990; and Ueltschy et al. 2007; Deng et al. 2010; Aref 2011; Bartikowski, Walsh and Beatty 2011). According to Zikmund and Babin (2009) “convenience samples are best for exploratory research when additional research will subsequently be conducted with probability sample”. Moreover limited time and financial constraints necessitates the use of convenience sampling. Another reason for using convenience sampling was supermarkets are common phenomenon in developing countries like, India. According to Rodriguez et al., 2002, “Economies of scale, tracking customer needs through customer relationship management, tight inventory control, availability of a broad selection of goods under a single roof at relatively low price has contributed towards the rapid growth of supermarkets”. For this we
made a selection of neighbouring major cities (Ludhiana, Jalandhar and Amritsar). Customers of different age and monthly income ensured a varied sample unit and encompass all segments to bring uniformity in the final decision making.

### 4.2.4 Sample Size

Main things that determine the sample size in any study are the nature and type of data analysis. In the first stage in which the primary purpose of analysis was pre-testing, we collected data from 250 respondents. We followed the rule of thumb recommended by Nunnally and Berntsein (1994) of having 5 subjects per item. As there were 46 items, so sample of 250 respondents was considered to be suitable for the pretesting study (46*5=230).

Sample size determination is also affected by the data analysis techniques. “Structural Equation Modeling (SEM) requires use of larger sample size to maintain the accuracy of estimates and to ensure representativeness” (Schumacker and Lomax 1996). According to Hair et al. 2006 “SEM has five factors that impact the sample size requirements: model complexity, multivariate procedure, estimation procedure, amount of missing data and average error variance of indicator”. SEM researchers suggest a minimum sample size 200 when comparatively simple models are tested (Schumacker and Lomax 1996; Ullman 2001; Tomarken and Waller 2005). For more complex models involving more than six factors, sample size requirement may exceed 500 (Hair et al. 2006). Considering all these guidelines, the final questionnaire was administered to a sample of 600 customers. Of this random sample, 558 usable questionnaires were obtained; an effective response rate of 93 percent.

### 4.3 Questionnaire Development

Data for present study was collected using questionnaire. The questionnaire was drafted after carrying out an extensive literature survey and conducting unstructured interviews with several customers. Three marketing academics served as judges to evaluate the content and wording of items. Based on their comments, several statements were added or deleted and some statements
were rephrased. Section I incorporates the statements measuring customer-based corporate reputation, commitment, affective loyalty, intentional loyalty, social risk, temporal risk, customer participation in service delivery. Each statement was measured on a ‘7-point Likert scale’ ranging from ‘Very Strongly Agree’ to ‘Very Strongly Disagree’. We started with 46 six items and eleven constructs. In refinement process 7 items were deleted. Section II consists of questions relating to respondent’s demographic characteristics. These questions included information about their age, gender, education level and family income. Regarding general information about supermarkets, respondents were asked about the name of supermarket, their association with particular supermarket, frequency and purpose of visit.

4.4 Respondent Characteristics

This section describes respondents’ demographics information. Table 4.1 deals with a summary of respondents’ demographic characteristics. The analysis indicates that there were 297 male and 261 females participants. Forty per cent of respondents were in the age group of 26 to 40 years of age, followed by 29 per cent in 18 to 25 age group; 19.4 per cent in age group of 41 to 56 and just 11 per cent in the age category of 56 or above.

With regard to education, 46 per cent of respondents are postgraduates, followed by 31 per cent above postgraduates, 17 per cent graduates and only 5.4 per cent undergraduates.

In terms of monthly income, 22.8 per cent respondents were in the in the group 20,001 to 30,000 income category, immediately followed by 22.2 per cent in group of less than 20,000, 21.3 in group of 30,001 to 40,000 income group, 14 per cent in group of 50,001 to 60,000 and 13.4 per cent in income group of 40,001 to 50,000 and merely 6.3 per cent in income group above 60,001.
**Table: 4.1 Demographics Characteristics of the Sample**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>297 (53.2%)</td>
</tr>
<tr>
<td>Female</td>
<td>261 (46.8%)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>18-25</td>
<td>162 (29%)</td>
</tr>
<tr>
<td>26-40</td>
<td>225 (40.3%)</td>
</tr>
<tr>
<td>41-56</td>
<td>108 (19.4%)</td>
</tr>
<tr>
<td>Above 56</td>
<td>63 (11.3%)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>30 (5.4%)</td>
</tr>
<tr>
<td>Graduate</td>
<td>95 (17%)</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>259 (46.4%)</td>
</tr>
<tr>
<td>Postgraduate and above</td>
<td>174 (31.2%)</td>
</tr>
<tr>
<td><strong>Income (Monthly in Rs.)</strong></td>
<td></td>
</tr>
<tr>
<td>Less than 20,000</td>
<td>124 (22.2%)</td>
</tr>
<tr>
<td>20,001-30,000</td>
<td>127 (22.8%)</td>
</tr>
<tr>
<td>30,001-40,000</td>
<td>119 (21.3%)</td>
</tr>
<tr>
<td>40,001-50,000</td>
<td>75 (13.4%)</td>
</tr>
<tr>
<td>50,001-60,000</td>
<td>78 (14%)</td>
</tr>
<tr>
<td>Above 60,001</td>
<td>35 (6.3%)</td>
</tr>
</tbody>
</table>

*Note: Figures in parentheses denotes the percentages*
### Table 4.2 Respondents Information Regarding Visit to Supermarkets

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No. of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration of Association</strong></td>
<td></td>
</tr>
<tr>
<td>• Less than 1 year</td>
<td>207 (37.1%)</td>
</tr>
<tr>
<td>• 1-3 years</td>
<td>251 (44.98%)</td>
</tr>
<tr>
<td>• 3-5 years</td>
<td>84 (15.04%)</td>
</tr>
<tr>
<td>• 5-7 years</td>
<td>16 (2.88%)</td>
</tr>
<tr>
<td>• 7-10 years</td>
<td></td>
</tr>
<tr>
<td>• Over 10 years</td>
<td></td>
</tr>
<tr>
<td><strong>Frequency of Visit</strong></td>
<td></td>
</tr>
<tr>
<td>• Twice a week</td>
<td>90 (16.1%)</td>
</tr>
<tr>
<td>• Once a week</td>
<td>137 (24.6%)</td>
</tr>
<tr>
<td>• Once in 15 days</td>
<td>125 (22.4%)</td>
</tr>
<tr>
<td>• Once in a month</td>
<td>172 (30.8%)</td>
</tr>
<tr>
<td>• Once in 6 months</td>
<td>29 (5.2%)</td>
</tr>
<tr>
<td>• Once in year</td>
<td>5 (0.9%)</td>
</tr>
<tr>
<td><strong>Purpose of Visit</strong></td>
<td></td>
</tr>
<tr>
<td>• Shopping</td>
<td>397 (71.15%)</td>
</tr>
<tr>
<td>• Window Shopping</td>
<td>135 (24.19%)</td>
</tr>
<tr>
<td>• Entertainment</td>
<td>184 (32.97%)</td>
</tr>
<tr>
<td>• Any other</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Figures in parentheses denotes percentages. In case of purpose of visit customers have opted for two or more options.*
4.5 Statistical Techniques Used

In this section we briefly describe the various statistical techniques used to perform the final analysis. To refine the scale items the procedure as suggested by Churchill Jr., Ford and Walker (1974); Ruekert and Churchill (1984) and Netemeyer, Bearden and Sharma (2003) were followed. In this study to validate the scale and to test the proposed relationship the approach as suggested by Anderson and Gerbing (1988) was used.

4.5.1 Pearson Product Moment Correlation

It deals with correlation between an optimally weighted linear combination of predictors and criterion (Nunnally and Bernstein 1994). It is commonly symbolized as $r$, specifies the magnitude of linear relationship between two variables. “Correlation coefficient (r) of two variables is obtained by dividing the sum of the product of the corresponding deviations of the various items of two series from their respective means by the product of their standard deviations and number of pair of observations.

$$r = \frac{\sum Zx*Zy}{N}$$

where, $Zx$ is the standardized score of variable x and $Zy$ is the standardized score of variable y”

According to Nunnally and Bernstein (1994) the advantages of ‘r’ is that it (1) permits the variance of each of the two measures to be partitioned into meaningful components; (2) may also be used to predict one variable from one other variable (linear regression); (3) is the foundation for predicting one variable from several other variables (multiple regression); and (4) serves as foundation for many complex methods of correlation analysis such as multiple correlation, partial correlation and factor analysis.

Pearson product moment correlation was done for item deletion and assessment of reliability for scale refinement. Ruekert and Churchill 1984 are of view that “Each item in the construct was correlated first with its own total score and then with total score of other constructs to determine whether each item correlated primarily with one dimension and in those cases where an item correlated with one dimension, whether that correlation made conceptual sense”.
4.5.2 Corrected Item-to-total Correlation

“Corrected item-to-total correlation is correlation between the score on the item and the sum of all other items making up the dimension to which the item was assigned” (Parasuraman, Zeithmal and Berry 1988). In the words of Netemeyer, Bearden and Sharma 2003 “These correlations reflect the extent to which any one item is correlated with the remaining items in a set of items under consideration”. In the words of Nunnally and Bernstein 1994, “Items with high corrected item-to-total correlation have more variance relating to what the items have in common and add more to the test’s reliability than items with low corrected item-to-total correlation”. According to Netemeyer, Bearden and Sharma (2003) items with low corrected item-to-total correlation are candidate for deletion.

Past studies on scale development and refinement have used different decision rules for deleting items using corrected item-to-total correlation criteria. For example Bearden, Netemeyer and Teel (1989), in the refinement of scale for measuring ‘Consumer Susceptibility to Interpersonal Influences’, deleted items that had item-to-total correlation below 0.30 in purification stage for developing an instrument to measure satisfaction with transaction specific service recovery. Bearden, Hardesty and Rose (2001) retained items having corrected item-to-total correlation above 0.35 in the development of scale for measuring Consumer Self-confidence”. Tian, Bearden and Hunter (2001) in the item refinement for measuring their scale ‘Consumers’ Need for Uniqueness’, deleted those items having corrected item-to-total correlation below 0.50. Sin et al. (2005) in the refinement stages of developing a scale for measuring Relationship Marketing Orientation retained those items that had corrected item-to-total correlation above 0.30. Nunnally and Bernstein (1994) recommend that items with corrected item-to-total correlation above 0.30 are to be retained. In accordance with Bearden, Hardesty and Rose (2001), we retained items that have corrected item-to-total correlations more than or equal to 0.35 and rest were deleted.

4.5.3 Means and Variances

Mean and variances is another technique for deletion of items and assessment of internal consistency. The means and variances can provide hint about which items will to be retained and which will be of not use in a given study. According to Kumar and Beryerlien 1991; Kline 2005,
“Generally the higher the variability of the item and the more the mean of the item is at the centre point of distribution, the better the item will perform”. We rejected item that had both mean below scale midpoint and limited variance. This was done to maximize variance and minimize skewness (Kumar and Beyerlein 1991; Thomson, MacInnis and Park 2005).

4.5.4 Factor Analysis

Factor analysis is a multivariate statistical method used primarily for the data reduction and summarization after identifying the dimensions and the structure and determining the extent to which each variable is explained by theses dimensions (Hair et al. 2006). “Factor analysis allows the identification of a relatively small number of factors that can be used to represent the relations among a set of interrelated variables, such as a set of items on a measure or a set of instruments” (Goodwin 1999). According to Nunnally and Bernstein (1994), factor analysis can be used to determine:

1. The arrangement of relations among variables;

2. Grouping or clustering of variables, which variables belong to which group and how strongly they belong;

3. How many dimensions are needed to explain the relations among the variables?

The two most widely used forms of factor analysis are principal components and common factor analysis. The former analysis is used when the objective is to summarize most of the original information (variance) in a minimum number of factors for prediction purpose. “This technique is suggested when the primary concern is to determine the minimum number of factors that will account for minimum variance in data” (Malhotra 2007). The second type of analysis is used primarily to recognize essential factor or dimensions that reflect what the variables share in similar (Hair et al. 2006).

Exploratory factor analysis was performed to check the unidimensionality of the items. Since the primary purpose of using factor analysis was data summarization, we used principal component
analysis method with varimax rotation for establishing the construct validity of the scale (Malhotra 2007).

4.5.5 Structural Equation Modelling (SEM)

SEM also known as ‘path analysis with latent variables’. Is used for the specification and analysis of interdependencies among observed variables and underlying theoretical constructs, often called latent variable (Hwang et al. 2010). “SEM is a technique to specify, estimate and evaluate models of linear relationships among a set of observed variables in terms of generally smaller number of unobserved variables” (Shah and Goldstein 2006). SEM is statistical methodology that takes a confirmatory approach to the analysis of a structural theory bearing on some phenomenon (Steenkamp and Baumgartner 2000; Byrne 2001). “It can be viewed as combination of factor analysis, regression and path analysis” (Hox and Bechger 1998). SEM provides researchers with the skill to study and assess the multiple interrelated dependence relationships in a particular/single model. Its closest analogy is multiple regression analysis, which can estimate a single relationship. “But SEM can estimate many equations at once, and they can be interrelated, meaning that the dependent variable in one equation can be an independent variable in other equations. This allows the researcher to model complex relationships that are not possible with other multivariate techniques” (Hair et al. 1998).

Confirmatory factor analysis plays important role in SEM for model validation in path (Jackson, Gillaspy and Purc-Stephenson 2009). Thompson (2004) claimed that without CFA it makes little sense to relate constructs within an SEM model of the factors specified as part of the model are not worthy of further attention. Advantages of SEM compared to multiple regression include more flexible assumptions (particularly allowing interpretations even in the face of multicollinearity), use of confirmatory factor analysis to reduce measurement error by having multiple indicators per latent variable, graphical modeling interface, the desirability of testing models overall rather than coefficients individually, the ability to test models with multiple dependent, the ability to model mediating variables, the ability to model error terms and the ability to handle difficult data. There are six stages in SEM (Hair et al. 2006), which are explained below:
Step 1: Defining and Operationalizing Individual Constructs

This step involves defining the constructs that provides basis for the selection and designing of individual indicator item (Hair et al. 2006). The scale items can be operationalized in format such as Likert scale or semantic differential scale. The items can be derived from the previous research studies or can develop new construct measures (Hair et al. 2006).

Stage 2: Developing and Specifying the Measurement Model

In the words of Schumacker and Lomax 1996, “Model specification refers to the initial theoretical model the researcher formulates”. In this stage, “the researcher specifies how all of the individual constructs will come together to form an overall measurement model” (Hair et al. 2006). The researcher defines the relations between the observed and unobserved variables (Byrne 2001). The key issues involved in this stage are: (i) determining the items per construct and (ii) model identification.

(i) Determining the items per construct

Every latent variable should be measured by multiple manifest variables but the number of manifest variables to be used is less clear (Shah and Goldstein 2006). On the one hand, researcher prefers a multiple indicators to fully represent a construct and maximize reliability (Hair et al. 2006). “On the other hand, large number of indicators per latent variable makes it difficult to parsimoniously signify the measurement structure underlying a set of observed variables and to find a model fit that fits the data” (Baumgartner and Homburg 1996). In words of Shah and Goldstein 2006, “More items per indicators are disadvantageous as it means more parameters to estimate and this requires a large sample size for adequate power”. A small number of items per construct are cause of anxiety because it leads to non-convergence and improper solutions (Anderson and Gerbing 1984). Single indicators constructs should be used when measurement reliability is not an issue (Baumgartner and Homburg 1996; Shah and Goldstein 2006).
(ii) Model Identification

“Identification is concerned with whether the parameters of the model are uniquely determined or not” (Long 1983; Kline 1998). The identification issue deals with whether enough information exists to identify a solution to a set of structural equations (Hair et al. 2006). In the words of Baumgartner and Homburg 1996, “A necessary condition for identification is that the number of parameters to be estimated should not exceed the number of distinct elements in the variance-covariance matrix of the observed variables”.

The degrees of freedom are calculated as:

\[ \frac{1}{2} v (v+1) - q \]

Where \( v \) is the number of manifest variables and \( q \) is the number of distinct parameters to be estimated. In the words of Kline “Degrees of freedom are function of model specification in terms of the number of equations and the effective number of parameters that need to be estimated”. Structural model may be just-identified, over-identified or under-identified.

**Just-identified Model** – According to Shah and Goldstein 2006, “When the effective number of parameters is exactly equal to the number of equations, the model is said to be just-identified”. A just-identified model contains just enough degrees of freedom to estimate all free parameters (Hair et al. 2006). In the just-identified model, all the parameters are uniquely determined and there is enough information in the matrix (Schumacker and Lomax 1996). “The just-identified model is not scientifically interesting because it has no degrees of freedom and therefore can never be rejected” (Byrne 2001).

**Under-identified Model** – “An under-identified model is the one in which more parameters are to be estimated than there are item variance and covariance” (Hair et al. 2006). The model contains insufficient information to uniquely estimate the parameters (Schumacker and Lomax 1996; Shah and Goldstein 2006). In the words of Byrne 2001; Shah and Goldstein 2006 “An infinite number of solutions are possible for an under-identified mode making it difficult for the researcher to choose among the various solutions because each is equally valid”.

**Over-identified Model** – “An over-identified model is one that has more unique covariance and variance terms than parameters to be estimated” (Hair et al. 2006). The degrees of freedom are
one or greater (Shah and Goldstein 2006). The aim in SEM is to specify a model such that it meets the criteria of over-identification (Schumacker and Lomax 1996).

**Stage 3: Designing a Study to Produce Empirical Results**

The key issues to be addressed in this stage are: (i) what type of data to be analysed (ii) what is the sample size.

(i) **Type of Data to be Analysed**

There are two options for the format of the input data-correlation and covariance matrix. According to Schumacker and Lomax 1996 “A variance-covariance matrix is made up of variance terms on the diagonal and covariance terms on the off-diagonal”. The covariance matrix has the benefit of providing valid comparisons between diverse sample or populations (Hair et al. 1998). On the other hand the key advantage of using correlation input for SEM lies in the fact that the default resulting parameter estimates are standardized (Hair et al. 2006). Researchers recommend using covariance matrix for SEM (Schumacker and Lomax 1996).

(ii) **Sample Size**

A noteworthy influence on the reliability of parameters estimates and model fit is by adequacy of sample size (Schumacker and Lomax 1996; Shah and Goldstein 2006). SEM requires large sample size. Five factors impact the sample size requirements:

1. **Model complexity** – Kline 1998, is of view “More complex models involve the estimation of more statistical effects, and thus larger samples are necessary in order for the results to be reasonably stable”. Model complexity in SEM leads to the need for larger samples because larger samples mean less variability (Hair et al. 2006).

2. **Multivariate Procedure** – According to Hair et al. 1998, “As the data deviate more from the assumption of multivariate normality, the ratio of respondents to parameter need to increase with generally accepted ratio of 15 respondents for each parameter”.
3. **Estimation Procedure** – “The type of estimation algorithm used in the analysis also affects sample size requirements” (Kline 1998). According to Baumgartner and Homburg 1996, “All methods for the estimation and testing of structural equation models are based on asymptotic theory and the sample size has to be large for the parameter estimates and test statistics to be valid”.

4. **Amount of Missing Data** – In the words of Hair et al. 1998, “Missing data can have profound effect on calculating the input data matrix and its ability to be used in the estimation process”. If more than 10 percent missing data is expected, the researcher should use larger sample size to offset any problem of missing data (Hair et al. 2006).

5. **Average Error Variance of Indicators** – According to Hair et al. (2006)” larger sample sizes are required as communalities become smaller. Communalities represent the average amount of variation in the measured variables explained by the measurement model. SEM models with communalities less than 0.5 require larger sample sizes for convergence and model stability”.

**Stage 4: Assessing Measurement Model Validity**

The next stage involves assessing measurement model’s validity. Byrne 2001 considered that “This concerns the extent to which a hypothesized model describes the sample data”. Three issues are considered in it: (i) parameter estimates, (ii) model fit and (iii) evaluation internal structure of model fit.
(i) **Parameter Estimates**

With respect to the fit of individual parameters, three aspects of concern are feasibility of parameter estimates, appropriateness of standard error and statistical significance of parameter estimates.

- **Feasibility of Parameter Estimates** – Parameter estimates should exhibit the correct sign and size and be consistent with the underlying theory (Byrne 2001). Negative variances, correlations greater than one and extremely larger parameters estimates are examples of parameters exhibiting unreasonable estimates (Bagozzi and Yi 1988).

- ** Appropriateness of Standard Errors** – Standard errors that are excessively large are indicators of poor fit (Byrne 2001). Shah and Goldstein 2006 are of view that “A large standard error indicates an unstable parameter that is subject to sampling error”.

- **Statistical Significance of Parameter Estimates** – This involves an examination of parameter estimates and the accompanying tests of significance (Bagozzi and Yi 1988). No significant parameters should be deleted from the model (Byrne 2001).

(ii) **Overall Model Fit**

According to Mulaik et al. 1989; Hair et al. 1998 “Goodness-of-fit (GOF) measures the correspondence of the actual or observed input (covariance or correlation) matrix with that predicted from the proposed model”. These measures are categorized into groups’ namely *absolute measures, incremental measures and parsimony measure*, which are explained below –

### I. Absolute Fit Measures -

According to Shah and Goldstein 2006, “Absolute fit measures indicate the degree to which the hypothesized model reproduces the sample data”. They do not compare the GOF of a specified model to any other model (Mulaik et al. 1989; Hair et al. 2006). The most essential absolute fit index is the $\chi^2$ statistic. *Other generally used absolute measures include goodness-of-fit index (GFI). Adjusted goodness-of-fit index (AGFI) and root mean square error of approximation (RMSEA)*
• **Chi-Square (χ²) statistic** - “The most popular index used to assess the overall goodness of fit has been χ² statistic. The χ² statistic tests the null hypothesis that the estimated variance-covariance matrix deviates from sample variance-covariance matrix only because of sampling error” (Baumgartner and Homburg 1996). The χ² statistic is inherently biased when the sample size is large (Shah and Goldstein 2006). That is, as the sample size increases, the chance of rejecting a model (whether true or false) also increases (Bagozzi and Yi 1988).

• **Goodness of fit Index (GFI)** – “GFI indicates the proportion of the observed covariance explained by the model-implied covariance” (Kline 1998). In the words of Byrne 2001, “It is a measure of the relative amount of variance and covariance in S that is explained ∑”. The GFI is based on a ration of the sum of the squared differences between the observed and reproduced matrices to the observed variances (Shumacker and Lomax 1996). The GFI typically ranges from zero to 1.00, with values close to one indicating better fit (Byrne 2001). “The usual rule of thumb for this index is that .95 is indicative of good fit relative to the baseline model, while values greater than .90 are usually interpreted as indicating an acceptable fit” (Schermelleh-Engel, Moosbrugger and Muller 2003).

• **Adjusted Goodness of Fit Index (AGFI)** – AGFI adjusts the GFI index for the degree of freedom of a model relative to the number of observed variables (Shumacker and Lomax 1996). AGFI also addresses the issue of parsimony by incorporating a penalty for the inclusion of additional parameters (Byrne 2001). “A rule of thumb for this index is that .90 is indicative of good fit relative to the baseline model, while values greater than .85 may be considered as an acceptable fit” (Anderson and Gerbing 1988; Schermelleh-Engel, Moosbrugger and Muller 2003).

• **Root Mean Square Error of Approximation (RMSEA)** – “RMSEA is related to the difference in the sample data and what would be expected if the model were assumed correct” (Dion 2008). In the words of Schemelleh-Engel, Mososbrugger and Muller 2003,” It is a measure of approximate fit in the population and is therefore concerned with the discrepancy due to approximation”. RMSEA values less than .05 indicate good fit, values between 0.05 and 0.08 as an adequate fit, and value between 0.08 and 0.10 as a
mediocre fit and values greater than .10 indicate poor fit (Browne and Cudeck 1993; MacCallum, Browne and Sugawara 1996).

II. Incremental fit Measures-

“Incremental measures compare a proposed model with a null model” (Schumacker and Lomax 1996). According to Hair et al. 2006, “These measures differ from absolute fit indices that they assess how well a specified model fits relative to some alternative baseline model”. The generally used incremental measures are Comparative Fit Index (CFI), Normed Fit Index (NFI) and Tucker Lewis Index (TLI).

- **Normed Fit Index (NFI)** – It is a ratio of the difference in the $\chi^2$ value fitted model and a null model divided by the value $\chi^2$ for the null model (Hair et al. 2006). NFI values range from 0 to 1, with higher values indicating better fit (Shumacker and Lomax 1996). A rule of thumb for this index is that .97 is indicative of good fit relative to the independence model. Whereas values greater than .95 may be interpreted as an acceptable fit (Schermelleh-Engel, Mososbrugger and Muller 2003).

- **Comparative Fit Index (CFI)** – CFI is modified version of NFI and is less affected by sample size (Kline 1998). CFI is normed so that values range between 0 and 1, with higher values indicating better fit (Hair et al. 2006). A rule of Thumb for this index is that .97 is indicative of good fit relative to the independence model, while values greater than 0.95 may be interpreted as an acceptable fit (Schermelleh-Engel, Mososbrugger and Muller 2003).

- **Tucker Lewis Index (TLI)** – The TLI measure is used to compare alternative models against a null model (Schumacker and Lomax 1996). TLI show how effective the model is compared to a null model (Dion 2008). Models with good fit have values that approach 1, and a model with higher values suggests a better fit than a model from zero to one, with higher values indicating better fit (Schermelleh-Engel, Mososbrugger and Muller 2003).
III. Parsimony Fit Measure –

Parsimony refers to “the number of estimated coefficients’ required achieving a specific level of fit” (Schumacker and Lomax 1996). It a criterion for selecting between alternative models (Shermelleh-Engel, Moosbrugger and Muller 2003). Hair et al. 2006 opined that “Parsimony indices provide information about which model among a set of competing models is best, considering its fit relative to its complexity”. The most usually used parsimony measures are *Parsimony Normed Fit Index (PNFI) and the Parsimony Goodness-of-Fit Index (PGFI).*

- **Parsimony goodness-of-fit Index (PGFI)** – PGFI is a modification of the GFI that takes into account the degrees of freedom available for testing the model (Hooper, Coughlan and Mullen 2008). The PGFI takes into account the complexity of the hypothesized model in the assessment of the overall model fit (Byrne 2001). PGFI values range between zero and one, with higher value indicating a more parsimonious fit (Schermelleh-Engel, Mossbrugger and Muller 2003).

- **Parsimony Normed fit Index (PNFI)** – the PNFI adjusts the normed fit index by multiplying it times the parsimony ration (Hair et al. 2006). The PNFI coefficient value ranges from zero to one, with higher values indicating better fit (Schermelleh-Engel, Mossbrugger and Muller 2003).

(iii) Assessment of Fit of Internal Structure of Model

Once the overall model fit has been evaluated, the next step is to assess the reliability of the measurement of each construct (Hair et al. 1998). “Three types of reliabilities – individual item reliability, construct reliability and average variance extracted can be examined” (Bagozzi and Yi 1988).

**Individual Item Reliability** – “Item reliability signifies the amount of variance in an item due to the underlying constructs rather than to error” (Suh and Han, 2002).

It is defined as
According to Bagozzi and Yi 1988, “The reliability of a measure is equal to its true score variance divided by the total variance”. The reliability of individual items may also be assessed by squaring their respective loadings (Segars 1997). “Values greater than 0.50 are considered adequate, which indicate that items explain more variance than is explained by the error” (Segars 1997; Bollen 1989).

**Construct or Composite Reliability** – The composite reliability of a construct is determined as follows:

\[
\text{Composite Reliability} = \frac{(\sum \text{standardized loadings})^2}{(\sum \text{standardized loadings})^2 + \sum \varepsilon_j}
\]

Where \(\varepsilon_j\) is the measurement error for each indicator

“High construct reliability depicts that the internal consistency exists. In other words, that the measures all consistently represent the same latent construct” (Hair et al. 2006). Composite reliability values greater than .60 are considered adequate (Bagozzi and Yi 1988).

**Average Variance Extracted (AVE)** – the average variance extracted (AVE) measure is calculated as follows:

\[
\text{AVE} = \frac{(\sum \text{standardized loadings})^2}{(\sum \text{standardized loadings})^2 + \sum \varepsilon_j}
\]

Where \(\sum \varepsilon_j\) is the measurement error for each indicator

“AVE is the average percentage of variation explained among the items” (Hair et al. 2006). “It measures the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error” (Fornell and Larcker 1981). High variance extraction specify that items are truly representative of the latent construct (Hair et al. 1998). An AVE value of 0.5 or higher indicates adequate convergence (Bagozzi and Yi 1988; Hair et al. 2010). According to Hair et al. 2006 “An AVE of less than 0.50 indicates that the variance due to measurement is larger than the variance captured by the construct”.

\[p_i = \lambda_i^2 \frac{\text{Var } T}{\lambda_i^2 \text{Var } T + \theta_{ii}}\]
Assessment of Construct Validity

The next step is to assess the validity of each construct. Construct validity refers to “how well a measure actually measures the construct it is intended to measure” (Netemeyer, Bearden and Sharma 2003). Two types of validities can be examined: convergent validity and discriminant validity.

Convergent Validity – “Convergent validity is the extent to which the scale correlates positively with other measures of the same construct” (Malhotra 2007). Netemeyer, Bearden and Sharma 2003 are of view that “Convergence is found if the two measures of the same construct are highly correlated”. Convergent Validity can be assessed in the following ways:

Factor Loadings – According to Anderson and Gerbing 1988; Hair et al. 2010, “Standardized loading estimates should be 0.5 or higher”

Average Variance Extracted (AVE) – “To suggest adequate convergent validity AVE should be 0.5 or greater” (Fornell and Larcker 1981; Hair et al. 2010).

Reliability – To indicate adequate convergence or internal consistency construct reliability should be 0.7 or higher (Nunnally and Bernstein 1994; Hair et al. 2010).

Discriminant Validity – In the words of Malhotra 2007, “Discriminant validity is the extent to which a measure does not correlate with other constructs from which it is supposed to differ“. Kline 1998 opines that “If the correlations of the factors that underlie sets of indicators that are supposed to measure different constructs are not excessively high, then there is evidence of discriminant validity”. There are two methods of assessing discriminant validity:

- Discriminant validity can be assessed by fixing the correlations between any two constructs as equal to one and then performing a chi-square difference test (Anderson and Gerbing 1988). A significant lower \( \chi^2 \) value for the model in which trait correlations are constrained to unity would indicate that the traits are not perfectly correlated that discriminant validity is achieved (Bagozzi and Philips 1982).
Discriminant validity can also be assessed by comparing the average variance extracted (AVE) squared correlation between constructs (Fornell and Larcker 1981). The AVE values should be greater than squared correlation estimate, this suggests that the constructs have more extracted variance than variance shared with other constructs (Ping 2004; Hair et al. 2010).

Stage 5: Specifying the Structural Model

According to Hair et al. 2010, “This stage involves specifying the structural model by assigning relationships from one construct to another construct based on the proposed theoretical model”. The researcher specified the manner by which concerned latent variable directly or indirectly influences changes in the value of certain other latent variables in the model (Byrne 2001).

Stage 6: Assessing the Structural Model Validity

According to Hair et al. 2010, “This stage assesses the validity of the structural model and its corresponding hypothesized theoretical relationships”. This stage includes the following:

a) Structural Model Fit –

To assess the overall fit of the structural model the same criteria as the measurement model is used.

b) Comparing Competing Models –

Comparing the fit indices of the hypothesized and rival model provides a reliable test of the hypothesized model. By using a competing model strategy, the researcher can verify “whether the proposed model, regardless of overall fit, is acceptable. It is so, because no other similarly formulated model can achieve a higher level of fit” (Hair et al. 2010).
c) Testing Structural Relationships –

The researcher must examine the individual parameters that represent each of the specific hypotheses. The individual parameters are examined to see whether they are in the predicted direction and statistically significant.

4.5.6 Higher-Order Factor Analysis

In the words of Hair et al. 2008 “Higher-order CFAs most often test a second-order factor structure that contains two layers of latent constructs. They introduce a second-order latent factor(s) that causes multiple first-order latent factors, which in turn cause the measured variables (x)”. Theoretically this process can be extended to any number of multiple layers.

Hair et al. (2008) and Koufteros, Babbar and Kaighobadi (2009) contended that higher-order measurement models are appropriate and applicable if it satisfies the conditions:

- Higher-order factor model exhibit adequate fit.
- It predicts other conceptually related constructs adequately and as expected.
- Comparing to a lower-order factor model, the higher-order model exhibit equal or better predictive validity.

4.6 Summary

This chapter is concerned with the methodology of research used in the study. This chapter has outlined the sampling methodology; issues related to target population; sampling unit, sample size and sampling technique. Development of questionnaire and statistical approaches to the data analysis were also explained.