KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, KUMASI

DEVELOPING A MODEL FOR CUSTOMER SWITCHING FOR BANKS (A CASE STUDY OF AHANTAMAN RURAL BANK)

BY

BRIGHT KARIM-ABDALLAH

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FUFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF M.PHIL MATHEMATICAL STATISTICS

SEPTEMBER 2016

DECLARATION

I hereby declare that this submission is my own work towards the award of the M. Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgment had been made in the text.

Bright Karim Abdallah	....	....
Student	Signature	Date

Certified by:

Dr. Richard K. Avuglah	....	....
Supervisor	Signature	Date

Certified by:

Dr. Richard K. Avuglah	....	....
Head of Department	Signature	Date

DEDICATION

This work is wholly dedicated to the management of Ahantaman Rural Bank for their immense support.

ABSTRACT

This study examined the Customer Switching Data of Ahantaman Rural Bank with particular interest in predicting customers who are very likely to switch bank in the near future. The data used was a primary data which was obtained through interviewing of customers of the bank. The objective of the study was to design a model to predict likelihood of customer switching banking services. Logistic regression model was considered as an appropriate statistical tool for analyzing the data since the variables were categorical in nature. Analysis of the data revealed that customer satisfaction, pricing, reliability and employee competence constructs made significant contribution to customer switching intentions in Ahantaman Rural Bank in Ahanta West District in Ghana.

The results showed further that 51.3% of customers who do banking transactions at the head office of Ahantaman Rural Bank have intended to switch services with the bank.

ACKNOWLEDGMENT

A study of this nature cannot be completed without guidance and advice from others.

First of all I thanked Almighty God for his guidance, grace, mercy and protection, without Him I would not have come to this far.

My heartfelt gratitude also goes to my Supervisor, Dr. Richard K. Avuglah, Senior lecturer at Mathematics Department, Kwame Nkrumah University of Science and Technology, for his patient, good sense of direction, critical comments and suggestions during the period of this work.

Special thanks and gratitude go to Mr. Emmanuel Harris, Senior lecturer at the Mathematics Department, KNUST for providing me with guidance and advice. I say thank you to all those who helped in diverse ways for the success of this research, may God richly bless you all.

CONTENTS

LIST OF ABBREVIATION

MANOVA .... Multivariate Analysis of Variance

FINSAP .... Financial Sector Adjustment Programme

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background of the study

Ghanaian banking industry has passed through stages of financial deregulations with the main purpose of making the industry very vibrant and competitive. According to Bawumia et al. (2005) these reformations, apart from competition have changed the market structure coupled with the type of financial services and products that are delivered. The sector during the 1990's was full of government interference. According to Brownbridge and Gockel (1996) majority of the banking institutions in this period were owned by government with a few foreign ones and entry of foreign ones were restricted with Bank of Ghana controlling interest rates. A lot of reforms were introduced in the past to address these restrictions: Financial Sector Adjustment Programme (FINSAP) in 1989; New Banking Act in 1989; Universal Banking License in 2003.

The government in 1989 implemented Financial Sector Adjustment Programme (FINSAP) in collaboration with the IMF and World Bank. The main purpose of the FINSAP was to make the financial industry more resilient and responsive. According to Brownbridge and Gockel (1996) most of the hindrances were removed by FINSAP, for example the interest rates controlled by the Central Bank, allowing the market to determine the equilibrium interest rates. The New Banking Act in 1989 empowered the Central Bank to determine the Capital requirement for banking institutions, issues new license and revokes license of banks when laws are violated. These deregulations stimulated and stabilized the banking industry. From 1994 to 2012 the list of commercial banks increased from 14 to 24 (Ghana Banking Survey, 2012). The Universal Banking License was introduced in 2003 and Banking Act was re-enacted in 2004 leading to the establishment of more banks (Ghana banking Survey, 2008). This resulted in establishment of Nigerian banks in the country and emersion of 447 microfinance establishments by 2014.

The industry today is very competitive and only way to survive is to retain your customers. According to Hull (2002) the competition in banking sector is so keen that, banks are not only facing internal competition; but also with micro finance institutions and mobile money operators.

The primary function of a bank is to accept deposits and then lends the same to earn certain profit, the performance of this function is so similar that the only thing that will retain or avoid customer switching is customer loyalty and satisfaction. That is the way and manner a bank carries these services.

In order to retain customers, customer satisfaction and loyalty become a crux issue to bank management. Previous research in bank customers' satisfaction suggest that customers' satisfaction lead to a better retention of customers, increasing sales profit (Levesque and McDougall (1996); Kish (2000)). It also urges banks to improve their services, introduce innovative products, and efficient bank management. It is critical that banks deliver quality services which in turn result in customer satisfaction in today competitive banking environment.

According to Beckett et al. (2000) in order to match your competitors, the bank management has to continually interact with customers to have a fair knowledge of their problems to avoid switching. It has been established that, customer loyalty is enhanced and improved if a bank carefully implement customers' view and let customers to perceive that their views are important in day to day running of the bank (Andreasen (1988); Lees et al. (2007)).With the emergence of mobile money and hustle-free nature of their activities, in a next few years the competition between them and banks would be a threat for banks survival. According to Bank of Ghana by the end of the year 2016, the mobile money operators will be offering interest on the money in their customer's mobile account and this will reduce the profit base of the banks. A bank customer has to go through a lot of activities in the banking hall to get his own money deposited in his account and this does not exist in the mobile money business. A mobile money customer can transact business even on Sundays and in the night when the banks are not there. So, if banks want to survive it is very important for them to continually find out from their customers what makes them to continue doing banking business with them.

Due to financial deregulations, competitive nature of the industry and technological advancement, a bank customer can now decide to switch bank with least provocation. Therefore, it is very important that banks spend some of their resources to study their customers and find out factors that stimulate a customer to switch banks.

1.2 Profile of Ahantaman Rural Bank

Ahantaman Rural Bank Limited was established in 1984 to do banking business in Ghana. The Bank aspires to be and remains the leading Rural Bank in Ghana meeting stakeholder expectations, to strengthen stakeholder relationship by providing the right solutions that combine technologies, expertise and financial strength, create customer loyalty and shareholder value and employee satisfaction. The Bank currently operates in four districts in the Western Region with a network of sixteen connected Branches. The Bank's Board of Directors is made up of 9 highly dedicated development-oriented persons of divers background, with committed and experienced management team and staff. It offers wide range of services to its customers: fixed deposit accounts; savings and current accounts; loans and overdrafts to salaried workers; domestic funds transfer; outboard motor assistance; hire purchase; international funds transfer. Their Susu Savings and Credit Scheme remain outstanding since 1995 assisting diverse range of economic units across five districts. The vision of Ahantaman Rural Bank Limited is to be the preferred financial institution in Ghana offering utmost satisfaction to all its stakeholders. The mission of the bank is to be the leading rural bank in the country. The Core Values of the bank are as follows: teamwork, integrity, result oriented, customer satisfaction, confidentiality, and innovation. For the past years, the bank has helped in improving education in the Ahanta West District and in partnership with District Office of Education organizing BECE mock examination for both public and private schools. In 2013 the bank provided financial support to 106 brilliant but needy students.

The bank is now competing with two other banks and other microfinance institutions in the district, namely Agricultural Development Bank and GN Bank Limited. Ahantaman Rural Bank has been dominating in the district till the establishment of these two Banks about four years ago which have made a lot of their customers to switch their banks to either Agricultural Development Bank or GN Bank.

1.3 Statement of problem

Clearly, if the bank can sustain their vision, the management has to continuously research into the factors that make customers unhappy and understand customer switching behavior. In order to minimize customer switching and enhance customer loyalty, researchers are now researching into factors that influence customer switching behavior (Keaveney (1995); Colgate and Hedge (2001); Matthews and Murray (2007)).
However, most banking studies on customer retention or customer switching behavior focused narrowly on customer loyalty, customer satisfaction and other switching factors without attempting to link them in a model to further explore or explain customer retention and predict switching intention (Colgate et al. (1996); Lymperopoulos et al. (2006)). If retention criteria are not well managed, customers may still leave their banks, regardless of how hard bankers try to retain

them.

Therefore, it is against this background that the study seeks to design a model that have power to predict a customer in terms of switching so that tailored or customized services can be designed for the customers.

1.4 Rational for the study

It is the expectation of every banking institution to reduce customer switching and gain more customers thereby maximizing profit. This expectation can be achieved if the banks understand and know what their customers want and ready to welcome inputs or suggestion offer by their customers.

This research is therefore aimed at assessing factors that determine customer retention or switching at Ahantaman Rural Bank from customer's point of view and making predictions for future switching in order for management to prepare for the future and most importantly classify customers into likely switchers or unlikely switchers.

1.5 The objectives of the study

The objectives of the study are to use logistic regression to:

i. examine factors that influence bank customers' switching behavior in Ahantaman Rural Bank.

ii. model customer switching intentions in Ahantaman Rural Bank.

iii. predict customer switching.

1.6 Justification

The research is intended to interview customers of Ahantaman Rural Bank and it is intended to help management of the bank to understand the impact of service attributes on customer satisfaction, customer loyalty and decision makers in the banking industry on how to predict customers.

1.7 Scope and Limitations

This study covers the operations of Ahantaman Rural Bank at Ahanta West District. A research of this nature cannot be carried out without some limitations. Because of financial constraint, the study took a critical look at only the customers of Ahantaman Rural Bank in the Ahanta West District in the Western Region. In view of this, the findings from this research may not apply to other Banks in the District since they may have their own peculiar characteristics.

1.8 Methodology of the study

In this study, both descriptive and inferential research types were adopted. A questionnaire was used to collect data after the ethical clearance was given by the bank. In all 476 customers were interviewed of which 75% of the sample was used as training set and 25% used as validation set. The main statistical tools used in analyzing the data is the logistic regression. Bivariate and multivariate distributions of customer switching intentions were carried out to examine the relationship between customer switching intention and their covariates. Classification performance of logistic regression model for the training set and validation were compared. Respondents' data was analyzed using SPSS version 20. Libraries and Internet were used to get vital information on related research and publications..

1.9 Organization of the study

The study is organized into five chapters.

Chapter 1 covers the background of the study, statement of the problem, objectives of the study, methodology, justification of the study and organization of the thesis. Chapter 2 contains review of literature and methods. Chapter 3 is devoted to the research methodology of the study. Chapter 4 contains the data analysis and results and Chapter 5 examines conclusions and recommendations.

1.10 Summary

Logistic regression is one of many statistical tools that can be used to classify cases into groups. One of the objectives of logistic regression is to use explanatory variables to predict a new case. The method is preferred to other methods by many researchers because it is very robust in nature. For this study, the tool was used to predict switching behavior of Ahantaman Rural Bank customers.Until recently, the bank has been the only bank in the Ahanta West District of Ghana, with the emergence of Agricultural Development Bank and GN Bank, most of their customers are exiting to these two new banks in the district.The bank was used as a case study because of keen competition it now faces and most of its customers are complaining bitterly about their services. The goal of this study is to use statistical model to find out what factors influence the customers to exit and use the same tool to predict the customers so that tailored or customized services can be designed for the them in order to reduce the number of customers switching services with the bank.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This section is made up of two parts. The first part delves into concepts, factors and theories that are relevant to this study. Some of these factors were used as explanatory and response variables in the research. A brief explanation of the various abstract concepts that the study explores as operational indicators will help make the issues raised in the study clear. It focuses on factors like customers' satisfaction, customer loyalty, the employee-customer relationship, bank customers' switching behavior among others.

The second part consists of empirical studies on the topic; views and findings of different writers on similar research as documented in books and Journals.

2.2 Exploration of basic concepts

2.2.1 Switching Behavior and its Impact

According to Stewart (1998) switching behavior is customer defection or exit. Switching behavior involves a customer stopping to do business with particular service provider based on factors best known to him or her. In banking sector, customer switching simply means that a customer moves from one bank to another for banking transactions. The banking industry is so competitive that, the era where bankers sit in their comfort zones for customers to walk in and open accounts and transact business is no more. The industry today is a survivor of the fittest, for a bank to survive today you need to reduce customer switching and increase your customer base. According to Reichheld (1996) banks' earnings and profits decrease when customers switch, old customers generally generate more business than new ones. Sustaining your customers for longer period can minimize the negative effects of defection (Matthews and Murray (2007)). Many researchers have come out with factors that influence or determine customer switching behavior, price (Mavri and Ioannou (2008)), reputation, service quality, customer satisfaction (Taylor and Baker (1994)), responses to service failure (Hirschman (1970)), customer commitment (Dube and Shoemaker (2000)), advertising (Blanchard and Galloway (1994)).

2.2.2 Customer Satisfaction

Customer satisfaction is now an anthem or popular phrase in most business mission statements and has been defined in many ways.

According to Hunt (1977) the pleasure of the experience is not satisfaction, it is an evaluation of experience that matches what it is supposed to be. According to Taylor and Baker (1994) customer satisfaction is often seen as a main motivator of customers' next purchase intention. When a customer is satisfied repeat purchase is guaranteed. According to Fornell (1992) customer satisfaction stops one from exiting, enhances customers' loyalty and lower the bank's switching rate. In banks, the customers compare the services of various banks by asking the customers of other banks, when they perceive that they are not getting value of their money, they then decide whether to stay or exit. A customer is always satisfied when he/she perceives that he or she is getting value of his or her money.

2.2.3 Service Quality

According to Avkiran (1994) establishment of many banks these days and the same nature of banking services have made service quality to come out as a very important factor for combating customer switching in today's competitive banking environment. According to Kamilia and Jacques (2000) service quality is the difference between perceptions for the service received by the customer and customer expectations of the service. The quality is high when the expectation exceeds the perception. Parasuraman et al. (1988) said service quality is the discrepancy between service perception and its expectation. In their definition they identified assurance, empathy, responsiveness, reliability and tangibles as the five drivers of service quality and have been used in comparing banking service quality. According to Bitner et al. (1994) service reliability is considered as the single most important factor used by customers to judge quality of service. According to Oliva and Sterman (2001), customers consider many factors when evaluating service quality. They established that customers look at the process or how services are delivered and rendered, and service is considered quality when perception exceeds expectation.

2.2.4 The Employee-Customer Relationship

A strong influence on customer satisfaction is assured when an employee is in constant contact with external customers. According to Rust et al. (1996) the relationship between employee satisfaction and good customer-employee relation is directly related, because when employee is well treated employee-customer relationship is enhanced thereby increasing satisfaction. Empowering employees to establish good relationship with customers is very important because bank employees deal directly with the customers. It is the duty of every service provider to know what their customers want and how to deliver it for maximum satisfaction, feeling of satisfaction and accomplishment is attained when an employee knows that he has been able to fulfill it. This can be achieved only when there is good employee-customer relationship. Good employee-customer relationship is established when a customer perceives that he or she has been treated specially.

2.2.5 Effective Complaint Management

Consumers' decision making on re-purchase is enhanced when complaints are managed effectively. According to Wimmer (1985) Effective Complaint Management includes the result-oriented planning, execution plan, and control of all avenues available to a business for dealing with customer complaints. Its purpose is to resolve issues of dissatisfied customers. In reality, through complaints management of banks become aware of issues that bring customer dissatisfaction, motivates customers to come out with the reason of dissatisfaction, and customer confidence restored through implementation of specific rules and regulations. When customers are well satisfied, they become loyal customers and the risk of losing them to competitors is reduced. Effective Complaint Management is concerned with improving customer satisfaction, motivating customers to recommend the institution to others, and reducing customer switching which contribute positively to the firm's profit and margin goals. Addressing customer complaints is interpreted as a service that involves continuous interaction, where organizational psychological factors and individual largely determine the quality of the service process and its results.

2.2.6 Price

According to Gerrard and Cunningham (2004) price undoubtedly is a major factor that influences switching in banking sector. Price is an attribute that must be paid to obtain certain kinds of services or products (Zeithaml (1998)). In the banking sector, perceived 'high price' is multifaceted dimension that needs to be addressed properly. When a customer perceives that he is paying 'high price' for a service that can be obtained somewhere lesser, that customer is likely to switch (Beckett et al., 2000). For example, in the banking industry, price includes fee implementation, bank charges, interest rates charged and paid (Gerrard & Cunningham, 2004).

It has been established empirically that, price is a vital factor that influenced customers' switching intentions (Stewart, 1998; Keaveney, 1995). Gerrard and Cunningham (2004) show that pricing seems to influence switching behavior among bank customers more than customers of other services. According to Colgate and Hedge (2001) the major factor that influences switching behavior in the banking sector is pricing.

2.2.7 Customer Loyalty

According to Teich (1997) loyalty is enhanced over a period of time from regular meeting with a customer. It has been established that keeping a loyal customer is cheaper than attracting or getting a new customer (Kotler et al. (1999)). It takes a lot of expensive adverts and sale promotions to lure or entice a new customer. Gremler and Brown (1996) explained customer loyalty as the extent to which a customer shows repeat purchasing intentions, exhibits a positive attitudinal habit toward a service provider, and considers using only this provider in a near future. Loyalty to a bank is seen as continuing patronage over time even when there is a problem. Customer loyalty is important to every bank if they want to combat customer switching since service failure is inevitable in banking industry. It is a customer loyalty that will make a customer to stay when the inevitable service failure occurs.

2.2.8 Empirical Studies on Customer Switching Behavior in the Banking Sector

Customer switching intention is a growing research area in competitive marketing. Although it has been documented that pricing and customer satisfaction are major factors influencing customer switching intent, researchers are now moving away from finding underlying factors that influence switching to the use of model to solve the problem.

Keaveney (1995) examined customer switching behavior in banks using generalized model. Eight constructs: pricing construct, inconvenience construct, core service failure construct, service encounter failure construct, response to service failure construct, ethics construct, competition construct, and involuntary switching were used in modeling and found four factors: service failure response, pricing, ethics and competition to be significant. Keaveney's (1995) research has faced a lot of criticism. Mittal et al. (1998) established that the certain characteristics of switching intention in banking may be hidden when generalized models are used directly. According to Colgate and Hedge (2001), Keaveney's (1995) switching model overestimated the relative weight of his factors on a customer's intentions to switch banks or service providers. Therefore, further research need to be carried out to assess the essence of using Keaveney's (1995) generalized switching model to the banking sector.

Mavri and Ioannou (2008) used survival analysis to investigate switching intentions in Greek banking industry. Their main purpose was to find factor that influence customer switching behavior. They examined a lot of predictors which represent characteristics of the customers and banking services and products. They used life tables to estimate the partial contribution of each explanatory factor or variable in customers' switching behavior in different periods of time. In their findings, hazard proportional model was used to determine the risk of switching behavior, which is the end-result of all the examined factors. The method employed by Mavri and Ioannou (2008) lacked predicting power since hazard proportional model performed better when used to model risks. However, their findings were appropriate for banking industry when they want to determine specific factors that influence switching behavior.

In a similar research conducted by Lariviere and Van den Poel (2004), the same proportional hazard model was used to analyze customer attrition for European financial services. Their main purpose was the same as what Mavri and Ioannou (2008) investigated; investigating explanatory of churn or switching incidence as part of customer relationship management. They combined several different types of independent variables into constructs and analyzed to see the relationship between the covariates and actual customer behavior. They concluded that: (1) demographic characteristics, environmental changes and stimulating 'interactive and continuous' relationships with customers are of major factors that influence customer switching; (2) the explanatory variables partially contributed to attrition in terms of total products owned as well as the inter purchase time. However, statistical tool used by Lariviere and Van den Poel (2004) lacked predictive power and would be more appropriate in calculating risks.

In another studies carried out by Licata and Chakraborty (2009), their motive was to find the relation between price, customer satisfaction, the cost of switching and switching in the banking sector using MANOVA. They examined the differential influence of three dimensions on customer switching. The three dimensions were price, customer satisfaction, and the cost of switching service providers which were used as independent variables. Commitment of people for the service, behavioral response and institution's commitment were the factors of customer switching used as response variables. An additional objective is to investigate whether any of these variables, age or depth of service relationship is affected by the pattern of influence. In their findings all three factors showed significant influence on customer switching. The commitment to people dimension was influenced by all three drivers. The behavioral response dimension was influenced by stake and satisfaction. The commitment to the institution dimension was influenced by value of switching and satisfaction. In their conclusion customer satisfaction was found to be the major factor impacting on customer switching. The approach or method used is exploratory in nature, all that MANOVA does is to combine the multiple response variables in a linear manner to produce a combination which best separates the explanatory variable groups. It establishes relationship between group of dependent variables and independent variables.

Ghouri et al. (2010) investigated factors that influence switching intentions in Private Banking Sector of Pakistan. Their study combined six factors (Price, Service Quality, Involuntary Switching, Effective Advertising Competition, Reputation, and Switching Cost) of customer switching which affects retail banking operations in Pakistan. The primary goal of their research was to find the determinants of customers' switching, and determine major and minor factors that influence customers' switching behavior. A total of 302 customers were interviewed and their findings revealed that all variables impacted significantly on customer switching. However, 'service quality' and 'reputation' identified as major and minor influential factors respectively on customer switching.

In a research conducted by Santonen (2007) on the topic "Price sensitivity as an indicator of customer defection in retail banking", he analyzed how customer switching intentions in retail banking can be influenced by different dimensions of perceived service loyalty including price sensitivity. In his approach nearly 1,750 customers were interviewed in Finland and responses analyzed using exploratory data analysis. The findings of Santonen (2007) supported previous suggestion that purchase construct, word-of-mouth communication construct, price sensitivity construct and complaining behavior influenced service loyalty. The study revealed that only the price sensitivity construct impacted on defection in the case of low-price and limited product range driven sales offers. However, their research did not come out with any model and cannot be used to predict like others.

Garland (2002) used Juster's probability scale to model customer switching in personal retail banking; the study examined a sample of 881 customers who had declared their intention to defect from their current bank. These customers were contacted after 12 months to see whether they have changed their switching intentions and comparison made between intended and actual defection. Garland's methodology combined Juster scale's performance in a "subscription" type market and identification of those customers with a predisposition to switch banks. The customers were asked to declare their likelihood of switching services with their current in the next twelve months. The sample estimate was 10 percent while the actual defection result was 5 percent. This approach has been condemned by a lot of researchers that it is inaccurate and tends to over-estimate switching behavior, but its estimates have consistently proved superior previously to those derived from attitudinally based intention-to-buy scales.

In a popular study conducted by Colgate and Hedge (2001), 694 banking customers in New Zealand and Australia were sampled and interviewed. Their objective was to find out why customers switch banks by investigating problems that have impact on both customer switching and complaints made before they switch. The factors for switching were put into 3 constructs: services denial construct; service failure construct; pricing construct. Their research revealed that pricing construct was major determinant of customers' intentions to switch. They also established that customers complain about service failures before exiting. In one of their submissions, they said customers at times keep mute of their problems that influenced them switch services.

2.2.9 Summary

In the light of the reviewed literature above, it is sufficient to conclude without doubt that customer switching behavior is determined by certain factors that are likely to cause a decrease in customer attitudes. Most of the factors used by researchers in determining switching behavior in the reviewed literature above are service related and may not be applicable to some banking customers because apart from these factors, there are other unmeasured factors like geographical location which makes it impossible to apply their findings to other banks. However, it helps researchers to have a fair knowledge on how to choose the explanatory or independent variables when researching into customer switching.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The review of statistical or mathematical methods employed in the analysis of the customer switching behavior data is presented in this chapter. However, before we discuss the methodology, it is vital that one acquires the best of knowledge and understanding about the theoretical and conceptual framework of the statistical method that is used to analyze the data efficiently and effectively. Therefore this chapter focuses on the theoretical and conceptual framework on reliability of the measuring instrument, bivariate analysis, logistic regression and multinomial regression.

3.2 Reliability of Measuring Instrument

Assessing the degree of consistency between multiple response variables, reliability checks proved to be superior. According to Cooper and Schindler (2006) one of numerous ways to check internal consistency in items is by the use of reliability test. Cronbach's Alpha test has been used widely to measure internal consistency of questionnaire items, and according to Churchill (1979), Cronbach's Alpha is considered significant if its value exceeds 0.60. Churchill (1979) advised that consistency in survey instrument should be checked before any further data analysis. The study employed Cronbach's Alpha test to check the internal consistency of the questionnaire.

3.3 Bivariate Analysis

In bivariate analysis, "association" and “causality” are tested. Association refers to relationship between variables which will help one variable to explain the other variable. In other words how two variables correlated or related. A measure of association helps researchers to get a clear picture of relationship that exist in variables. The range of the measure of association is between -1 and 1. The sign of the integer represents the "direction" of association and an integer closer to the limits represent perfect association.

3.4 Logistic Regression

Logistic regression is one of many multivariate statistical methods or tools available for analyzing data with categorical dependent variable. In statistics, the method is used to analyze data where the dependent variable is categorical in nature and independent variables can be categorical and metric in nature or both. The purpose of logistic regression is to group outcome on bases of some variables by estimating probabilities using a logistic function, which is the cumulative logistic distribution. In logistic regression the response variable can be binary, ordinal or multinomial in nature. In binary logistic regression the dependent variable has two possible outcomes (success or failure, presence or absent). In multinomial logistic regression the response variable is nominal in nature and has more than two unordered possible outcomes. Ordinal logistic regression deals with ordered response variable.

The binary logistic regression model the natural log of the odds ratio called logit (p).

$$logit(p) = \ln(\frac{p}{1-p}) \tag{3.1}$$

The logit is a function of the probability p. In the simplest model, we assume that the logit graphs as a straight line in the predictor variable X so

$$logit(p) = \ln(\frac{p}{1-p}) = \beta_0 + \beta_1 X_1 + ...\beta_k X_k \tag{3.2}$$

In other words, the log odds are linear in the predictor variable. Because it is easier for most people to think in terms of probabilities, we can convert from the logit or log odds to the probability p. By first exponentiating equation (3-2) we obtain

$$\frac{p}{1-p} = e^{\beta_0 + \beta_1 X_1 + ...\beta_k X_k} \tag{3.3}$$

Next solving for p, we obtain

$$p(t) = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}} \tag{3.4}$$

which describes a logistic curve. Equation (3-4) is called the logistic function. P (t) is interpreted as the probability of a case of interest (success).The relation between p (t) and the predictors 't' is not linear but has an S-shaped graph as illustrated in Figure 3.1

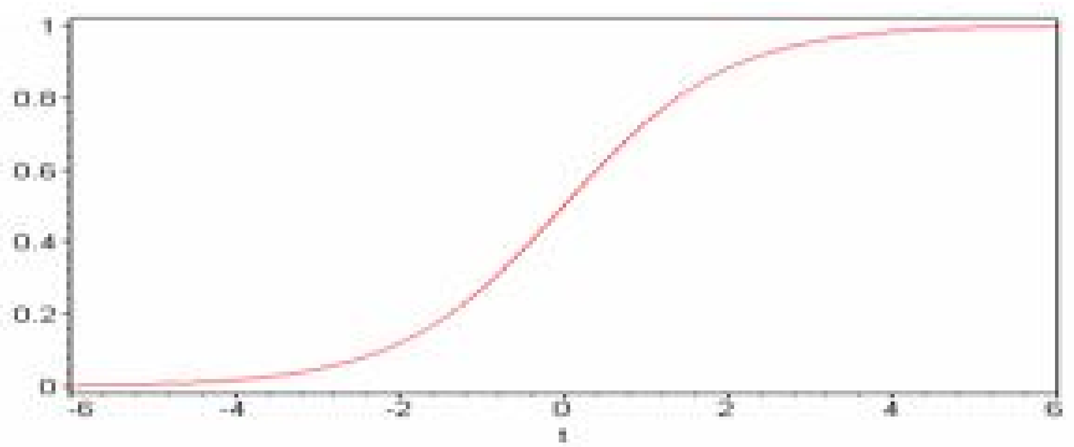

Figure 3.1: The logistic function, with t on the horizontal axis and p (t) on the vertical axis The input is t and the output is p (t).

The logistic function is useful because it can take as an input any value from negative infinity to positive infinity, whereas the output is confined to values

between 0 and 1. The variable t represents the exposure to some set of independent variables, while p (t) represents the probability of a particular outcome, given that set of explanatory variables. The variable t is a measure of the total contribution of all the independent variables used in the model.
The variable t is usually defined as

$$t = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k$$

where $B(0)$ denotes the "intercept" and Bk the logistic coefficients of Xk The intercept denote the value of t when the values of all predictor or independent variables are zeros. Each of the logistic coefficients represents the partial contribution of that variable to the model. A positive logistic coefficient denotes that the explanatory or predictor variable increases the probability of the outcome, while a negative logistic coefficient means that the variable decreases the probability of that outcome; a large logistic coefficient shows that the risk factor strongly influences the probability of that outcome, while a near-zero logistic coefficient shows that the risk factor has little influence on the probability of that outcome.

3.5 Sample Size-Dependent Efficiency

In logistic regression when the sample size is small the beta coefficients or logistic coefficients tends to be systematically overestimated. The sample size is inversely related to magnitude of overestimation, and the estimated odds ratio asymptotically approaches the true population value when the sample size increases.
According to Hosmer and Lemeshow (2000), minimum of ten cases per variable or parameter can solve any problem associated with overestimated and sample size. Even though the sample size satisfy the minimum number of events per parameter as suggested by Hosmer and Lemeshow (2000), in order to make generalizations

and draw appropriate inferences, there was the need to conduct the study using optimal sample size, that is, a sample size that can represent the population of the bank customers. Due to constraints of time, efforts and resources it became virtually impossible for one to use the entire population.

The researcher considered the optimal sample size that fulfilled the requirements of efficiency, representativeness, reliability, and flexibility. A method proposed by Yamane (1996) was used to compute the sample size based on the accessible population of 833 total customers of the bank in two branches in Agona Nkwanta in the Ahanta West district. Sample size of 476 was obtained from the calculation;

$$n = \frac{N}{1 + e^2} \tag{3.5}$$

n denotes the sample size, N represents the size of population and e represents standard error, always .03 is used. The sample size meet the requirement of logistic regression. The detailed sample size calculation is shown at the appendix B.

3.6 Formal Mathematical Specification

Logistic regression analyzes binomially distributed data of the form

$$Y_i \sim \beta(n_i, p_i), fori = 1, , , , , m. \tag{3.6}$$

where the number of Bernoulli trials ni are known and the probabilities of success pi are unknown. The model proposes for each trial i there is a set of predictor variables that might inform the final probability. These predictor variables can be thought of as being in a k-dimensional vector Xi and the model then takes the form

$$p_i = E(\frac{y_i}{n_i}|Xi) \tag{3.7}$$

The logit is modeled as a linear function of the Xi.

$$logit(p_i) = \ln(\frac{p_i}{1-p_i}) = \beta_0 + \beta_1 X_{1i} + ...\beta_k X_{ki} \tag{3.8}$$

$$where\ odds\ ratio = \frac{p_i}{1-p_i} \tag{3.9}$$

Note that a particular element of Xi can be set to 1 for all i to yield an intercept in the model.

The unknown parameters Bj are usually estimated by maximum likelihood using a method common to all generalized. The model has an equivalent general formulation

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + ...\beta_k X_{ki})}} \tag{3.10}$$

3.6.1 Assumptions and Data Requirements of Logistic Regression

1. The logit of the independent variables and dependent variables are assumed to be linearly related.
2. The large sample is needed to avoid the problem of overestimation.
3. Normal distribution is not assumed for the dependent variable and errors
4. The independent variables need not be metric or continuous
5. Logistic regression is sensitive to multicollinearity. Under normal circumstances the dependent variable should be correlated with independent variables.
6. Logistic regression is very sensitive to outliers.

3.6.2 The Hosmer-Lemeshow Test

Hosmer and Lemeshow (1998) proposed grouping based on the values of the estimated probabilities. The test compares the observed event rates with expected event rates in subgroups of the model population. Models for which expected and observed event rates in subgroups are similar are called well calibrated. The Hosmer–Lemeshow test statistic is given by:

$$H = \sum_{g=1}^{n} \frac{(O_g - E_g)^2}{N_g \pi_g (1 - \pi_g)} \tag{3.11}$$

where Og, Eg, Ng, and g denote the observed events, expected events, observations, predicted risk for the gth risk decile group, and n is the number of groups. The statistic follows a Chi-square distribution with *n-2* degrees of freedom.

3.6.3 Likelihood Ratio Test

Likelihood ratio test is a statistical test that compares how two competing models fit a given data where one of the model (the null model) should be part of the other model (the full model). The test is based on the likelihood ratio, which expresses how likely the data are under one model than the other. The ratio can be used to compute a p-value to be used to compare with a critical value to decide whether to reject the null model in favor of the alternative model. The probability distribution of this test statistic, assuming that the null model is true, can be approximated using Wilks' theorem. In Likelihood Ratio Test, the model with constant only and the full model, is fitted to the data separately and the log-likelihood recorded. The test statistic (often denoted by D) is twice the difference in these log-likelihoods:

$$D = -2\ln\left(\frac{likelihood\ for\ null\ model}{likelihood\ for\ alternative\ model}\right) \tag{3.12}$$

$$D = -2\ln(likelihood\ for\ null\ model) + 2\ln(likelihood\ for\ alternative\ model)$$

In many cases, the probability distribution of the test statistic is approximately a chi-squared distribution with degrees of freedom equal to df2 - df1, if the nested model with fewer parameters is correct. The degree of freedom for models 1 and 2, the null model and the alternative model are denoted by df1 and df2 respectively.

3.6.4 Maximum Likelihood Estimation of the Logit Model

The coefficients of logit model is estimated using the Maximum Likelihood Estimation. The function (L) measures the probability of observing the particular set of dependent variable values (P1, P2, ..., Pn) that occur in the sample:
L = Prob (P1* P2* * * Pn).
Maximum Likelihood Estimation involves getting the coefficients (,) that makes the log of the likelihood function as large as possible or finds the coefficients that make -2 times the log of the likelihood function (-2LL) as small as possible.

3.6.5 The Logit Model with One Explanatory Variable

Let (Y1, X1),..., (Yn, Xn) be a random sample from the conditional logit distribution:

$$Pr[Y_j = 1/X_j] = \frac{1}{1 + exp(-\alpha_0 - \beta_0 X_j)} \tag{3.13}$$

$$Pr[Y_j = 0/X_j] = 1 - Pr[Y_j = 1/X_j]$$

$$= \frac{exp(-\alpha_0 - \beta_0 X_j)}{1 + exp(-\alpha_0 - \beta_0 X_j)}$$

where the Xj's are the independent variables and alpha and betta are unknown parameters to be estimated. This model is called a Logit model, because

$$P_r[Y_j = 1/X_j] = F(\alpha_0 + \beta_0 X_j) \tag{3.14}$$

where

$$F(X) = \frac{1}{1 + exp(-X)} \tag{3.15}$$

is the distribution function of the logistic (Logit) distribution. The conditional probability function involved is

$$f(y/X_j, \alpha_o, \beta_0) = Pr[Y_j = y/X_j]$$

$$= F(\alpha_0 + \beta_0 X_j)^y (1 - F(\alpha^0 + \beta_0 X_j))^{1-y}$$

$$= \begin{cases} f(\alpha_0 + \beta_0) & \text{if} y = 1, \\ 1 - F(\alpha_0 + \beta_0 X_j) & \text{if} y = 0 \end{cases} \tag{3.16}$$

Now the conditional log-likelihood function is

$$\ln(Ln(\alpha, \beta)) = \Sigma_{j=1}^{n} \ln(f(Y_j/X_j\alpha, \beta))$$

$$= \Sigma_{j=1}^{n} Y_j \ln(F(\alpha +_j)) + \Sigma_{j=1}^{n} (1 - Y_j) \ln(1 - F(\alpha + \beta X_j))$$

$$= -\Sigma_{j=1}^{n} (1 - Y_j)(\alpha + \beta X_j) - \Sigma_{j=1}^{n} \ln(1 + exp(-\alpha - \beta X_j)) \tag{3.17}$$

$$E[\ln(Ln(\alpha, \beta))/X_1, ...X_n] \leq E[[\ln(Ln(\alpha_0, \beta_0))/X_1, ...X_n]$$

By maximizing In(Ln(α, β)) $\rightarrow$ α β : $\ln(L_n(\hat{\alpha}, \hat{\beta}))$=$\max_{\alpha,\beta} \ln(L_n(\alpha, \beta))$, $\alpha_0 and \beta_0 are estimated.$ These coefficients can be estimated by $most$ statistics $software.$

3.6.6 Pseudo t-Values

If the sample size n is large then it can be shown that

$$\sqrt{n}(\hat{\alpha} - \alpha_0) \sim N(0, \sigma_\alpha^2), \sqrt{n}(\hat{\beta} - \beta_0) \sim N(0, \sigma_\beta^2)$$

Given consistent estimators $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_\beta^2$ of the unknown variances σ_α^2 and σ_β respectively (which can be computed by software), we then have

$$\frac{\sqrt{n}(\hat{\alpha} - \alpha_0)}{\hat{\sigma}_\alpha} \sim N(0,1), \frac{\sqrt{n}(\hat{\beta} - \beta_0)}{\hat{\sigma}_\beta} \sim N(0,1) \tag{3.18}$$

These results can be used to test whether the coefficients alpha and beta are zero or not. In particular the null hypothesis

$$\beta_0 = 0$$

is of interest, because this hypothesis implies that the conditional probability

$$Pr[Y_j = 1 | X_j]$$

does not depend on Xj. Under the null hypothesis

$$\beta_0 = 0$$

we have

$$\hat{t}_\beta = \frac{\sqrt{n}\hat{\beta}}{\hat{\sigma}_\beta} \sim N(0,1) \tag{3.19}$$

For example at alpha level of 5%, the null hypothesis

$$\beta_0 = 0$$

will not be accepted at the 5% significance level in favor of the alternative hypothesis

$$\beta_0 \neq 0 \; if |\hat{t}_\beta| > 1.96$$

and accepted if

$$|\hat{t}_\beta| \leq 1.96$$

The statistic

$$\hat{t}_\beta$$

denotes the pseudo t-value of

$$\hat{\beta}$$

because it is used in the same way as the t-value in linear regression and

$$\hat{\alpha}_\beta \; is \; called \; the \; standard \; error \; of \hat{\beta}$$

3.6.7 The General Logit Model

The general Logit model takes the form

$$Pr[Y_j = 1|X_{1j}, ..., X_{kj}] = \frac{1}{1 + exp(-\beta_1^0 X_{1j} .....\beta_k^0 X_{kj})} = \frac{1}{1 + exp(-\sum_{i=1}^k \beta_1^0 X_{kj})}$$

where one of the Xij equals 1 for the constant term, for example, let Xkj = 1, and the beta(0i)'s are the true parameter values. This model can be estimated by ML in the same way as before. Thus, the log-likelihood function is

$$\ln(Ln(\beta_1, ...\beta_k)) = -\sum_{j=1}^n (1 - Y_j) \sum_{j=1}^k \beta_i X_{1j} - \sum_{j=1}^n \ln(1 + exp(-\sum_{i=1}^k \beta_i X_{ij})) \quad (3.20)$$

and the ML estimators *B1,...Bk* are obtained by maximizing

$$\ln(Ln(\beta_1, ..., \beta_k)) :$$

$$\ln(Ln(\hat{\beta}_1, ..\hat{\beta}_k) = \beta_1, ...\beta_k max \ln(Ln(\beta_i, ...\beta_k))$$

Again, it can be shown that if n is large then for i = 1, ..., k,

$$\sqrt{n}(\beta_i - \beta^0) \sim N(0, \sigma_i^2)$$

Given consistent estimators

$$\hat{\sigma}_i^2\ of\ the\ variances \sigma_i^2,\ it\ follows\ then\ that$$

$$\frac{\sqrt{n}(\beta_i - \beta_i^0)}{\hat{\sigma}_i} \sim N[0, 1]$$

3.6.8 Testing Joint Significance

Now suppose you want to test the joint null hypothesis

$$H_0 : \beta_1^0 =, \beta_2^0 = 0, ..., \beta_m^0 = 0, \tag{3.21}$$

where m <k. These can be done in two ways. One way is similar to the F test in linear regression: Re-estimate the Logit model under the null hypothesis:

$$\ln(Ln(0, ...0, \hat{\beta}_k)) = \beta_{m+1}, ...\beta_k max \ln(0, ...0, \beta_{m+1}, , ...\beta_{kj})$$

and compare the log- likelihoods sqaure. It can be shown that under the null hypothesis (3-21) and for large samples,

$$LR_m = -2\ln(\frac{L_n(0, ...0, \hat{\beta}_{m+1}, ...\hat{\beta}_k)}{L_n(\hat{\beta}_1, ...\hat{\beta}_k)}) \sim \chi_m^2,$$

where the degrees of freedom m corresponds to the number of restrictions imposed under the null hypothesis. This is the so-called likelihood ratio test, which is conducted right-sided.

3.6.9 Interpretation of the Coefficients of the Logit Model(Using Marginal Effects)

Consider the Logit model (3-13). If

$$\beta_0 > 0 \; then$$

$$Pr[Y_j = 1|X_j] = F(\alpha_0 + \beta_0 X_j)$$

is an increasing function of Xj :

$$\frac{dp[Y_j = 1|X_j]}{dX_j} = \beta_0 F^I(\alpha_0 + \beta_0 X_j)$$

where F prime is the derivative of

$$F(x) = \frac{1}{1 + exp(-x)} \; in \; (3 - 13) \; above$$

$$F_I(X) = \frac{exp(-x)}{(1 + exp(-x))^2} = \frac{1 + exp(-x)}{(1 + exp(-x))^2} - \frac{1}{(1 + exp(-x))^2}$$

$$\frac{1}{1 + exp(-x)} - \frac{1}{(1 + exp(-x))^2} = F(x) - F(x)^2 = F(x)(1 - F(x))$$

Therefore, the marginal effect of Xj on Pr [Yj = 1|Xj] depends on Xj :

$$\frac{dP[Y_j = 1|X_j]}{dX_j} = \beta_0 F(\alpha_0 + \beta_0 X_j)(1 - F(\alpha_0 + \beta_0 X_j)) \tag{3.22}$$

which renders the interpretation of *B0* difficult. However, the coefficient *B0* can be interpreted in terms of relative changes in odds.

3.6.10 Odds and Odds Ratios

The odd's is the ratio of the probability of true cases divided by the probability of not true cases. Thus, in the Logit case (3-2),

$$odds(X) = \frac{Pr[Y_j = 1|X_j]}{Pr[Y_j = 0|X_j]} = \frac{F(\alpha_0 + \beta_0 X_j)}{1 - F(\alpha_0 + \beta_0 X_j)} = exp(\alpha_0 + \beta_0 X_j) \qquad (3.23)$$

The odds ratio is the ratio of two odds for different values of Xj, say Xj = x and Xj = x + Δx :

$$\frac{Odds(x + \Delta x)}{Odds(x)} = \frac{exp(\alpha + \beta x + \beta \Delta x)}{exp(\alpha + \beta x)} = exp(\beta \Delta x)$$

where Δx *is a small change in* x. *Then*

$$\lim_{\Delta x \longrightarrow 0} \frac{1}{\Delta x}(\frac{Odds(x + \Delta x) - Odds(x)}{Odds(x)}) = \lim_{\Delta x \longrightarrow 0} \frac{exp(\beta_0 \Delta x) - 1}{\Delta x}$$

$$= \beta_0 \lim_{\beta_0 \Delta x \longrightarrow 0} \frac{exp(\beta_0 \Delta x) - 1}{\beta_0 \Delta x} = \beta_0 \frac{dexp(u)}{du}|u = 0$$

$$\beta_0 exp(0) = \beta_0$$

Thus, *B0* may be interpreted as the relative change in the odds due to a small change (delta x) in Xj:

$$\frac{Odds(x + \Delta x) - Odds(x)}{Odds(x)} = \frac{Odds(x + \Delta x)}{Odds(x)} - 1 \approx \beta_0 \Delta x_j \qquad (3.24)$$

For example, if Xj is a binary variable itself, Xj = 0 or Xj = 1, then the only reasonable choices for x + delta(X) and x are 1 and 0, respectively, so that then

$$\frac{Odds(1)}{Odds(0)} - 1 = \frac{Odds(1) - Odds(0)}{Odds(0)} = exp(\beta_0) - 1$$

Only if *B0* is small we may then use the approximation, exp (*B0*) - 1 approximately to ***B0***. If not, one has to interpret *B0* in terms of the log of the odds ratio involved:

$$\ln \frac{Odds(1)}{Odds(0)} = \beta_0$$

The interpretation of the coefficients,

$$\beta_i^0 i = 1......k - 1$$

in the general Logit model (3-20) is similar as in the case (3-24):

$$\frac{Odds(X_{1j}, ...X_{i-1},, X_{ij} + \Delta X_{ij}, X_{i+1j}, .., X_{kj})}{Odds(X_{1j}, ...X_{i-1j}, X_{ij}, X_{i+1j}, .., X_{kj})} - 1 \approx \beta_i^0 \Delta X_{ij}$$

3.7 Model Validation

Another way of reducing biasness in model and evaluating the performance of classification model is through model validation. In model evaluation, the model is validated to see the extent to which it measures what was intended. There are many methods available in validating a model: Partitioning the sample, The Hold-out method, The Cross-Validation method, Leave-one-out cross-validation method, etc.

3.7.1 Partitioning the Sample

In this method the data is partitioned into two: training and validation set. The training sample is used to create the classification model. The model is then ran on the validation sample to calculate the apparent error rate. The resulting error rate in this method is unbiased because two different samples are used. For this method to obtain optimal result, large data set is needed so that part can be used as training set and validation set. To minimize the variance of our error rate estimate it is appropriate to use the entire data set to construct the classification function.This method was adopted for the study because of its easiness to apply and available in SPSS version 20.

3.7.2 Leave-One-Out Cross-Validation

This method involves using one case from the original sample as the validation set, and the remaining observation as the training set. The model is then ran on the validation set, this is repeated such that each observation in the sample is used once as the validation data. An error rate is calculated for each case and average error is computed and used to evaluate the model. The actual classification rule for future observations would be based on all observations in the sample. SPSS has this in one of its output for both logistic regression and linear discriminant analysis.

3.7.3 The Cross-Validation Method

In this method, *n combination k set* are obtained. Each combination is designed by choosing k of the observations as a training set, and its error rate is estimated using the remaining (n-k) observations as validation set. This process is repeated for all distinct choices of k patterns and the average of the error rates is computed. The average of the error rate of each combination is therefore an estimate of the error rate of the combination. This type of validation is, of course, more expensive computationally, but useful when the most accurate estimate of a combination's error rate is required. A popular choice for the value of k is k=1, yielding the well-known leave-one-out method.

3.7.4 The Hold-Out Method

The holdout method is a special case of partitioning the sample method. In this procedure, for example, 2/3 of the data is used as the training set to design the classification rule and the remaining 1/3 as the validation set is used to estimate the error rate and validating the classification rule. There are a problems associated with holdout method. For this method to be perfect in classifying cases, a large sample size is required which is hard to come by in real world situation, and there is a problem associated with the appropriate relative size of the training sample to the validating sample.

3.8 Method of Analysis

Logistic regression was employed in analyzing customer switching intention data. The first stage of data analysis was going through the questionnaire to check errors and any missing values that can affect analysis. The next stage was coding of the questionnaire items by assigning numbers to them. Exploratory data analysis was used to check whether all the data sets were valid. Reliability and Internal Consistency checks were carried out and the result satisfies the reliability criteria which Churchill (1979) recommended. Multicollinearity check was carried out to find out whether there is interaction or correlation between the independent variables. Outliers in the data were also checked using Standardized residuals.

After satisfying multicollinearity and removing excessive outliers, the next task was to estimate the parameters of the models. The model was assessed for fitness and validated using the validation set. The software used for the analysis were SPSS version 20. Again, all the tests were undertaken or variables measured at an alpha level of 0.05 or 5% level of significance.

3.9 The Data

The data used for the analysis were primary data obtained through a well-structured questionnaire, administered to only customers of Ahantaman Rural Bank at the Ahanta West District of Ghana. The survey randomly sampled 476 customers. The customers were interviewed in the banking hall of Ahantaman Rural Bank at their head office after the ethical clearance has been given. The customers who could read and write were given the questionnaire to tick the appropriate responses, and those who could not read, the questions in the questionnaire were read out to them and their responses were recorded by the interviewers. In all 476 customers were interviewed comprising 255 males and 221 females, six (6) cases were removed from the data because their standardized residuals for outliers exceeded cut-off point of 2.58. The entire data was divided into 2 set: training and validation sets. 353 cases of the data representing 75% were used as the training set to model the customer switching intentions and the remaining 117 cases representing 25% were used as a validation set to validate the classification model.

The sample partitioning in the SPSS was done using the following steps: click on Transform on the menu$\longrightarrow$ *Random* Number *Generators* $\longrightarrow$ select set starting point $\longrightarrow$ select fixed value and type 9191972 as the value $\longrightarrow$ click ok .

The above steps set the random seed that allows one to replicate the random selection of cases in this analysis. The next step is to create the selection variable for validation using the following: Choose Transform on the menu $\longrightarrow$ Compute Variable. Type validate in the Target Variable. $\longrightarrow$ text box$\longrightarrow$ type rv.bernoulli(0.7) in the numeric expression text box. The percentage of the sample to be used for modelling is now set to 70 %. The next steps set the values of validate to be randomly generated. Still in compute variable windows click on if$\longrightarrow$ Select Include if case satisfies Condition. Type MISSING(SwitchingBank)=0 as conditional expression.Click on continue and

Click OK in the Compute Variable dialog box. The switching bank in the MISSING bracket is the name of the dependent variable. This ensures that validate is only computed for cases with non-missing values for the dependent; that is, for respondents who are very likely to switch services with the bank or not. Approximately 70 percent of the sample will have a validate value of 1. These customers or respondents will be used to create the model. The remaining customers will be used to validate the model results. In running the analysis for logistic regression, select validate as the selection variable and click Rule.Type 1 as the value for selection variable.

CHAPTER 4

DATA ANALYSIS AND RESULTS

4.1 Introduction

This chapter looks at the analysis of Bank Customer Switching behavior data. The chapter considers personal details or demography of customers; age of a customer, gender, marital status, educational level, occupation, number of years a customer has spent with the bank and customer rating of bank's services and products; employee competence construct, reliability construct, product innovation construct, pricing construct, convenience construct, customer satisfaction construct, physical evidence construct, effective advertising competition construct that were used as predictor variables.

4.2 Coding of Responses

The five-point Likert-scale ranging from strongly disagree (1) to strongly agree (5) was used to rate bank's services and products. In coding responses for the Likert-scale; 1 was used to code for strongly disagree, 2 for disagree, 3 for undecided, 4 for agree and 5 for strongly agree. To make analysis and explanation of the logistic model easy, the five-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5) was reduced to two; "Agree" and "Disagree". SPSS interface of variable view is shown at appendix A. The questionnaire consists of eight (8) constructs: Employee competence, Reliability, Product innovation, Pricing, Convenience, Customer satisfaction, Physical evidence, and Effective advertising competition. Each construct has two to four questionnaire items under it. These constructs and respondents demographics: Age, Gender, Marital status, educational level, years spent with the bank, were used as predictor variables. The mean score for each construct was derived used in estimating the parameters of logistic regression. The dependent or response variable is customer's intention of switching bank in a near future, which is made up of two categories, "very likely to switch" and "very unlikely to switch". In coding the dependent variable zero (0) was used as a code for very unlikely to switch responses and one (1) for very likely to switch.

4.3 Exploratory Data Analysis

Preliminary data exploration is very essential for any data analysis since it helps one to clean the data set of any error, reveals pattern and relationship that may exist. The data was screened for outliers and figure 4.1 revealed that 6 cases: case 2, case 3, case 6, case 12, case 19 and case 180 were outliers. The standardized residuals (ZResid) of these outliers were greater than cut-off point of 2.58 so they were removed from the data. Their inclusion in the data analysis may affect the results significantly.

Case	Selected Status[a]	Observed	Predicted	Predicted Group	Temporary Variable	
		Do you intend switching bank in a near future?			Resid	ZResid
2	S	n**	.885	yes	-.885	-2.770
3	S	n**	.973	yes	-.973	-6.013
6	S	n**	.907	yes	-.907	-3.123
12	S	y**	.033	no	.967	5.378
19	S	n**	.960	yes	-.960	-4.868
180	S	y**	.135	no	.865	2.533

Figure 4.1: Casewise List for Outliers

The five point Likert-scale used in the questionnaire was checked for internal consistency and reliability.

Table 4.1: Reliability Test

Cronbach's Alpha	N of Items
0.65	21

In Table 4.1 Alpha Cronbach's test statistics generated from SPSS was 0.65 which satisfies the reliability criteria, which is above 0.6 that Churchill (1979) recommended. This shows that the internal consistency and data reliability are high for the research. Independent variables were also tested for multicollinearity. From table 4.2 the Variance Inflation Factor (VIF) and Tolerance value for each construct were less than 10 and above .10 respectively indicating the absence of significant interaction among the independent variables that might affect the results in the analysis. This implies that all the eight constructs and demographics are fit to be used in developing the model for banking customer switching behavior.

Table 4.2: Collinearity Test of Independent Variables

Model	Tolerance	VIF
Age	0.884	1.131
Gender	0.862	1.159
Marital Status	0.966	1.035
Education	0.877	1.140
Year spent with the bank	0.885	1.130
Employee Competence	0.375	2.666
Reliability	0.285	3.154
Product Innovation	0.923	1.083
Pricing	0.243	4.114
Convenience	0.808	1.237
Physical Evidence	0.785	1.275
Customer Satisfaction	0.209	4.791
Effective Advert Competition	0.800	1.250

4.4 Descriptive Statistics of Respondents Characteristics

Figure 4.2 presents descriptive statistics of customer switching intent compare with demographics of respondents. Customer switching intent was grouped into

Variables		Do you intend switching bank in a near future?		Total
		No, very unlikely	Yes, very likely	
Age of Respondent	below 40 yrs.	142 (30.2%)	179 (38.1%)	321 (68.3%)
	40 yrs. and above	87 (18.5%)	62 (13.2%)	149 (31.7%)
Sex	male	153 (32.6%)	98 (20.9%)	251 (53.4%)
	female	76 (16.2%)	143 (30.3%)	219 (46.4%)
Marital Status of Respondent	single	117 (24.9%)	119 (25.3%)	236 (50.2%)
	married	112 (23.8%)	122 (26%)	234 (49.8%)
Educational Level of Respondent	low education	22 (4.7%)	13 (2.8%)	35 (7.4%)
	Higher education	207(44%)	228 (48.5%)	435 (92.6%)
Number of years spent with the bank	≤ 4yrs	98 (20.9%)	102 (21.7%)	200 (42.6%)
	> 4 yrs.	131 (27.9%)	139 (29.6%)	270 (57.4%)
		229 (48.7%)	**241 (51.3%)**	**470 (100%)**

Figure 4.2: Characteristics of Response to Customer Switching Intent by Bank Customers

two (2) categories: very likely to switch and very unlikely to switch. Out of the 470 valid respondents, 51.3 % (241) reported that it is very likely that they will switch services with the bank and 48.7 % (229) said it is very unlikely that they will switch services with the bank. 68 % of the respondents interviewed were below 40 years representing the majority. 53.4 % of customers interviewed were males and 46.6 % being females. 92.6 % of customer interviewed were 'high educated' and 7.4 % of them being 'low educated'. The high education group were customers with SHS certificate and above, and those with JHS and below were the low education group. 57.4 % of customers interviewed have spent more than four years doing banking business with the bank.

4.5 Results of the Logistic Regression Model

The performance of the model was conducted on the following independent variables, customer satisfaction construct, employee competence construct, pricing construct, reliability construct, product innovation construct, convenience construct, effective advertising construct, physical evidence construct, sex, education, number of years spent with the bank, age and marital status. The

Table 4.3: Logistic Regression Predicting Likelihood of a Customer Switching Services with Ahantaman Rural Bank.

IV's	B	S.E.	Wald	df	Sig.	OR	low (OR C.I.)	Up(OR C.I.)
Gender(1)	-.030	.536	.003	1	.956	.971	.340	2.775
Ma St.(1)	-.031	.502	.004	1	.951	.970	.363	2.592
Edu(1)	-1.639	.784	4.373	1	.057	.194	.042	.902
Yrs.Bnk(1)	.545	.513	1.129	1	.288	1.724	.631	4.708
EmpCom(1)	2.770	.535	26.846	1	.000	15.957	5.596	45.497
Reliab.(1)	2.176	1.078	4.073	1	.044	8.808	1.065	72.868
Pro.Inno(1)	-.525	.545	.926	1	.336	.592	.203	1.723
Pricing(1)	3.003	.609	24.328	1	.000	20.146	6.109	66.441
Conveni(1)	.057	.537	.011	1	.916	1.058	.369	3.033
Phys.Evi(1)	.301	.548	.302	1	.583	1.351	.461	3.959
Cust Sat(1)	2.438	.890	7.505	1	.006	11.445	2.001	65.457
EffAdv(1)	.007	.509	.000	1	.988	1.007	.371	2.733
Const	-6.744	1.337	25.423	1	.000	.001		

sources:Researchers Computation based on observed data

training set sample was used to design the model and the model tested on the validation set. The fitted logistic regression for customer switching intent in Ahantaman Rural Bank in the Ahanta West District are presented in table 4.3. The significance of the variables are assessed by the pvalues (represented in the table by "sig"), the Wald's statistics values or the odd ratios. The values of the Standard Errors in column 3, ranging from 0.121-2.351 is within the acceptable range, that is, must be between 0.001-5.000 as recommended by Chan (2004) and this shows the absence of multicollinearity in the independent variables. This implies that the model is statistically stable. When the column named Sig. is studied carefully, we have four variables: employee competence(EmpCom) p = 0.000, reliability(Reliab.) p = 0.044, pricing p = 0.000 and customer satisfaction(Cust Sat) p = 0.006 having their p-values less than .05, these are variables that contribute significantly to the model. This implies that, the major variables influencing whether a customer will switch services with the bank are: customer satisfaction, employee competence, pricing and reliability. Product innovation, convenience, effective advertising competition, physical evidence and demographics did not have any impact on the model statistically. The B values are the values to be used to construct logistic equation to calculate probability for classifying new cases. The B values are also used in calculating odds ratio. The exponents of B values are the odds ratio values. Using the odds ratio we can say that, customers who disagreed with reliability construct were 8.008 times more likely to switch services with the bank. In other words, it is more likely that you are "very likely to switch customer" than "very unlikely to switch" if you disagreed rather than agreed with reliability construct. The confidence interval column(low OR C.I. and Up OR C.I.) represent the 95 % confidence interval for the odds ratio. This tell us that we are 95 % confidence that the interval contains the true value of odds ratio. Critical look at the confidence interval of the significant variables show that, the intervals do not contain the value 1, if the interval had contained the value 1 we would conclude that, there is equal probability of the responses (very likely to switch/ very unlikely to switch).

Logistic regression model for predicting very likely to switch customer is defined as;

$$logit(p) = 2.770X_1 + 2.176X_2 + 3.003X_3 + 2.438X_4 - 6.744$$

And by making 'p' the subject, we obtained

$$p = \frac{e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}{1 + e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}$$

where X(1) is employee competence(agree)

where X(2) is reliability (disagree)

where X(3) is pricing (agree)

where X(4) is customer satisfaction (disagree)

And 'P' is probability of a customer who is very likely to switch a bank.

4.6 Assessing the Fit of the Model

The fit from logistic regression model was assessed for adequacies. In this study, the adequacy of the logistic regression model is assessed based on the Omnibus Test of Model Coefficients table, Model summary table, Hosmer and Lemeshow Test and classification.

Table 4.4: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig
	Step	526.408	12	.000
Step 1	Block	526.408	12	.000
	Model	526.408	12	.000

Table 4.4 tests the hypothesis that, the model without the predictors is a good fit, it tests the following:

Ho: The model with constant only is a good fitting model.

H1: The model is not a good fitting model (i.e. the predictors have a significant effect).

The test reveals that the model chi-square has 12 degrees of freedom, a value of 526.408 and probability of $p < 0.000$ is highly significant, indication that the predictors have effect on the model.

Table 4.5: Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	124.844	.674	.898

Table 4.5 indicates that 67.4 % of the variation in the customer switching intentions is accounted for by the logistic regression model.

Nagelkerke R-Square statistic of 0.898 indicates a very strong relationship of 89.8 % between the independent variables and the prediction.

Table 4.6: Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.447	8	.964

Table 4.6 test the null hypothesis, Ho: The model with predictors fit the data, against the alternate hypothesis that, HA: The model does not fit the data. Based on the table, on the Sig. column, the test statistic has a significance of .999 which implies that it is not statistically significant and therefore we refuse to reject the null hypothesis, implying that the model is quite a good fit.

Observed			Predicted					
			Training Set[b]			Validation Set[c]		
			Do you intend switching bank in a near future?		% Correct	Do you intend switching bank in a near future?		% Correct
			no, very unlikely to switch	yes, very likely to switch		no, very unlikely to switch	yes, very likely to switch	
Step 1	Do you intend switching bank in a near future?	no, very unlikely to switch	168	9	94.9	48	4	92.3
		yes, very likely to switch	3	173	98.3	5	60	92.3
	Overall Percentage				96.6			92.3

Figure 4.3: Final Classification table

Figure 4.3 is a final classification table for logistic regression model. The training set sample (353) was used to design the logistic regression model and the model was tested on the validation set sample (117) to assess its classification power. The study using the training set for classification, 98.3 % cases were correctly classified for yes, very likely to switch group and 94.9% cases correctly classified for the no, very unlikely to switch group. Overall 96.6 % were correctly classified with sensitivity and specificity of 98.3% and 94.9% respectively.

Using the same sample that was used to design a classification model to test the classification power of the model has been criticized by researchers. The model was validated using the validation set, and 92.3% of yes cases were correctly classified as yes cases and 92.3% no cases were correctly classified as no cases.

The validation method used was Hold-out method. The overall correct classification for the validation set is 92.3%, which is a bit lower than the final classification using the training set.
The overall classifications in both training and validation sets are high imlying the model is good.

4.7 Prediction and Classification of Customer using the equation

By making 'P' the subject of the logistic regression model we have;

$$p = \frac{e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}{1+e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}$$

where X(1) is employee competence
where X(2) is reliability
where X(3) is pricing
where X(4) is customer satisfaction
And 'P' is probability of a customer who is very likely to switch a bank.

The predictive equation can be used in this way to classify a new case. Suppose a customer who disagreed with reliability construct, agreed with employee competence construct, agreed with pricing construct and disagree with customer satisfaction construct. Would he/she be predicted as very likely to switch customer?

We put 1 in place of the variables in the equation:

$$p = \frac{e^{2.770(1)+2.176(1)+3.003(1)+2.438X(1)-6.744}}{1+e^{2.770(1)+2.176(1)+3.003(1)+2.438(1)-6.744}}$$

$$= \frac{37.07711749}{38.07711749} = 0.974$$

Therefore, the probability that a customer who customer who disagreed with reliability construct, agreed with employee competence construct, agreed with pricing construct and disagree with customer satisfaction construct will switch services with the bank is 97.4%, or 97.4% of such individuals will be very likely to switch customers.

4.8 Discussion of Main Findings

From the results obtained, it shows clearly that 51.3% of customers who do banking transaction at the head office of Ahantaman Rural Bank have intended to switch services with the bank. This gives an indication that some of the customers are not happy with the services of the bank and the bank might risk of losing customers to two competing banks; Agricultural Development Bank and GN Bank. The results of the study revealed that the following factors: reliability construct, customer satisfaction construct, employee competence construct and pricing construct impacted greatly on customer intentions to switch bank.

The study revealed that pricing construct which consists of three questionnaire items, was the major factor that influenced customers' decision to switch banks. The customers perceived that the bank charged high fees on its transaction and interest rates for loans were high. The study identified employee competence construct as the second most important factor influencing customers decision to switch bank. This might be as a result of the way customers were treated by some of the bank employees. Customer satisfaction and reliability constructs were the third and fourth factors respectively influencing customers' decision to switch bank. Some customers perceived the bank as unreliable.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

The study examined the factors that influence customers' switching intentions in Ahantaman Rural Bank and classify customers into likely to switch and unlikely to switch customers. Critical examination of customer demographics and ratings of bank services and products was carried out using logistic regression. The method was used to establish the significance of the relationship between the customer demographics, ratings grouped under eight constructs used as independent variables and likelihood of customer switching bank used as dependent variable.

5.2 Conclusions

The study found four out of eight constructs to influence customers' decision to switch bank, that is, employee competence, customer satisfaction, pricing and reliability constructs. The other insignificant factors or constructs were Product Innovative, Convenience, Physical Evidence and Effective advertising competition. Pricing construct which consists of three questionnaire items, had the greatest impact on a customer's intentions to switch or change bank. Employee competence construct was the second vital factor which have effect on customers' intentions to switch bank. Customer satisfaction and reliability constructs were the third and fourth factors respectively influencing customers' decision to switch bank.

The logistic regression model for predicting "very likely to switch customer" is given as;

$$logit(p) = 2.770X_1 + 2.176X_2 + 3.003X_3 + 2.438X_4 - 6.744$$

or

$$p = \frac{e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}{1 + e^{2.770X_1+2.176X_2+3.003X_3+2.438X_4-6.744}}$$

where X(1) is employee competence
where X(2) is reliability
where X(3) is pricing
where X(4) is customer satisfaction
And 'P' is probability of a customer who is very likely to switch a bank.

The model was tested on the validation set to classify cases and 92.3% of cases were classified correctly.

5.3 Recommendations

The following recommendations were made based on the findings and conclusion above:
Bank management may adopt the research model of switching intentions designed in this study to investigate the reasons why some customers are switching bank whiles others are not. The model may also provide management with an insight on how to predict a customer so that a customized service may be tailored for a potential switchers thereby reducing switching rate and increasing profitability.

Bank management should develop strategies that increase customer satisfaction. Strategies that focus on avoiding service related problems from happening, effectively handling of customers who are not happy with banking services as early as possible, and addressing complaints of customers. By so doing, the customers would be motivated enough to recommend the bank to their friends.

For reliability, management should make sure all transactions are clear, accurate and all instructions to customers are easy to follow and easily understood. The bank must be seen to deliver its promises to customers and ready to accept customers' recommendations. The impressions should be created in customers' mind that their submissions are needed to make the banking services attractive. The customers should be sensitized and made to understand that what they are receiving at the bank they will not find it in any other bank.

For pricing, the current charges used by the bank should be boldly displayed in the banking hall for the customers to see. Most often it is the hidden charges that create the wrong perception about pricing. The customers should be made to understand that the bank charges are lower than the other competing banks.

On employee competence, the employees should be educated on modern ways of attending to customers' needs by creating friendly customer-employee relationship before serving the customers. To enhance employee competence, the customers should be encouraged to rate the employees of the bank and the best worker of a given period should be adjudged based on customers' ratings and picture of the worker boldly displayed in the banking hall.

On how to adopt and use the research model, at any point in time, the management should sample some of the customers and ask them to respond to questions in the questionnaire. Their scores on significant constructs should be inserted into the equation to classify each customer into very likely to switch and very unlikely to switch customers. In other to retain very likely to switch customers, customized or tailored services should be designed for them or find out what exactly is the problem that they want to exit.

5.4 Limitation and Future Research

The study only interviewed the customers without involving the employees. Further research should seek the opinion of bank employees to obtain their views on why customers switch banks. The employees and customers' view could then be compared for further analysis.

Finally, the research was conducted using the customers of one branch, Agona Nkwanta, and this may limit inferring the findings to other branches of the bank

REFERENCES

Andreasen, A. R. (1988). *Consumer complaints and redress: What we know and what we dont know.* Columbia: MO.

Avkiran, N. K. (1994). Developing an instrument to measure customer service in branch banking. *International Journal of Bank Marketing*, 12:10–18.

Bawumia, M., Belnye, F., and E, O. M. (2005). The determinants of banks interest spread in ghana: An empirical analysis of panel data. In *Bank of Ghana Working Paper No 2005/09*.

Beckett, A., Hewer, P., and Howcroft, B. (2000). An exposition of consumer behaviour in the financial services industry. *The International Journal of Bank Marketing*, 18:231–248.

Bitner, M. J., Booms, B. H., and Mohr, L. A. (1994). Critical service encounters: The employee's viewpoint. *Journal of Marketing*, 58:95–106.

Blanchard, R. F. and Galloway, R. L. (1994). Quality in retail banking. *International Journal of Services Industry Management*, 16:5–23.

Brownbridge, M. and Gockel, A. F. (1996). The impact of financial sector policies on banking in ghana.

Chan, Y. H. (2004). Logistic regression analysis. *Singapore Med J*, 45:149.

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16:64–73.

Colgate, M. and Hedge, R. (2001). An investigation into the switching process in retail banking services. *International Journal of Bank Marketing*, 19:201–212.

Colgate, M., Steward, K., and Kinsella, R. (1996). Customer defection: A study of student market in ireland. *International Journal of Bank Marketing*, 14:23–29.

Dube, L. and Shoemaker, S. (2000). *Brand switching and loyalty for services in Swartz, T.A. and Dawn, I. (Ed.), Handbook of Services Marketing and Management.* CA : Sage.

Fornell, C. (1992). A national customer satisfaction barometer: The swedish experience. *The Journal of Marketing*, 56:66–21.

Garland, R. (2002). Estimating customer defection in personal retail banking. *International Journal of Bank Marketing*, 20:317324.

Gerrard, P. and Cunningham, J. B. (2004). Consumer switching behavior in the asian banking market. . *Journal of Services Marketing*, 18:215–223.

Ghouri, A. M., Khan, N., Siddqui, U. A., Shaikh, A., and Alam, I. (2010). Determinants analysis of customer switching behavior in private banking sector of pakistan. *Interdisciplinary Journal of Contemporary Research in Business*, 2:96–110.

Gremler, D. D. and Brown, S. W. (1996). Service loyalty: Its nature, importance and implications. *Proceedings American Marketing Association*, pages 171–180.

Hirschman, A. O. (1970). *Exit, voice and loyalty.* Cambridge, MA: Harvard University Press.

Hosmer, D. W. and Lemeshow, S. (1998). *Applied logistic regression.* New York: Wiley.

Hosmer, D. W. and Lemeshow, S. (2000). *Applied logistic regression.* New York: Wiley.

Hull, L. (2002). Foreign-owned banks: Implications for new zealands financial stability. discussion paper series, dp2002/05. Technical report.

Hunt, H. K. (1977). *CS/D Overview and future research directions, in Hunt, H. Keith, (Ed.), conceptualization and measurement of consumer satisfaction and dissatisfaction.* Cambridge, MA: Marketing Science Institute.

Kamilia, B. and Jacques, N. (2000). A reliable and valid measurement scale for the perceived service quality of banks. *The International Journal of Bank Marketing*, 18:18.

Keaveney, S. M. (1995). Customer switching behaviour in service industries: An exploratory study. *Journal of Marketing*, 59:71–83.

Kish, J. (2000). Before your customers leave. *Bank Marketing*, 32:30.

Kotler, P., Ang, S. H., Leong, S. M., and Tan, C. T. (1999). *Marketing Management: An Asia perspective.* New Jersey: Prentice Hall.

Lees, G., Garland, R., and Wright, . M. (2007). Switching banks: Old bank gone but not forgotten. *Journal of Financial Services Marketing*, 12:146–157.

Levesque, T. and McDougall, G. (1996). Determinants of customer satisfaction in retail banking. *International Journal of Bank Marketing*, 14:12–20.

Licata, J. W. and Chakraborty, G. (2009). The effects of stake, satisfaction, and switching on true loyalty: A financial services study. *International Journal of Bank Marketing,*, 27:252–269.

Lymperopoulos, C., Chaniotakis, I. E., and Soureli, M. (2006). The importance of service quality in bank selection for mortgage loans. *Managing Service Quality*, 16:365–379.

Matthews, C. and Murray, D. (2007). Helping bank customers switch: A case study. *Journal of Financial Services Marketing*, 11:360–370.

Mavri, M. and Ioannou, G. (2008). Customer switching behaviour in greek banking services using survival analysis. *Journal of Managerial Finance,*, 34:186–197.

Mittal, B., Ross, W., and Baldasare, P. P. (1998). The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. *Journal of Marketing*, 62:33–47.

Oliva, R. and Sterman, J. D. (2001). Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, 47:894–914.

Parasuraman, A., Zeithaml, V. A., and Berry, L. (1988). Servqual: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64:12–40.

Reichheld, F. F. (1996). *The loyalty effect.* Boston, MA: Harvard Business School Press.

Rust, R. T., Zahorik, A. J., and Keiningham, T. L. (1996). *Services marketing.* New York, NY: Harper Collins.

Santonen, T. (2007). Price sensitivity as an indicator of customer defection in retail banking. *International Journal of Bank Marketing*, 25:3955.

Stewart, K. (1998). An exploration of customer exit in retail banking. *International Journal of Bank Marketing,*, 16:6–14.

Taylor, S. A. and Baker, T. L. (1994). An assessment of the relationship between service quality and customer satisfaction in the formation of consumers purchase intentions. *Journal of Retailing*, 70:163–78.

Teich, I. (1997). Holding on to customers: The bottom-line benefits of relationship building. *Bank Marketing*, 29:12–13.

Yamane, T. (1996). *Statistics: An Introductory analysis, 2nd Ed.* New York: Harper and Row.

Zeithaml, V. A. (1998). Customer perceptions of price, quality and value: A means end model and synthesis of evidence. *Journal of Marketing*, 52:2–22.

APPENDIX A

*Untitled1_1 no outliers - VALIDATION.sav [DataSet1] - IBM SPSS Statistics Data Editor

View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

id	Age	Gender	Marital Status	Edu	occupation	YearsWitBank	Switch Bank	EmployeeCompetence	Reliability	ProductInnovation	Pricing	Convenience	Physical Evidence	Customer Satisfaction	EffectiveAdvertCompetition
001	0	1	1	1	1	0	0	0	1	0	0	1	1	1	1
004	0	0	0	1	1	0	1	0	0	1	1	0	1	0	1
005	0	1	1	1	1	1	0	1	0	0	1	1	1	0	0
007	0	1	1	1	1	0	0	1	0	1	1	1	1	0	1
008	0	1	1	1	1	1	0	1	0	0	1	1	0	0	0
009	0	0	1	1	1	1	0	0	0	0	1	1	1	0	1
010	1	0	1	1	1	1	0	1	0	1	0	0	1	1	1
011	0	0	1	1	1	1	0	0	0	0	0	1	1	0	1
013	0	1	0	1	1	1	0	0	0	0	0	1	1	1	1
014	1	0	1	1	1	1	0	1	0	0	0	1	1	0	1
015	1	0	1	1	1	0	0	0	0	1	0	1	1	1	1
016	1	0	0	1	1	0	0	0	0	1	0	1	1	1	1
017	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1
018	1	0	0	1	1	0	0	0	0	0	0	1	1	0	1
020	1	0	1	1	1	1	0	1	0	0	1	1	1	0	0
021	1	1	1	1	2	0	0	0	0	1	0	0	1	1	1

APPENDIX B

<u>Sample Size Calculation</u>

$$n = \frac{N}{1 + N(e)^2}$$

$$n = \frac{833}{1 + 833(0.03)^2} = \frac{833}{1.7497}$$

$$n = 476$$