

UNIVERSITY OF NAIROBI

SCHOOL OF ECONOMICS

**DETERMINANTS OF COST EFFICIENCY LEVEL OF COMMERCIAL
BANKS IN KENYA**

**SUBMITTED BY:
KARIUKI JOSHUA
X50/69306/2011**

**A RESEARCH PAPER PRESENTED TO THE SCHOOL OF ECONOMICS IN PARTIAL
FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF
MASTER OF ARTS IN ECONOMICS OF THE UNIVERSITY OF NAIROBI.**

**SUPERVISORS:
DR. THOMAS N. ONGORO
DR. FREDRICK O. SULE**

RESEARCH PROPOSAL

DECLARATION

This research paper is my original work and has not been presented for a Masters Degree in any other university. All information from other sources is acknowledged appropriately.

SIGNED:.....

DATE

KARIUKI JOSHUA

X50/69306/2011

This research project has been submitted for examination with our approval as the supervisors.

SIGNED:.....

DATE

DR. THOMAS N. ONGORO

SIGNED:

DATE

DR. FREDRICK O. SULE

TABLE OF CONTENTS

| | |
|-----------------------------|-----------|
| List of tables..... | ix |
| List of figures..... | x |
| Abbreviations..... | xi |

CHAPTER 1 INTRODUCTION

| | |
|-------------------------------------------------------------|----|
| 1.0 Introduction..... | 1 |
| 1.1 Background..... | 2 |
| 1.1.0 Preview of Commercial Banks in Kenya..... | 2 |
| 1.1.1 Meaning of Efficiency and Measurements..... | 4 |
| 1.1.2 Conventional verses Frontier Efficiency Approach..... | 5 |
| 1.1.3 Framework of Efficiency..... | 7 |
| 1.1.4 Technical, Allocative and Cost Efficiency..... | 7 |
| 1.1.5 Economies of Scale..... | 10 |
| 1.2 Problem Statement..... | 12 |
| 1.3 Research Questions..... | 13 |
| 1.4 Objectives of the study..... | 14 |
| 1.5 Significance of the study..... | 14 |
| 1.6 Limitations of the study..... | 14 |
| 1.7 Organization of the study..... | 15 |

CHAPTER 2 LITERATURE REVIEW

| | |
|--------------------------------------------|----|
| 2.0 Introduction..... | 16 |
| 2.1 Theoretical Literature Review..... | 16 |
| 2.1.1 Conventional Theory of the Firm..... | 16 |

| | |
|------------------------------------------------|----|
| 2.1.2 The Managerial Theories of the Firm..... | 18 |
| 2.1.3 Behavioral Theories of the Firm..... | 20 |
| 2.1.4 The X-Efficiency Theory of the firm..... | 22 |
| 2.2 Empirical Literature Review..... | 24 |
| 2.3 Literature Overview..... | 28 |

CHAPTER 3 RESEARCH METHODOLOGY

| | |
|----------------------------------------------|----|
| 3.0 Introduction..... | 30 |
| 3.1 Concept of Inefficiency..... | 31 |
| 3.2 Theoretical Model..... | 32 |
| 3.2.0 Stochastic Cost Frontier Approach..... | 32 |
| 3.3 Empirical Model..... | 36 |
| 3.4 Data Source..... | 41 |

CHAPTER 4 PRESENTATION AND DISCUSSION OF RESULTS

| | |
|----------------------------------------------------------|----|
| 4.1 Introduction..... | 42 |
| 4.2 Cost Frontier Estimates..... | 42 |
| 4.3 Model Specification..... | 46 |
| 4.4 Efficiency Level..... | 47 |
| 4.5 Empirical results of the tobit regression model..... | 47 |
| 4.6 Analysis of the average bank efficiency scores..... | 51 |
| 4.7 Summary and conclusion..... | 53 |

CHAPTER 5 CONCLUSION AND POLICY RECOMMENDATION

| | |
|-----------------------|----|
| 5.1 Introduction..... | 55 |
|-----------------------|----|

| | |
|-------------------------------------|-----------|
| 5.2 Conclusion of the study..... | 56 |
| 5.3 Policy Implications..... | 57 |
| 5.4 Areas for further research..... | 59 |
| Reference list | 42 |
| Appendix 1 | 65 |

DEDICATION

To my parents, Stephen Kibungi and Rachel Kibungi for their support through my academic life and my valued friend and soul mate Beatrice Makeri without whom I would not have finished this paper on time. I am forever grateful.

ACKNOWLEDGEMENT

I would want to express my deepest gratitude to my research supervisors Dr. Thomas N. Ongoro and Dr. Fredrick O. Sule for their excellent guidance and supervision. Their support was unconditional and aided me in the completion of this research paper on time.

I would also want to thank the School of Economics, University of Nairobi for the continued support through my master's degree study. I will forever be indebted for the scholarship opportunity that was extended to me which facilitated the completion of master's degree.

The support that was provided by the staff of the Capital Markets Authority Library is also greatly acknowledged. I also want to extend my gratitude to my classmates and colleagues, with a special mention of Mr. Fredrick Ayot who kept encouraging me when things seemed tough.

ABSTRACT

The Kenyan banking sector has experienced tremendous structural changes over the years. Initially, during the 1980s to 1990s the banking sector was characterized by high level of undercapitalization, increased number of non-performing loans and weak corporate governance. This led to a high aggregate inefficiency level of the banking sector and costs associated with operation of banks that existed. In addition, corruption was also a major factor that led to the collapse of some of the major banks. It is evident that efficiency is an important element of enhancing a stable banking sector. Therefore, cost efficiency becomes an important study to inform proper policy formulation and the establishment of sound supervisory framework.

In recent years commercial banks in Kenya have been grappling with increased costs attributed to the rising employee and interest expenses amongst other factors. The banking sector has been expanding in terms of size and network. This has increased competition and capitalization in the strategic expansion. Therefore, it becomes important to determine the factors that affect bank cost efficiency level Kenya. This research paper applies the Stochastic Frontier Analysis (SFA) model where the intermediation approach of determining bank efficiency is used. The intermediation approach is superior to the production approach as it is characterized by fewer data problems. Secondary data is extracted from the balance sheets and income statements of commercial banks in Kenya listed in the NSE for the period of 2002-2011.

The cost efficiency level in Kenya is found to be 99.2% on average. The local banks are found to be more efficient than both the local private and foreign banks. The

parameter estimates of branch size, government securities, advanced loans and inflation have positive and significant effects on the efficiency level.

These and other results of the empirical findings suggest that in order to enhance banking efficiency in Kenya, the government through the CBK should continue implementing bank reforms. Particularly, the banking markets should be opened to foreign competition, formulate bank risk management and corporate governance policies. This should be done in an attempt to encourage bank expansion in Kenya.

Keywords: Cost efficiency, Stochastic Frontier Analysis, banking.

LIST OF TABLES

| | |
|---------------------------------------------------------------------------------------|----|
| Table 4.0 Maximum Parameter Estimates for Stochastic Frontier Cost Function... | 43 |
| Table 4.1 Statistical cost efficiency in summary form..... | 47 |
| Table 4.2 Parameter Estimates of Random Effect Tobit regression..... | 48 |
| Table 4.3 Average Bank Efficiency Scores in Kenya..... | 52 |

LIST OF FIGURES

| | |
|---------------------------------------------------------------------------------------|----|
| Figure 1.0 Technical, Allocative and Cost Efficiency..... | 8 |
| Figure 1.1 Cost Curves and Economies of Scale..... | 11 |
| Figure 4.0 Average cost Efficiency scores of banks in Kenya from 2002-201..... | 53 |

ACRONYMS AND ABBREVIATIONS

- AE** – Allocative Efficiency
- CE** – Cost Efficiency
- CBK** – Central Bank of Kenya
- CMA** – Capital Markets Authority
- CRS** – Constant Returns to Scale
- DEA** – Data Envelopment Analysis
- DFA** – Distribution Free Approach
- DMU** – Decision Making Unit
- ERS** – Economic Recovery Strategy
- FDH** – Free Disposal Hull
- GDP** – Gross Domestic Product
- LACs** – Long-run Average Cost Curves
- MES** – Minimum Efficient Scale
- ML** – Maximum Likelihood
- NSE** – Nairobi Securities Exchange
- ROA** – Return On Assets
- SACs** – Short-run Average Cost Curves
- SAPs** – Structural Adjustment Programs
- SFA** – Stochastic Frontier Analysis
- SSA** – Sub Saharan Africa
- TE** – Technical Efficiency
- TFA** – Thick Frontier Analysis

CHAPTER ONE

1.0 INTRODUCTION

Commercial banks play a critical role in the financial intermediation process. Financial intermediation involves the channeling of funds between surplus and deficit agents in the economy. In this regard, the banks consolidate deposits received from their customers and use these funds to issue out loans. The efficient allocation of scarce resources in the economy thus becomes of critical importance in the intermediation process. Extensive global research on financial efficiency has been directed towards the bank intermediation process since the period of the great depression that occurred between 1929 and 1933.

In order to sustain the intermediation role in the financial sector, commercial banks need to be profitable. Therefore, the intermediation role and the eventual performance of the commercial banks have critical implications in terms of economic growth and development of countries. This is seen in the reduction of transactional costs, information asymmetry, risk diversification and moral hazard. Therefore, sound financial performance acts as incentive to the bank principals (i.e. shareholders) to invest more as it increases their profit levels and the return to their investments.

In the SSA countries there exists a huge gap between the demand and supply for the bank services due to the shallow financial deepening. An economy with shallow financial depth is characterized by a narrow range of financial assets that reduce the supply of these assets in the economy. Therefore, the demand for banking services far outstrips the number of banks that exist. This leads to reduced bank competition and

the eventual increase in the interest rate spread. The reduced bank competition leads to inefficiency in the allocation of the scarce resources in the economy.

1.1 BACKGROUND OF THE STUDY

1.1.0 PREVIEW OF COMMERCIAL BANKS IN KENYA

Although the Kenyan financial system in comparison to the regional standards is well developed and diversified, it faces major structural impediments that restrict it from reaching its full potential. Nevertheless, the Kenyan financial system enjoys a higher level of its credit being channeled to the private sector and a higher deposit level in the financial institutions than other Sub-Saharan Africa and low-income countries.

In this regard, the World Bank report (2006), argues that commercial banks in Kenya have tremendously improved their asset portfolio and liquidity over the years. This has had a positive impact on the economy as it has made the economy more resilient to both internal and external shocks. In comparison to her neighbors, the Kenyan banking sector has been credited for its diverse size and its innovative product diversification.

In their research, Beck et al. (2005) found that the standard indicator of financial development (i.e. the private credit to GDP ratio) was at 23.7% in 2008 as compared to a median of 12.3% for the SSA. The quality of lending over this period also improved significantly as concluded from their findings of the increased ratio net of loans relative to GDP. Nevertheless, the banking sector in Kenya has dismally failed to provide for financial inclusion of the majority of the population. This is true given that majority of the population lack access to various banking financial services. They

are expensive as evidenced from the high interest rate spreads and account fees charged (Beck et al., 2005).

According to the World Bank (2006), the financial sector in Kenya faced a major crisis during the period of early 1980's to late 1990's due to the high level of undercapitalization, increased number of non-performing loans and the weakness in corporate governance in the country. It further observed that Kenya suffered a systemic banking crisis in 1992. The financial and corporate sector experienced a large number of defaults and the various financial institutions and corporations experienced a lot of difficulty in repaying contracts in time. As a result the non-performing loans increased sharply and this led to the exhaustion of the aggregate capital banking system.

According to the Economic Recovery Strategy (ERS) Paper on Wealth Creation and Employment in 2007, the Kenyan banking sector experienced difficulties that undermined the objectives created by the ERS that was formulated by the Kenyan Government in 2003. These included: a high ratio of non-performing loans, reduced bank competition, a wide bank interest rate spread, the absence of vibrant institutions that offer long term finance, weak legal arrangements that led to delays in contract enforcement and dispute resolutions.

Recently, Kenya has adopted financial liberalization in its economy (World Bank, 2006). This was greatly advocated by the Structural Adjustment Programs (SAPs) in the early 1990's. The SAPs had a substantial effect on financial deepening and progress. In addition, it has led to the increased efficiency of the banking sector. The increased

competition has led to the marginal decline of the interest spread (Beck et al., 2009). This is attributed to the fact that there have been reduced loss provisions of overhead costs and lower profit margins from the increased market competition. Mwega (2011) reported that the Kenyan financial system is faced with tremendous challenges of fragmentation, where many small banks serve specific market niches, but do not effectively contribute to the competition in the banking sector.

Financial efficiency in any economy is an important element of a successful economy and is a critical ingredient of enhancing a stable banking sector. Therefore, financial efficiency is considered a fundamental element to economic growth, technical change of the banking sector productivity and a significant effect to the productivity of the economy that ensures its eventual convergence in the long run.

Financial institutions efficiency determination becomes important both at the microeconomic and macroeconomic level (Berger and Mester, 1997). From the microeconomic level, an efficient banking system is considered crucial as it allows for increased competition and the continuous improvement of the financial institutions. In addition, it allows for the formulation of proper regulatory and supervisory frameworks. Conversely, from the macroeconomic perspective, the banking sector efficiency level influences the cost of financial intermediation and the eventual soundness of the financial market.

1.1.1 MEANING OF EFFICIENCY AND MEASUREMENT

Broadly in microeconomics there are there are two types of efficiencies, namely: the economic efficiency and economies of scale. Firstly, the economic efficiency (also

referred to as cost efficiency) consists of two components. They are the technical and allocative efficiencies. Secondly, the economies of scale is the other measure of efficiency that deals with the size of production units being considered. It is founded on the concept of economies and diseconomies of scale in the production process. From the economies of scale, technical efficiency can be investigated and decomposed into pure technical and scale efficiencies.

Generally in microeconomics, there are two measures of efficiency i.e. the non-parametric and parametric techniques. The non-parametric techniques consist of the Data Envelopment Analysis (DEA) and the Free Disposal Hull (FDH) approaches. These are mathematical programming techniques that are used to estimate the efficiency levels of firms.

Conversely, the parametric techniques consist of the Stochastic Frontier Analysis (SFA), Distribution Free Approach (DFA) and the Thick Frontier Approach (TFA). These approaches usually consist of a pre-specified functional form for the best practicing frontier (i.e. cost, profit or production). A firm is considered as being inefficient if its costs are higher or profits lower than the best practice firm after removing the random error.

1.1.2 CONVENTIONAL VERSES FRONTIER EFFICIENCY APPROACH

There are two ways of measuring the performance of a firm. Firstly, there is the classical approach that is based purely on the simple profit-cost model analysis. This approach is represented by the use of conventional performance ratios that usually concentrate on the examination of financial ratios such as the Return on Equity

(ROE), Return On Assets (ROA), Capital Asset Ratio, growth rate of total revenue and the cost-income ratio. Managers and consultants usually use these variables to evaluate the financial performance of different firms. Nevertheless, these conventional performance ratios usually fail to control for the influence of input price, output price and other exogenous market factors. These shortcomings constrain the standard performance ratios from reaching closer to the true estimates for the performance of a given firm.

In the last quarter century, a lot of academic research has been focused on another approach of measuring the efficiency of any given firm. This is the frontier efficiency (also known as the X-efficiency) approach. The frontier efficiency measures deviations in performance from that of the best practicing firms on the efficient frontier. Therefore, it controls for the effects of exogenous factors e.g. the prices being experienced in the local market. Therefore, frontier efficiency method measures how well the financial institution performs relative to the predicted performance of the best firms facing the same market conditions in the industry.

It is the ability of the management to control costs and use resources to provide output. According to Weiss (2000), the frontier efficiency measures are used to summarize a firm's performance in a single statistic (also known as the efficiency score). This controls the differences among firms in a sophisticated multidimensional framework that is founded on economic theory. Therefore, frontier efficiency is considered superior to the conventional performance ratios since it gives better estimates of the underlying performance of firms.

1.1.3 FRAMEWORK OF EFFICIENCY

The concept of efficiency is derived from the microeconomic theory of the firm. It is based on the ideas advanced by Debreu (1951) and Farrell (1957) in their seminal work. They built the standard framework of productive efficiency, also known as the production frontier. The overall economic efficiency can be decomposed into scale efficiency, scope efficiency, pure technical efficiency and allocative efficiency. Theoretically, a firm is considered to be fully efficient if it produced an output level that maximizes profits and minimizes possible costs.

1.1.4 TECHNICAL, ALLOCATIVE AND COST EFFICIENCY

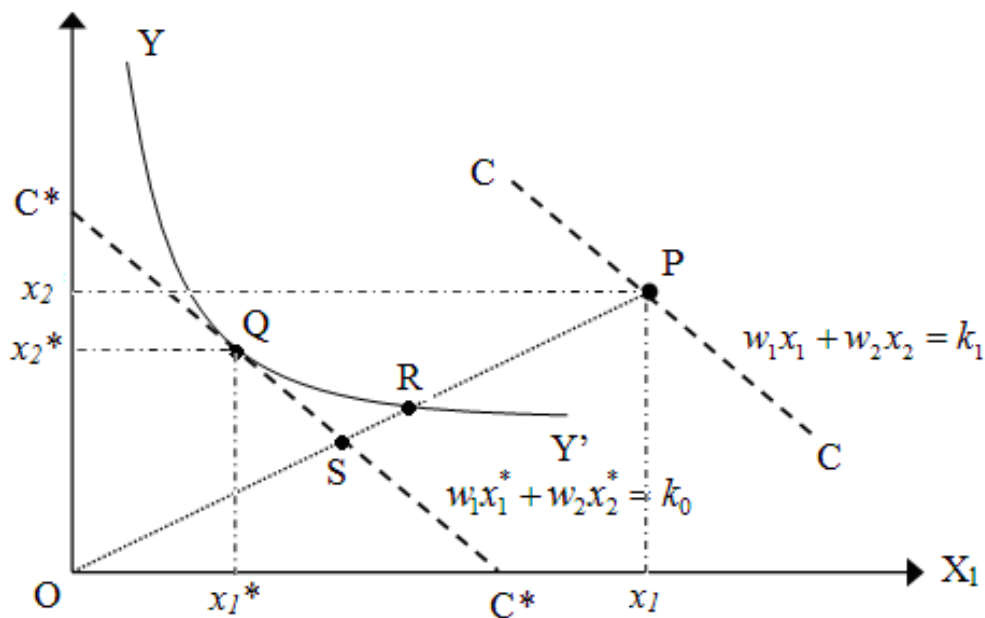
During his seminal work, Farrell (1957) proposed a method of measuring productive efficiency. It used an efficient isoquant estimated as part of the convex hull of observed points. He assumed that the production function was homothetic. A homothetic function is a monotonic transformation of a homogeneous function in which the marginal rate of technical substitution is constant along a ray drawn from the origin (Varian, 1993).

Using Varian (1993) argument of the homothetic function, to illustrate this analogy let the production function $f(X_1, X_2)$ be homogeneous of the first degree in X_1 and X_2 . Assuming that the isoquant of this homogeneous production function is an efficient isoquant, an increasing monotonic transformation of a homogeneous production function yields a homothetic production function in $F(X) = g[f(X_1, X_2)]$. Where g represents the strictly increasing monotonic transformation. A series of homothetic isoquants can be derived from the original (efficient) isoquant by appropriate scaling up. Therefore, a proportional increase or decrease of all inputs does not affect the

marginal rate of technical substitution along the isoquants. According to Clemhout (1968), a comparison between the efficient isoquant and any other isoquant for a given output level would indicate a departure from full efficiency level.

The efficiency analysis carried out by Farrell (1957) for a single output and two inputs case in a unit isoquant diagram as shown below:

Figure 1.0 Technical, Allocative and Cost Efficiency



Source: Cooper et al. (2007, p.258)

Farrell (1957) initially assumed Constant Returns to Scale (CRS) in the production process. This depicted an efficient production function or production frontier. The technological set was fully described by the unit isoquant YY' that captured the combination of inputs (x_1, x_2) by which a given firm could produce a certain output when it is perfectly efficient. Therefore, YY' showed the minimum combination of

inputs that were needed to produce a unit of output. Under this framework, every combination of inputs along an isoquant was considered as technically efficient while any point above and to the right of it, for example point **P** was defined as a technically inefficient point. Here it was argued that the producer used an input combination that was more than enough to produce a unit of output. Hence, the distance **RP** along the **OP** line was considered to measure the technical inefficiency of a producer located at point **P**. The distance **RP** thus represents the amount by which all inputs can be reduced without decreasing the amount of output.

Cooper et al. (2007) argue that geometrically the technical inefficiency level associated with package **P** can be expressed by the ratio **RP/OP** and, therefore, the technical efficiency (TE) of the producer under analysis is given by the ratio **OR/OP**. It takes a value of **0** and **1**. A value of **1** implies that the firm is fully efficient.

Allocative efficiency involves the selection of an input mix that allocates factors to their highest value, uses and introduces the opportunity cost of factor inputs to the measurement of productive efficiency (Cooper et al., 2007). Therefore, allocative efficiency can also be derived from the unit isoquant plotted in figure 1.1. Given that the market prices of inputs are (w_1, w_2) , the isocost line **CC** through **P** is associated with $w_1x_1 + w_2x_2 = k_1$ and the slope of this line reflects the input price ratio.

However, moving the line parallel until it is tangential to the isoquant at point **Q** can further reduce this cost. The coordinates of **CC** then give: $w_1x^*_1 + w_2x^*_2 = k_0$, achieving the minimal cost at the determined output level. The relative distance between **S** and **R** can be used to obtain the ratio **OS/OR**. The above ratio with respect

to the least cost combination of inputs given by point **Q**, indicates the cost reduction that a producer would be able to achieve if it is moved from a technically but not allocatively efficient input combination (**R**) to both a technically and allocatively efficient point (**Q**). Thus, the allocative efficiency of the producer is given by the ratio **OS/OR**.

Cooper et al. (2007) further argue that there is another measure of cost efficiency or economic efficiency. It is represented by the ratio of minimum cost (**wx***) to actual cost (**wx₀**) i.e. $wx^*/wx_0 = OS/OP$. A cost efficient firm will choose its inputs and mixes according to their prices so as to minimize the total costs. Cost inefficiency can arise from different sources. One is the deficiency in application of technology (i.e. technical inefficiency) and another one is the sub-optimal allocation of resources (allocative inefficiency). Therefore, the total overall cost efficiency can be presented as the product of technical efficiency and allocative efficiency. Thus, **OS/OR * OR/OP = OS/OP**

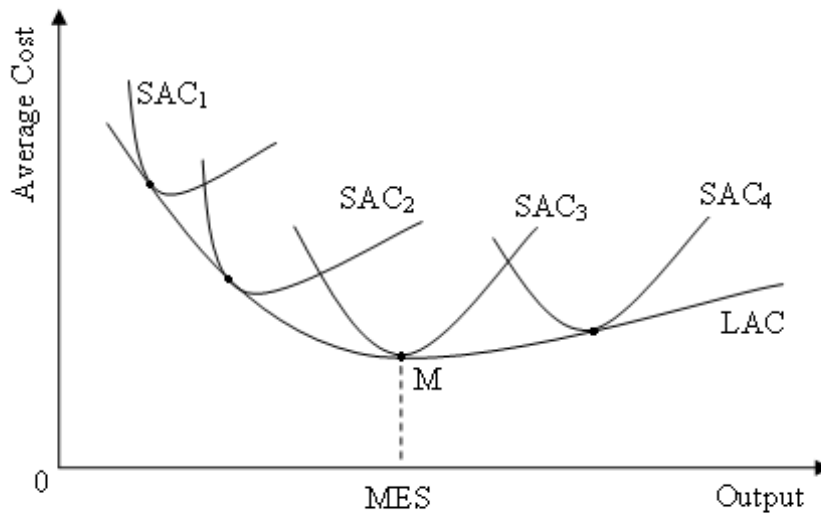
1.1.5 ECONOMIES OF SCALE

Economies of scale (referred also as the returns to scale) is the rate at which output changes with varying factor quantities (Molyneux et al., 1996). It measures whether firms with similar production and managerial technologies are operating at the optimal size. Increasing returns to scale exists where a proportional increase in a given firm's output would lead to a less than proportionate increase in its total costs.

Conversely, diseconomies of scale exists when a proportionate increase in a given firm's output would lead to a more than proportionate increase in its total costs.

Constant returns to scale occurs when a proportionate increase in a given firm's output leads to the same proportionate increase in its total costs.

Figure 1.1 Cost Curves and Economies of Scale



Source: *Humphrey (1990, p.38).*

According to Molyneux et al. (1996), economies of scale are based on the shape of the average cost curves. This is diagrammatically presented in figure 1.1. The figure shows short-run average cost curves (SACs) and a long-run average cost curve (LAC). Each of the short-run average cost represents the average cost of different sized firms during a short period of time. The firm chooses the size that yields the lowest average cost for that particular level of output. The average long-run cost curve is traced out from the SACs where each point of the LAC is to a point of tangency with a corresponding short run cost curve. This shows the least cost method of production for any level output. According to Humphrey (1990), scale economies appear as the slope of an average cost curve indicating how costs vary with output.

The downward sloping LAC reflects economies of scale as the average costs of production declines with output increase. This cost characteristic exists only up to a certain size of the firm known as the minimum efficient scale (MES). The firm achieves the lowest attainable average cost at point M and experiences constant returns to scale at this point. Beyond point M, the upward sloping LAC indicates diseconomies of scale.

1.2 PROBLEM STATEMENT

Commercial banks in Kenya have been grappling with increased cost in recent years. The increased costs are attributed to the rising employee and interest expenses amongst other factors. For example, according to the Central Bank of Kenya (CBK) in 2012 there was 13% increase in employment costs, which outpaced the annual rate of inflation at 9.4% in the banking sector. The total employment costs for the banks in Kenya was recorded to be 17.2% of the total income earned. The high cost is not a problem of the big banks only, but also for the small banks.

There was also increased demand for higher interest rates by the institutional investors who formed the large depositors for many of the Kenyan banks. Institutional investors are organizations that pool large sums of money and invest in securities, real property and other financial assets. According to the CBK, in 2011 commercial banks offered retail investors interest rates as low as 4%, compared to the institutional investors who demanded rates as high as 12-13%. This led to increased interest expenses in the various banks financial statements. For example in the last quarter of 2012, the top five banks paid out a total of Ksh.2.5 billion in interest expense compared to Ksh.600 million in the same period in 2011.

In addition, the banking sector has been expanding tremendously over the years as a result of increased competition and capitalization in strategic expansion of banks. The competitive environment has been an ex-post of continuous innovation through the various banking products and the efficient management in resource allocation of different commercial banks. Nevertheless, there is need for banks to survive and make decent returns by being effective and efficient in bank operations.

Currently there are 43 commercial banks in Kenya and each bank tries to be unique as compared to its competitors so as to achieve a higher market share. Therefore, it becomes imperative to determine the efficiency level of banks in Kenya and determine the factors that affect the bank efficiency level. Costs have also been increasing as a result of increased capitalization directed towards the strategic expansion of banks through their branches countrywide. Specifically, other banks continue to establish countrywide money agents in an attempt to capture and expand their markets.

1.3 RESEARCH QUESTION

- i. What are the factors that can lead to cost efficiency of commercial banks in Kenya?

1.4 OBJECTIVES OF THE STUDY

The general objective of this research paper will be to identify the major determinants of cost efficiency of commercial banks in Kenya. The specific objectives will be as follows:

- i. To determine the factors that can lead to cost efficiency of commercial banks in Kenya.
- ii. Offer policy recommendations based on the results of the study.

1.5 SIGNIFICANCE OF THE STUDY

The research paper findings will be useful in the advancement of the present frontier of knowledge on the determinants of the cost efficiency level of commercial banks, with special reference to commercial banks in Kenya. In addition, they will act as a source of information for policy makers and commercial bank management.

1.6 LIMITATIONS OF THE STUDY

The study will be limited by the existing available data sample for econometric analysis. The comprehensive panel data for commercial banks in Kenya with the necessary variables for analysis is spread out over a short period of time i.e. from 2002-2011. In addition, the concept of cost efficiency analysis of the Kenyan commercial banks is a fairly new area of study with a limited number of studies carried out to date. Most of the studies done have concentrated on the productivity and profit efficiency of the banking industry. Therefore, this study may be constrained in terms of scope and may recommend future empirical studies to be done on the various methodologies of measuring cost efficiency for further analysis.

1.7 ORGANISATION OF THE STUDY

Following the introduction part of this research paper represented by chapter one, chapter two reviews literature on cost efficiency level of commercial banks all over the globe. It summarizes what has been done in both the theoretical and empirical

literature. Chapter three that contains the research methodology then follows. Chapter four contains the presentation and discussion of the findings and finally, chapter five ends with the conclusion and policy recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.0 INTRODUCTION

There has been tremendous studies in the area of bank efficiency around the globe, most of which have concentrated on the Western economies and recently in the South East Asian economies. In contrast, there are few if any any, empirical studies on the commercial bank cost efficiency in Kenya.

This chapter majorly consists of theoretical and empirical literature review. It concludes with an overview of the literature reviewed.

2.1 THEORETICAL LITERATURE REVIEW

The theoretical literature on bank efficiency is founded on the Neo-Classical Microeconomic theory, which mainly concentrates on resource allocation and utilization. The theoretical literature reviewed in this segment is broadly divided into: the conventional theory of the firm, the managerial theories of the firm, behavioral theories of the firm and the X-efficiency theory of the firm. They are comprehensively discussed below.

2.1.1 CONVENTIONAL THEORY OF THE FIRM

The concept of efficiency was derived from the microeconomic theory of the firm. The Neo-Classical theory mainly concentrates on the efficient allocation and utilization of resources in the production process. The conventional Neo-Classical theory treats the firm as an organization that transforms resources into consumable

goods. This transformation of inputs into output is described by a production function or production possibility set.

According to Cohen and Cyert (1975), the Neo-Classical theory of the firm assumes that the firms operate in a perfectly competitive market. The firms aim at maximizing profits that is simultaneously accompanied by revenue maximization and cost minimization. Therefore, a competitive general equilibrium is achieved by equating the marginal rate of substitution for all firms between any two variables i.e. inputs or output. The competitive equilibrium allows for all the firms to earn a normal profit. This is where the firms only earn money necessary to cover their economic costs.

Nevertheless, in the short run it is possible for some individual firms to make abnormal profits. The existence of these abnormal profits attracts other firms to enter the market and compete with incumbent firms. This allows for competition between the firms and the market price is driven down until all the firms earn a normal profit in the long run. Therefore, if any firm is unable to make normal profits due to inefficient operation in the long run, it is either acquired by the efficient firms or withdrawn from the market.

According to the Neo-Classical theory of the firm, a firm can efficiently allocate resources to produce the maximum level of output for a given level of input to survive. Nonetheless, the firm that fails to reach the efficient frontier is forced out of the market leaving the efficient ones in business. Empirical research suggests, however, that not all firms operate on the efficient frontier and a number of firms do not produce at the point where the long run average costs are minimized but still

survive in the market. Therefore, micro economists have developed alternative theories of the firm.

Demsetz (1997) observed that firms in the Neo-Classical theory reflect on the imperatives of the price system. In this regard, if the price mechanism is fully efficient, resources are well allocated. Nevertheless, the conventional Neo-Classical theory pays little attention to the internal workings of the firm and provides no analysis of the decision making process. In addition, there are no explanations on the factors that determine whether a firm will either succeed or fail. Given these shortcomings, the Neo-Classical theory of the firm has been challenged by other alternatives theories of the firm such as the managerial theories, behavioral theories and the X-efficiency theory.

2.1.2 THE MANAGERIAL THEORIES OF THE FIRM

The managerial theories of the firm argue that the managers who are agents of their principals (i.e. shareholders or owners) possess the controlling rights of the firm. They pursue their own interests and utility rather than maximizing the profits of the firm. Nevertheless, they are always subject to the profit constraint. According to Brewster (1997), a given firm manager may want to seek objectives from which he obtains prestige, power and greater personal monetary reward. In so doing, the costs may not be minimized and some sort of organizational slack results in the firm.

Baumol (1959) introduced the sales maximization model that argued that the managerial objectives (i.e. income, power and prestige) were highly correlated to the sales revenue. Therefore, Baumol suggested that the prime objective of the

management was to maximize the sales revenue after achieving some minimal level of profit necessary to satisfy shareholders. In addition, Marris (1964) developed a dynamic model of the firm by assuming that the managerial objective was to concentrate on the maximization of the growth over time of a given firm.

Williamson (1964) formulated the general managerial utility maximizing model. It argued that the managers attempted to maximize their own utility rather than that of the firm profit. He argued that the managers enhanced an expense preference. This meant that the managers achieved their own objectives by spending some of the firm profits for unnecessary purposes, thus increasing managerial satisfaction or utility.

In the 1970s the managerial theories of the firm concentrated on the principal-agent analysis. This form of analysis stemmed from two sources. Firstly, the work of Spence and Zeckhauser (1971) and Ross (1973) that concentrated mainly on the problem of arranging contracts with imperfect and asymmetric information. Secondly, the other approach that was developed by Jensen and Meckling (1976) and Fama (1980) concentrated on the agency theory. They argued that the principal-agent analysis was a nexus of contracts between the firm, principal and the agent.

Spence and Zeckhauser (1971) argued that the principal (shareholders) assign a group of agents (managers) certain tasks so as to maximize the firm value. Nevertheless, the principal lacks full knowledge and information about the firm operation and performance capabilities. However, the agents have more information of the firm operation thereby, introducing information asymmetry. The existence of information asymmetry introduces the element of uncertainty, which leads to the problem of

hidden action or moral hazard. The agents pursue their own interests such as a high salary, better working conditions, on-the-job leisure and job security amongst other utilities. But the principal is unable to observe the actions of the agent. In order to monitor the behavior of the agents, the principal has to spend additional costs. In addition, the principal tries to affect and motivate the agents' behavior in his own interest by creating an incentive compatible reward structure and remuneration package (Spence and Zeckhauser, 1971). In this regard, the principal-agent problem reduces a given firm's profit and it induces inefficiency.

2.1.3 BEHAVIORAL THEORIES OF THE FIRM

The behavioral theories of the firm question the ability of the firm to optimize on its profit maximization and cost minimization objective. This is so because it is characteristic of firms to be faced by uncertainty and asymmetry of information. Given these circumstances, Simon (1959) developed the theory of the firm that emphasized on the satisfying behavior and the bounded rationality in the decision making process. This was a contravening argument against the profit maximization or cost minimization goal. According to Brewster (1997), individuals in the firm want to act rationally, but are unable to do so because they possess cognitive limitations in solving complex problems and in possessing information. Therefore, there exists bounded rationality in decision making. Decision makers exhibit some form of satisfying behavior, which depend on some aspiration level, rather than optimizing behavior. As a consequence, the firm that operates in this manner will not minimize costs and this results in productive inefficiency.

Cyert and March (1963), building on the seminal work of Simon (1959), argued that the firm is an organization but not a unified structure. It is a coalition of various participants such as the owners, managers, workers and customers. Each of these groups or individuals possesses some varying interests and objectives. This is guided by the five objectives of the firm. They include: production, inventory, sales, market share and profit. Some of these objectives may conflict with each other. Consequently, the decision making process of the firm is a continuous process of bargaining and aspiration level. Here side payments are made to ensure compliance or entice individuals into some sub-grouping. There could be a disparity between the available resources to the firm and the payments required to keep the factors of production in place. The difference between the total resources and payments is defined as the organizational slack. For example, the wages in excess of those required maintaining labor may be paid. This organizational slack increases unnecessary costs and reduces the overall efficiency of the firm.

According to Dobbs (2000), he argues that in a stable environment the factor payments may converge towards the aspiration levels, thus pushing the organizational slack close to zero. Nevertheless, in practice the environment is usually not stationary. This is explained by the dynamics of business cycles and the onward surge of technological progress, which ensures that the firm continues to strive at the threshold efficiency frontier. Therefore, this makes it possible for inefficient firms to survive in the market, only if they are not too inefficient.

2.1.4 THE X-EFFICIENCY THEORY OF THE FIRM

The X-efficiency theory of the firm tries to link the behavioral theory to that of managerial utility theory. It was formulated in a succession of Leibenstein's seminal papers. According to Demsetz (1995), X-efficiency describes the general efficiency level of the firm in transforming inputs at minimum cost into maximum output. Leibenstein criticized the assumption of the Neo-Classical theory that firms maximized profits since many of them maximized managerial utility instead.

Leibenstein (1966) argued that his study proposed that firms and economies do not operate on an outer bound production possibility surface consistent with their resources. But rather they worked on a production surface that is within the outer bound. This meant that individuals and organizations neither work as hard nor as effective as they should. In situations where there is less competition pressure, many individuals faced a tradeoff between the disutility of greater effort, search and the control of other individuals' activities for the utility of feeling less pressure and better interpersonal relations. But in high competitive pressures, the costs of such trades became high, thus individuals exchanged less of the disutility of effort for the utility of freedom from pressure.

From his findings, Leibenstein (1966) identified two sources of inefficiency. Firstly, the divergence between price and marginal cost, defined as the allocative inefficiency. He argued that one of the causes of allocative efficiency would be due to monopoly, tariffs and other impediments to competitive output rates. Secondly, the other was X-inefficiency, which emanated from the failure of businesses to achieve the lowest possible cost functions for the production of goods. These accounted for wasted

resources. X-inefficiency strived at analyzing the intra-firm behavior and relations. It also entailed the interaction of people within a firm rather than the mechanics of the price system. Leibenstein (1966) concluded that the inefficiency that is derived from X-inefficiency is significant in comparison to the inefficiency derived from the allocative inefficiency.

Leibenstein (1978) identified the non-maximizing behavior of the firm as the key idea of X-efficiency. The degree of X-inefficiency was primarily determined by the effort of the individuals within the firm. In addition, due to the incomplete contracts between the principals and the agents, the agents could evade the consequences of cost overruns, as they had no motivation to keep the costs down. Therefore, in this case the firms would be more X-inefficient.

Nevertheless, the concept of X-efficiency is very controversial as its meaning is considered abstract. Peel (1974) opposes Leibenstein's argument by proposing that X-efficiency can be attained with a lazy or hardworking individual. This is so only when we consider the effort and the efficiency dimension of the inputs being translated into maximum output. In addition, Stigler (1976) argues that the high minimum cost level is attributed majorly to the rationality of the agents who maximize their utility strategies rather than the X-efficiency. This criticism of X-efficiency proposes that the firm always strive at producing output at the highest level of the efficiency of the production frontier given a specified level of resource constraints.

2.2 EMPIRICAL LITERATURE REVIEW

Ferrier and Lovell (1990) applied both the Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) as they tried to measure the cost efficiency level of USA banks. They conducted an explicit comparison between the two methods and discussed the results obtained. They analyzed the cost structure of 575 banks and concluded that the two methods generally drew similar conclusions on the average level of cost efficiency. Nevertheless, the DEA cost efficiency score was higher than that of the SFA. This result was explained by their suggestion that the DEA frontier was considered sufficiently flexible to envelop the data more closely than the SFA frontier. When they decomposed the cost inefficiency into the technical and allocative inefficiency, both techniques led to different conclusions. The magnitudes of both the technical and allocative inefficiency scores gave different results. In addition, the rank correlation coefficients between the DEA and SFA technical efficiency and cost efficiency were 0.014 and 0.017 respectively. They were found not to be significantly different from zero. Therefore, it was concluded that the efficiencies derived from both the DEA and SFA did not give consistent rankings.

Resti (1997) further provided evidence on the efficiency levels of the Italian banking sector. He formulated a sample of 270 Italian banks and applied the use of both the parametric (SFA) and non-parametric (DEA) frontier techniques to examine the banks cost efficiency level. It was observed that the mean efficiency score ranged between 66% and 76% under both the DEA and SFA. In addition, there was a very high positive correlation for the score rankings between the two methods. Therefore, Resti (1997) concluded that both the DEA and SFA methods do not differ substantially. Nevertheless, it was reported that there was an efficiency gap that existed especially

when the efficiency values were grouped into different classes of geographical area and bank size.

Drake and Weyman (1996) used both the parametric (SFA) and non-parametric (DEA) techniques to estimate the cost efficiency level of 46 British building societies. It was observed that the cost inefficiency level in the British building society sector were of the order of 12% to 13% for parametric and non-parametric techniques respectively. In addition, there was a high rank correlation order between the two sets of efficiency scores.

In a comprehensive study of bank efficiency level in the USA, Bauer et al. (1998) applied the use of four frontier approaches. They included: the Stochastic Frontier Analysis (SFA), Distribution Free Approach (DFA), Thick Frontier Approach (TFA) and the Data Envelopment Approach (DEA). They formulated a panel of 683 large US banks over a period of 12 years from 1977-1988. Bauer et al. (1998) found out that the mean cost efficiency of the parametric models were higher in comparison to the non-parametric models. It was observed that the parametric models had an average of 83% and the non-parametric models averaged at 30%. In addition, the average rank correlation between the DEA and the parametric techniques was 10%. This suggested that the parametric and non-parametric approaches gave very weak consistency in their efficiency scores.

Bauer et al. (1998) also observed that the identification that was done between the best and the worst practicing banks was not consistent between the DEA and the parametric techniques. However, there were similarities in the distributional

characteristics between the efficiency scores and the efficiency rank correlation. It was generally observed that all the approaches were stable over time although the DEA approach had slightly better stability than the parametric techniques. However, it was concluded that the parametric technique appeared more consistent with competitive conditions in the market. This was argued to be true as economic theory stipulates that competition drives out of the market the most inefficient banks. Therefore, the banks that are left in the market in the long run should be reasonably efficient. Finally, it was argued that the parametric methodologies seemed to be consistent with the standard non-frontier performance measures (i.e. the financial ratios). Nevertheless, the non-parametric measures were in contrast to this observation, as they were weakly related to the financial ratio performance measures.

In estimating the cost efficiencies of a sample of 254 large banks in the US, Eisenbeis et al. (1999) applied the use of both a Stochastic Frontier Approach (SFA) and Linear Frontier Approach (DEA). They applied the two techniques in order to compare the robustness of the results obtained. The study period was from 1986-1991. It was observed that the DEA inefficiency scores were two or three times larger than those generated by SFA. The average of those generated by the DEA was 30% against 15% for SFA. When the banks were classified into various size-based quartiles, they found out that the level and variation of small banks inefficiency scores were on average higher than those of large banks. Inefficiencies were observed to persist over time. Nevertheless, it was observed that the persistence results for the linear programming estimates (DEA) were significantly greater than those of the econometric estimates (SFA). Moreover, the efficiency rank-order correlations between the two approaches ranged from a low of 0.44 to a high of 0.58.

Eisenbeis et al. (1999) concluded that there were significant differences between the efficiency measures derived from both the DEA and SFA techniques. Both techniques examined the relationship between bank cost efficiency and their risk taking behavior, managerial competence and stock price return behavior. It was concluded that both the parametric and non-parametric efficiency estimates produced reasonable and informative efficiency scores. Nevertheless, it was argued that the SFA estimates should be given more weight in the assessment of the banking efficiency than those provided by the DEA methodology.

Recently, Casu and Girardone (2004) while determining the cost characteristics, profit efficiency and productivity change of the Italian conglomerates applied the use of three frontier models. They included: the Data Envelopment Analysis (DEA), Stochastic Frontier Analysis (SFA) and Distribution Free Approach (DFA). They used a sample of Italian banking groups and established an unbalanced panel from the year 1996-1999. It was observed that the efficiency measures from the stochastic and deterministic frontiers were reasonably similar in magnitude and variation in the efficiency level. The standard deviation was around 10%. In addition, the DEA trend of cost efficiency was increasing between 1996 and 1998. It was also observed that there was a sharp decline in cost efficiency in 1999. Nevertheless, the SFA estimates exhibited a steady improvement in the cost efficiency. It was also observed that the DFA efficiency estimates were consistent with the DEA scores rather than the SFA. In addition, the DFA displayed a decreasing trend in efficiency level.

Becalli et al. (2006) measured the cost efficiency of the stock market listed European banks in year 1999 and 2000. They investigated the link between the efficiency

measures and the market performance of financial institutions by the use of the mean measures of SFA and DEA. They found out that the percentage changes in stock prices reflected in the cost efficiency, particularly the percentages derived from the DEA technique. Furthermore, the SFA efficiency scores were slightly higher than the DEA scores. This was 85% against 83% respectively. The DEA efficiency scores were also more dispersed than those of the SFA technique.

Majority of the empirical papers discussed on SFA technique have applied the Battese and Coelli (1995) model for data analysis. It is a translog cost function that is considered to be a suitable frontier estimation technique that captures the dynamic changes of the cost function over time. The optimal cost function derived represents a long run equilibrium relationship between total costs on one-hand and outputs and input prices on the other. In this regard this research paper extends the use of the SFA model using the Battese and Coelli (1995) model.

2.3 LITERATURE OVERVIEW

In the more recent studies on cost efficiency of the banks, it has been generally observed that the SFA efficiency scores are higher in comparison to those of the DEA. This may be as a result of the SFA model treatment of stochastic noise and the ability to control for heterogeneity that enables this model to give unbiased estimates of bank cost efficiency. In this regard, the SFA is superior to the DEA. Nevertheless, the different results of the reviewed literature generally generate mixed results for the different frontier techniques. Some studies have found a strong relationship between the different techniques (i.e. parametric versus non-parametric) whilst others have reported a lack of consistency between the two approaches.

Nonetheless, there exists some consensus in the literature reviewed. These studies have aptly demonstrated that neither the parametric nor non-parametric techniques have an absolute advantage over the other. But in certain specific situations depending on the sample size or the amount of noise (i.e. the unexplained random variation of a given sample) and data inefficiency, some estimation techniques out-perform others. Given the above conclusion, this research paper proceeds with the empirical analysis of the determinants of bank cost efficiency in Kenya. The study will apply the Stochastic Frontier Analysis (SFA) to the data derived from the Nairobi Securities Exchange (NSE) listed banks. This will be done to fill on the knowledge gap that presently exists on the cost efficiency level of the commercial banks in Kenya.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 INTRODUCTION

This chapter comprises of four segments, i.e. the concept of efficiency, theoretical framework, empirical model and the data sources. The main objective of this chapter is to describe the modeling framework that will measure the determinants of cost efficiency level of the commercial banks in Kenya. There is no consensus on the preferred method of measuring bank cost efficiency or performance. This is so because there is extensive use of both the DEA and SFA techniques on this subject of inquiry.

Nevertheless, this study will use the parametric technique of SFA due to its convenience and ease of application in the data econometric analysis. The SFA approach is chosen over the DEA approach because it allows for the use of simultaneous estimation of both the cost function and the inefficiency model, a feature not supported by the DEA linear programming approach. In addition, the SFA is superior to the DEA technique due to the fact that it better deals with the problem of statistical noise (i.e. unexplained random variation of a sample) with the use of panel data, thus allowing for the inclusion of additional information from the multiple time periods into the estimation.

This study applies the intermediation approach in determining the banking efficiency level. The intermediation approach is superior to the production approach because the former is characterized by fewer data problems than the latter (Rao, 2002). The intermediation approach is thus considered to be effective in evaluating the entire

banking industry, as it is inclusive of the interest expenses that accounts for 50-60 percent of the total costs in banks.

3.1 CONCEPT OF EFFICIENCY

Farrell (1957) in his seminal study laid the foundation of the measure of efficiency and productivity at the micro level. His study contributed in the definition of efficiency and productivity. The fundamental assumption used by Farrell (1957) is considered a departure from the perfect input-output allocation, as it allows for inefficient operations in a given firm. Therefore, inefficiency is defined as the distance from an accepted benchmark frontier production function of a given firm. In this regard, a firm that produces on the benchmark frontier is considered to be perfectly efficient. If the firm produces below the benchmark frontier, then it is considered to be inefficient, with the ratio of the actual to potential production being used to define the level of efficiency of a given individual firm.

The frontier is either established in terms of cost or technical efficiency (where the maximum output is derived for a given input) or allocative efficiency (given output for a given minimum input). There are basically two methods of measuring efficiency. These are: Parametric Stochastic Frontier Analysis (SFA) and a Non-Parametric Development Envelope Analysis (DEA). Over time the output and input functions have become highly sophisticated. Therefore, this has led to the application of translog functions i.e. the flexible functional forms in preference to the more traditional Cobb-Douglas production function. The SFA approach uses econometric techniques and imposes a priori for the frontier and distribution efficiency. The DEA

approach on the contrary relies heavily on the Linear Programming technique to obtain the benchmark of optimal cost and production factor combinations.

The Stochastic Frontier Analysis (SFA) applies a two-stage format procedure of measuring bank efficiency level. In the first stage, efficiency is measured from the benchmark frontier level; where total cost is a function of vector inputs and outputs. In the second stage of analysis, the calculated measures of efficiency are treated as the dependent variables with a given selection of determinants of efficiency that are treated as independent variables.

Therefore, this research paper will strive at applying the parametric methods (i.e. the SFA) of estimating the frontier with statistical inference. Therefore, it will impose an implicit functional form for both the frontier and the deviations from it that will form the inefficiency level of a given bank. The difference between an inefficient observed Decision Making Unit (DMU), i.e. the observed bank and an efficient reference DMU on the frontier is attributed to the deterministic construction of the frontier. Thus the estimation of the frontier allows for the random noise being considered in the analysis. This involves the estimation of a stochastic frontier.

3.2 THEORETICAL FRAMEWORK

3.2.0 STOCHASTIC COST FRONTIER APPROACH

From the economic theory, the cost or profit function is a dual of the production function. Therefore, this allows for the treatment of multiple outputs, quasi-fixed inputs, behavioral objectives and the analysis of the economic efficiency levels. This condition from microeconomic theory between the production and cost functions

ensures that both contain the same information about the firm's production possibility function. Therefore, the cost function allows for the economic determination of the technically efficient combinations of factors of production, which minimize the total cost of the various output levels. This aspect of cost minimization is referred to as the allocative efficiency. Shephards (1953) demonstrated how the cost functions are derived from the production function and the input prices of factors of production. A general cost minimization function (also referred to as the cost frontier) can be expressed as follows:

$$TC_i \geq TC^* = f(Q_i, W_i; \beta) \dots \dots \dots \text{Equation 3.0}$$

$$i = 1, 2, 3, \dots \dots \dots N$$

Where TC_i is defined as the total observed cost of individual bank i ; Q_i is a vector of outputs of bank i ; W_i is the input price vector of bank i ; $f(Q_i, W_i; \beta)$ is the cost frontier common to all banks representing the minimum cost of producing outputs Q_i when the bank face input prices W_i and β is a vector of technology parameters to be estimated and N represents the total number of banks i.e. firms under study. According to Coelli et al. (2005), the cost function should satisfy the following assumptions.

- Non-negativity constraint: $f(Q_i, W_i) > 0$ for $Q > 0$ and $W > 0$. This means that it is possible to produce a positive output without incurring any costs.
- Non-decreasing in output, Q : if $Q^0 > Q^1$ then $f(Q^0, W) > f(Q^1, W)$. This implies that an increase in input prices will not lead to a decrease in costs.

- Non-decreasing in input prices, \mathbf{W} : if $\mathbf{W}^0 > \mathbf{W}'$ then $f(\mathbf{Q}^0, \mathbf{W}) > f(\mathbf{Q}, \mathbf{W}')$.

This implies that an increase in input prices will not lead to a decrease in costs.

- Homogeneity of degree one in input price, \mathbf{W} : $f(\mathbf{Q}, k\mathbf{W}) = k f(\mathbf{Q}, \mathbf{W})$ for $k > 0$. This means that a proportional increase or decrease of all input prices will cause the same proportional change in total costs.
- Concavity in input prices.

$$\mathbf{W}: f[\mathbf{Q}, \theta\mathbf{W}^0 + (1-\theta)\mathbf{W}'] > \theta f(\mathbf{Q}, \mathbf{W}^0) + (1-\theta) f(\mathbf{Q}, \mathbf{W}') \text{ for all } 0 \geq \theta \leq 1$$

The Cost Efficiency (CE) is measured relative to the efficient cost frontier. It is derived as the ratio of the minimum cost to the actual cost incurred. Therefore, the total cost incurred in producing a given output level is \mathbf{TC} and the technically efficient combination of factors of production which minimizes costs for the output level is \mathbf{TC}^* , then the cost efficiency of the firm will be $\mathbf{CE} = \mathbf{TC}^* / \mathbf{TC}$. Failure to attain the cost frontier may be due to either technical or allocative inefficiency (or both). As the cost frontier is deterministic, such a formulation ignores measurement errors and other sources of statistical noise. All the deviations from the frontier are attributed to inefficiency.

To measure this anomaly, Aigner et al., (1977) and Meeuse et al. (1977) proposed the stochastic frontier model (SFA). Their models added a symmetric error term to the deterministic frontier that accounted for statistical noise, which is the unexplained random variation in a sample. The original models were defined as stochastic production frontiers, but the same framework was used to define the stochastic cost frontier. Aigner et al., (1977) proposal of the use of the Stochastic Frontier Analysis (SFA) was motivated by the idea that the deviations from the production frontier

might not be totally under the control of the Decision Making Units (DMUs) that are being studied. Therefore, it was concluded that SFA allows for the random errors associated with the functional form choice that results to a stochastic frontier model. To illustrate this argument, a single-equation of a stochastic cost function model for a panel data is presented below:

$$\ln TC_{it} = f(Q_{it}, W_{it}; \beta) + V_{it} + U_{it} \dots\dots\dots \text{Equation 3.1}$$

$i = 1, 2, 3, \dots, N$

$t = 1, 2, 3, \dots, M$

Where $\ln TC_{it}$ is the logarithm of the total cost of bank i at time t ; $f(Q_{it}, W_{it}; \beta)$ is the deterministic kernel of the cost frontier; Q_{it} and W_{it} are the vector outputs and input prices in the logarithmic form of bank i at time t ; N and M represent the total number of firms and period considered in the study respectively; V_{it} is a two sided normal disturbance term with mean of zero and variance, δ_v^2 . This represents the effects of noise, and U_{it} is a non-negative random disturbance term capturing the effects of the cost efficiency and is usually assumed as a half-normal distribution, $N+(0, \delta_u^2)$. V_{it} and U_{it} are independently distributed from each other. As the cost function is specified as being stochastic, the appropriate measure of cost efficiency becomes:

$$CE_{it} = f(Q_{it}, W_{it}; \beta) \exp(V_{it}) / f(Q_{it}, W_{it}; \beta) \exp(V_{it} + U_{it})$$

$$= \exp(-U_{it}) \dots\dots\dots \text{Equation 3.2}$$

The value of U_{it} cannot be observed directly from the above equation; only the composite error term $\epsilon_{it} = V_{it} + U_{it}$ can be observed. A solution to this problem is

obtained by using the distribution of the inefficiency term condition on the estimation of the composite error term. For the half normal case, Battese and Coelli (1995) proposed the appropriate point estimator for the cost efficiency, which involves the conditional expectation of $\exp(-U_{it})$ given the entire error term. This is as shown below:

$$CE_{it} = E [\exp(-U_{it}) / \varepsilon_{it}] = \{ [1 - \Phi(\delta_* - \varepsilon_{it}\gamma/\delta_*)] / 1 - \Phi(-\varepsilon_{it}\gamma/\delta_*) \} \cdot \exp(-\varepsilon_{it}\gamma + 0.5\delta_*^2) \dots \dots \dots \text{Equation 3.3}$$

Where $\Phi(\delta_* - \varepsilon_{it}\gamma/\delta_*)$ is the standard normal cumulative distribution function and $\delta = (\delta_v^2 + \delta_u^2)^{1/2}$, $\delta_* = \delta_v^2 \delta_u^2 / \delta^2$ and $\gamma = \delta_u^2 / \delta^2$. The value of γ must lie between 0 and 1. The value of 1 indicates that the deviation from the frontier is due to cost inefficiency, while a value of 0 indicates that the deviation is explained by pure noise. According to Greene (1991), the estimates of efficiency are unbiased but inconsistent; this is because the variation associated with the distribution of the estimator (U_i / ε_i) is independent of i and remains non-zero. The efficiency measure of equation 3.3 takes values over the interval $(1, \infty)$ and a value equal to 1 means that the firm is fully efficient. Hence, the cost efficiency score can be calculated as $1/CE_{it}$.

3.3 EMPIRICAL MODEL

According to Battese and Coelli (1995), the optimal total cost function as derived in equation 3.4 shows a long run equilibrium relationship between total costs on one hand and inputs and outputs and bank specific factors on the other hand. In addition, it has all the properties of a cost function as according to the microeconomic theory.

Therefore, the analysis of the cost efficiency approach can be extended by the use of panel data.

In order to estimate the coefficients of the cost stochastic frontier function and efficiency model, equation 3.4 and 3.5 are estimated by the use of the method of maximum likelihood (ML) and the tobit model respectively. Equation 3.4 is given by the transcendental logarithmic (translog) cost function. This is so as the translog function is considered to be a suitable frontier estimation technique as according to Battese and Coelli (1995). Given that it is assumed that the cost function must be linearly homogenous in input prices, the assumption of homogeneity of degree one in input prices is introduced in the model. This is derived from the normalization of the total costs and input prices by the input price of funds, i.e. w_2 , before the log transformation is effected on the model. According to Berger and Mester (1997) this transformation is used to solve the problem of heteroskedasticity. In addition, the symmetric restrictions of the parameters are imposed i.e. $\sigma_{12}=\sigma_{21}$ and $\beta_{13}=\beta_{31}$. The estimated translog econometric cost function is as shown below in equation 3.4:

$$\begin{aligned}
 \text{Ln}(TC/w_2)_{it} = & \beta_0 + \beta_1 \text{Ln}(w_1/w_2)_{it} + \beta_3 \text{Ln}(w_3/w_2)_{it} + 0.5 \beta_{11} [\text{Ln}(w_1/w_2)]^2_{it} + 0.5 \beta_{33} \\
 & [\text{Ln}(w_3/w_2)]^2_{it} + \beta_{13} \text{Ln}(w_1/w_2)_{it} \cdot \text{Ln}(w_3/w_2)_{it} + \theta_1 \text{Ln}(\text{Sec})_{it} + \theta_2 \\
 & \text{Ln}(\text{Loans})_{it} + 0.5 \theta_{11} [\text{Ln}(\text{Sec})]_{it}^2 + 0.5 \theta_{22} [\text{Ln}(\text{Loans})]_{it}^2 + \lambda_{11} \\
 & \text{Ln}(w_1/w_2)_{it} \cdot \text{Ln}(\text{Sec})_{it} + \lambda_{12} \text{Ln}(w_1/w_2)_{it} \cdot \text{Ln}(\text{Loans})_{it} + \lambda_{13} \\
 & \text{Ln}(w_3/w_2)_{it} \cdot \text{Ln}(\text{Sec})_{it} + \lambda_{32} \text{Ln}(w_3/w_2)_{it} \cdot \text{Ln}(\text{Loans})_{it} + \theta_{12} \\
 & \text{Ln}(\text{Sec})_{it} \cdot \text{Ln}(\text{Loans})_{it} + \varphi_1 \text{Ln}(\text{Fundrisk})_{it} + \varphi_2 \text{Ln}(\text{Branch})_{it} + \\
 & \epsilon \dots\dots\dots \text{Equation 3.4}
 \end{aligned}$$

Where:

- **TC** – Represents the sum of total interest and the operating expenses obtained from the banks consolidated annual income statements.
- **w₁** – Represents the labor cost, defined as the proportion of all personnel expenses to total assets.
- **w₂** – Represents the funding costs, defined as the proportion of all total interest expenses to all total deposits and other borrowed funds.
- **w₃** – Represents the capital costs, defined as the proportion of all other expenses to all the fixed and other assets.
- **Sec** – Represents the investments in government securities. It captures the output from investments in risk free assets.
- **Loans** – Represents the net loans, overdrafts and interbank loans.
- **Fundrisk** – Represents the proportion of total deposits to gross loans. This attempts to capture the effects of funding risks of the commercial banks in Kenya.
- **B₁, B₃, B₁₁, B₃₃ and B₁₃** represents the parameter estimates for input price vectors.
- **λ₁₁, λ₁₂, λ₁₃, λ₃₂ θ₁, θ₂, θ₁₁, θ₁₂ θ₂₂, φ₁ and φ₂** represent parameters of output variables and various variables of the total cost frontier function.

The efficiency model is given as follows:

$$(1-U_{it}) = \phi_0 + \phi_1 \text{Ln}(\text{Sec})_{it} + \phi_2 \text{Ln}(\text{Branch})_{it} + \phi_3 \text{Ln}(\text{Inflation})_{it} + \phi_4 \text{Ownershipd1} + \phi_5 \text{Ownershipd2} + \phi_6 \text{Ownershipd3} + \phi_7 \text{Ln}(\text{Contass})_{it} + \phi_8 \text{Ln}(\text{loans})_{it} + \epsilon_{it} \dots \text{Equation 3.5}$$

Where:

- **U_{it}** – Represents the unobserved inefficiency score of bank i at time period t.
- **1-U_{it}** – Represents the unobserved efficiency score of bank i at time period t.
- **Sec** – Represents the investments in government securities. It captures the output from investments in risk free assets.
- **Loans** – Represents the net loans, overdrafts and interbank loans.
- **Branch** – Represents the size of a given bank from its number of branches. To model the banks heterogeneity from the scale effects on operating costs this research paper applies the number of bank branches.
- **Contass** – Represents a measure of asset concentration. The proportion of the government debt investments to the private bank investments presents this.
- **Inflation** – Represents inflation over a specified period of time. It captures the effects of macroeconomic uncertainty and policy stance on bank inefficiency.
- **Ownershipd1** - Represents a dummy variable included to evaluate the impact of local public banks on bank efficiency (i.e. the binary value of 1 for local state banks and 0 for otherwise).
- **Ownershipd2** - Represents a dummy variable included to evaluate the impact of local private banks on bank efficiency (i.e. the binary value of 1 for local private banks and 0 for otherwise).

- **Ownership_{d3}** Represents a dummy variable included to evaluate the impact of foreign banks on bank efficiency (i.e. the binary value of 1 for foreign banks and 0 for otherwise).
- $\phi_0, \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7$ and ϕ_8 represents the parameters of the various variables of the determinants of bank cost efficiency.

Although equations 3.4 and 3.5 are estimated by both the maximum likelihood method and tobit model respectively, it is important to highlight the significance of the inefficiency correlates. There may be relationship between inefficiency estimates and possible independent determinants. Battese and Coelli (1995) argue that both the frontier and inefficiency model could have all or some variables in common as drivers of costs may influence the bank efficiency performance.

Accordingly, for the purpose of the study, bank output is defined as a sum of net loans, overdrafts and interbank loans. This will reflect the gain from the financial intermediation process. Investments in government securities (sec) capture the output from investments in risk free assets. The input prices will include: labor costs (w_1), funding costs (w_2) and capital cost (w_3). The coefficients of input prices and output variables are expected to be positive, implying that the total bank costs will rise with increased use of inputs in the production process.

3.4 DATA SOURCE

This study will apply secondary data for the purpose of analysis. Data will be extracted from the balance sheets and income statements of the commercial banks in Kenya listed in the NSE that fall within the sample for the period of 2002-2011. The financial reports will be collected from the individual bank registers or the respective annual NSE handbooks. Another source of the secondary data will be the CBK statistical documents, the banking surveys for the years of analysis and finally the Kenya Bureau of Statistics amongst others.

CHAPTER FOUR

PRESENTATION AND DISCUSION OF RESULTS

4.1 INTRODUCTION

This chapter comprises of five segments, namely: the cost frontier estimates, model specification, statistical approximations of the cost efficiency level, empirical results and discussion of the tobit regression model and finally a brief summary of the chapter.

4.2 COST FRONTIER ESTIMATES

The Stochastic Frontier cost model is estimated using the maximum likelihood technique that is based on Stata. It follows the parameterization developed by Battese and Coelli (1995) model. The maximum likelihood estimates are presented in table 4.0. The model assumes that the cost inefficiency term follows a truncated normal distribution. Although the translog cost function is more flexible than the functional forms, multicollinearity may exist among variables thus resulting in inconsistent parameter estimates. Nevertheless, multicollinearity may not be a serious problem if the obtained efficiency scores are purely used for forecasting purposes.

Given that the model is fairly complex due to the many interactions between output and input price variables of the translog functional form, some of the estimated individual coefficients may not be directly interpretable. In addition, the normalization of the variables allows that the estimated first-order parameters of the translog function to be directly interpreted as the approximation of cost elasticity estimates.

Table 4.0 shows that the parameter estimates of output quantities and input price are positive and highly significant from zero. This results support the theoretical requirements advanced for a valid cost function that assumes that the cost function is non-decreasing in outputs (**Q**) and input prices (**w**). The empirical results of the translog cost function as summarized in table 4.0 are consistent with the priori hypothesis as output and input price variables have the expected positive signs. Furthermore, this suggests that multicollinearity is not a serious problem in this research paper's empirical estimates of the cost function.

Table 4.0 Maximum Parameter Estimates for Stochastic Frontier Cost Function

| | <u>Coefficient</u> | <u>Parameters</u> | <u>Standard Error</u> | <u>t-statistic</u> | <u>p-value</u> |
|----------------------------------------|--------------------|-------------------|-----------------------|--------------------|----------------|
| Intercept | β_0 | -13.6985 | 1.5041 | -9.1074 | 0.0000*** |
| $\ln(w_1/w_2)$ | β_1 | 0.1479 | 0.0233 | 6.3476 | 0.0000*** |
| $\ln(w_3/w_2)$ | β_3 | 1.8607 | 0.3026 | 6.1490 | 0.0000*** |
| $0.5(\ln(w_1/w_2))^2$ | β_{11} | -0.3153 | 0.0555 | -5.6811 | 0.0000*** |
| $0.5(\ln(w_3/w_2))^2$ | β_{33} | 0.2313 | 0.0777 | 2.9768 | 0.0030*** |
| $\ln(w_1/w_2) \cdot \ln(w_3/w_2)_{it}$ | β_{13} | 0.0799 | 0.1038 | 0.7698 | 0.4410 |
| $\ln(\text{Sec})$ | θ_1 | 2.3807 | 0.2955 | 8.0565 | 0.0000*** |
| $0.5 \ln(\text{Sec})^2$ | θ_{11} | -0.0391 | 0.0134 | -2.9179 | 0.0040*** |
| $\ln(w_1/w_2) \cdot \ln(\text{Sec})$ | λ_{11} | 0.0695 | 0.0166 | 4.1868 | 0.0000*** |
| $\ln(w_3/w_2) \cdot \ln(\text{Sec})$ | λ_{13} | 0.2948 | 0.0423 | 6.9693 | 0.0000*** |
| $\ln(\text{Loans})$ | θ_2 | 2.1457 | 0.3802 | 5.6436 | 0.0000*** |
| $0.5(\ln(\text{Loans}))^2$ | θ_{22} | 0.0351 | 0.0301 | 1.1661 | 0.2440 |
| $\ln(w_1/w_2) \cdot \ln(\text{Loan})$ | λ_{12} | 0.1246 | 0.0186 | 6.6989 | 0.0000*** |
| $\ln(w_3/w_2) \cdot \ln(\text{Loan})$ | λ_{32} | 0.0516 | 0.0451 | 1.1441 | 0.2530 |

| | | | | | |
|--------------------|---------------|---------|--------|---------|-----------|
| Ln(Sec). Ln(Loans) | θ_{12} | -0.2058 | 0.0248 | -8.2984 | 0.0000*** |
|--------------------|---------------|---------|--------|---------|-----------|

Control variables

| | | | | | |
|--------------|-------------|--------|--------|--------|-----------|
| Ln(fundrisk) | φ_1 | 0.3086 | 0.0508 | 6.0748 | 0.0000*** |
|--------------|-------------|--------|--------|--------|-----------|

| | | | | | |
|------------|-------------|--------|--------|---------|-----------|
| Ln(Branch) | φ_2 | 0.2822 | 0.0270 | 10.4519 | 0.0000*** |
|------------|-------------|--------|--------|---------|-----------|

Diagnostics

| | |
|-------------------------|---|
| Log Likelihood function | 0 |
|-------------------------|---|

| | |
|-----------------|----------|
| Wald Chi square | 23796.46 |
|-----------------|----------|

| | |
|--------------|---------------|
| σ_u^2 | $2.34e^{-21}$ |
|--------------|---------------|

| | |
|--------------|--------|
| σ_v^2 | 0.0016 |
|--------------|--------|

| | |
|------------|--------|
| σ^2 | 0.0016 |
|------------|--------|

| | |
|----------|---------------|
| γ | $1.48e^{-18}$ |
|----------|---------------|

Wald test

Chi(2) = 139.81

Prob > Chi(2) = 0.0000

.....
Significance Level: ***, ** and * indicate 1%, 5% and 10% respectively.

Source: *Author's own calculation*

The coefficient estimate of ln(Sec) θ_1 suggest that, on average a 1% increase in the amount of money invested in government securities will increase the total cost by 2.38%. Similarly, the cost elasticity with respect to loans advanced θ_2 is 2.15%. This means that on average, a 1% increase in the amount of loans advanced to customers by commercial banks results in a 2.15% increase in the total costs. The estimated coefficient of government securities is higher than that of the loans advanced. This

implies that the amount of money invested in government securities have a significant effect on total costs than the amount advanced to the bank customers as loans.

In addition, on average a 1% increase in a unit of labor cost (w_1/w_2) results in a 0.15% increase in the total cost level. Contrary, the cost elasticity in respect to the unit capital costs (w_3/w_2) is 1.86. This means that on average, a 1% increase in the capital unit cost results in a 1.86% increase of the total costs. The capital unit cost coefficient is observed to be higher than that of the unit labor cost. This implies that the amount of expenditure used as capital costs have a significant effect on the total cost level than those associated with unit labor cost.

The control variables are introduced in the model to isolate the effect of some bank specific and environmental factors on the total costs. The results suggest that fundrisk and bank branch expansion account for a substantial increase on the overall total costs. Specifically, their estimated cost elasticities are 0.3086 and 0.2822, all of which are statistically significant at all conventions. This means that on average, a 1% increase in the fundrisk will increase the total costs by 0.3086%. The funding risk variable was defined as the proportion of total deposits to gross loans. Therefore, higher costs of the intermediation process results in the increase of the total costs of the banks. Therefore, in order for commercial banks to reduce costs associated with fund mobilization, they should strategize to attract high value and low interest bearing demand deposits.

In addition, on average, a 1% increase in the branch size on average results in 0.2822% increase of the total costs. It is generally observed that maintaining a large

branch network has a significant effect on the cost level of banks. Giokas (2008) observes that in order to get closer to their customers, banks usually expand their branches at a cost of increased operating expenditure. Grigorian and Manole (2006) argue that at times of high economic inflationary pressures banks tend to increase the number of branches. However, this branch expansion of banks tends to reduce the cost efficiency level associated with raising overhead expenses. Berger et al., (1997) argue further that branch expansion raises the revenue levels associated with increased customer convenience but comes at a cost that is disproportionately reflected by a high level of X-inefficiency.

4.3 MODEL SPECIFICATION

In order to determine the best-specified cost frontier model, this research paper adopts the Wald test instead of the Likelihood Ratio test. The advantage of the Wald test over the Likelihood Ratio test is that it is convenient to use, as it requires only one model to be estimated. Given that the model used is quite complex and nesting a model out of the model used results in the lack of convergence of the iterations, the Wald test is considered to be the best test for model specification.

The Wald test tries to show whether the parameters of interest are simultaneously equal to zero. If the parameters are considered to be equal to zero, this suggests that removing them from the model will not substantially reduce the fit of the model (see appendix 1 for further details). Based on the reported P-value, we are able to reject the null hypothesis, indicating that the coefficients for $\text{Ln}(\text{fundrisk})$ and $\text{Ln}(\text{branch})$ are not simultaneously equal to zero. This means that including these variables creates a statistically significant improvement in the fit of the model.

4.4 EFFECIENCY LEVEL

In this section, the various efficiency levels derived from the model are compared. Table 4.1 provides a statistical summary of both the inefficiency and efficiency scores of all the banks in the model. It presents the mean, standard deviation, the lowest and the highest levels of the same. The maximum efficiency score is very high suggesting that heterogeneity across banks is an important driver to cost differences. Therefore, it is evident that controlling for heterogeneity becomes of critical importance.

Table 4.1 Statistical cost efficiency in summary form

| VARIABLE | OBSERVATION | MEAN | STANDARD DEVIATION | MINIMUM | MAXIMUM |
|------------------------|-------------|----------------------|----------------------|----------------------|----------------------|
| Inefficiency score (u) | 60 | 7.54e ⁻⁰⁷ | 1.15e ⁻⁰⁷ | 5.98e ⁻⁰⁷ | 9.67e ⁻⁰⁷ |
| Efficiency score (ce) | 56 | 0.992217 | 0.0167886 | 0.931606 | 0.999999 |

Source: Author's own calculations

4.5 EMPIRICAL RESULTS OF THE TOBIT REGRESSION MODEL

The results of the multivariate tobit model of efficiency is as presented below in Table 4.2. Many studies including Maudos (2002), Casu and Molyneux (2003) and Weill (2003) have used a censored tobit model rather than the OLS to investigate the determinants of commercial bank efficiency. The tobit model is considered to be superior to OLS as it takes into account the censored nature of the efficiency scores (i.e. the dependent variable). The cost efficiency estimates are bounded between 0 and 1. Therefore, this research paper applies the tobit regression procedure with a left censored bound of zero and right censored bound of one to regress the cost efficiency

scores against set factors (i.e. ownership structure, government securities, loans, branches, inflation and contass).

Table 4.2 Parameter Estimates of Random Effect Tobit regression

| | <u>Coefficient</u> | <u>Parameter</u> | <u>Standard Error</u> | <u>t-Statistic</u> | <u>p-Value</u> |
|---------------|--------------------|------------------|-----------------------|--------------------|----------------|
| Constant | ϕ_0 | 0.6655 | 0.5701 | 11.67 | 0.0000*** |
| Ln(Sec) | ϕ_1 | 0.0026 | 0.0031 | 0.840 | 0.4000 |
| Ln(Loans) | ϕ_2 | 0.0448 | 0.0085 | 5.300 | 0.0000*** |
| Ln(Branches) | ϕ_3 | 0.0354 | 0.0101 | 3.51 | 0.0000*** |
| Ln(Inflation) | ϕ_4 | 0.0134 | 0.0024 | 5.710 | 0.0000*** |
| Ownership2 | ϕ_5 | -0.0126 | 0.0057 | -2.200 | 0.0280** |
| Ownership3 | ϕ_6 | -0.0061 | 0.0040 | -1.550 | 0.1200 |
| Ln(Contass) | ϕ_7 | 0.0003 | 0.0008 | 0.370 | 0.710 |

.....
 Significance level: * p<0.10, ** p<0.05, *** p<0.01

Source: *Author's own calculations*

The role of ownership structure in determining cost efficiency level is important. Therefore, the commercial banks in Kenya in this model are divided into three categories, namely: local public, local private and foreign banks. These variables are presented by a group of three dummy variables. After regressing the model, the local public bank efficiency level variable (i.e. ownershipd1) is dropped to avoid problems of the dummy trap. Thus interpretation of other bank ownership dummy variables on the cost efficiency level is measured relative to the local public bank efficiency level. The negative sign on both the private and foreign banks suggest that on average, both

the private and the foreign banks exhibit lower efficiency measures than the local public banks.

It is observed that the coefficient of local private banks is statistically significant at 5% level of significance. Nevertheless, the foreign bank's coefficient is observed to be statistically insignificant. The negative sign of the local private banks suggest that, on average domestic private banks exhibit lower cost efficiency than their local public counterparts. Therefore, holding other factors constant, the mean cost efficiency elasticity of private banks is 1.26% lower than their local public counterparts. This means that the local public banks are more efficient than their local private competitors. Kenya local public owned banks benefit more in comparison to their local private competitors as most government subsidies and reforms focus on them.

Moreover, local public banks are big in size and have the financial muscle for institutional lending. Thus local public banks experience lower default and bankruptcy risks. They would likely attract deposits at lower rates of interest on borrowing than their local private competitors. This saves the public banks a considerable amount of money and improves their cost efficiency levels.

Priori hypothesis suggest that bank size is highly correlated to bank performance. Therefore, the branch size is considered as a proxy of bank size. The branch size coefficient from table 4.2 is positive and significant at all conventional levels. Thus holding other factors constant, on average a 1% increase in branch size increases the probability of a bank being efficient by 0.0354%. The positive relationship between the branch size and cost efficiency can be explained by numerous reasons. Firstly,

large banks usually experience economies of scale promoted by growth and joint production of all the branches, thus leading to higher efficiency levels. Secondly, large banks attract more professionals and specialized management teams, which have greater ability to control costs and ensure increased revenues. Thirdly, larger banks have the financial capacity to diversify their portfolio investments and credit risks in an uncertain environment (Cole and Gunther, 1995). This reduces the risk levels of bank operations by the management team. Lastly, large banks enjoy market power; thereby proportionately pay less for the inputs than their small bank counterparts in absolute terms.

The parameter estimates of Contass and government securities are positive and insignificant at all conventional levels. These findings indicate that although both estimates have positive cost efficiency implications, their estimated magnitude is small to affect the bank efficiency performance.

The rate of inflation is considered as a viable measure of macroeconomic uncertainty and policy stance in any economy. The coefficient of inflation is both positive and statistically significant at all conventional levels. Specifically, all other factors held constant, on average a 1% increase in the level of inflation in Kenya results in the probability of the bank efficiency level increasing by 0.0134%. This is an ambiguous observation because a volatile macroeconomic environment is considered to be counterproductive to bank efficiency level improvement. When inflation levels increase economic agents convert their money holdings into assets to reduce erosion of their purchasing power. Therefore, their demand for money increases and money held in bank accounts decline (i.e. demand deposit).

The parameter of loans is positive and significant at all conventional levels. Therefore, holding other factors constant, on average a 1% increase on loans results in the probability of bank efficiency level increasing by 0.048%. This is explained by the concept of money creation. When more money is created the economies of scale in money creation allows for increased bank cost efficiency level. There is an increase in revenue from interest charged on loans advanced to customers by the banks. Lending for banks is the most profitable activity because interest rates realized on loans are more profitable than other investments. Thus interest received from loans is the main source of income for banks.

4.6 ANALYSIS OF THE AVERAGE BANK EFFICIENCY SCORES

Table 4.3 presents the analysis of average bank level efficiency scores. These are estimated bank-specific cost efficiency indices estimated from the translog frontier cost function. The efficiency estimates show that for the whole banking sector, banks operate below the frontier, with mean cost efficiency of 0.9925. Nevertheless, this estimated score indicates that on average, the banks in Kenya over the ten periods under study operated near the efficiency cost threshold. In addition, it is generally observed that the mean cost efficiency score of all the banks has been increasing over the sampled period.

This means that the management of these banks has been efficient in cutting down on their operating costs although with increasing income revenue. This can be explained by the various strategies adopted by the managers of these banks to cut on costs. For example, in 2011 Barclays Kenya (foreign owned bank) announced that it was cutting

200 management level staff, which accounted for 5% of its workforce, to cut on its costs. This also has been the case in Equity bank, Co-operative bank and Kenya Commercial bank amongst others. These banks are highly cost efficient, as their revenue-cost ratios have increased over the years. Thus increasing the profit level of these firms over the years. Many banks opt to trim down bloated management teams as a measure of restructuring and cutting down of high salary staff budgets.

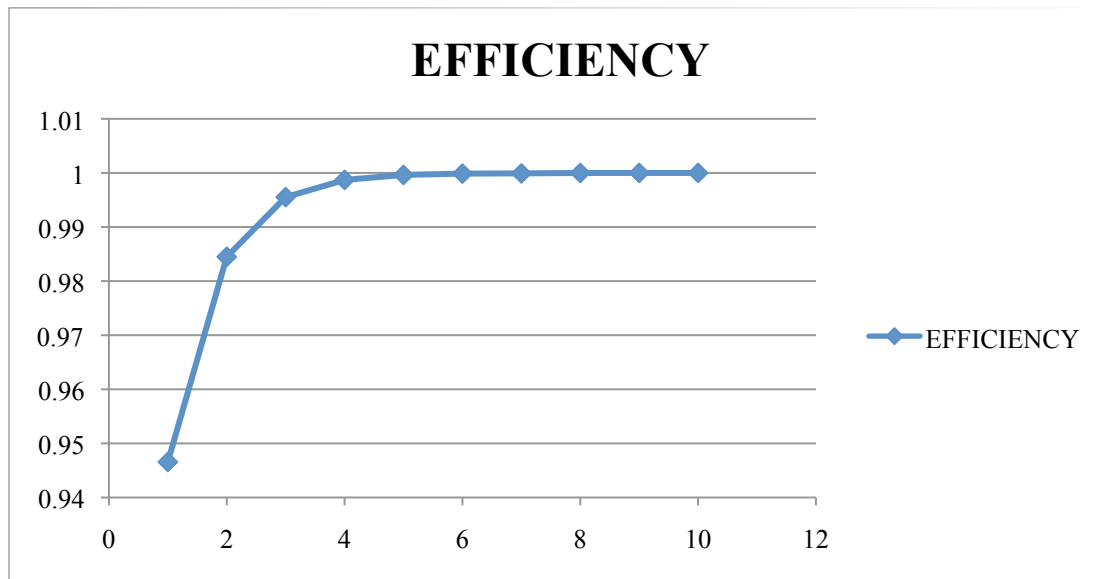
Table 4.3 Average Bank Efficiency Scores in Kenya

| YEAR | EFFICIENCY |
|-------------|-------------------|
| 1 | 0.9465700 |
| 2 | 0.9844960 |
| 3 | 0.9955300 |
| 4 | 0.9987100 |
| 5 | 0.9996300 |
| 6 | 0.9998750 |
| 7 | 0.9999150 |
| 8 | 0.9999920 |
| 9 | 0.9999976 |
| 10 | 0.9999993 |

Source: *Author's own calculations*

The average efficiency scores have been increasing steadily from the year 2002-2011 as shown on figure 4.0 below. This can be attributed to sound management strategies aimed at cutting down on operation costs as discussed above. The cost efficiency trends over the ten period is as shown below:

Figure 4.0 Average cost Efficiency scores of banks in Kenya from 2002-2011



Source: Author's own calculations

4.7 SUMMARY AND CONCLUSION

This chapter investigates and reports on the determinants of cost efficiency level of commercial banks in Kenya. The efficiency scores are obtained from the Stochastic Frontier Analysis (SFA) model. The results obtained indicate that when the SFA model is used for analysis, the average cost efficiency score of commercial banks in Kenya ranges from 94.66% in 2002 up to 99.99% in 2011. The rising efficiency cost level is associated with prudence in management strategies, innovation, portfolio diversification, increased competition and a stable macroeconomic environment in the country. Although it is generally argued that the operational costs have increased over the years, there has been an equal increase in the revenue level of banks. This has increased the revenue-cost ratio indicating increased efficiency over the sampled period.

This chapter investigates whether government securities, branch size, inflation, ownership structure and asset concentration affect the cost efficiency level of commercial banks in Kenya. The results show that the local public banks are more efficient than both the local private and foreign owned banks, although the effect of foreign owned banks is observed to be insignificant. Nevertheless, the results indicate that the differences in ownership structure does affect performance in cost efficiency terms of commercial banks in Kenya. It is further observed that the increase in branch size has a positive and significant impact on the cost efficiency level. In addition, asset concentration and investment in government securities have insignificant but positive effect on the cost efficiency level of commercial banks in Kenya. Both the loans advanced to customers and level of inflation have a positive and significant effect on the cost efficiency level of banks in Kenya.

Finally, it is generally deduced that the average cost efficiency level of commercial banks in Kenya has been increasing over the period of analysis. This is attributable to the increased revenue-cost ratios occasioned by prudence in the management of these banks over the period under review.

CHAPTER FIVE

CONCLUSION AND POLICY RECOMMENDATION

5.1 INTRODUCTION

The Kenyan banking sector has experienced tremendous changes over the last twenty years. A comprehensive program of financial liberalization was implemented in the early 1990s. Under the financial reform process implemented in the banking sector in Kenya, the government through the CBK has enhanced liberalization, deregulation and financial innovation as the main forces determining the efficient performance of the banking sector. Therefore, it becomes of critical importance to examine the bank cost efficiency level. In this regard, this study seeks to investigate the determinants of cost efficiency level of commercial banks in Kenya.

In more recent times, during the global financial meltdown, the commercial banks in Kenya continued to report growth in profitability levels. This observation is greatly attributable to the country's stable macroeconomic environment. The government through the CBK has been effective in employing discretionary policy instruments to enhance bank efficiency. Kenya adopted expansionary fiscal and monetary policies in response to the effects of the global financial crisis. These policies were effective as the bank cost efficiency increased as observed in Kenya over this period (i.e. from 2007-2009).

As at 2009 the banking sector in Kenya had maintained a high capital adequacy ratio, adequate liquidity level and low performance loans in relation to the gross loans. The total shareholders funds, deposits and assets expanded by 25.4%, 9.5% and 15% respectively. Liquidity remained strong, with the ratio of liquid assets to total deposit

liabilities being at 40.9%. This is above the statutory minimum requirement of 20%. The overall performance of the banking sector has since remained strong and sound. In addition, the banks have in recent times continued to out perform other firms in other sectors, thus reporting increased profit streams despite the high interest rates and inflationary level.

5.2 CONCLUSION OF THE STUDY

In consideration of the above developments, the main aim of this research paper is to provide an empirical analysis of cost efficiency of the Kenyan banking system over the period from 2002-2011. Following the extensive literature reviewed, the paper employs the SFA technique analysis to estimate the cost efficiency of Kenyan banks. In addition, the main determinants of the Kenyan banking cost efficiency are investigated using the tobit model. The study employs an unbalanced panel, which covers a sample 10 of the listed commercial banks in Kenya and totals 100 observations.

In order to reach the best-specified stochastic cost frontier model for measuring the total cost function, the research paper employs the Wald test specification testing procedure. The findings show that the results incorporate the control variables, thus accounting for bank heterogeneity on key parameter estimates. The results further show that accounting for heterogeneity across banks is crucial. Therefore, the preferred model specification for the sample is the one stage SFA model that includes: input prices, outputs and control variables.

This study stipulates that on average, commercial banks in Kenya exhibit a high level of cost efficiency. In this regard, it is established that using the SFA model, the average cost efficiency level of commercial banks in Kenya over the period from 2002-2011 is around 99.25%. It is further observed that the efficiency level has been increasing over the period under study. This efficiency improvement is in response to the concerted efforts of cutting down on the operating costs by the management of banks, explained by the various strategies adopted.

The empirical analysis carried out in this research paper further investigates the main determinants of cost efficiency level in commercial banks in Kenya using the tobit model. This is aimed at providing important insights on internal bank management improvement and informs policy. The results show that the local public banks are more efficient than the local private and foreign banks (based on the efficiency scores obtained from the SFA model). It is further observed that bank size is positively related to cost efficiency. The economies of scale concept is advanced to explain this observation. Both the asset concentration and government securities coefficients are observed to be positive and insignificant. Nevertheless, both the inflation and loans coefficients are observed to be positive and significant. The inflation coefficient sign is considered as being ambiguous as a volatile macroeconomic environment is considered to have a negative effect on bank efficiency level.

5.3 POLICY IMPLICATIONS

The empirical findings from this research paper have pointed on to the increasing cost efficiency level of commercial banks in Kenya despite a background of reported

proportional increase of the general cost levels. The policy implications of this research paper are as summarized as follows:

Firstly, foreign owned banks are not sufficient enough to be significant in ensuring that they affect the general efficiency level of commercial banks in Kenya. This is supported by this research paper's empirical findings that report that the foreign bank's parameter is statistically insignificant in explaining the general cost efficiency level. Economic theory of comparative advantage stipulates that the foreign banks enjoy superior risk management, advanced technology and vast experience through international exposure. Nevertheless, it becomes important that the Kenyan banking sector should be conducive enough to attract foreign investments. This is in terms of the increase of foreign bank presence in the country so as to increase competition and transfer this advantages effectively to the local financial sector. Therefore, opening up of the Kenyan banking sector to foreign banks is critical to ensure that there is improvement of cost efficiency levels and increased innovation.

Secondly, it is observed that larger banks are more cost efficient than smaller banks. This finding suggests that the Kenyan banks can improve their cost efficiency by increasing their size. This can be achieved from increased branch networks of the various banks, mergers and acquisitions. In addition, the government through the CBK should provide incentives for increased branch networks to ensure that there is financial inclusion for the majority of the unbanked population. This is in terms of provision of security, infrastructural development and social amenities especially in the sprouting cities and urban centers.

Finally, the bank management should improve on the current risk assessment techniques to ensure lower loan default rates and continue to invest in risk free assets such as the government securities. In addition, the government through the CBK should reinforce and ensure full operation of the recently established credit reference bureau to reduce moral hazard in the loan advancement process by the various banks.

5.4 AREAS FOR FURTHER RESEARCH

This study employs only the SFA technique in the empirical analysis of the determinants of the cost efficiency level of commercial banks in Kenya. Not all the available efficiency frontier techniques are used in this study. For example there is the Data Envelopment Analysis, the Distribution Free Approach and Thick Frontier Approach. To ensure robust results in data analysis, it might be interesting that further research in the future employs all these techniques to estimate the relative cost efficiency levels of commercial banks in Kenya. In addition, this would ensure proper analysis of the procedures used and allow for the objective assessment and comparison of the estimated parameters adding value to present literature on the Kenyan banking sector.

The SFA model used in this research paper incorporates the control variables to control for the heterogeneity problem. From the analysis it is possible that the model does not properly employ the heterogeneity variables in the regression thus making the regression incomplete. This may result in omitted variable problem and create potential biases in the obtained efficiency estimated scores. Green (2005) proposes a solution to this problem; it entails the integration of an additional stochastic term in

the prescribed SFA model. This distinguishes all the time invariant unobserved heterogeneities from the inefficiency term.

Finally, this research paper was also constrained by the number of observations included in the data sample. This is greatly attributable to the fact that the panel was constrained to the listed commercial banks in the NSE over a period of ten years. Nevertheless, gradually as the years proceed the data set is increasing and becoming more readily available through the CMA. Future research can take advantage of the larger sample size to provide a more comprehensive study on the determinants of bank cost efficiency in Kenya.

REFERENCE LIST

- Aigner, D., Lovell C. and Schmidt, P. 1977, "Formulation and Estimation of Stochastic Frontier Production Function models", *Journal of Econometrics*, Vol. 6, pp. 21-37.
- Battese, G. and Coelli, T. J. 1995, "A model of Technical Inefficiency Effects in a Stochastic Frontier Production function for Panel Data", *Empirical Economics*, Vol. 20, pp. 325-332.
- Bauer, P. W., Berger, A. N., Ferrier, G. D. and Humphrey, D. B. 1998, "Consistency conditions for Regulatory Analysis of financial Institutions: A comparison of Frontier Efficiency methods", *Journal of Economics and Business*, Vol. 50 (2), pp. 85-114.
- Baumol, W. J. 1959, *Business Behavior, Value and Growth*, Macmillan: New York.
- Beccali, E., Casu, B., and Girardone, C. 2006, "Efficiency and Stock Performance in European Banking", *Journal of Business and Accounting*, Vol. 33 (1-2), pp. 245-262.
- Beck, T., Cull, R. and Jerome, A. 2005, "Bank Privatization and Performance: Empirical Evidence from Nigeria", *Journal of Banking and Finance*, Vol. 29 (8-9), pp. 2355-2379.
- Berger, A. N. and Mester, L. J. 1997, "Inside the Black Box: What Explains the Differences in the Efficiencies of Financial institutions?" *Journal of Banking and Finance*, Vol. 21, pp. 895-947.
- Brewster, D. (ed.) 1997, *Business Economics-Decision-Making and the Firm*, the Dryden Press, London.
- Casu, B. and Molyneux, P. 2003, "A Comparative Study of Efficiency in European Banking", *Applied Economics*, Vol. 19, no. 6, pp. 1865-1876.
- Casu, B. and Girardone, C. 2004, "Financial Conglomeration: Efficiency, Productivity and Strategic Drive", *Applied Financial Economics*, Vol. 14 (10), pp. 687-696.
- Central Bank of Kenya, Annual Reports, Various Issues.
- Clemhout, S. 1968, "The Class of Homothetic Isoquant Production Functions", *The Review of Economics Studies*, Vol. 35 (1), pp. 91-108.
- Coelli, T. J., Prasada Rao, D. S. O'Donnell, C. J., and Battese, G. E. (eds.) 2005, *An Introduction to Efficiency and Productivity Analysis*, Second Edition, Springer.
- Cohen, K. J. and Cyert, R. M. (ed.) 1975, *Theory of the Firm: Resource Allocation in the Market Economy*, Prentice Hall, New Jersey.
- Cole, R. A and Gunther, J. W. 1995, "Separating the Likelihood of Timing of Bank Failure", *Journal of Banking and Finance*, Vol. 19, no. 6, pp. 1073-1089.

Cooper, W. W., Seiford, L. M. and Tone, K. 2007, *Data Envelop Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Second edition, Springer, New York.

Cyert, R. and March, J. 1963, *Behavioral Theory of the Firm*, Blackwell, Oxford.

Debreu, G. 1951, "The Coefficient of Resource Utilization", *Econometrica*, Vol. 19, pp. 273-292.

Demsetz, H. (ed.) 1995, *The Economies of the Business Firm: Seven Critical Commentaries*, Cambridge University Press, Cambridge: New York and Melbourne.

Demsetz, H. 1997, "The Firm in Economic Theory: A Quite Revolution", *American Economic Review*, Vol. 87 (2), pp. 426-429.

Dobbs, L. 2000, *Managerial Economics: Firms, Markets and Business Decision*, Oxford University Press, New York.

Drake, L. and Weyman-Jones, T. G. 1996, "Productive and Allocative Inefficiencies in UK Building Societies: A comparison of Non-Parametric and Stochastic Frontiers Techniques", *The Manchester School of Economic and Social Studies*, Vol. 64, pp. 229-245.

Eisenbeis, R. A., Ferrier, G. D. and Kwan, S. H. 1999, "The Informativeness of Stochastic Frontier and Programming Frontier Efficiency Scores: Cost Efficiency and other Measures of Bank Holding Company Performance", *Working Paper no. 99 (23)*, Federal Reserve Bank of Atlanta.

Fama, E. F. 1980, "Agency Problems and the Theory of the Firm", *Journal of Political Economy*, Vol. 88 (2), pp. 288-307.

Farrell, M. J. 1957, "The Measurement of Productive Efficiency", *Journal of Royal Statistical Society*, Vol. 120, pp. 253-281.

Ferrier, G. D and Lovell, C. A. 1990, "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence", *Journal of Econometrics*, Vol. 46 (1-2), pp. 229-245.

Giokar, D. 2008, "Assessing the Efficiency in Operations of a Large Greek Bank Branch Network adopting different Economic Behaviors", *Economic modeling*, Vol. 25, pp. 557-574.

Green, W. H. 2005, "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model", *Journal of Econometrics*, Vol. 126, no. 2, pp. 269-303.

Griogorian, D. and Manole, V. 2006, "Determinants of Commercial Bank Performance in Transition: an application of Data Envelopment Analysis", *Comparative Economic Studies*, Vol. 48, no. 3, pp. 497-522.

Humphrey, D. B. 1990, "Why Do Estimates of Bank Scale Economies Differ?" *Economic Review*, Vol. 76, pp. 38-50.

- Jensen, M. C. and Meckling, W. H. 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, Vol. 3 (4), pp. 305-360.
- Leinbenstein, H. 1966, "Allocative Efficiency Verses X-Efficiency", *The American Economic Review*, Vol. 56 (3), pp. 392-415.
- Leibenstein, H. 1978, "On the Basic Proposition of X-efficiency Theory", *The American Economic Review*, Vol. 68 (2), pp.328.
- Marris, R. (ed.) 1964, *The Economic Theory of Managerial Capitalism*: Free Press of Glencoe, New York.
- Maudos, J. A 2002, "Cost and Profit Efficiency in European Banks", *Journal of International Financial Markets, Institutions and Money*, Vol. 12, no. 1, pp. 33-58.
- Meeusen, W. and Van de Broeck, J. 1977, "Efficiency Estimation from Cobb-Douglas Production Function with composed error", *International Economic Review*, Vol. 18 (2), pp. 435-444.
- Molyneux, P., Altunbas, Y. and Gardener, E. 1996, *Efficiency in European Banking*, John Wiley and Sons, Sussex.
- Mwega, F. 2011, "The Competitiveness and Efficiency of the Financial Services Sector in Africa: A Case Study of Kenya", *African Development Review*, Vol. 23 (1), pp. 44-59.
- Peel, D. 1974, "A note of X-Inefficiency", *The Quarterly Journal of Economics*, Vol. 88 (4), pp. 687-688.
- Rao, A. 2002, Estimation of Efficiency, Scale and Scope and Productivity measures in UAE Banks: *Unpublished seminar paper*, USA.
- Shephard, R. W. 1953, *Cost and Production functions*, Princeton University Press.
- Simon, H. A. 1959, "Theories of Decision Making in Economics and Behavioral Science", *The American Economic Review*, Vol. 49 (3), pp. 253-283.
- Spence, M. A. and Zeckhauser, R. 1971, "Insurance, Information and Individual Action", *American Economic Review*, Vol. 61 (2), pp. 254-268.
- Stigler, G. 1976, "The existence of X-Efficiency", *American Economic review*, Vol. 66, pp. 213-216.
- Varian, H. R. 1992, *Microeconomic Analysis*, W. W. Norton and Company Inc.: New York.
- Weill, L. 2003, "Banking Efficiency in Transition Economies", *The Economics of Transition*, Vol. 12, no. 1, pp. 33-58.
- Weiss, M. A. 1986, Analysis of Productivity at the Firm Level: An application of Life Insurers. *Journal of Risk and Insurance*, Vol. 58 (3), pp. 452-479.

Williamson, O. E. 1964, *The Economics of Discretionary Behavior: Managerial Objectives in the theory of the Firm*, Prentice Hall, Englewood Cliffs, N. J.

World Bank 2006, "A Review of World Bank Assistance for Financial Sector Reforms", *Washington DC: Independent Evaluation Group*, World Bank.

APPENDIX 1

WALD TEST

In order to test for the best specified cost frontier model, the Wald test is used in this research paper. It is considered to be superior to the Likelihood ratio test as it is convenient to use because it requires only one model to be estimated. This is critical as the model used is quite complex and nesting a model out of the original model results in the lack of convergence in the iteration process. Therefore, the Wald test is considered the best test for this model's specification. The Wald test tries to test whether the parameters of interest are simultaneously equal to zero. If they are considered to be, this suggests that removing them from the model will not substantially reduce the fit of that model.

The Wald test hypothesis is as shown below:

H₀: Ln(fundrisk) = Ln(branch) = 0

H_a: Both Ln(fundrisk) and Ln(branch) are not equal to zero

Results Chi 2 (2) = 139.81

Prob > Chi 2 = 0.0000

Based on the above findings (i.e. **P-value** is less than the **Chi 2 (2)**) the null hypothesis is rejected, indicating both the coefficients for Ln(fundrisk) and Ln(branch) are not simultaneously equal to zero. This means that including these two exogenous variables create a statistically significant improvement of the model fit.