CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

This chapter details the methodology used to empirically examine the theoretical model established in this chapter. Following the introduction, the methodological overview of the study is discussed. This is followed by the quantitative approach, scale development and measurement of constructs. Finally, the method of statistical analyses used in this study is discussed.

3.2 The Proposed Theoretical Model

As discussed in Chapters One and Two, this study is concerned with understanding customer relationships and switching intentions. Furthermore, for the first time linkages between corporate image, service quality, customer perceived value, relationship quality and switching intention have been integrated into one relationship model. In addition, this study also seeks to contribute to existing literature by investigating alternative attractiveness as a moderator variable in the proposed model.

Based on the preceding literature review in Chapter Two, the proposed theoretical framework of the present study is shown in Figure 3.1. The model includes hypotheses which will be tested and reflect the influence of each type of relational variable - corporate image, service quality and customer perceived value on relationship quality and switching intentions respectively.

This is followed by the hypothesis to examine the mediating effect of relationship quality in the model. The last hypothesis proposes the linkage between the moderator variable, alternative attractiveness and its relationship between relationship quality and switching intention.

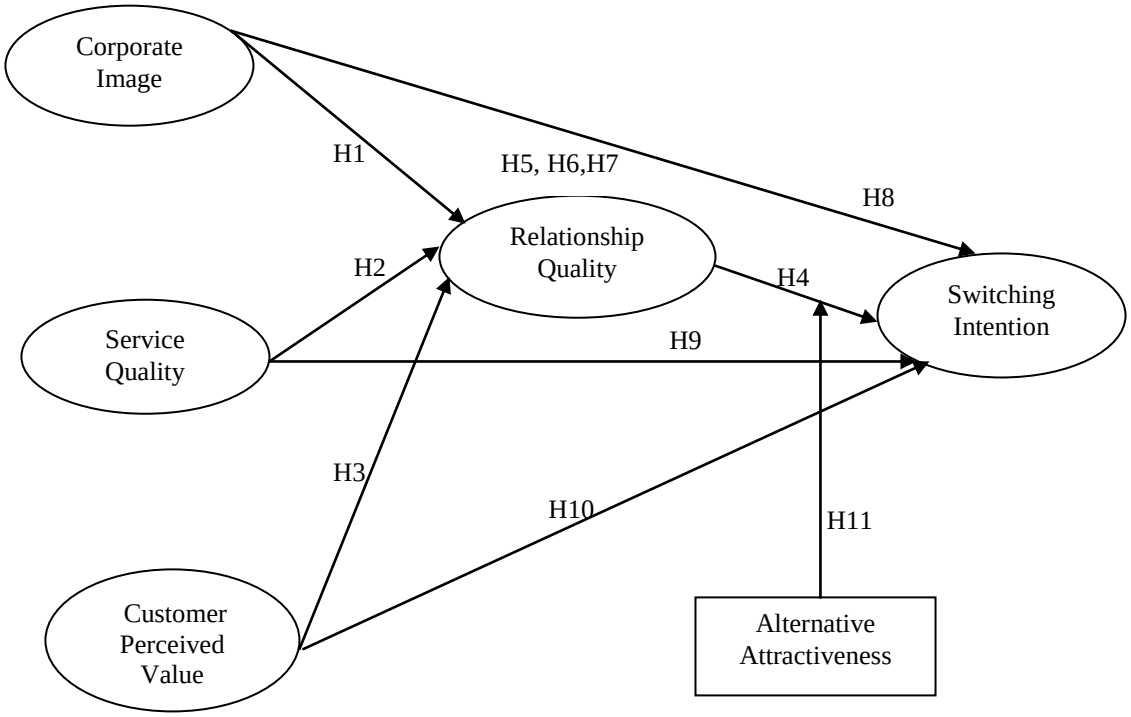

Figure 3.1 The Proposed Theoretical Model

3.3 Hypotheses Development

Based on the literature review and theoretical foundations discussed in Chapter 2, eleven hypotheses have been developed. These hypotheses focus on the influence of corporate image, service quality and customer perceived value on relationship quality and the consequences of relationship quality, specifically switching intention. In addition, the moderating impact of alternative attractiveness on relationship quality and switching intention is proposed.

3.3.1 Corporate Image, Service Quality, Customer Perceived Value and Relationship Quality

Within the service marketing literature, corporate image was early identified by several authors as an important factor in the overall evaluation of services and the companies (Bitner, 1991; Grönroos, 1984; Gummesson and Grönroos, 1988). According to Kennedy (1977) a successful image will involve an association, an emotion and an expectation in the consumer's mind and it is image that can be used to change the perception of the customers in a real or profound way. At the same time an inconsistent corporate image weakens and diminishes the bank' competitiveness.

Research shows the role of image becomes even more significant when competing services are perceived as virtually identical on performance, price, and availability (Andreassen and Lindestad, 1998). In addition, corporate image can be treated as an outcome from accumulated attitude derived from experience and/or direct or indirect market communication (Andreassen and Lindestad, 1998). Supporting this view, De Ruyter and Wetzels (2000) state that corporate image is an information cue that consumers use to judge matters such as credibility, perceived quality and purchase intentions.

For complex services such as financial service providers or those more difficult to understand, elements of previous experiences or credibility will be of great importance. It was found that aspects such as the image and the reputation of the firm play a major part in the decision making process (Roig et al, 2006).

Some researchers agree that bank customers lack measurable attributes to form an organization's image and must resort to extrinsic cues associated with the organization or to elements encountered during the delivery process (Bitner, 1992; Nguyen and LeBlanc, 2002). Empirically, service quality has been found to have an impact on satisfaction and commitment, two dimensions of relationship quality used in this study. Many researchers agree the pursuit of delivering quality service is considered to be essential and an appropriate strategy for success in today's intense competitive and dynamic business environment (Parasuraman et al, 1985; Reichheld and Sasser, 1990; Zeithaml et al, 1990, 1996).

A review of selected literature on service quality in Islamic banking context find that service quality dimensions used by Islamic banking customers is different from those used by customers of conventional banking. Several authors found service quality dimensions used by Islamic banking customers to include image, values and compliance other than the five SERVQUAL dimensions that comprise tangibles, responsiveness, reliability, assurance and empathy (Othman and Owen, 2000; Ahmad Tamimi and Al-Amiri, 20003; Jabnoun and Khalifa, 2005; Izah Mohd Tahir and Wan Tahir, 2005).

In like manner, service quality has been known to contribute to market share and customer satisfaction (Anderson and Zeithaml, 1984; Parasuraman et al, 1985; Zeithaml, 2000). Research on Islamic banking customers in Malaysia (Amin and Isa, 2008) and Kuwait (Othman and Owen, 2001) found a strong link between service quality and satisfaction, and high levels of service quality lead to high levels of satisfaction.

Likewise, in the retail banking sector in China, it was reported that among the five service quality dimensions, reliability and assurance were important predictors of customer satisfaction (Zhou, 2004). Other than satisfaction, Fullerton (2005) also found service quality to positively affect commitment. By the same token, studies on customer perceived value has been linked to satisfaction and customer retention (Rust et al, 1995) and its impact on market share. Similarly, Ennew and Bink (1996) examined the links between customer defection and service quality and found that the cost of retaining customers is generally much less than the cost of acquiring new customers.

Several authors agree customer perceived value has been found to have a direct impact on satisfaction (Ravald and Grönroos, 1996; Zeithaml, 1988 and Fornell et al, 1996). Likewise, McDougall and Levesque (2000) found that perceived value contributes directly to customer satisfaction which in turn, leads to future intentions. In summary, based on the review of the literature in Chapter 2 and considering the above discussion, to some extent there is a relationship between corporate image, service quality, customer perceived value and relationship quality. Accordingly this study assumes that corporate image, service quality and customer perceived value will have an impact on relationship quality. Therefore the following hypotheses are proposed:

H1 : Corporate image positively affects relationship quality for Islamic banking customers in a dual banking environment

H2 : Service quality positively affects relationship quality for Islamic banking customers in a dual banking environment

H3 : Customer perceived value positively affects relationship quality for Islamic banking customers in a dual banking environment

3.3.2 Consequences of Relationship Quality

Within the service marketing literature, behavioral consequences can be categorized into two groups, favorable and unfavorable intentions (Zeithaml, 1996). As such in her study of consequences of service quality on behavioral intentions, four unfavorable intentions were proposed, say negative things, switch to another company, complain and do less business with the company. Similarly, within the relationship marketing literature, Crosby et al, (1990) say relationship quality has been suggested to be a good indicator of the future well-being of long-term service relationship.

In most relationship marketing studies, consequences of behavioral intentions result in positive outcomes, with repeat purchase and word-of-mouth recommendation generally used as the dependent variable in marketing models. However, repeat purchase has been viewed as an indicator of whether or not a customer will maintain a relationship with the company (Zeithaml et al, 1996).

Several studies have used relationship quality as a mediator in relationship marketing models (Crosby et al, 1990; Hennig-Thurau and Klee 1997; Shamdasani and Balakrishnan, 2000; Hennig-Thurau, 2006; Wong and Sohal, 2006; Chen et al, 2008). These studies found that relationship quality is a multidimensional construct and some of the dimensions for relationship quality in the studies include satisfaction, commitment, trust, similarity and expertise. There does not appear to be a consensus on the dimensions for relationship quality, but the common dimensions used in most relationship marketing research are satisfaction, commitment and trust.

Thus, it appears that relationship quality is a mediator in the link between relational variables and relationship marketing outcomes. Based on the above discussion, the following hypotheses are proposed:

H4: Relationship quality is negatively associated with switching intention

H5: Relationship quality mediates the relationship between corporate image and switching intention

H6: Relationship quality mediates the relationship between service quality and switching intention

H7: Relationship quality mediates the relationship between customer perceived value and switching intention

H8: Corporate image is negatively associated with switching intention

H9: Service quality is negatively associated with switching intention

H10: Customer perceived value is negatively associated with switching intention

3.3.3 The Moderating role of Alternative Attractiveness

The use of moderating variables is quite rare in relationship quality studies and Athanasopoulou (2009) found only two studies which have investigated moderators in their research. Gender and age of the relationship, corporate culture and uncertainty are the moderators that were used in previous relationship quality studies. The use of alternative attractiveness has been highlighted by research on consumer decision making and interpersonal relationships. As defined in the literature review in Chapter 2, alternative attractiveness is conceptualized as the client's estimate of the likely satisfaction available in an alternative relationship (Ping, 1993; Rusbult, 1980).

Related literature suggests customers may decide to terminate the current relationship and seek new service providers if they perceive the alternative to offer better service, the proximity of location and the availability of a full range of services (Sharma and Patterson (2000). Similarly, related switching literature found that positive characteristics of competing service providers positively influences consumers' intentions to switch (Jones et al, 2000). Consequently, when a customer perceives low service quality, low satisfaction and low trust and commitment to a service provider, he or she is more likely to feel pushed to switch (Bansal et al, 2005). Thus, the following hypothesis is developed.

> H11: Alternative attractiveness moderates the relationship between relationship quality and switching intentions

3.4 Methodological Overview

This section provides an overview of the methods undertaken in this thesis to answer the research questions in Chapter One, and to test hypotheses that shall be proposed in this chapter.

A quantitative survey methodology using self-administered questionnaires has been adopted to collect data for the constructs proposed in the theoretical model. These constructs are corporate image, service quality, customer value, relationship quality, alternative attractiveness and switching intentions. These constructs were operationalized by multi-item measures using 6-point Likert scales, and the items used to measure them were adopted and adapted from previously tested scales.

An English version and a Bahasa Malaysia version of the questionnaire are developed separately. To ensure that the wording of this questionnaire was clear and understandable and the equivalence of the instrument was achieved, a pre-test was conducted prior to conducting the final survey. A pre-test is deemed necessary to discover any problems in the instrument, and to determine face validity of the measures.

Following the pre-test procedures, the final survey was conducted. In total, 1,000 questionnaires were distributed to respondents who have an Islamic banking account with a conventional bank or any Islamic bank in Malaysia. The questionnaires were distributed to all states in Peninsular Malaysia and East Malaysia (e.g. Sabah and Sarawak). The respondents for the study must be Malaysian Muslim or non-Muslim, 18 years and older and who have been Islamic banking customers for at least a year.

To analyze the data, two statistical techniques were adopted. The Statistical Package for the Social Sciences (SPSS) version 16 was used to analyze the preliminary data and provide descriptive analyses about the thesis sample such as means, standard deviations, and frequencies. Structural Equation Modeling (SEM using AMOS 16) using Confirmatory Factor Analysis (CFA) was used to test the measurement model.

SEM was conducted using the two-step approach recommended by Anderson and Gerbing (1988). The first step includes the assessment of the measurement model, while the second step includes assessment of the structural model. The measurement model stage in this thesis was conducted in two steps. This involves the assessment of the unidimensionality, followed by the assessment of reliability and validity of the underlying constructs.

Reliability was investigated using both the internal consistency measures of Cronbach's alpha and CFA. Validity criterion construct, including convergent, discriminant and external validity was also assessed. Once the scale was developed in stage one, the hypotheses developed in Chapter Two was tested in stage two (the structural model).

3.5 Quantitative Approach

This section justifies the use of a survey methodology using self-administered questionnaires as being appropriate for collecting data for the research. According to Amaratunga et al. (2002) quantitative methods help the researcher to establish statistical evidence on the strengths of relationships between both exogenous and endogenous constructs. They also emphasize that the statistical results provide directions of relationships when combined with theory and literature.

Cavana (2001) and Amaratunga et al, (2002) also point out that quantitative methods can be used to verify the hypotheses and provide strong reliability and validity. Likewise, this methodology has been used in similar studies of relationship marketing (Henning-Thurau et al, 2002; Wong and Sohal, 2002), and particularly those in a retail bank setting (Beerli, 2002, Ndubisi and Wah, 2005) have also widely used this approach. In this respect, because the objective of this thesis is to empirically investigate casual relationships among the underlying constructs, this methodology was deemed to be appropriate (Churchill, 1995).

3.5.1 Self-Administered Questionnaire

Data collection can be gathered in a number of ways. Some of the methods used include personal interviews, telephone interview, and self-administered questionnaires.

Self-administered questionnaire used in this thesis is described as a data collection technique in which the respondent reads the survey questions and records his or her own responses without the presence of a trained interviewer (Zikmund, 2003). Self-administered questionnaires to some extent present a challenge in which they rely on the clarity of the written word more than on skill of interviewers (Zikmund, 2003) with the following advantages:

1) the population in this theses includes a large number of respondents, and thus self-administered questionnaires can be used to survey quickly and economically compared with other methods such as personal interview or telephone interview
2) the questionnaire can be completed whenever respondents have time; and
3) it reaches a geographically widespread sample with lower cost because the researcher is not required (Zikmund, 2003).

In addition, studies in the domain of relationship marketing have utilized self-administered questionnaires (Morgan and Hunt, 1994; Shamdasani and Balakrishnan, 2000; Liang and Wan, 2004; Wang et al, 2006)

3.6 Scale Development

This section of the chapter explains the selection of scale items that are used to measure the constructs in this thesis.

These are corporate image (corporate reliability and reputation); service quality (tangibles, responsiveness, assurance, empathy, reliability); customer perceived value (functional value, functional value for contact personnel, social value and emotional value); relationship quality (satisfaction and commitment), alternative attractiveness (7 items) and switching intentions (comprising of three items). The items chosen for this study have been selected from the literature to be the most representative of customers' perceptions as end users of buyer-seller relationships. All scales used have been adopted from studies with valid and reliable measures of corresponding constructs. The scales used in this thesis have been developed from a review of the relevant literature. In total seventy two items were used to measure the constructs in the model.

Table 3.1 shows a summary of the number and source of the items used to test each construct.

Table 3.1: Total Number of Scale Items Used in this Study

Constructs	Number of items	Sources
Corporate image	**14 items**	Newell and Goldsmith (2001), Sweeney and Swait (2007), Othman and Owen (2001), Dusuki and Abdullah (2007)
Reliability	7 items	
Reputation	7 items	
Service Quality	**22 items**	
Responsiveness	4 items	Cronin and Taylor (1992)
Tangibles	4 items	
Reliability	5 items	
Assurance	4 items	
Empathy	5 items	

Table 3.1: Total Number of Scale Items Used in this Study (continuation)

Customer perceived value	**15 items**	Sweeney and Soutar (2001),Roig et al (2006), Wang et al,(2004)
Functional value	4 items	
Functional value of contact personnel	4 items	
Social value	4 items	
Emotional value	3 items	
Switching intentions	**3 items**	Antreas Athanassopoulos, Spiros Gounaris and Vlassis Stathakopoulos (2001)
Alternative attractiveness	**7 items**	Sharma and Patterson (2000), Jones et al, (2000), Ping (1993)

Each scale is measured using the scale of 1 to 6 with 1 being strongly disagree and 6 = strongly agree. The Likert-scales were selected because they take less time and are easy to answer (Churchill, 1995; Frazer and Lawley, 2000).

3.7 Pretest of the Questionnaire

A pretest of the questionnaire was conducted to assess face validity or content validity of measurement scales. Face validity can be 'evaluated by a group of judges, sometimes experts, who read or look at a measuring technique and decide whether in their opinion it measures what its name suggests' (Judd, Smith, and Kidder 1991, p.54).

In the pre-test process, the questionnaire was reviewed by three academics and three professionals who are specialists in Islamic banking and retail banking to elicit comments on the content and wording, to modify scale items to suit the specific industry context and to assess questions for face validity. Valuable feedback was gained and changes were made to the document before it was finalized and used in the actual field work. Respondents took an average of 25 minutes to complete the questionnaire.

The questionnaire was originally prepared in English and then translated into Malay by using backward translation method with assistance from linguists to reduce translation bias and error. Malhotra et al, (1996, p.24) commented, "if a translator is not fluent in both languages and familiar with two cultures, direct translation of certain words and phrases may be erroneous". Minor discrepancies between the original instrument and its back translated version were resolved by the translator.

Pilot Test

There is a wide agreement among marketing scholars that pre-testing is an integral part of the questionnaire development process. A Pre-test is defined as a "trial run with a group of respondents used to screen out problems in the instructions or design of a questionnaire" (Zikmund, 2003, p.229). The pilot test was also conducted to test the effectiveness of the instrument with regard to ease in understanding questions as well as to test the reliability of the various measures in the instrument.

The sample size for the pre-test in this study constitutes fifty working post-graduate students, a number that corresponds with the population to be studied. Based on the pretest on fifty respondents, several observations are made. Most of the respondents preferred to use the Bahasa Malaysia version of the questionnaire, as they better understood the questions in the Bahasa Malaysia version. Face to face administration of the survey is favorable as respondents receive immediate feedback if they faced problems in understanding the questions. Furthermore, face to face administration also ensured a better rate of return of the questionnaires.

3.8 Sampling technique

A convenience sampling technique was used for this study. Convenience sampling is a type of non-probability sampling which involves the sample being drawn from that part of the population which is close at hand. That is, a sample population selected because it is readily available and convenient (Hair et al, 2006). Respondents were selected from Islamic banking customers of banks offering Islamic banking services.

Sample size

A total of 1,000 questionnaires were distributed to potential respondents. In this study the respondents comprised Muslim and non-Muslim Malaysians, aged 18 years and above who have held an Islamic banking account for at least a year. Potential respondents were initially approached and screened to ensure they fulfilled the afore-mentioned criteria. Respondents were selected from customers who visit the sampled banks on various days for either a week or a month. The purpose of the questionnaire was explained to all.

Avenues were pursued to ensure that the questionnaires distributed at banks in major cities located in Peninsular Malaysia formed a reliable geographic representation. The northern cities of Pulau Pinang and Alor Setar, the southern cities of Johor Bharu and Melaka and the two cities of Kota Bharu and Kuala Terengganu representing the east coast of Peninsular Malaysia were chosen. Shah Alam and Kuala Lumpur were the central cities chosen while questionnaires were distributed to Kuching in Sarawak and Kota Kinabalu in Sabah which represent the East Malaysian cities.

The selection of bank branches across different cities aims to strengthen the generalizability of the findings and representation of Malaysian Muslims and non-Muslims customers. The selection of these branches for data collection is based on the accessibility of locations and the availability of potential respondents. Selected banks located at major cities and towns were used to fulfill these criterion. The survey was conducted from October 2007 until February 2008.

3.9 Measurement of Constructs

An even numbered six-point scale was used in order to avoid clustering of responses at the centre point, which will make the result unreliable (Ling, 1998). Most respondents use a neutral response as a dumping ground when they prefer not to choose, do not care or do not have an opinion. Therefore, the validity of the question will be improved by using a 6-point scale. Although Kinnear and Taylor (1996) conclude that there is no significant difference in the results between scales, a six-point scale is more reliable than five or three-point scales as scale reliability increases with the number of intervals (Osuagwu, 2001).

The first section of the questionnaire required the respondents to identify the bank that they utilize for their Islamic banking services. The second section included the corporate image items followed by service quality items in section three. Section four has items from the perceived customer value items and relationship quality items, followed by seven items for the alternative attractiveness and three switching intention items. The last section of the questionnaire contains demographic questions. A set of the questionnaire is presented in Appendix A.

3.8.1 Corporate Image

The first three constructs to be discussed are corporate image. To measure corporate image, items were taken from corporate credibility (Newell and Goldsmith, 2001) and brand credibility scale (Sweeney and Swait, 2007). To measure corporate image, the fourteen items from Newell and Goldsmith (2001) and Sweeney and Swait (2007) scale have been used for the following reasons. First, the seven items on corporate credibility have been validated compared to other scales which have not been validated (Schumann et al, 1991, Newell and Goldsmith, 2001). The seven items are:

1. The bank delivers what it promises
2. The bank's claims about its services are believable
3. The bank is competent and knows what it is doing
4. The bank has a name you can trust
5. The bank doesn't pretend to be what it isn't
6. The bank is skilled in what they do and
7. I trust my current bank reflect the expertise of the organization.

The corporate credibility scales used also reported high reliability scores $\alpha = 0.73$. Also the use of this scale across groups of consumers, namely Muslims and non-Muslims and retail banking service providers operating in a dual banking environment will enrich the study of corporate image.

Meanwhile, the reputation items were adopted from previous studies of Othman and Owen (2001) and Dusuki and Abdullah (2007) as the items reflect the Islamic retail banking setting. The seven items are:

1. The bank is a socially responsible bank
2. The bank operates in an Islamic working environment
3. The bank complies with Shariah rules and principles capture the compliance of the organization
4. The bank behaves responsibly towards the people in the communities where it operates
5. The bank is a financially sound bank
6. The bank is a reliable and trustworthy bank, and
7. The bank supports good causes

Consequently, the reputation scales have also been used on Islamic banking customers in Kuwait and Malaysia with reliability scores $\alpha = 0.70$ and above. Therefore the scales are deemed appropriate to use for this study.

3.8.2 Service Quality

The service quality construct comprise five dimensions; tangibles, reliability, responsiveness, assurance and empathy. In total twenty two perceptions of performance items have been used in this study. Service quality was measured using the perceptions of performance items of the SERVPERF scale by Cronin and Taylor (1992).

The SERVPERF scale has been extensively applied in numerous retail banking studies because of its good psychometric properties (Yaves et al, 1997; Bloemer et al, 1998, Beerli et al, 2004). Although this scale has a high number of items, the items are not reduced because the items are considered important and the scale was adopted since it is still the most widely used instrument for retail banking studies.

The first four items concern the tangibles of the organization and all items were taken from SERVPERF scale with no modification. The items are:

1. The bank has up to date equipment
2. The bank's physical facilities are visually appealing
3. The bank's employees are well dressed and appear neat
4. The appearance of the physical facilities of the bank is in keeping with the type of services provided

Next five items all capture the reliability dimensions. The items are:

1. When the bank promised to do something by a certain time, it does so
2. When there is a problem, the bank is sympathetic and reassuring
3. The bank is dependable
4. The bank provides its services at the time it promises to do so
5. The bank keeps its records accurately

The following four items are for the responsiveness dimension. The four items are:

1. The bank tells its customers exactly when services will be performed
2. You receive prompt service from the bank
3. The bank employees are always willing to help customers
4. The bank employees are not too busy to respond to customer request promptly

Finally the last nine items were for the assurance and empathy dimensions.

For the assurance dimension four items were used. The items are:

1. The bank employees are trustworthy
2. You feel safe when conducting transactions with the bank's employees
3. The bank employees are polite
4. There is adequate support from the bank so employees can do their job well

Finally, for the empathy dimension, five items were used in the study. The items are:

1. The bank gives individual attention
2. The bank employees give you personal attention
3. The bank employees know what your needs are
4. The bank has my best interest at heart
5. The bank has convenient operating hours

Studies that have used the performance only scales have reported reliabilities in excess of $\alpha = 0.7$ norm (Nunnally, 1978). For example, Cronin and Taylor (1992) reported $\alpha = 0.96$ using the 22 item scale for measuring perceived service quality for banks. Zhou used a similar scale and reported $\alpha = 0.93$ in the China bank retail setting.

3.8.3 Customer Perceived Value

The customer perceived value construct has been conceptualized as comprising four dimensions (functional value, functional value contact personnel, emotional value and social value). The scales used to measure these dimensions are discussed below. The eight functional value items were taken from Roberts et al, (2003) and Wang et al, (2004) study that developed the perceived value scale specifically to be used in the retail banking context.

Some minor modifications were made to two items from the functional value dimension to reflect the Islamic banking setting of this study and the respondents who are Islamic banking customers of conventional and Islamic banks. Similarly, some wordings from the original items from the emotional value dimension are replaced. The functional items are:

1. The bank always delivers superior service
2. The Islamic banking products and services of the bank
3. The quality of the service of this bank is consistently good
4. The Islamic banking products and services of this bank make me feel confident

The following four items are used for the functional value contact personnel. The items are:

1. The personnel know their job well
2. The personnel's knowledge about Islamic banking products and services is up to date

3. The information provided by the personnel has always been very valuable to me
4. The personnel has knowledge of all the Islamic banking services offered by the bank

Three items were used for the emotional dimensions. The items are:

1. The Islamic banking service of this bank is the one that I would enjoy
2. The Islamic banking service of this bank is the one that I would feel relaxed about using it
3. The Islamic banking service of this bank would give me great pleasure

For the social value dimension, four items were used. The items are:

1. Using Islamic banking services would improve the way I am perceived by others
2. Using Islamic banking services would make a good impression on other people
3. Using Islamic banking services would give its owners social approval
4. Using Islamic banking services would help me feel accepted by others

The values for functional value and functional value contact personnel showed reliability in excess of 0.7 ($\alpha = 0.827$ and 0.918). The emotional value items and social value items were adopted from Sweeney and Soutar (2001) scale which reported reliability scores of above 0.70 ($\alpha = 0.95, 0.91$). The scales have also been used in previous studies by Wang et al, (2004), and thus it is deemed appropriate to use in this study.

3.8.4 Relationship Quality

The relationship quality construct used in this study consists of two dimensions, commitment and satisfaction. The scales used to measure these dimensions are discussed below. Commitment was measured by affective commitment and calculative commitment. In total there are six items for commitment.

The first three items of affective commitment are:

1. I feel emotionally attached to my current bank
2. My relationship with my current bank is important to me
3. I have a strong sense of loyalty toward my current bank

The following items are used for calculative commitment. The items are:

1. It pays off economically to be a customer of this bank
2. This bank has location advantages versus other banks
3. It would be too costly for me to switch from this bank right now

Commitment in this study has been regarded as affective commitment and calculative commitment. These were drawn from Morgan and Hunt (1994) and Hennig-Thurau (2002). Although Morgan and Hunt (1994) measured commitment in the business to business market, it has been subsequently widely used in the context of business to customer research, and is a well-established measure of commitment.Liang and Wang (2005) and Roberts et al,. (2003) used this scale in their study of business to customer relationship and is appropriate to be used in this study.

The affective commitment items have a Cronbach's alpha reliability of 0.79 (Roberts et al, 2003) and the commitment items is directed toward the customers as end-users. Another dimension of relationship quality is satisfaction and four items were used to measure satisfaction. The items are:

1. I am satisfied with my decision to deal with this bank
2. My choice to use this bank was a wise one
3. This is one of the best banks I could have chosen
4. Using this bank has been a good experience

The satisfaction scale was adopted from Sweeney and Swait (2001), Hennig-Thurau (2006) and Ndubisi and Wah (2005). The use of these satisfaction items has been validated in the study of Hennig-Thurau (2006) and Ndubisi and Wah (2005) who measured satisfaction as one dimension of relationship quality. Both studies reported a high reliability score of 0.924 and 0.89 respectively. As Ndubisi and Wah (2005) used the satisfaction items in Malaysia, it may be assumed that the items could also provide high validity in this study.

3.8.5 Switching Intention

The switching intention scale consists of three items and was modified to reflect the intention to switch to conventional banking from Islamic banking. The items were taken from a study of retail bank customers in Greece (Athanassopoulos et al, 2001) to determine switching behavior. The three switching intention items from the study showed high validity ($\alpha = 0.92$ to 0.96).

The items are:

1. In the last year I have thought very seriously to switch to conventional banking
2. I have decided to switch to conventional banking that offers better service
3. I am likely to switch to another competing conventional bank during the next year

3.8.6 Alternative Attractiveness - Moderating Variable

When a relationship between variables is changed due to the change level of an additional variable, a moderating or contingency effect is said to exist in that relationship. The additional variable that causes the change effect is called a moderating variable (Hayes, 2005). According to Hair et al, (2006) a moderating variable can be metric or non-metric.

The alternative attractiveness items used address the perception of lower cost availability of a full range of services, proximity of location and more overall service satisfaction available in an alternative relationship. The seven items are:

1. Other banks would be less costly than my present bank
2. A new bank would provide a full range of services than my current bank
3. A new bank is located closer to me as compared to my current bank
4. A new bank would benefit me more than my current bank in achieving my goals
5. I would be more satisfied with the services of a new bank than I am with my current bank
6. If I needed to change banks, there are other good banks to choose from
7. I would probably be happy with the products and services of another bank

The alternative attractiveness items are adopted from Sharma and Patterson (2000) and Jones, Mothersbaugh and Beatty (2000). The use of these items has been validated and show high reliability score in excess of 0.7 (α = 0.85 to 0.92) and and thus it is deemed appropriate to use in this study.

3.10 Data Analysis Methods

To test the relationships of this thesis a relevant model has been developed and proposed. The results of the investigation will be generated by several statistical tests, using Statistical Package for Social Sciences (SPSS, version 16) to test the preliminary data and Structural Equation Modeling (SEM) using confirmatory factor analysis to test the hypothesized model discussed in Chapter 4. This section describes and justifies the use of these statistical techniques.

3.10.1 Preliminary Data Analysis

In order to analyze quantitative data gathered from the questionnaires, Statistical Package for Social Sciences (SPSS) version 16 was used. This software has largely been used and accepted by researchers as a data analysis technique (Zikmund, 2003). Therefore, this technique has been used to screen data of this study in terms of coding, missing data, outliers (i.e. using Box and Whisker, normal probability plot) and normality (i.e. using skewness and kurtosis).

SPSS was also employed to conduct preliminary data analysis including frequencies, mean and standard deviation. These analyses were conducted for each of the variables to gain preliminary information about the sample.

3.10.2 Data Editing and Coding

Following the collection of data from Islamic banking customers, editing of the data was undertaken in order to ensure the omission, completeness, and consistency of the data. Editing is considered as a part of the data processing and analysis stage (Zikmund, 2003).

Coding was used to assign numbers to each answer (Malhotra, 2004) and allow the transference of data from the questionnaire to SPSS. In this study, the coding procedure was performed by establishing a data file in SPSS, and all question items were all pre-coded with numerical values. Data editing procedures were undertaken after data were entered into the data file in order to detect any errors in data entry. Out-of-range values in the data file were corrected by referring to the original questionnaire.

3.10.3 Data Screening

As the first stage in the data analysis, screening for missing data, outliers, and normality was conducted. Data screening is useful in making sure that data have been correctly entered and that the distribution of variables that are to be used in the analysis, are normal. These preliminary analyses are discussed next.

3.10.4 Treatment of Missing Data

It is uncommon to obtain data without some missing data (Hair et al, 2006). During the period of data collection, the missing data was reduced as much as possible by checking all the questionnaires. When questions were found unanswered and uncompleted, it was immediately brought to attention of the related respondents to complete the missing information.

All the data was manually keyed into SPSS version 16. Then, a frequency distribution for each variable in the study, as well as missing values analysis was run to ensure that the data was 'clean'. The results indicate that there are no missing data in the data set.

3.10.5 Detection of Outliers

Hair et al, (2006) defined outliers as "the observations with a unique combination of characteristics identifiable as distinctly different from the other observations." In this study, outliers that required deletion came from a procedural error which included incorrect data entry or mistake in coding.

The data was cleaned by running the frequencies and obtaining the descriptive tables, which can be used to determine the extent of item non-responses, errors on terms of illegitimate responses, cases with extreme values or outliers. From the output descriptive tables, all the items in each section of the questionnaires were examined to ensure that responses were within the range of the items or scales, and that the extreme values were identified.

3.11 Assessment of the Normality

The scale data used in this study was assessed to determine normality of the distribution. Normality is used to describe a curve that is symmetrical and bell-shaped. Because of the assumption that factor analysis and structural equation modeling both require variables to be normally distributed, it was necessary to check the distribution of variables to be used in the analysis (Hair et al, 2006).

To diagnose the distribution of variables, Box and Whisker and stem and leaf plots were used to check for outliers. Outliers refer to the "observations with a unique combination of characteristics identifiable as distinctly different from other observations" (Hair et al, 2006). These outliers, according to Hair et al, (2006) might be very high or very low scores (extreme values), and could result in non-normality data and distorted statistics (Hair et al, 2006; Tabachnick and Fidell (2001).

In order to check any actual deviations from normality, a number of methods can be used. One method is to use skewness and kurtosis. By using this method, values for skewness and kurtosis should not be significant if the observed distribution is exactly normal. Skewness characterizes the degree of asymmetry of a distribution around its mean. For large sample size, 200 and over (Hair et al, 2006), even small deviations from normality can be significant but not substantive. Normal distribution produce a skewness statistic of about zero. Values of 2 standard error of skewness (ses) or more (regardless of sign) are probably skewed to a significant degree. The ses can be estimated roughly by using the following formula $\sqrt{6/N}$ ((Tabachnick and Fidell, 1996)

Another assumption is that the relationship between independent and dependent variable is linear and should exhibit homoscedasticity. The violation of these assumptions will underestimate the extent of the correlation between the variables and this will result in a degradation of the analysis (Tabachnick and Fidell, 2001). Linearity is assumed when a straight-line relationship is present between the two variables (independent and dependent variables) (Hair et al, 2006).

The fulfillment of this assumption leads to the existence of homoscedasticity, a desirable condition where dependent variables exhibit equal levels of variance across the range of predictor variables. The examination of these variables assumptions can be performed by conducting partial regression plots and regression standardized residuals (Pallant, 2005).

3.12 Multicollinearity

In investigating relationships between independent and dependent variables, the presence of multicollinearity can cause several problems including inaccurate results of regression coefficient estimation (Tabachnick and Fidell, 2001).

One of the ways to check for the presence of multicollinearity in the data is by assessing the tolerance and the variable inflation factor (Pallant, 2005). Tolerance is a value that measures the degree of the independent variable's variability that is not explained by the other independent in the model and it is computed by using the formula $1\text{-}R^2$ for each variable. Variance inflation factor (VIF) is the inverse of tolerance and is calculated simply by inverting the tolerance value (1 divided by Tolerance). An indication of multicollinearity is when the value of Tolerance is less than 0.10 and VIF is more than 10 (Hair et al, 2006).

3.13 Structural Equation Modeling

This study employs a structural equation modeling (SEM) as the statistical technique to analyze the hypothesized relationships.

The primary purpose of SEM is to explain the pattern of a series of inter-related dependence relationships simultaneously between a set of latent or unobserved constructs, each measured by one or more observed variables (Hair et al, 2006). SEM is based on the assumption of causal relationships where a change in one variable (x1) is supposed to result in a change in another variable (y1). According to Anderson and Gerbing (1988), SEM is a confirmatory method providing "a comprehensive means for assessing and modifying theoretical models'. Therefore, researchers in relationship marketing have found SEM to be an appropriate technique to examine their hypothesized models (Crosby et al, 1990; De Wulf et al, 2001; Roberts et al, 2003; Liang and Wang; 2005; Palmatier et al, 2006). AMOS 16 (Analysis of Moment Structures) was used to explore statistical relationships among the items of each factor and between the factors of independent (i.e., corporate image, service quality and customer perceived value) and dependent variables (i.e., relationship quality and switching intention).

Further, the researcher can specify, estimate, assess, and present the model in a causal path diagram to show hypothesized relationships among variables. The empirical model can be tested against the hypothesized model for goodness of fit. Any causal paths that do not fit with the original model can be modified or removed.

3.13.1 Two-Step Structural Equation Modeling

SEM can be performed by adopting the one-step approach or two-step approach. In the one step approach, the aim is to process the analysis with simultaneous estimations of both structural and measurement models.

The two step approach aims to process the measurement model first and then fix this measurement model in the second stage when the structural model is estimated. In this study, the two-step approach recommended by Anderson and Gerbing (1988) was adopted.

The first (measurement model) stage of the analysis by was conducted by specifying the causal relationships between the observed variables (items) and the underlying theoretical constructs (composite and latent variables). The purpose of this stage was to verify the unidimensionality of the composite and latent constructs in the first step. Unidimensionality has been defined as "an assumption underlying the calculation of reliability and is demonstrated when the indicators of a construct have acceptable fit on single-factor model" (Hair et al, 2006).

3.13.2 SEM Assumptions

Like any statistical method, a number of assumptions need to be met before conducting SEM. For example, SEM requires the sample size to be adequate, as covariance and correlations are less stable when estimated from small sample sizes (Tabachnick and Fidell, 2001). Some authors believe that SEM could be used for sample sizes as small as 50 (Anderson and Gerbing, 1984), but it has been generally accepted that 100 is the minimum sample size to ensure that appropriate use of maximum likelihood estimation (MLE) (Hair et al, 2006). While there is no agreement among the scholars about sample size, a number of 200 are considered ideal (Hair et al, 1995). The sample size of this study is 456, which is appropriate for using SEM.

3.14 Evaluating the Fit of the Model

In SEM, there are a number of goodness-of-fit indices, which indentify whether the model fits the data or not. There are many indices provided by SEM, although there is no agreement among scholars as to which fit indices should be reported. Anderson and Gerbing (1988) suggest that one or more overall goodness-of-fit indices should be used to assess how well the specified model accounts for data. Kline (1998) recommends at least four such as GFI, NFI, or CFI, NNFI and SRMR.

To reflect diverse criteria and provide the best overall picture of model fit, Hair et al, (1995) recommend the use of at least three fit indies by including one in each of the categories of model fit: absolute; incremental; and parsimonious. This study adopts those measures most commonly used in marketing research to evaluate modes in which the three categories are reflected. A summary of the goodness-of-fit indices are indicated in Table 3.2 followed by a more detailed explanation of each of the indices

Table 3.2: Summary of Goodness-of-Fit Indices

Name of the Index	Level of Acceptance	Comments
Absolute fit indices		
Chi-square ($\chi 2$)	p>0.05	This means the sample is sensitive to large sample sizes
Goodness-of-Fit (GFI)	0.9 or greater	Value close to 0 indicate poor fit, while value close to 1 indicates a perfect fit
Root Mean Square Error of Approximation (RMSEA)	Between 0.05 and 0.08	Value up to 1 and less than 0.05 is considered acceptable

Table 3.2: Summary of Goodness-of-Fit Indices (continuation)

Incremental Fit Indices		
Comparative Fit Index (CFI)	0.90 or greater	Value to 0 indicates a poor fit, while value close to 1 indicates a perfect fit
Parsimonious Fit Indices		
Normed chi-square ($\chi 2 / _{df}$)	$1.0 \leq \chi 2/df \leq 5$	Lower limit is 1.0, upper limit is 3.0 or as high as 5

Absolute fit indices

The chi-square χ^2 estimated the difference between the covariances produced by the proposed model and those the expected covariances based on theory (Satorra and Bentler, 1994). Although this type of statistical index is the most important one to evaluate fit of the model, it has been criticized for being too sensitive to sample size (Fornell and Lacker, 1981; Jöreskog and Sörborm, 1996), especially in cases where sample size is over 200 (Bagozzi and Yi, 1988; Hair et al, 1995). Thus, marketing researchers do not solely use the value of chi-square to reject or accept their models, but use it in conjunction with other indices to evaluate overall fit.

The second measure of absolute fit index used within this study is the Goodness-of-Fit Index (GFI). The GFI measure indicates the relative amount of variance and covariance together explained by the model. However, this measure is not adjusted for degrees of freedom (Hair et al, 2006). Ranging from 0 (indicating a poor fit) to 1 (indicating a perfect fit), where a recommended level of acceptance is 0.90 (Hair et al, 2006).

The third measure of absolute fit index used is Root Mean Square Error of Approximation (RMSEA). It measures discrepancy per degree of freedom. The closer to zero the value of RMSEA, the less the variances and covariances are left unexplained (Sanenz et al, 1999). RMSEA value ranging from 0 (indicating a poor fit) to 1 (indicating a perfect fit) and a suggested recommended level is 0.90 (Hair et al, 1995). However, it has been found that a value ranging from 0.05 to 0.08 is commonly acceptable (Hair et al, 2006).

Incremental fit indices

The Comparative Fit Index (CFI) for example compares the covariance matrix predicted by the model to the observed covariance matrix. The CFI value range from 0 (poor fit) to 1 (perfect fit) having a commonly recommended level of 0.90 or greater (Hair et al, 2006).

Parsimonious fit indices

The parsimonious fit indices tests the parsimony of the proposed model by evaluating the fit of the model to the number of estimated coefficient required to achieve the level of fit (Hair et al, 2006). In this measure, a range of acceptable value for the $\chi 2$ /df ratio have been suggested ranging from less than 2.0 (Hair et al, 2006; Tabachnick and Fidell, 2001), through less than 3.0 to more liberal limits of less than 5.0 (Eisen et al, 1999).

3.15 Reliability and Validity

As discussed in the earlier section, once unidimensionality has been established, the constructs of this study can be assessed for reliability and validity (Anderson and Gerbing, 1988; Hair et al, 2006). Here, a measure may be consistent (reliable) but not accurate (valid), and alternatively, a measure can be accurate but not consistent.

Therefore, in order to ensure the quality of the findings and conclusions of this study, both validity and reliability are assessed. Cronbach's coefficient alpha, construct reliability (CR) and Average Variance Extracted (AVE) are computed to assess reliability, while content, construct, criterion and external validity are examined for validity. Both reliability and validity assessments are discussed below.

3.15.1 Reliability

Zikmund (2003) defines reliability as the "degree to which measures are free from random error and therefore yield consistent results". That means reliability refers to the extent to which a scale produces consistent results if repeated measurements are made on the variables of concern (Malhotra, 2003).

One of the most commonly used indicators of reliability is the Cronbach's alpha (Nunnally, 1978; Sekaran, 2000). This technique estimates the degree to which the items in the scale are representative of the domain of the construct being measured. An alpha coefficient of 0.7 and over is often recommended as the minimum acceptable level. Therefore, this cut-off point (0.70) has been used as the minimum for determining internal consistency of scales for this study.

Accordingly, in this study Construct Reliability (CR) and Average Variance Extracted (AVE) has been used. CR measures the internal consistency of a set of measures rather than the reliability of a single variable to capture the degree to which a set of measures indicates the common latent construct (Holmes-Smith et al, 2006). The AVE estimate is a more conservative indicator of the shared variance in a set of measures than construct reliability. Hence, the variance-extracted estimate reflects the overall amount of variance in the items accounted for by the latent construct. Bagozzi and Yi (1988) recommended that CR should be equal to or greater than 0.60 and AVE should be equal to or greater than 0.50. Thus, this threshold has been used in this study.

3.15.2 Validity

Reliability alone is not sufficient to consider that an instrument is adequate (Anderson and Gerbing, 1988; Hair et al, 2006). Therefore validity is required to validate the constructs of this study. According to Zikmund (2003) validity means "the ability of a scale to measure what is intended to be measured". In this study, content and construct validity (convergent and discriminant validity) have been examined".

3.15.3 Content Validity

Content validity is the extent to which there is a need for the adequate coverage of all the domains of the constructs being examined (Cooper and Schindler, 2001). Content validity cannot be examined using statistical analysis and thus, a thorough exploration of the literature and an extensive search of measures used in the literature must be applied.

Moreover, pre-testing will be used to check on the validity of the constructs. In this study, the measures used will be reviewed by experts, academicians and professionals on the relevancy and adequacy of the constructs (Zikmund, 2003).

3.15.4 Construct Validity

Construct validity is the "extent to which the constructs or a set of measured items actually reflect the theoretical latent construct those items are designed to measure (Hair et al, 2006). Therefore, construct validity deals with the accuracy of the measurement in which that item measures selected from a sample represent the actual true score that exists in the population (Hair et al, 2006). Likewise, Bagozzi, Youjae and Phillips (1991) posit that "without assessing construct validity one cannot estimate and correct for influences of random error and method variance, and the results of the theory testing may be ambiguous".

3.16 Exploratory Factor Analysis

Exploratory factor analysis (herein EFA) is used for data exploration in order to generate hypotheses. It is a technique that assists researchers to determine the structure of factors to be examined. In other words, it is a technique used when the relationship between latent and observed variables is unknown or uncertain.

The distinctive feature of EFA is that the factors are derived from theory, and these factors can only be named after factor analysis is performed. This means that EFA can be performed without knowing how many factors really exist or which variables belong with which constructs (Hair et al, 2006).

3.17 Discriminant Validity

Discriminant validity determines that the measures of a construct have not correlated too highly with other constructs (Sekaran, 2000). This means that items from one scale should not load or converge too closely with items from a different scale. Also, different latent variables which correlate too highly may indeed be measuring the same construct rather than different constructs (Garver and Mentzer, 1999). Therefore, relatively low correlations or no correlation between variables indicates the presence of discriminant validity. In this study, discriminant validity is examined by comparing the average variance extracted for any two constructs or more with the square of the correlation estimates (Hair et al, 2006).

3.18 Moderator Testing

According to Zhao and Cavusgil (2006) in path analysis, two-groups model was used in the core model which was tested for high and low groups. Baron and Kenny (1986) explained that for testing moderating effects, 'the levels of the moderator are tested as different groups '(p.1175). Rigorous pre-tests done verified that the changes in coefficients were truly due to group differences and not due to measurement errors.

To identify the moderator effects, a series of path analysis may be applied to the model. The first analysis was the parameter which was constrained to be equal across the group and the second analysis was not constrained. Thus, a difference in the chi-square value between two models determines the moderating effect on the independent and dependent relationship, only after obtaining evidence of the β coefficients compared across in the high and low group levels.

3.19 Chapter Summary

This chapter discusses the research methodology of this study. The discussion revolves on the issues of research design, research instrument, measurement of constructs and sampling. The data analysis methods used including preliminary data analysis and structural equation modeling were outlined in detail. The final section examined the validity and reliability assessment to ensure reliability and validity of the scale used in this study.