University of Fort Hare

Together in Excellence

AN ANALYSIS OF THE RELATIONSHIP BETWEEN BANK EFFICIENCY AND ACCESS TO BANKING SERVICES IN SOUTH AFRICA.

BY

ANDREW MAREDZA

A THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS OF THE DOCTOR OF COMMERCE (D.COM) DEGREE IN ECONOMICS

OCTOBER 2014

DEPARTMENT OF ECONOMICS

FACULTY OF MANAGEMENT & COMMERCE

UNIVERSITY OF FORT HARE

SOUTH AFRICA

SUPERVISOR

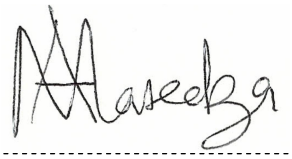
PROFESSOR SYLVANUS IKHIDE

DECLARATIONS

DECLARATION ON COPY RIGHT

I, the undersigned, **ANDREW MAREDZA**, student number **200808288**, hereby declare that the dissertation is my own original work, and that it has not been submitted, and will not be presented at any other University for a similar or any other degree award.

Date: 22 / 09 / 2014

Signature: 

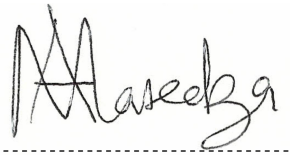
DECLARATION ON PLAGIARISM

I, **ANDREW MAREDZA**, student number **200808288**, hereby declare that I am fully aware of the University of Fort Hare's policy on plagiarism and I have taken every precaution to comply with the regulations.

Signature: 

DECLARATION ON RESEARCH ETHICS

I, **ANDREW MAREDZA** student number **200808288**, hereby declare that I am fully aware of the University of Fort Hare's policy on research ethics and I have taken every precaution to comply with the regulations. I have obtained an ethical clearance certificate from the University of Fort Hare's Research Ethics Committee and my reference number is the following:.....N/A.....

Signature: 

DEDICATION

This thesis is dedicated to my beloved wife Vongai Maredza and son, Eldad Ethan Maredza

Only believe, all things are possible

ACKNOWLEDGEMENTS

First and foremost, I am very grateful to the Almighty God for his grace and strength that has made all things possible. Nothing is impossible with God.

I express my profound gratitude to my supervisor, Professor Sylvanus Ikhide whose incisive analysis and enlightening guidance was of paramount value. Without his mentoring and valuable advice I would not have attempted or completed this thesis.

I am truly indebted to my precious and beloved wife, Vongai and cheerful son Eldad who have suffered due to my constant absence and long working hours. In particular, my wife's true love, care and inspiration gave me strength to press on. Indeed, she truly contributed much to this thesis in more ways than would be appropriate to list here.

The financial and moral support of University of Fort Hare (UFH) cannot go without mention, not forgetting the valuable comments and suggestions from many colleagues and friends at University of Fort Hare. I also thank the administrative support of the economics department at Fort Hare University and the research office at North-West University for their support in various capacities. I owe a debt of gratitude to the Vice rectors' office of North West University for awarding me such a very generous capacity development grant.

I also wish to specifically thank Anton for his efforts in providing bank data from their Bankscope database. His humility and willingness to assist others inspired me and will always be remembered.

I am grateful to my parents and friends who have also contributed to the successful completion of this thesis in many ways. Special mention goes to believers in Mafikeng and the saints in Chitungwiza pastored by William Mukuya for their unceasing prayers and love for me, without whose spiritual support I would not have accomplished this work.

Thank you to everyone who has helped to make this thesis a success.

ABSTRACT

The primary objective of this study is to investigate the nature of the relationship between bank efficiency gains and access to banking services in South Africa. The importance of making such an enquiry arises from the fact that various studies have identified access to financial services as an important vehicle for lifting the poor out of poverty. In particular, there is concern that banks` appetite for better scores on efficiency has the potential of reducing access to services for consumers particularly the low-income clients. The study attempted to answer two central research questions: Firstly, does the quest for banks to improve efficiency preclude access to banking services for some group of consumers? Secondly, do bank efficiency gains necessarily translate to improved accessibility to banking services? The researcher applied a two-stage methodology approach. In the first stage, the Hicks-Moorsteen aggregator functions were used to generate and decompose total factor productivity (TFP) into several efficiency measures for a panel of eight South African banks. First stage results revealed that the average banking sector total factor productivity efficiency (TFPE) was 59 percent implying that the observed TFP was 41 percent short of the maximum TFP possible using the available technology. A further comparison of performance revealed that large banks were better performing than small banks in terms of TFPE. Apart from estimating and decomposing TFP indices we needed to determine if there was a statistically significant change in the TFPE of South African banking system as a result of the global financial crisis. A general analysis of the generated scores showed that TFPE clearly decreased during 2008-2009, the period that coincided with the global financial crisis. We then used the Fixed Effects Model (FEM) in the second-stage analysis to examine the link between banking sector TFPE and access. The FEM was utilised to take account of bank-specific heterogeneity. The obtained results indicated existence of a positive and significant relationship between banking efficiency and access to banking services. This study suggests that banking sector efficiency plays a crucial role in promoting access to bank services in South Africa. We therefore underscore the need for all banks to attain and maintain high efficiency in order to augment government efforts towards improving accessibility for the unbanked South African people. We also found evidence similar to that reached by Kablan (2010) that an increase in the rural population is associated with a reduction in access to bank services. From this result, we speculated that banks are somewhat biased against providing their services to the general rural populace. Since the rural-population variable exerted the greatest marginal impact on access we suggested that perhaps investment in rural infrastructure would help broaden access and so improve financial inclusion on a larger scale. Finally we also investigated the link between banking sector efficiency and unemployment in South Africa. Of paramount importance in the second stage analysis was that we found a negative and significant association between banking sector efficiency and unemployment indicating that employment is influenced, inter alia, by the efficiency with which banks operate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ACRONYMS

ABSA	Amalgamated Bank of South Africa
AE	Allocative Efficiency
BASA	Banking Association South Africa
BCC	Banker, Charnes and Cooper
BEE	Black Economic Empowerment
BRICS	Brazil, Russia, India, China and South Africa
CAR	Capital Adequacy Ratio
CCR	Charnes, Cooper and Rhodes
CDO	Collateralized Debt Obligations
CE	Cost Efficiency
CIR	Cost to Income Ratio
CRS	Constant Returns to Scale
CRSTE	Constant Returns to Scale Technical Efficiency
DEA	Data Envelopment Analysis
DFA	Distribution Free Approach
DMU	Decision Making Unit
DPIN	Decomposing Productivity Index Numbers
DRS	Decreasing Returns to Scale
ECM	Error Component Model
EE	Economic Efficiency
EFFCH	Efficiency Change
EU	European Union
FAC	Financial Access Charter
FDH	Free Disposal Hull
FEM	Fixed Effects Model
FRB	FirstRand Bank
FSC	Financial Service Charter
GDP	Gross Domestic Product
GFC	Global Financial Crisis
GLS	Generalised Least Squares
HHI	Herfindahl Hicksman Index
HMTFP	Hicks-Moorsteen Total Factor Productivity

IME	Input Mix Efficiency
IMF	International Monetary Fund
IPS	Im, Pesaran & Shin
IRS	Increasing Returns to Scale
ISE	Input Scale Efficiency
ISME	Input Scale Mix Efficiency
ITE	Input Technical Efficiency;
LLC	Levin Lin & Chu
LP	Linear Program
LSDV	Least Square Dummy Variable
NIM	Net Interest Margin
NPLs	Non-performing loans
OLS	Ordinary Least Squares
OME	Output Mix Efficiency
OSE	Output Scale Efficiency
OSME	Output Scale Mix Efficiency
OTE	Output Technical Efficiency
POS	Point of Sale
REM	Random Effect Model
RISE	Residual Input Scale Efficiency
RME	Residual Mix Efficiency
ROA	Return on Assets
ROE	Return on Equity
ROSE	Residual Output Scale Efficiency
RSS	Residual Sum of Squares
RTS	Returns to Scale
SARB	South African Reserve Bank
SCP	Structural Conduct Performance
SECH	Scale Efficiency Change
SFA	Stochastic Frontier Approach
SME	Small Medium Enterprises
SSA	Sub-Saharan Africa
STATSSA	Statistics South Africa
ΔTECH	Technological Change

TE	Technical Efficiency
TECHCH	Technical Change
TFA	Thick Frontier Approach
TFP	Total Factor Productivity
TFPCH	Total Factor Productivity Change
TFPE	Total Factor Productivity Efficiency
VRS	Variable Returns to Scale
VRSTE	Variable Returns to Scale Technical Efficiency
WHO	World Health Organisation.

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

The banking sector is considered the backbone of every economy and plays a critical role in attaining growth and development. The efficiency and productivity of such an influential component of the economy is therefore essential to the wellbeing of the economy. Sufian and Majid (2007, p.21) states that "It is of public interest to know what firms can do to improve their efficiency so that scarce resources are allocated to their best uses and not wasted during the production of services and goods." Economic development and growth is therefore strongly dependent on the efficient functioning of the financial system. For the year 2012, the Banking Association South Africa (BASA) (2012) reported that the value of the total assets of the financial sector totalled over R6 trillion with the banking sector representing over 50 percent. As such, banks are the most important institutions within the financial system. The banking sector is regarded as the vital engine for growth in the sense that funds are channelled from savers to borrowers through the amenities of the banking system.

One of the contemporary issues in the South African financial landscape is the need to expand access to banking services to all South Africans. Hence, the debate around access to financial services has taken centre stage with organisations such as the Consultative Group to Assist the Poor (CGAP) advocating for provision of services to everyone including the rural poor. This is because increasing access to financial services for everyone has considerable benefits for the economy, including boosting economic growth, job creation, improving income distribution, alleviating poverty, and empowering women (National Treasury Policy Document, 2011). As Peachey and Roe (2004, p.9) put it:

> "...the issue of financial access in low-income countries is an integral part of the debate about how to address widespread poverty. It gets bracketed with issues of access to basic needs such as clean water and minimal education..."

In this light, access to banking services can be seen as a public good that is essential in our lives in a similar way as is access to clean water, primary education and health services.

This particular study seeks to contribute to current literature on banking by investigating the possibility of an inverse relationship between efficiency and access to bank services in South Africa. The importance of making such an enquiry arises from the fact that the author hypothesises that there is an inverse relationship between efficiency improvement and provision of bank services particularly for the rural and poor households. Fernando (2007) argues that given the lagging social development of the poor and the low income levels, there is little profit potential in the low end of the financial markets. Hence, banks find the impetus to "move up" the market more rewarding than "moving down" the market. Rational behaviour suggests that economic efficiency in the banking sector and elsewhere may be improved by relying more on the market segment that confers more output benefit per unit of the input resource employed. Scarce resources should be allocated to their best uses and not wasted during the production of services and goods. Hence, a dilemma exists when the same banks must expand services to low income groups which constitute a greater proportion of the population. According to Paulson and McAndrews (1998), the provision of banking services to clients of low-income status is problematic to do cost-effectively because of the inadequate balances kept on their deposit accounts. Their study explains that even if the required minimum balances are maintained these low-income customers generally make several minor withdrawals which drive up branch operational costs. Okeahalam (2006) states that banks argue that the expected returns from extending banking services to townships are low and that the average cost of production is high. The low net returns emanates from the fact that the nature of transactions in these townships is small in terms of returns and yet high in volume driving up operating cost.

Banks pursuit of good scores on efficiency has the potential of reducing access to services for consumers particularly the poor majority. Hence, the main research question this thesis attempts to answer is: do banks in their quest to improve efficiency compromise access to financial services for some group of consumers? On the other hand we ask the question: Do bank efficiency gains necessarily translate to enhanced accessibility to bank services for the people? In this context the mainstream argument is that efficiency gains which in essence is achieving the same level of outputs (volume of loans, interest income, non-interest income) with less resources (operating costs, fixed assets, deposits, number of employees) would impact accessibility positively via the operational costs (a proxy for service charges) channel. The World Bank (2006) noted that high minimum balances and monthly charges prevent a large proportion of the African population from accessing formal financial services.

Alternatively, the researcher is further inclined to expect that the benefit of a decrease in a bank`s operational cost could also be passed on to consumers in the form of reduced interest spread. Interest spread which is the difference between deposit and lending rate is a traditional measure of intermediation efficiency. Ikhide (2008), states that a basic and essential benefit of greater efficiency is a reduction in the difference between deposit and lending rates. This is expected to encourage greater demand for bank services in the form of new bank accounts or increased demand of loans for existing clients.

One important dimension of access to bank services is whether the present bank market structure affects access to services. Okeahalam (2000) argues that the existence of an oligopoly structure (imperfect competition) implies that the degree of competition necessary to stimulate greater efficiency may not exist. Napier (2005a) observed that South African banks operate as a complex monopoly, with perceived high barriers to entry. Okeahalam (2006) noted that 85% of the total deposits and assets were being controlled by the "big four" banks. Okeahalam further argued that the big four was inefficient and that it derived monopoly type rents from consumers. Mlambo and Ncube (2011) also found that for the period 1999 – 2008, the structure of the South African banking industry was characterised by monopolistic competition. The South African Reserve Bank`s Annual Report (2011) also confirmed that the four largest South African banks contributed 84.1 percent to the balance-sheet size of the total banking sector. The current banking structure raises various questions around efficiency and access. One of the basic theories of economics is that imperfect market structures are inherently inefficient due to the fact that they do not produce the socially desirable output. Most importantly in the context of banking, the critical question is whether such a structure can support policies designed to expand access to banking services.

Various types of efficiencies exist within the banking literature namely, technical, allocative, productive, profit, and cost efficiency. Efficiency in banking can both be technical and allocative. Technical efficiency implies that a particular bank produces a given set of outputs (such as interest income, loans, non-interest income) using the smallest possible amount of inputs (such as operating cost, employees, capital, and deposits). Ncube (2009) defines allocative efficiency as the extent to which a bank's resources are being allocated towards activities with the highest expected value. Depending on the perspective of analysis of the researcher, one can choose to analyse efficiency from the cost perspective or profit perspective. Cost efficiency provides a measure of how close a bank's actual cost is to what a

best-practice bank's cost would be for producing an identical output bundle under similar conditions. A bank is considered inefficient if its costs are higher than a best-practice bank. Similarly, profit efficiency measures the ability of a bank to generate the maximum attainable profit given prices of its inputs and outputs. A bank is categorised as inefficient if its profits are less relative to the profits of the best-practice bank. Another concept of bank efficiency which is ascribed to Harvey Leibenstein (1978) is x-efficiency. X-inefficiency results when technical efficiency is not being achieved as a result of lack of competitiveness and lack of incentives to reduce cost.

The efficiency with which banks operate is critical at both macro- and micro-level. At macro-level, banking institutions play the intermediary function by accepting deposits from the public and providing loans and other forms of credit to both the household and business sector. As such, an inefficient banking sector may fail to execute this role at the expense of key productive sectors. Ikhide (2008), states that bank efficiency gains in the form of narrowing of interest spreads stimulate an increased demand for loans for investment and greater mobilisation of savings. Moreover, Greenberg and Simbanegavi (2009), states that the banking sector acts as the primary channel of monetary policy and that any inefficiency from the banking sector diminishes the effectiveness and success of monetary policy. At micro-level, banks provide various services and products to consumers at a fee hence, the need to achieve optimum pricing. Higher bank service fees as a result of inefficiency generally act as a brake on efforts intended to increase access to bank services. Therefore, in the light of these facts, the need to investigate the efficiency of the banking industry of any economy cannot be overemphasised.

While access to financial access is a broader concept, this study adopts the idea of access to banking services. This is because access to banking services as approximated by a bank account, is the first necessary step in acquiring the ability to save and make payments as well as opening the door to a host of other financial services.

1.2 STATEMENT OF THE PROBLEM

Firstly, broader access to bank services for all remains an important issue in South Africa especially as one way of redressing the income inequalities of the apartheid regime and expanding economic opportunities. The statement of the problem is that the proportion of the population that does not have a bank account or easy access to bank services has not reached

the desired lowest levels in South Africa. FinScope (2011) reported that 37.2 percent of the adult population are unbanked and that 27 percent of the South African adult population are considered to be financially excluded and they do not use both informal and formal financial products to manage their financial lives. The set target for the banking industry is to increase financial inclusion by expanding access from the current 62.8 percent of the banked adult population to 70 percent by 2015. However, BASA (2012) reported that the National Development Plan is targeting 90 percent by 2030. According to World Bank (1995), access to a bank account is generally considered to be a starting point for one to make use of financial services. Hence, in the absence of a simple bank account, one finds it difficult to have access to various financial services namely, savings, insurance, mortgage loans, and other forms of consumer credit. This is because such services require e-payments or a debit order from a bank account. If the majority of people are financially excluded they also feel socially excluded which may result in many social ills or undesirable consequences to these individuals and to the society at large. As a result, there would be private and social costs imposed on the society as a direct consequence of financial exclusion. Such costs would however, be minimal if a significant portion of the financially excluded population were financially included. Berger et al (1993) argues that when banks attain greater efficiency it enables them to offer better quality services and better prices.

Secondly, unemployment is one of the crucial challenges facing South Africa with the official unemployment rate currently at 25.2 percent during the first quarter and more than half of the young people unemployed (STATSSA, 2013). South Africa's unemployment rate remains high by local and international standards, and evidently contributes towards much of the current social fragmentation and suffering experienced in South Africa. This study hypothesises that since banks play a significant role in the intermediation process, they hold great potential to influence employment generation in an economy via the efficiency transmission channel. However, the researcher argues that the extent to which a country contributes to economic growth and employment creation will depend inter alia on the extent to which the various sectors of the economy can gain access to affordable and efficient banking services. The 2011 annual budget highlighted unemployment as a problem which could have many adverse consequences if not curbed. National Treasury (2011, p.17) identified "improving efficiency across the economy" as one important job driver for accelerating growth and employment creation in achieving the targets of the New Growth Path. In a budget speech statement, Mr Gordhan (2013, p.7), the Minister of Finance, also

highlighted the need "to raise productivity, and diversify the economy, to create jobs and raise living standards." This study seeks to unpack the relationship between banking sector efficiency improvement and unemployment as well as provide strategic and forward-looking lessons for the banking sector. This is crucial in order to inform formulation of suitable banking sector policies that are supportive of employment creation in line with the targets of South Africa`s New Growth Path.

Thirdly, the SARB Supervision division currently uses financial ratios to evaluate performance of the various facets of the banking sector. Ncube (2009), states that financial ratios are simple and straightforward on one hand, but argues on the other hand that the exclusive use of ratios does not give enough insight on actual efficiencies. Productivity in banking has in earlier studies been examined using the Malmquist total factor productivity (TFP) index approach. However, following latest developments by O`Donnell (2010) the Malmquist has empirically been found to be inconsistent and biased when used to evaluate productivity under conditions of varying returns to scale (VRS) technologies. Coelli (1996) highlights the importance of imposing constant returns to scale (CRS) technology assumption when employing the Malmquist approach. A recent study by Kerstens and Van de Woestyne (2014) demonstrate the difference between the Hicks-Moorsteen and the Malmquist approach to be significant for all variable returns to scale technologies. Their study cautions that " If one wants to be on the safe side...it is probably wise to immediately opt for the Hicks-Moorsteen index" (Kerstens and Van de Woestyne 2014, p.756). In this particular study, the DPIN 3.0 software program which was developed by O`Donnell (2011) (which allows for the Hicks-Moorsteen Total Factor Productivity (HMTFP) approach) to redress the incompleteness of the Malmquist technique is utilised. This is an attempt to address the limitations entrenched in the literature on banking productivity with regard to previous studies. Hence, this thesis is a pioneering study in South African banking on the measurement and decomposition of total factor productivity using the non-parametric HMTFP. The application of DPIN 3.0 developed by O'Donnell (2011), a program for generating and decomposing productivity indices, is a new feature of productivity analysis.

Lastly, the failure of the financial system of the United States in 2007 and the recent 2012-2013 Cypriot financial crisis have generated a worldwide concern on the financial sector of which the banking system forms the major component. These banking crisis episodes showed how problems in the financial sector for a single country can be catastrophic for the entire

global economy. The banking system of South Africa is sophisticated and well-developed. Thus any irregularity resulting from inefficiency has great potential to cause devastating effects on the entire economy. There is therefore need to determine if there has been a change in efficiency and productivity of South African banks during the period of the global financial crisis. This is crucial in order to inform policy makers to provide precise and targeted policy measures as well as whether to tighten bank supervision and regulation.

1.3 OBJECTIVES OF THE STUDY

The primary objective of this thesis is to investigate the nature of the relationship between bank efficiency gains and access to banking services in South Africa. On the one hand, do efficiency gains necessarily translate to improved accessibility to banking services? On the other hand, do banks in their pursuit of efficiency compromise access to banking services for consumers particularly the low-income? The objectives of the study are:

- To employ the Hicks-Moorsteen total factor productivity (TFP) index approach to analyse TFP efficiency changes of the banks over the period 2004 to 2011.
- To determine if there has been a significant change in the total factor productivity efficiency of South African banking system during the period of the global financial crisis.
- To establish the nature of the relationship between banking sector efficiency improvement and unemployment in South Africa.
- To articulate the relevance of our results for policies on banking access in South Africa.

1.4 HYPOTHESES

The study tests the following hypotheses:

- There is a trade-off between improvement in efficiency and access to bank services.
- There has been an improvement in bank total factor productivity efficiency between the period 2004 – 2011.
- The global financial crisis adversely affected the efficiency and total factor productivity of the South African banking system.
- Efficiency improvement has a positive bearing on employment.

1.5 CONTRIBUTION OF THE STUDY

The contribution of this thesis is three fold. Firstly, this thesis contributes to banking literature by examining the nexus between bank efficiency and access to banking services. Access to financial services for all has gained prominence in recent years as a vehicle for elevating the poor from poverty and therefore fostering development. Broader access to banking services for the public remains an important issue in South Africa especially as one way of redressing the income inequalities of the apartheid regime and expanding economic opportunities. The current controversy is that banks may attain improvements in efficiency at the expense of a reduction in access to bank services for some consumers particularly the low-income consumers. Hence, the pursuit of efficiency by banks has the potential of reducing access particularly the poor majority. This thesis therefore argues that any mechanism or intervention focused on expanding access to bank services must not be evaluated from an efficiency perspective alone but must be inclusive taking into account the social benefits to the society as a whole. Traditionally, banks have been perceived as institutions whose sole objective is profit making and therefore lack social responsibility to the society. The most important objective of financial inclusion is to reach the poor and disadvantaged segments of the population. Therefore, the final goal of increasing access to services for low-income sectors of the populace is to increase social welfare.

The second distinguished significance of this study is entrenched in the researchers' quest to explore the nexus between bank efficiency and national unemployment. Among the high priority macro-economic challenges facing South Africa, the problem of high unemployment particularly youth unemployment is top of the agenda. During the second quarter of 2013, Statistics South Africa (STATSSA, 2013) reported a high unemployment rate of 25.6 percent. We argue that there is little chance of success for an economy if sustainable and efficient supporting services are not forthcoming from the banking sector. Therefore, a non-researched critical issue is whether changes in banking sector efficiencies are transmitted to national unemployment statistics. In other words, does the enhancement of banking sector efficiency confer positive gains as far as South Africa`s employment generation is concerned? This thesis therefore undertakes to establish the nature of the relationship between banking sector efficiency improvement and national unemployment. This is crucial in order to decisively inform formulation of suitable banking sector policies that are supportive of employment creation in line with the targets of South Africa`s New Growth Path.

Lastly, a distinguished contribution of this study is the novelty of the methodological approach to be employed. Earlier studies have examined banking productivity using the Malmquist TFP technique which Grifell-Tatje and Lovell (1995) have empirically proven is mostly appropriate for technologies exhibiting constant returns to scale. They argue that with variable returns to scale the Malmquist TFP index approach fails to accurately measure changes in productivity. Coelli and Rao (2005) also reached the same conclusion that under the VRS assumption the outcome of calculating the Malmquist TFP index will result in inaccurate measures of TFP changes. The CRS condition is only appropriate when all firms to be evaluated are operating at their optimal scale. In reality this is not the case, hence the need to allow for the assumption of VRS. This study is therefore the first in South Africa to use a software program DPIN[1] 3.0 (which allows for the HMTFP index approach) developed by O'Donnell (2011) to measure and decompose TFP changes in the South African banking for the period 2004 – 2011. Arjomandi and Valadkhani (2010, p.1) states that "the advantage of this approach over the prominent Malmquist index approach is that it is free from any assumptions concerning firm optimising behaviour, the structure of markets, or returns to scale". Furthermore, Hollingsworth and Wildman (2003, p. 497) state that "Malmquist techniques are unable to cope with unbalanced panel estimation procedures". However, a recent study by Kerstens and Van de Woestyne (2014) has shown that the Hicks-Moorsteen approach is always feasible in the presence of both balanced and unbalanced data.

[1] Decomposing Productivity Index Numbers (DPIN) is a software program which uses the Data Envelopment Analysis methodological framework and also provide for the application of the Hicks-Moorsteen TFP approach.

1.6 ORGANISATION OF THE STUDY

This thesis is divided into eight chapters. Chapter one introduces the whole study by looking at the background, statement of the problem, objectives, hypotheses and contribution of the study. Chapter two provides a general overview of the South African banking sector while underlining the main concerns of this study. Chapter three presents a comprehensive theoretical framework and literature review regarding banking productivity and efficiency. Various parametric and non-parametric methods of efficiency measurement are discussed in this chapter. The fourth chapter describes the data, sources of data and the non-parametric DEA methodology used to measure banking productivity and efficiency. The three main methods of panel data analysis are also discussed in the forth chapter. Chapter five starts off with financial ratio analysis to measure, describe and analyse bank efficiency. It presents the first-stage results of running the DPIN program to generate and decompose productivity indices into several efficiency measures for the period 2004 - 2011. The researcher further interprets and discusses the results of applying the fixed effects model to investigate the relationship between bank efficiency gains and access to bank services in chapter six. In chapter seven, the pooled GLS model is used to shed light on the nexus between banking sector efficiency and unemployment. Finally, chapter eight presents the key findings, policy recommendations, limitations and suggestions for further research. It is followed by a list of references and appendices.

CHAPTER TWO

AN OVERVIEW OF THE BANKING INDUSTRY IN SOUTH AFRICA

2.1 INTRODUCTION

South Africa has a well-developed banking system which parallels well with systems in many developed economies and this distinguishes South Africa from many other emerging economies. Banks, as alluded to in the first chapter, are engines of economic development. These crucial institutions provide financial opportunities by providing the general public and the business sector access to capital which enables them to contribute positively to economic growth and employment creation. Hence, ensuring access to financial services has gained universal recognition as a medium for inclusive growth, economic sustainability and development.

2.2 THE SOUTH AFRICAN BANKING INDUSTRY

2.2.1 Structure of South African Banking Industry

The banking system in South Africa is the largest and most sophisticated in the whole of Africa (AFD, 2011). According to Banking Association South Africa, BASA (2010), South Africa has a sound and well regulated banking system which is ranked among those of industrialised countries. South African banks are well managed and use sophisticated risk-management systems in conducting the business of a bank. Claassen and Brooks (2003, p.5) states that the "sophistication of the South African banking sector has its origin in the colonial history of the country whereby the country's banks were modelled according to European standards". This continued engagement with European counterparts have allowed South African banks to keep abreast with cutting-edge technology enabling the provision of better product suites.

Within the financial sector, commercial banks constituted the largest share with assets which represented the highest since 2000 of 139 percent of South Africa`s GDP in 2008 before decreasing to 114 percent in 2011, possibly reflecting the aftermath of the global financial crisis. According to SARB (2012), the banking sector currently has a composition of 17 locally controlled banks, 3 mutual banks, 14 branches of foreign banks, 1 co-operative bank

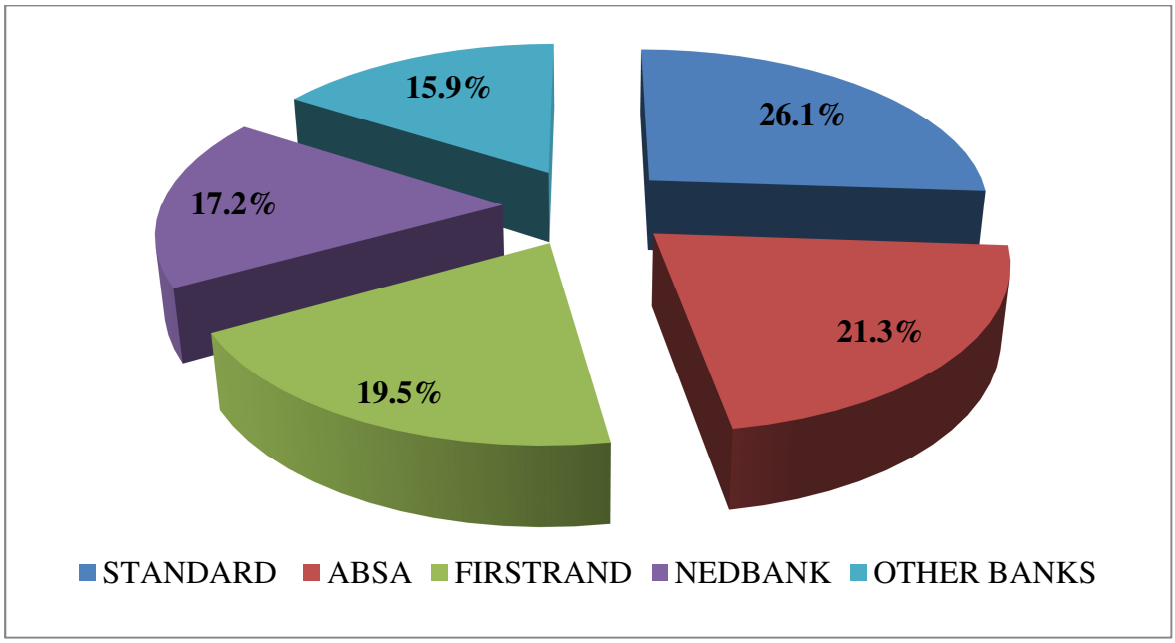
and 41 representative offices. However, the South African banking sector is dominated by four major banking institutions the so-called "Big Four" namely, the Amalgamated Bank of South Africa (ABSA), FirstRand Bank, Nedbank, and Standard Bank. These four largest banks, contributed 84.1 percent to the balance-sheet size of the total banking sector at the end of December 2011 (SARB Annual Report, 2011). The rest of the banks accounted for the remaining 15.9 percent indicating the high level of concentration in the banking market. Thus, the South African banking industry exhibits a high level of concentration. Concentration is defined as the extent to which most of the market's output is produced by a few firms in the industry. South Africa currently uses the total assets size (balance sheet size) of a bank as a proportion of the entire banking industry`s assets to approximate a bank`s market share. Table 2.1 reveals the dominance of the big-four within the banking market.

TABLE 2.1: MARKET SHARE BY TOTAL ASSETS, 2011

BANK	**TOTAL ASSETS (R millions)**	**PERCENTAGE OF INDUSTRY TOTAL ASSETS**
STANDARD BANK	889250	26.1
ABSA	725679	21.3
FIRSTRAND	665525	19.5
NEDBANK	585033	17.2
OTHER BANKS	541751	15.9
TOTAL	**3407238**	**100**

Source: SARB Bank Supervision Department, Annual Report 2011

FIGURE 2.1: PERCENTAGE OF THE BANKING INDUSTRY TOTAL ASSETS, 2011



Source: SARB Bank Supervision Department, Annual Report 2011

2.2.2 Banking entities registered in South Africa

The data in Table 2.2 provides a general overview of the growth of the South African banking industry. The number of registered or licensed entities since 2001 is shown in the table below. Overall, the number of banks has been declining over the years particularly local controlled banks. There are various reasons for this trend ranging from liquidation, mergers or amalgamation. The SARB (2002) noted that this marked decrease was a result of difficulties experienced within the small-bank sector which started with the placing of Fidelity Bank Limited (FBC) under curatorship in 1999 and the loss of sight by the management of Saambou bank limited in 2002, the then seventh largest bank. Consequently large deposit withdrawals were experienced among the majority of banks that were smaller than Saambou of up to 40 percent of the deposit base in some cases. This caused severe liquidity pressures and a general loss in depositor confidence among small-to-medium sized banks. Despite liquidity support from the Reserve bank, 22 banks eventually exited the banking industry during the period 2001 – 2005 as shown in Table 2.2. The largest decrease was recorded in 2002 and 2003 where in each of these years 9 banks had their registration cancelled making a sum of 18 banks (SARB, 2002). Other factors also contributed to this negative banking outlook namely, the 1997 Asian financial crisis, the 1998 Russian financial crisis and bank takeovers. Worth mentioning was the acquisition by Nedcor bank limited in 2002 of the Bank of England (BOE), the then sixth largest bank.

TABLE 2.2: SOUTH AFRICAN BANKING SECTOR: NUMBER OF ENTITIES REGISTERED

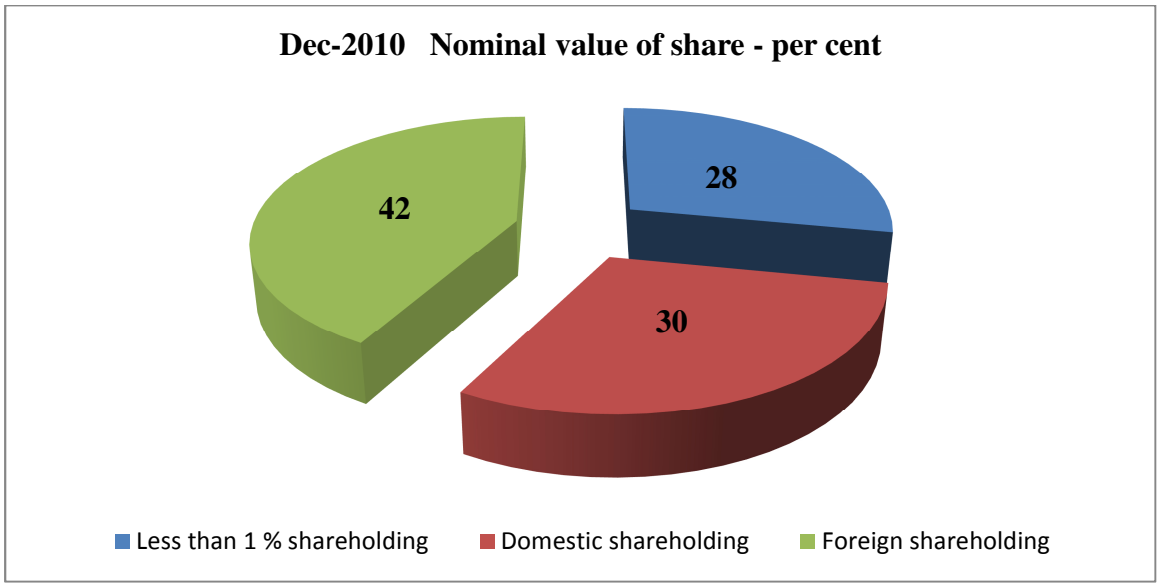
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Domestic Banks	41	30	22	20	19	19	19	19	18	17	17	17
Mutual banks	2	2	2	2	2	2	2	2	2	2	2	3
Branches of international banks	15	14	15	15	15	14	14	14	13	13	12	14
Representative offices	56	52	44	43	47	43	46	43	42	41	43	41

Source: Various SARB Quarterly Bulletins

2.2.3 Shareholding structure

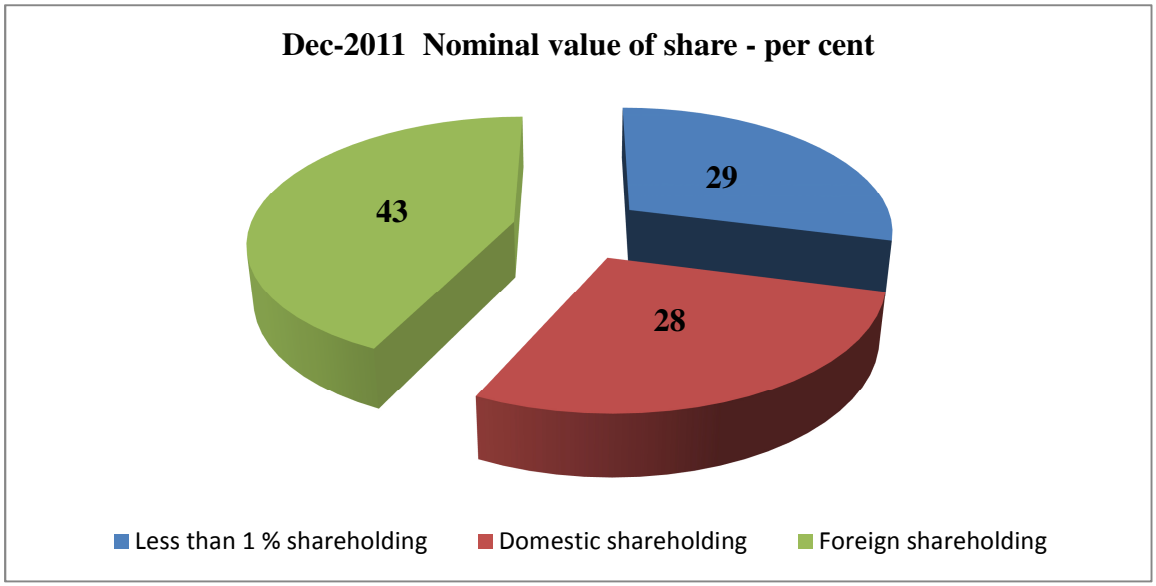
The shareholding structure of South African banks is set out in Figure 2.2 and Figure 2.3. In South Africa, foreign banks hold a large share of banking system assets. According to SARB Annual Report (2011), foreign shareholders held 43.2 percent of the nominal value of the total banking sector's shares in issue at the end of the year 2011, higher than the 42.3 percent recorded as at 31 December 2010. The high percentage of foreign shareholding in the South African banking industry is attributed to ABSA, the second largest bank in the country. Foreign shareholders represented 55.5 percent of the nominal value of ABSA`s total shares in issue at the end of 2011. Minority shareholders accounted for 29.3 percent while domestic shareholders represented 27.5 percent of the nominal value of banking-sector shares in issue at the end of December 2011. Domestic shareholders represent only *domestic* shareholders with *more than 1 percent* shareholding while minority shareholders refers to both *domestic or foreign* shareholders with *less than 1 percent* of the total nominal value of shares (SARB, 2012). Kirkpatrick et al (2008) in their study of Sub-Saharan African countries found that the degree of foreign bank penetration is inversely related to X-inefficiency, implying that foreign bank ownership in Africa has contributed positively to efficiency and better management of commercial banks. Bonin et al, (2005) and Cosset et al (2005) also found that foreign bank entry has beneficial implications in developing countries.

FIGURE 2.2: 2010 SHAREHOLDING STRUCTURE OF THE SOUTH AFRICAN BANKING SECTOR



Source: SARB Bank Supervision Department, Annual Report 2011

FIGURE 2.3: 2011 SHAREHOLDING STRUCTURE OF THE SOUTH AFRICAN BANKING SECTOR



Source: SARB Bank Supervision Department, Annual Report 2011

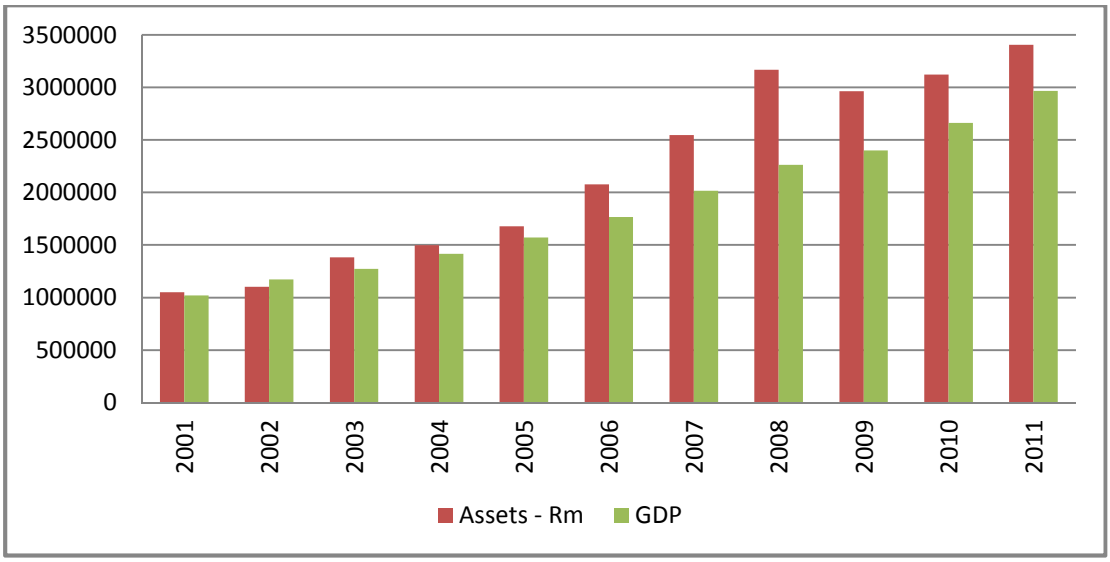
2.2.4 Banking-Sector Assets to Gross Domestic Product

The significance of the banking sector can be examined in terms of the size of the sector's assets. Figure 2.4, shows that the size of the banking industry's assets as a percentage of GDP has been increasing steadily over the years. By 2001, the value of banks' assets exceeded GDP for the first time. In 2008, the ratio of total banking sector assets to GDP reached its highest of 139 percent from 102 percent in 2001 (see figure 2.5 and 2.6.) According to Maheshwari (2009), for the period 2003-2006, asset growth recorded in the South African

banking industry was higher than that witnessed in the global banking industry. It could be argued that the size of these assets relative to GDP highlights the importance and potential influence of the banking sector to the South African economy. Banks' assets are predominantly loans and advances. Loans and advances as a percentage of total banking assets decreased from 80.3% in December 2002 (December 2001: 76.7%) to 71.4% in December 2003. This marked decrease in 2003 was mainly due to the cancellation of registered banks and the ultimate decrease of their total assets. The percentage of loans and advances to total assets subsequently increased over the years until 2007 where it averaged around 83%. This increase in total loans and advances to assets ratio during 2007 is ascribed mainly to the increase in overdrafts and loans (24.8 % year-on-year increase) and mortgage loans (24.7% year-on-year increase) (SARB, 2009). The ratio then drastically fell to 72.9% in 2008 before peaking up to 76% in 2009 during the height of the financial crisis. Due to the harsh economic environment which constrained loan demands combined with a low appetite for risk among banks, the ratio deteriorated to 74% in 2010 and further to 72.7% in 2011. According to BASA (2009), the economic recession in 2009 which was a consequence of the global financial crisis affected consumer affordability and therefore spending patterns, resulting in consumers being reluctant to take on more debt. This led to the increase in non-performing loans which had a huge impact on the banks' loan books. Thus, total assets and liabilities declined in 2009 but started to pick up as the country moved out of recession.

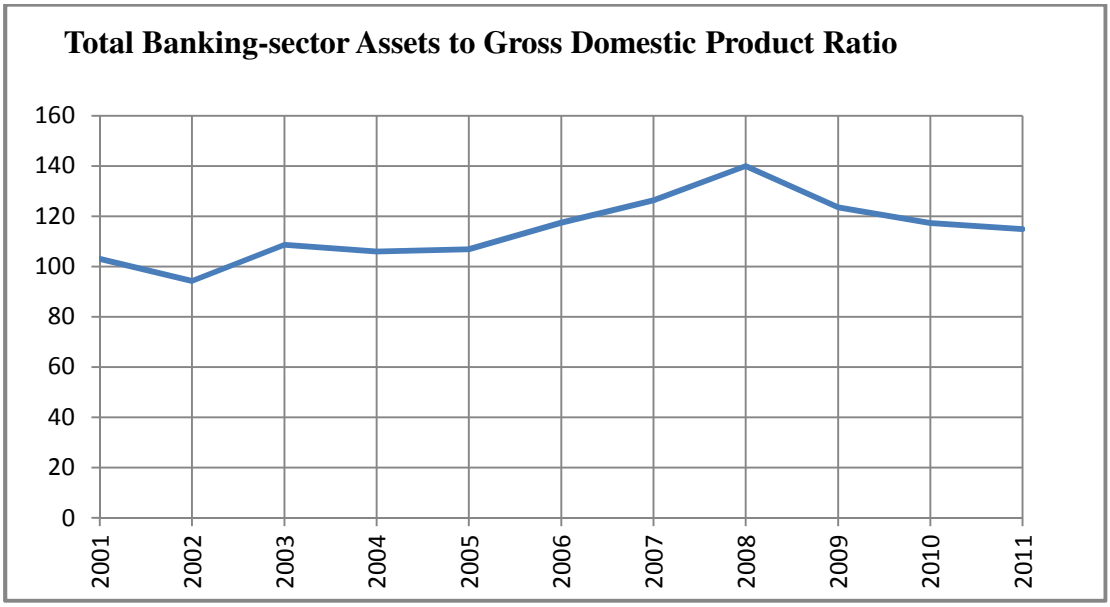
Total banking-sector assets amounted to R3 406 billion at the end of December 2011 (December 2010: R3 126 billion), representing a moderate year-on-year increase of 8.9 percent. As of December 2010, the banking sector assets represented an increase of 5.3 percent from a value of R2 967 billion in December 2009. Gross loans and advances, which represented, on average, 74 percent of banking-sector assets during 2010, increased marginally by 2.5 percent to R2 314 billion at the end of December 2010, mainly due to modest growth in home loans and higher overnight and interbank call loan balances (Bank Supervision Department Annual Report, 2010).

FIGURE 2.4: TOTAL BANKING-SECTOR ASSETS TO GROSS DOMESTIC PRODUCT



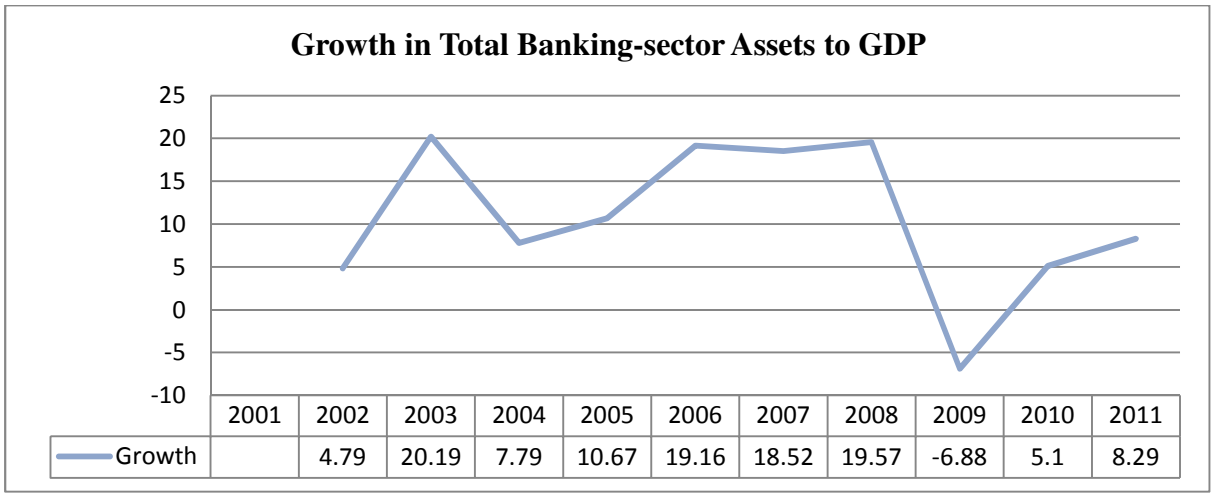
Source: Computed using data from SARB: www.quantec.co.za

FIGURE 2.5: TOTAL BANKING-SECTOR ASSETS TO GROSS DOMESTIC PRODUCT RATIO



Source: Computed using data from SARB: www.quantec.co.za

FIGURE 2.6: GROWTH RATE OF TOTAL BANKING-SECTOR ASSETS TO GDP



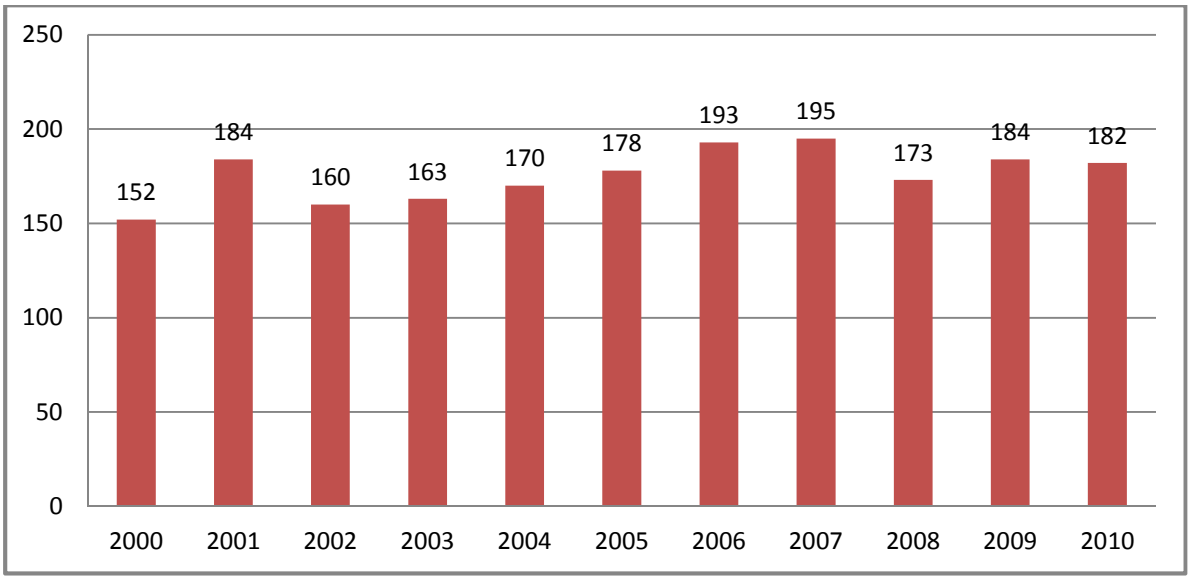
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Growth		4.79	20.19	7.79	10.67	19.16	18.52	19.57	-6.88	5.1	8.29

Source: Computed using data from SARB: www.quantec.co.za

2.2.5 Private Sector Credit provided by the Banking Sector

The share of private sector credit as a percentage of GDP provided by the South African financial sector is higher than in other countries in Sub-Saharan Africa, Brazil and India (see table 2.3 p. 22). As shown in Figure 2.7 below, the proportion of private sector credit from the banking sector increased to 195 percent in 2007 from 152 percent in 2000. In 2008, this proportion fell to 173 percent in 2008, most likely signifying the adverse effect of the 2008 financial crisis. It peaked in 2009 to 184 percent before it fell to 182 percent in 2010.

FIGURE 2.7: DOMESTIC CREDIT PROVIDED BY THE BANKING SECTOR (% OF GDP)

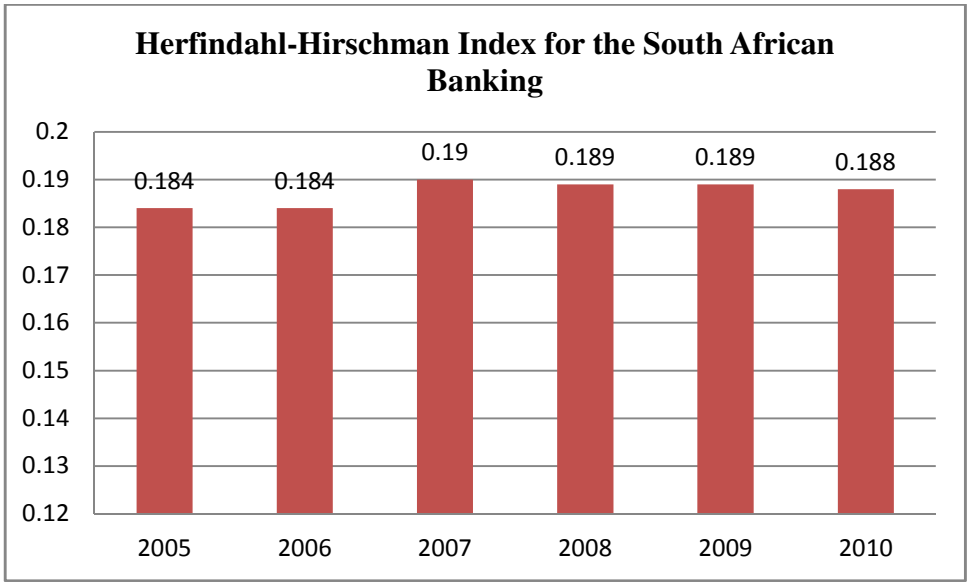


Source: Computed using data obtained from World Bank www.worldbank.org

2.2.6 Concentration in the South African Banking Industry

The SARB Supervision unit presently use the Herfindahl–Hirschman Index to measure the level of concentration in the banking market. The H-index is calculated by adding the squares of the market shares of each firm in the defined market. The index turns to zero when the market comprises a large number of firms of fairly equal size. However, the index increases following a decrease in the number of firms in the market as well as when the difference in size between these firms gets larger. Hence, a higher index is generally inferred to imply less competition in the market. As a rule of thumb, an index of less than 0.1 implies that there is no concentration in a particular market or industry. Average concentration is represented by an H-index within the range 0.1 to 0.18 while an excess of 0.18 represents a highly concentrated market. From the basic principles of economics, such a market structure is considered inefficient and associated with producing less of the socially desired output. CGAP (2011) states that banking in South African is heavily concentrated within the big-four banks which together account for more than 90 percent of retail banking market.

FIGURE 2.8: HERFINDAHL-HIRSCHMAN INDEX FOR THE SOUTH AFRICAN BANKING SECTOR



Source: Computed using data from SARB:www.quantec.co.za

Figure 2.8 indicates the level of market concentration in the South African banking sector, analysed using the Herfindahl index for the period 2005 to 2010. The index amounted to 0.188 at the end of 2010 compared with 0.189 at the end of 2009. The high index portrayed in Figure 2.8 could be attributed to the dominance of banking-sector assets by the largest four banks, which over the period represented above 80 percent of total assets. A recent study

conducted by Mlambo and Ncube (2011) found that for the period 1999 – 2008, the structure of the South African banking industry was characterised by monopolistic competition. This result is a clear revelation of the domination by four large banks, which together account for over 84 percent of total banking assets. Greenberg and Simbanegavi (2009) argue that this high market share has the potential to induce banks to engage in collusive behaviour for example, lowering their deposit rate and raising their lending rate.

Many empirical studies have shown that a high level of concentration is generally not good for consumers and the small players in any industry. In the case of the South African banking sector, this high concentration does not promote competition among banks in South Africa and as a result raises worry on issues of accessibility to banking services. Dominant players in any particular industry are a cause of concern as they might set their pricing in a manner that disadvantages consumers. Earlier efficiency studies that have focused on market structure and bank efficiency in South Africa have indicated a market structure between oligopoly and monopoly [Napier (2005a), Okeahalam (2006), Ncube (2009)]. Okeahalam (2000) argues that the existence of an oligopoly structure (imperfect competition) implies that the degree of competition necessary to stimulate greater efficiency may not exist. Therefore there is a need for the government to create an enabling environment and an inclusive regulatory system that promotes diversity in banking in order to realise sound levels of competition.

2.3 EFFICIENCY OF THE BANKING INDUSTRY

In sub-Saharan Africa (SSA), banks are the most important element of the financial system. Banks have three principal activities: taking deposits, making loans, and investing in securities. To do this they use labour (skilled and unskilled), physical capital, and financial capital (Kablan, 2010). In executing this crucial role banks needs to be efficient. Mr. Gordhan (2011), South Africa`s Minister of Finance, in his key note address at a banking summit highlighted the significance of the role of the banking sector in meeting the goal of the Government`s New Growth Path of generating 5 million jobs by the year 2020. He emphasised that being such a crucial role player and an important anchor for economic growth, it is important that the banking sector maintains its efficiency. Ikhide (2008) maintains that in the absence of a sound and efficient banking system, the economy cannot function well. Ikhide (2008, p.4) further argues that "when banks fail, the whole of a nation's payments system is thrown into jeopardy." It is because of this critical position of banks that

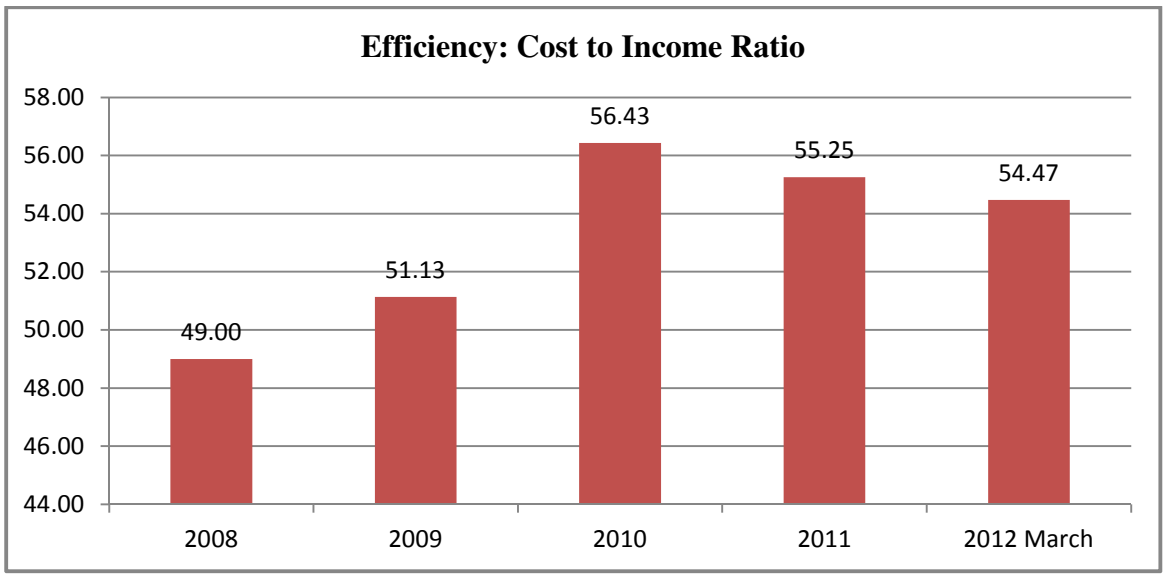
the efficiency of these instrumental institutions merits close attention. Efficiency is related to the ability to produce a result with minimum effort or resources. It measures how close a production unit gets to its production possibility frontier which is composed of sets of points that optimally combine inputs in order to produce one unit of output (Kablan, 2010). Efficiency is a way of generating more resources without necessarily looking for additional investment. The fact that resources are wasted in an inefficient system means that an improvement in efficiency is similar to additional resources that can be used in the system.

In recent years, there have been regular complaints that bank fees and charges are exorbitantly high in South Africa. In 2004, Deloitte (2004) also noted that South African banking fees were the world highest with roughly 2 percent of an average individual`s gross income paying bank charges. Dorsey and Jacobs (2005) observed that a simple and basic bank account could be so costly and have a complicated mix of charges, withdrawal and transaction fees and minimum balances. Their study also underlined the fact that most of the products and services of most banking institutions are not properly suited to the needs of the majority low-income people, serving only a well-to-do and educated minority. National Treasury (2011) highlighted that a Competition Commission study in South Africa also found that bank charges were unjustly high. It was also noted that the banking industry does not provide consumers with standardised interest rates that allow a comparison of risk-return and therefore an improved choice of financial products. In an investigation of the banking industry, the Competition Commission (2008) revealed evidence of abuse in the setting of certain fees and charges and found poor disclosure practices that made it difficult for customers to do interbank comparisons. However, many banks claim that their high fees are a direct result of the high administration costs generally related to low transaction volume and small account balances.

Traditionally efficiency has been measured using the cost to income ratio. Figure 2.9 provides an indication of the efficiency trends of the banking industry using ratio approach. However, Ncube (2009) states that while this approach is relatively simple and straightforward, ratios alone do not give sufficient insight on actual efficiency. Hence, there is need to use parametric and non-parametric methods. It is important to highlight the fact that the non-parametric technique is employed in this study as the main first-stage methodology. However, accounting ratios are used in detail in chapter five to supplement the

non-parametric approach and shed light on the possible channels through which changes in efficiency may possibly influence access and unemployment.

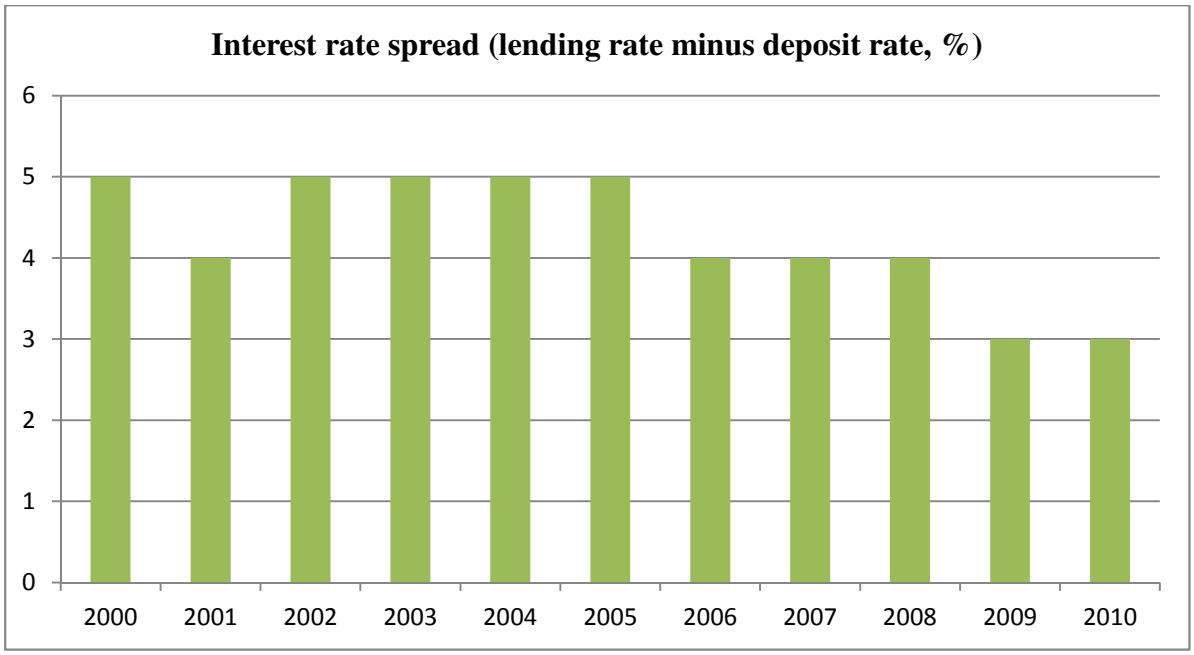
FIGURE 2.9: EFFICIENCY TRENDS IN BANKING AS MEASURED BY COST-TO-INCOME RATIO



Source: Computed using data obtained from SARB: www.quantec.co.za

Banking sector cost efficiency deteriorated as the cost-to-income ratio increased from 51.1 percent in 2009 to 56.4 percent in 2010. However, from 2010 until March 2012 the cost-to-income ratio has been decreasing implying that banking efficiency has been improving. Interest rate spread is used as a measure of intermediation efficiency. Interest spread is the difference between the lending rate and the deposit rate and represents a growing share of gross income for banks. Most banking institutions of developing countries have wider interest spreads. Wide spreads affect banks` basic function of intermediation and distort prices thus slowing down the role of the banking system in contributing to economic growth (Ikhide, 2008). For borrowers, a reduction in spreads is generally expected to encourage a general increased demand of loans for investment and consumption thereby contributing positively to accelerated growth. For savers, it encourages a culture of saving through formal institutions such as banks. Ikhide (2008) states that historically, government regulations and policies were held responsible for such wide spreads. However, he argues that with the advent of financial liberalisation in many African countries, efficiency studies in banking have attributed these wide interest spreads to be a consequence of inefficiency.

FIGURE 2.10: EFFICIENCY TRENDS IN BANKING AS MEASURED BY INTEREST RATE SPREAD



Source: Computed using data obtained from World Bank www.worldbank.org

TABLE 2.3: SOUTH AFRICA, SUB-SAHARAN AFRICA, BRAZIL AND INDIA

	South Africa		Sub-Saharan Africa		Brazil		India	
	2000	2008	2000	2008	2000	2008	2000	2008
Interest Spread (%)	5	4	13	10	40	36	-	-
NPL / Gross Loans (%)	-	4	-	-	8	3	13	2
Private sector credit / GDP (%)	134	145	61	59	33	31	29	51

Source: World Bank Indicators, 2009.

Table 2.3 indicates that interest rate spread was lower for the period 2000 to 2008 in South Africa. This suggests that compared to Brazil and other Sub-Saharan Africa (SSA) countries, the financial system in South Africa is efficient in transferring capital to investment. The most recent data from World Bank (see Figure 2.10) indicates that the interest rate spread decreased to 3 percent in 2009 and 2010.

2.3.1 Enhancing Bank Efficiency, Good or Bad?

Falkena et al (2004) describes a technically efficient firm to be one that employs the least possible amount of inputs to produce a given amount of outputs. According to Ncube (2009) allocative efficiency measures the ability of a firm to allocate its resources to activities with the greatest expected value. The economic customer profile of the low-income households is such that it does not promote efficiency thereby suggesting the possibility of a trade-off between access and improvement in bank efficiency. Given the relatively higher costs normally related with serving clients of low-income status, banks consider it rational to serve the high-income clients as the profit potential is greater. The question this thesis attempts to address is: does improving bank efficiency preclude the provision of banking services particularly to the low-income households? Based on the results to be obtained, this thesis argues that although enhancing bank efficiency is an important policy objective, it is certainly not the only priority. Access to banking services and consumer protection are equally important policy priorities. Therefore, one cannot sacrifice one policy goal for the other. There is need for a holistic approach to the problem of access in order to generate solutions that achieve both efficiency and accessibility in a balanced fashion.

According to a World Bank study by Paulson and McAndrews (1998), banks find it difficult to serve low-income clients profitably because of their failure to maintain minimum balances on their bank accounts. Their study explains that even if required minimum balances are maintained these low-income clients may take many small withdrawals which drive up branch operational costs. Okeahalam (2006) states that banks argue that the expected returns from extending banking services to townships are low and the average cost of production is high. The low net returns emanate from the fact that the nature of transactions in these townships is small in terms of returns and yet high in volume, driving up operating cost.

This thesis makes the argument that any form of intervention designed to enhance access to banking services to reach those financially excluded cannot be appraised from an efficiency perspective alone but must be inclusive taking into account the social benefits to the society as a whole. Traditionally, banks have been perceived as institutions whose sole objective is profit making and therefore lack social responsibility to the society. The basic and most important objective of financial inclusion is to touch the lives of the poor and disadvantaged segments of the population by giving them the basic opportunity to have access to financial services. Therefore, the final goal of expanding access for low-income sectors of the

populace is to increase social welfare. On the one hand this thesis argues that given the socio-economic characteristics of the low-income poor people it is only rational for banks not to extend their services beyond the high-income segment of the market. This begs the question if enhancing bank efficiency is good or bad?

2.3.2 Arguments against serving the low-end segment of the market

The reasons behind the reluctance by banks to serve the low-income segment of the market are twofold, the demand side and supply side. This section discusses the arguments from the supply side perspective. As Richardson cited by Fernando (2007, p. vii) once stated "the biggest challenge for developing economies is to get the banks to the unbanked, rather than to get the unbanked to the banks".

A basic and fundamental cause of supply-side constraints stems from the very nature of the prospective market comprising low-income and poor people. Firstly it is argued that given the socio-economic profile of the poor in general, there is little profit potential in serving such a market. As a result, Fernando (2007) argues against the role of the private sector or market-oriented solutions as feasible ways to expand access to low-income people. Secondly, it is argued that due to the low socio-economic profile of low-income clients, financial services are best provided through programs of the government and non-governmental organisations (NGOs).

Another important reason for inadequate supply is that banks are generally not geared to serve the low-income financial markets because they have not been organized and established to serve such a market. Their products, costs structures and organizational structures, are designed to serve the high-income clientele. Given this background, banking institutions find that it's relatively more rewarding moving "up-market" than moving "down-market". Therefore, it is usually not in the interest of banks to serve markets in which a large population are low-income households because the relative cost of setting up branch networks in such marginalised areas is substantially higher. There is therefore, need for banking institutions to innovate in order to produce better priced products and services that are tailor-made to suit the socio-economic characteristics of poor people.

Banks need sellable collateral security which most households of low-income status are unable to offer thereby constraining them from obtaining the much needed credit. With

respect to opening and maintaining savings accounts, most banks impose prohibitive requirements regarding minimum balances and withdrawal frequency.

The lack of formal financial institutions near points of actual demand is another important supply side factor. A significant proportion of people in South Africa live in marginalised rural and remote parts where banking amenities are not easily accessible within a reasonable short distance. As a result, the majority of people become more reliant on informal financial services which are relatively expensive compared to formal services in developed urban areas. According to financial access survey data compiled by IMF (2012) for all countries in the world, South Africa had 3 commercial bank branches per 1000km^2 and 17.3 ATMs per 1000km^2 in 2011. When compared to other BRICS[2] countries for which data was available, South Africa was better than Russia which had 2.7 bank branches per 1000km^2 and 11.2 ATMs per 1000km^2. However, South Africa was considerably below Brazil (7.9 and 20.6) and India (30.4 and 25.4) in terms of the number of bank branches per 1000km^2 and ATMs per 1000km^2 respectively. Nevertheless, Fernando (2007) argues that the availability of such facilities closer to the places of actual demand does not benefit much if financial service providers are not earnestly dedicated to serve the down-market. In other words, proximity does not matter much if the bank`s main activities do not involve the provision of services and products that low-income people need at better prices. Fernando, (2007, p.11) states that "geographic access does not necessarily mean economic access." He argues that the mismatch of services and products produced by financial service providers with the specific requirements of people of low-income status has exacerbated the challenge of access. It is argued that even low-income people who live near financial service providers cannot have access to financial services because of the problem of incompatibility. This incompatibility result from several reasons namely, that the offered products may have certain characteristics that are not adapted to the socioeconomic profile of consumers. For example, people with poor literacy skills may need a custom-designed insurance product whose terms and conditions are less complicated but readily simple and straightforward. A loan repayment arrangement for a bank client who earn a weekly income may need to be adjusted to correspond to weekly instalments as opposed to standard monthly instalments.

[2] BRICS represents **B**razil, **R**ussia, **I**ndia, **C**hina and **S**outh Africa.

This study has thus far emphasised the argument that geographic access does not necessarily amount to actual economic access unless the business models of banks are innovated or custom-designed to serve those at the base of the pyramid. However, compelling evidence from empirical studies have argued that the amount of savings, for instance, depends partly on how widespread the saving institutions are. Porter (1966, p.349) states that saving is "institution elastic" arguing that easy access to bank branches promotes savings and so develops a habit of banking. Therefore the same can be said that the extension of accessible bank branches to rural and remote areas, *ceteris paribus*, develops a habit of banking and enabling one to undertake a wide-range of banking services such as saving, investment, credit, electronic transfers etc. Evanoff (1988) also found empirical evidence of a positive impact of broader bank branching on service accessibility in rural counties. Basing his argument inter alia on cost and capital requirements, he justified a branch office as a better way to improve accessibility as opposed to unit banking. Within the framework of unit banking, a new bank is needed to increase accessibility. The author found that service accessibility improved by 53 percent when rural areas in which branching was done were compared with those rural areas having only one unit banks. In a study of African economies, Ikhide (1996, p.118) also argued in favour of "extending branch offices of banks to rural areas so as to increase accessibility and hence convenience". He noted that an efficient financial sector that provides attractive and reliable services will help to mobilise real domestic savings thereby reducing the poor savings culture in most African countries. The author then suggested a policy of rural banking emphasising the need for a balance between feasibility and social considerations in allocating branches.

This section has discussed the main arguments from the supply side perspective behind the reluctance of banks to serve those at the base of the economic pyramid. It highlighted the desperate need for a collective and inclusive approach in efforts geared towards expanding access to the unbanked. Emphasis was placed on encouraging the development of banks that are custom-designed to specifically serve clients of low-income profile. It also discussed the importance and positive role of simply extending bank branches to rural areas in promoting the expansion of access to banking services.

2.4 POST INDEPENDENCE BANKING SECTOR INITIATIVES[3]

After gaining independence in 1994, South Africa inherited a highly sophisticated financial system and yet the majority South Africans suffered lack of access to financial services. FinMark Trust (2005) reported that more than a decade after the apartheid system 42 percent of the South Africans had never had a bank account. The government of South Africa since independence has intervened in various ways to foster an enabling financial environment. The purpose of this section is to discuss the South African experience since gaining independence regarding interventions to increase access to those financially excluded.

2.4.1 The Financial Service Charter

2.4.1.1 What is a Financial Access Charter?

Napier (2005b) of FinMark Trust defines a financial access charter (FAC) as an agreement in which the concerned parties solemnly make a commitment to put their efforts together in order to alter a certain level of access to financial services for the poor. The parties or key stakeholders include the government, insurers, banks, labour and civil society organisations. A charter is not legally enforced but it is generally followed and adhered to as if it were legally imposed. Napier (2005b, p.8) outlines that a FAC should encompass the following:

Firstly, a FAC acknowledges the commercial as well as the social function of financial intermediation and that the many varied services be accessible to everyone. Hence, providers of such services should target to provide services that are tailor-made to meet the needs of wealthier as well as of poorer people. More importantly, the role of the government should be central in building an enabling environment in which commercial financial service providers develop.

Secondly, the financial access charter recognises the insufficient state of access in the country quantified as a proportion of people of a certain category or classification that are utilising financial services and products. Targets are then set within a particular time frame for improving this poor level of access. Categories of financial products would be defined according to particular niche market characteristics and placed under general functional headings of insurance, savings, credit and transaction. Lastly, it is essential within the

[3] Initiatives related to financial inclusion or access to banking services since 1994

framework of an access charter to have an establishment such as, a council or a monitoring board that makes decisions on access definitions, sets targets for achievement by relevant participants, and evaluates progress. Hence, a financial access charter projects the level of financial access to be attained at a specific date in the future and provides a coherent strategy on how that target is to be reached. The current set target for South Africa is to expand access from the current 62.8 percent of the banked adult population to 70 percent by 2015 and 90 percent by 2030.

2.4.1.2 Benefits of a Charter

Box 2.1: Benefits of Adopting a Financial Access Charter

Napier (2005b, 10) outlines two primary benefits to a government adopting a financial access charter as follows:

Shared Vision
Firstly, a charter helps to develop a vision for attaining a defined desirable state of financial access at a future point in time and therefore facilitates the development of a collective plan by stakeholders for accomplishing that goal. Due to the specified time frame, charters are essentially forward looking. The adoption of an access charter highlights the significance of financial inclusion and so promotes financial access as an important priority in government planning.

Framework for Engaging the Private Sector
Secondly, the adoption of a charter provides the government with a concrete framework for engaging various stakeholders. Following an agreement with regard to targets and time frame, the actual dialogue will now be on how the agreed targets will be realised in each of the several sub-sectors namely, insurance, savings, mortgage finance, corporate finance, banking, and so on.

Basis for Government Supporting Role
A charter also provides a platform for identifying areas where it may be best for the government to perform a supportive role. For example, the government could contribute by presenting economic incentives, developing basic infrastructure, reducing banking risk in certain areas. Napier (2005b) maintains that a charter offers the needed platform for a partnership approach in which commercial providers of different services commit themselves to developing the financial market. In turn this partnership approach would increase the market size, contribute infrastructure, financial resources, and skills.

Adapted from Napier, M. 2005. "Engaging the Private Sector: The Case for Financial Access Charters in Sub-Saharan Africa." **http://www.uncdf.org/mfmatters**

2.4.1.3 The South African Financial Service Charter (FSC)

The development of the Financial Service Charter (FSC) was a consequence of the acknowledgment by the government of South Africa and the financial service sector that the lack of basic financial services, particularly in rural areas, was a major hindrance to economic growth and poverty alleviation. Thus in 2004 the FSC was adopted in order to advocate for financial inclusion. This was carried out by compelling various important financial institutions to cooperate diligently with the government, and communities in order to transform the financial sector to better serve the poor and ultimately the nation at large. The financial sector recognised that access to first-order retail financial services was central to black economic empowerment and to the economic development (Banking Council of South Africa, 2003). Hence there was a commitment to extensively expand access to retail financial services for the low-income groups by 2008. Key participants to the charter included the government, representative bodies of insurers and banks, the joint investment industry, and representatives of labour and civil society.

One of the significant initiatives that came from the FSC process was the Mzansi account. The Mzansi account was launched in 2004 by the largest four South African banks also known as the big four and the Post bank. It was intended as an entry-level transaction account with common features to address the impediments that deterred consumers from being an essential constituent of the formal banking sector. The next section contains a detailed discussion of the Mzansi initiative that resulted in a considerable decrease of the population of the unbanked South Africans.

The South African Financial Service Charter includes defined targets for improvement in financial access. At the introduction of the FSC, it was acknowledged that the emphasis of the charter was to be centered on accessibility. Therefore the charter embraced achieving agreed targets in certain areas of low-income housing finance, small to medium initiatives, infrastructure, savings, affordable banking, and insurance. Box 2.2 below contains the summarised main principles of the financial service charter of South Africa.

Box 2.2: The Main Principles of the Financial Service Charter of South Africa

Main Charter Principles:

The targets outlined in the charter relates to six main empowerment principles as outlined by Sutton, C.N., and Jenkins, B., (2007):

Human Resource Development. The charter aims to increase the participation of black South Africans at all levels of management, with a target to have black women occupy one third of positions in each level of employment. In training and development spending, each firm commits to spend 1.5% of total basic payroll per year training black employees.

Procurement and Enterprise Development. The charter seeks to substantially increase financial sector procurement from black economic empowerment enterprises. It stipulates that 50% of all procurement should be from BEE-accredited companies by 2008 and 7% by 2014. The charter also commits the financial service sector to create new BEE entities via equity investments and joint ventures.

Access to Financial Services. Financial services firms aim to increase accessibility and affordability of banking and savings accounts. The charter also commits firms to increase lending to low-income borrowers.

Empowerment Financing. The charter commits the sector to provide 73.5 billion rand to development finance in the infrastructure, SME, rural development, and low-cost housing industries. It also commits firms to participate in BEE deals.

Ownership and Control. According to the charter, it was agreed that each institution must target minimum black ownership of 25% by 2010 and recruit board and executive level leaders.

Corporate Social Investment. According to the charter, each firm was to make a commitment to channel 0.5% of post-tax profit to corporate social investment projects, such as educational and training, youth development and job creation programs.

Adapted from Sutton C.N., and Jenkins, B., 2007. The Role of the Financial Service Sector in Expanding Economic Opportunity.

2.4.2 Mzansi Initiative

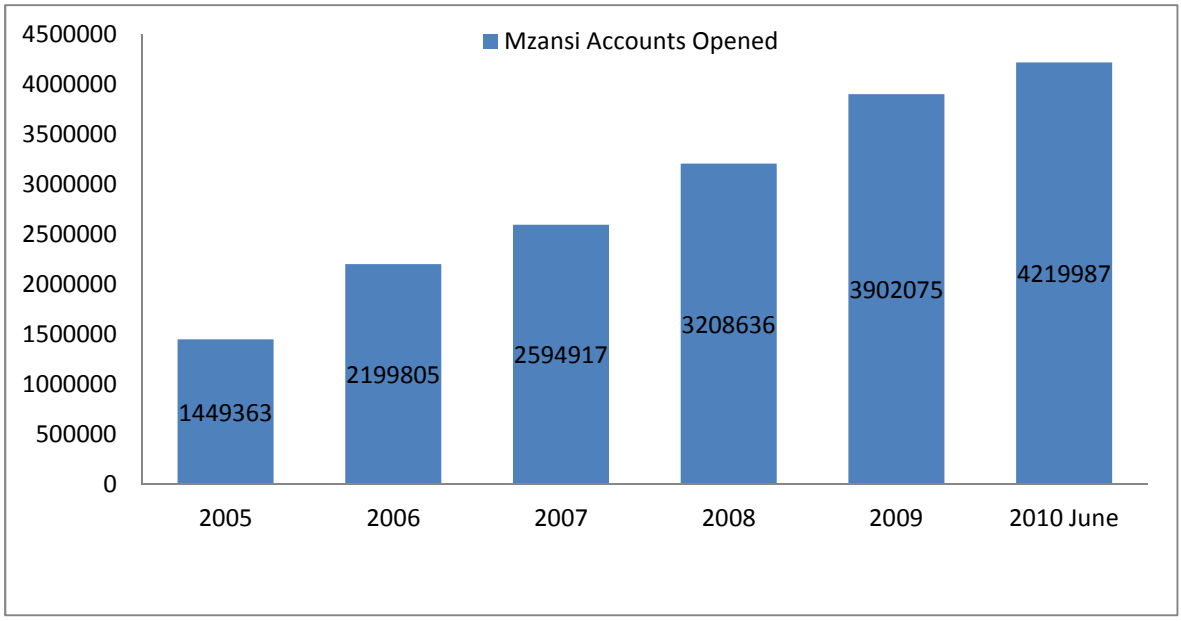
The Mzansi account is an initiative that was launched in 2004 to provide on a large scale transaction bank accounts to the unbanked majority South Africans. This initiative was based upon the premise that sustainable and inclusive economic growth and development is dependent on enhancing financial access for the rural poor and vulnerable South African people. These Mzansi accounts were initiated by the banking industry and launched jointly by the big four banks of South Africa together with the government-owned Post bank in October 2004. Over six million Mzansi accounts were opened by 2008. At present it is generally acknowledged that at least 10 percent of the South African adult population have a Mzansi account and that one in six banked persons are active Mzansi customers. Therefore, banks have made significant progress in providing their services to the previously unbanked members of the population, partly through these low-cost Mzansi bank account.

FinScope (2010) reported that despite the relative achievement of the Mzansi initiative, by 2010 there was still 37 percent of the 33 million South African adult population without a bank account and that only 40 percent had a recognised long-term insurance product. During the first half of 2010, the four major banks opened a total of 317 912 Mzansi accounts, bringing the total number of accounts opened since 2004 to 4 219 987 (Figure 2.12). This, however, was 11 percent lower than the number of accounts opened during the first half of 2009. Statistics on Mzansi accounts indicated that uptake of these accounts declined during the first half of 2011 with only one bank reporting an increase in the number of accounts. Recently there has been a concern that the total number of dormant accounts has been trending upwards. For instance, the total number of Mzansi accounts declined by 19 percent during the first half of 2011. To address this dilemma, banks have responded by introducing their own innovative products which provide a wider choice of banking services to customers since the functionality of the Mzansi account is limited to basic savings and related services (SARB Annual Economic Report, 2011).

A number of reasons have been suggested for the decline in these accounts. This decline has mainly been ascribed to the closure of dormant accounts by three of the big four banks. Moreover, a number of banks have introduced a number of products targeted at the low-income portion of the market, which also explain the declining interest in the Mzansi account. A study conducted by FinScope (2011) found that a number of people in this market dropped out of the banking system in the past three years due to the increased financial strain brought

about by the economic recession. On the whole, the Mzansi initiative has been successful in that it has provided a large portion of the previously unbanked population with access to financial services. According to BASA (2011), the Mzansi paved a way for financial inclusion in South Africa as banks reported that about 90 percent of Mzansi account holders were previously unbanked. Further, Mzansi has been used as a stepping stone to other financial products as some accountholders have migrated to nearest equivalent accounts (BASA, 2011).

FIGURE 2.11: TOTAL NUMBER OF ACCOUNTS OPENED



Source: Banking Association South Africa (BASA), (2010).

2.4.3 A Proposal to introduce Deposit Insurance

In 2011 National Treasury released a policy document entitled, A Safer Financial Sector to Serve South Africa Better, in which it highlighted the possibility of introducing a deposit insurance scheme for deposit-taking institutions in South Africa. However, little progress was made in the launching of this deposit insurance system in South Africa. A draft deposit insurance bill was circulated to interested parties for comments and deliberations among the concerned parties and is still continuing and no time frame for finalisation of the proposals has been set.

According to National Treasury (2011), deposit insurance is intended to safeguard bank depositors partially or totally in the incidence of a bank failure. The proposal of deposit

insurance is particularly crucial for promoting new competitors to the sector. It is generally acknowledged that lack of a deposit insurance system presents a strong argument for customers to favour larger banks such as the South African big-four banks as they are regarded to be “too-big-to-fail” and therefore safer. Thus it has been argued that this “too-big-to-fail” perception may prevail although smaller banks are offering cheaper and better quality products.

Given that South Africa's big four banks hold about 84 percent of total deposits in the banking sector, it is argued that in the event of a severe banking crisis these big banks would require government support. In developing countries, there is need to expand access of formal banking services and increase loanable funds by minimising depositor doubts about the banking system’s capacity to redeem depositor`s funds when needed. An insurance deposit has been suggested to eliminate the rationale for depositors to discriminate between banks on grounds of risk. Demaestri et al (2006) argues that the introduction of a deposit insurance system contributes to greater confidence in financial institutions generally, and thus lead to a higher level of investment, incomes and savings.

In most cases, the general purpose of a deposit insurance scheme is to promote the stability of the financial system of a country. A deposit insurance essentially removes or at least decreases the incentive for depositors to withdraw their funds from a bank unexpectedly. Thus such a scheme would help to preserve stability, especially during turbulent economic times, and therefore serve to reduce the possibility of a systemic crisis. Hence, by instilling and preserving the needed confidence in the financial system, financial crises are either prevented or the severity thereof diminished.

Deposit insurance is regarded as one direct instrument for alleviating poverty, as it mainly serves to enhance the safety of savings of the poor people and improve their overall financial security. In addition, there is a view that a deposit insurance system serves to improve the status of small depositors compared to larger depositors through encouraging better treatment of small depositors and promoting the participation of more financial institutions and improving access and competitive benefits for poorer individuals. Garcia (1996) also noted that in certain cases, limited deposit insurance coverage may help small depositors to obtain preferential treatment relative to large depositors. This argument stems from the view that

small depositors are more likely to have a higher proportion of their funds protected in the event of bank failure.

Most importantly, it is generally claimed that the growth prospects for an economy in which the banking sector has a likely default risk will be lower than one where there is some measure of confidence within the banking sector. This is essentially because a potentially insecure banking system will mobilise fewer savings. Demaestri (2006) maintains that an efficient deposit insurance system may instil and preserve public confidence in the banking system. As a result the volume of intermediation will expand improving the availability of financial resources for investment and resulting in economic growth. Thus a deposit insurance system reinforces depositors' faith in the stability of the banking system of a country and therefore mobilises household savings leading to efficient ways of investing such savings (Demaestri, 2006).

Finally, Garcia (1996) argues that a deposit insurance scheme may level the business environment in a competitive banking market for small banks leading to a much broader participation by smaller financial institutions thereby reducing lending costs and increasing deposit rates. Consequently, this particular system encourages smaller depositors, improving the state of access to banking services and opening a door of opportunities for the majority of poor people. Again since larger institutions are perceived to be safe and secure as they are considered "too big to fail", lack of a deposit insurance system may encourage depositors to choose these large banks at the expense of smaller banks. Ironically, one can argue that too much competition in the banking system may shrink profit margins thereby causing less stability in the banking sector.

At this juncture, it is important to take into account the fact that small depositors are not essentially the poorest. By contrast those who are indeed poor are those who cannot have access to financial services. However, in the presence of a stable and efficient financial sector in which depositors funds are safe and secure, poor people are on average more persuaded to become participants than in one where there is no protection.

However, the role of deposit insurance has not been without its criticism. Some researchers, academics and policy makers have argued against deposit insurance schemes stating that they have a potential to create a moral hazard problem by freeing both banks and depositors from

the consequences of their actions. On the part of consumers deposit insurance presents the problem of moral hazard by reducing the careful monitoring by depositors of financial institutions. On the part of banks especially in the absence of a sound regulatory system a deposit insurance may provide an incentive for banks to engage in more risky activities than would normally be the case. This study therefore highlights that although deposit insurance can preserve the stability of banking system, it may introduce instability through the moral hazard problem.

Demirguc-Kunt and Detragiache (1998) assert that on average the likelihood of a banking crisis happening increases under a deposit insurance system due to the problem of moral hazard. Their study revealed that the adverse effects on the stability of a bank under a deposit insurance scheme are more devastating in economies with weak institutions due to their failure to handle the moral hazard risk. Such countries need prudential and sound regulation and consistent supervision of the banking sector.

2.4.4 The Reinforcement of the Post bank

Post bank is a financial services division of Post Offices around the world. It aims to give mostly isolated and rural communities access to vital financial and related services. Among the various financial institutions in the country, the post bank is one of those well suited to take a leading role in providing financial services to the low-income, poor and rural societies in South Africa. Given this essential role of the Postbank, the government decided to bring its banking activities under the legal and regulatory ambit of the financial services sector. This process resulted in the creation of the South African Postbank Limited Act which created a legislative framework that resulted in the establishment of the Postbank. Generally, the whole idea was to develop and enlarge the products and services of the Postbank mainly for the rural and low-income markets including communities that had minimum or no access to banking services.

The restructuring of Post bank was part of government plan to provide a broader choice of accessible, affordable and relevant financial services and products to those who did not have bank accounts and low-income earners. Post bank took the lead with Mzansi accounts to 40 percent of the total market in 2006 and currently has the biggest share of these Mzansi accounts since its introduction in 2004. The post bank is expected to finally operate as a self-sufficient bank competing with the four largest banks known as the big-four.

Among other aims, the focus of the restructuring of the post bank with regard to access to financial services were to: (i) Expand the range of banking services and to develop into a bank of first choice, particularly to the marginalised rural and poor markets as well as communities that had less or no access to commercial banking services, (ii) Promote universal and affordable access to banking services, (iii) Ensure lending to rural and lower income markets and (iv) Ensure that the rates and charges take into consideration the needs of the people in the lower income market.

2.5 ACCESS TO FINANCIAL SERVICES

2.5.1 What is Financial Inclusion or Access to Financial Services?

Financial inclusion underlines the importance of accessing a range of financial services, namely, savings, loans, insurance, payments and credit (which are generally provided through banks) that are essentially aimed to help the poor people out of poverty. Fernando (2007) defines access to financial services in terms of the ability of an individual to use formal or semiformal financial services in the right and suitable form at an affordable price whenever the need arises. Thus financial inclusion entails the extensive provision of banking and other financial services to the low-income, rural and poor segments of the society at a reasonable cost. Bhandari (2009) states that the level of access in a region is normally indicated by the number of persons with access to a bank account. He argues that a bank account enables one to undertake essential functions such as investment, savings, insurance, obtaining a loan and other forms of credit. As such an ordinary bank account opens a door to a wide-range of many other financial services.

2.5.2 Access to Financial Services? Which ones?

Access to financial services contributes significantly to the developmental process via its impact on economic growth, poverty- and income inequality reduction. An inclusive financial system equips the poor with the needed financial resources to cushion themselves in the event of an external economic shock such as illness, deaths, accidents, theft, and unemployment (Demirguc-Kunt and Klapper, 2011). It enables poor people to save and borrow to build their assets and to make educational and entrepreneurial investments thereby enhance their standard of living. Inclusive finance is particularly important to benefit the underprivileged groups: rural societies, youth, women, and the poor. It is because of these reasons that

financial inclusion has received widespread recognition in current socio-economic discourse as it touches the lives of the poor people.

Egwuatu (2008) defines access to finance as sustainable financial services that enable the poor to invest, create wealth, increase income, and reduce their exposure to external shocks. One of the crucial questions that have been frequently asked regarding financial access has been: which services are to be considered essential? The World Bank (1995) identified four important areas namely, transaction banking, insurance credit and savings. The next sections will discuss in detail these key areas of financial services.

2.5.2.1 Banking

Hawkins (2010) acknowledges that the use of an entry-level bank account is a primary proxy for access to financial services, with the number of accounts an obvious initial indicator for access and usage. Access to banking services is important for at least the following reasons: Firstly, for emerging and developed economies that are evolving from a cash-type of payments to an era of e-payments, it is problematic and costly for a person who can only make payments in cash. Secondly, a basic bank account is considered a necessary gateway to a portfolio of financial services that can be provided by a bank such as personal loans, mortgages, vehicle finance or overdrafts. For this reason, it is argued that if a person has any financial product whatsoever, it is usually a bank account. Thirdly, lacking a simple bank account makes it difficult for one to have access to various other services such as insurance as such services require e-payments or a debit order from a bank account.

Access to banking services in general, presents people with at least one of the following benefits: (i) Banks perform a critical role in ensuring an efficient payment system for the economy. It facilitates receiving of regular and systematic electronic payments such as salaries, social grants, and pensions (ii) Banks convert cheques into cash, (iii) Banks provide safekeeping of depositors money and make available whenever a withdrawal is made, (iv) Banks facilitates payments for goods and services via several means besides cash, (v) Banks enable one to pay bills electronically, and (vi) Banks facilitates remittances. With the exception of the first three activities which can be done with a simple deposit account, the rest of the services can only be accessed by someone with a transaction banking account.

There are three categories based on the degree to which people are involved with banking namely; the un-banked, under-banked and fully-banked. The "*unbanked*" refers to those people in society who have no relationship with banking at all. The types of people within this category do not have any bank account and therefore do not use the services of banking institutions to manage their financial lives. Dorsey and Jacobs (2005) observed that in South Africa the unbanked generally tend to be those of low-income status. The "*under-banked"* are those with a simple bank account that does not include services for electronic payments or making remittances or even a cheque book. This category sometimes includes those persons who have such amenities at their disposal but do not make use of them. Kempson (2006) defines the "*fully banked"* category as referring to those persons who have access to a broad range of transaction banking services that are suitable and adapted to their specific needs and socio-economic status. It must be noted that since access to banking services is so crucial in everyday living, policy makers commonly use it as a rough indicator of financial inclusion.

The percentage of the unbanked population has not reached the desired minimum levels. As a result, there are private and social cost imposed on the society as a direct consequence of financial exclusion. Such costs would however be minimal if a significant portion of the excluded population were banked. The government of South Africa is currently targeting 70 percent by 2015 and 90 percent by 2030 from its current 62.8 percent of the banked population (BASA, 2012). Ironically, Deloitte (2004) reported the findings of one study as revealing that one in three South Africans felt that one could live his or her life without the need for a bank account. The study indicated that some individuals felt that it was in fact more problematic to live with a bank account than living without one. The case study identified the major reason given by the unbanked for failing to have a bank account as being the absence of a job or the lack of money. In addition, accessibility was also a commonly cited cause for not having an account. Some of the South Africans living in rural areas indicated that they did not have a bank nearby. Hence for the rural poor, lack of physical access to banks is still a challenge. This is because rural households have to incur transport costs apart from the banking charges for them to access financial services as these banks are mostly found in the developed urban areas.

2.5.2.2 Savings

Prinsloo (2002, p.73) defines saving by the household sector "as that part of current income, after the payment of direct taxes, that is not consumed or transferred as part of household

current expenditure." By contrast dissaving takes place when current expenditure exceeds current income. There is a very small chance on average that a person will manage even the smallest financial shock if the culture of savings is not adopted. The concern is to reduce the vulnerability of the poor in the face of uncertainty. There is also a growing need for the ageing population of South Africa to arrange financial provisions for their retirement. Those who keep cash savings at home rather than at a formal institution are usually at risk to theft and forfeit the benefit of interest payments. Rosenzweig (2001) argues that apart from serving as a way to accumulate balances in the long term, savings can help to smoothen one`s consumption behaviour and so enables one to cope with monthly expenses on a regular basis.

The poor like the rich need safe, accessible, flexible saving opportunities that provide them with positive returns. They need saving facilities to build a buffer of cash (Sadoulet, 2006). The poor need buffers to deal with variable income streams while still meeting their regular consumption requirement. Moreover, other expenditures require access to resources that are much greater than the amounts of money that households typically have available on-hand from income flows (Sadoulet, 2006).

Rutherford (2000b) classifies the need for savings into three general categories of events. The first reason is for life-cycle needs or predictable events and includes events in life that are relatively predictable, yet require resources that far exceed one`s current income. The financing of these events thus require accumulating income over time. Some of these events can be relatively short-term, such as consumption smoothing between periods of plenty and periods of deficit and purchases of durable goods. Others, such as social celebrations or gatherings such as childbirth, marriage, or funerals or other life-cycle needs namely, education and retirement. The latter usually require accumulation over a longer length of time.

The second reason for the need for savings is categorised under emergencies or unpredictable events. These emergencies are either personal such as sickness, theft, and loss of employment or wide-spread in nature such as drought, floods, wars and epidemics, which require a sudden and unanticipated need for money to meet expenditures which often correlated with a concurrent loss of current income (Sadoulet, 2006).

The third reason for savings is the need to take advantage of opportunities. Without broader access to finance, only the financially capable will be competent to make better use of economic opportunities. Rajan and Zingales (2004, p.113) makes the point that, "finance cannot create opportunities. It only makes it easier to exploit them." Thus the poor face high opportunity costs in terms of the foregone income because of incapability to exploit possible high-return investment opportunities. Accumulating savings, in the form of money or other assets, thus helps prepare the required buffer for these expenditures. However, saving at home can be affected by a number of risks such as the theft or loss of their cash savings or financial risks, such as inflation. Rutherford (2000) noted that savings services must encompass these four properties: safety, accessibility, flexibility and positive returns.

Safety: As mentioned earlier saving in kind or at home is considered risky. For example if one saves their money at home there is danger that these little savings may compete against other pressing claims usually from friends and relatives.

Accessibility: Small transactions mean that the poor need to be able to save tiny amounts, when they have these at their disposal at low transaction costs. Thus an important feature of savings products is to be able to harness these small income surpluses by having access to regular and frequent opportunities to save small amounts. Accessibility includes proximity and convenient opening hours, and speed of processing.

Flexibility: Rutherford (2000) argues that having a buffer is not of much use unless it can be accessed in case of need. This means that savings must be close at hand, available, and easily convertible into cash if they are in another form. However, ease of withdrawal does not necessarily mean that savings must be extremely liquid at all times. Liquidity refers to the ease with which saved funds may be accessed when need arises.

Positive returns: The opportunity cost of saving one`s funds at home is the positive real return that could have been earned should these funds have been saved in a bank or formal financial institution. Savings depletion stems from negative returns, among other things. An attractive feature for savings is that they do not dwindle through the effect of inflation or service charges but should accumulate over the duration of the time. Hence, when credible saving schemes exist, positive interest rates on savings products increase their attractiveness, which stimulates the amount saved (CGAP 2003). Therefore the attractiveness of savings

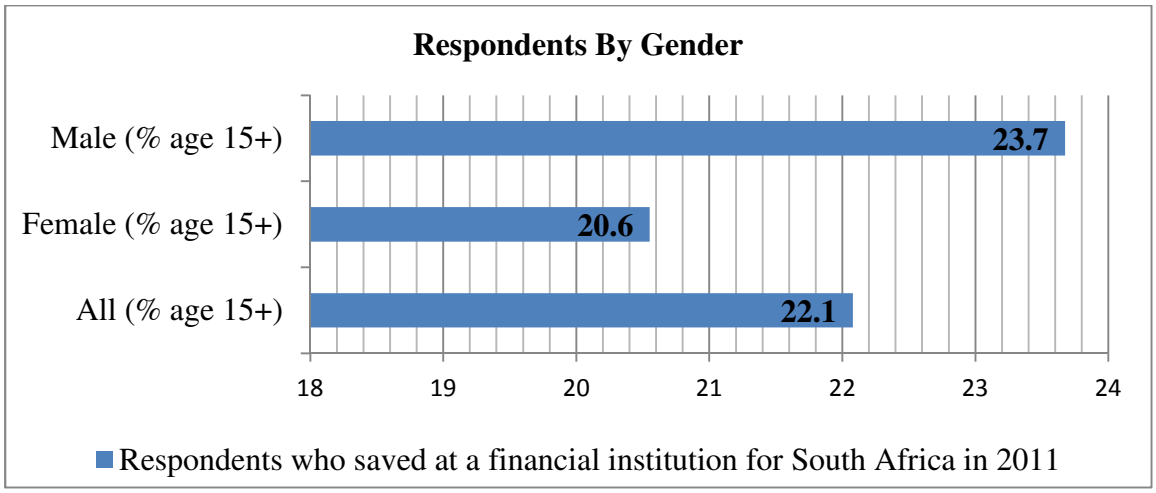
facilities should be evaluated on the basis of whether there are safe, accessible, flexible, or beneficial. Sadoulet (2006) states that savings facilities are either unsafe, inaccessible, inflexible or expensive.

Finally, geographic proximity of savings institutions is a critical factor which primarily affects those who are located in remote areas. Since the poor need to be able to save small amounts frequently, deposit facilities should be local, available, and less complicated with paperwork. There is therefore, need to ensure that bank branches be extended to rural and remote areas in order to enhance access to these services. As Lewis (1955, p.229) puts it:

> "Experience shows that the amount of savings depends partly on how widespread these facilities [i.e. Savings institutions] are; if they are pushed right under the individual`s nose, ... people save more than if the nearest savings institution is some distance away."

In his keynote address at the 10th Anniversary of the South African Savings Institute (SASI), Mr. Gordhan (2011) - the Minister of Finance, noted that South Africa`s low savings rate compared to its international peers was holding back the country. According to latest data from SARB Quarterly bulletin (2013), South Africa`s household savings to disposable income remained unchanged at 0% for the period 2012 and 2013. The figure indicated an improvement from – 0.6% and – 0.3% in 2010 and 2011 respectively. Using World Bank Global Financial Inclusion (Global Findex) (2012) data, we analysed the behaviour of saving by characteristics such as gender, income groups, rural or urban residence and age groups as depicted in figure 2.12 – 2.15 for South Africa for the period 2011. Demirguc-Kunt and Klapper, (2012) argues that analysis of financial access data by individual characteristics is necessary in order to allow policy makers and researchers to identify groups that are financially excluded enabling a better understanding of particular financial behaviours. The saving indicator variable used in this analysis is defined by the Global Findex (2012) as the percentage of "respondents who report saving or setting aside any money by using an account at a formal financial institution such as a bank, credit union, microfinance institution, or cooperative in the past 12 months".

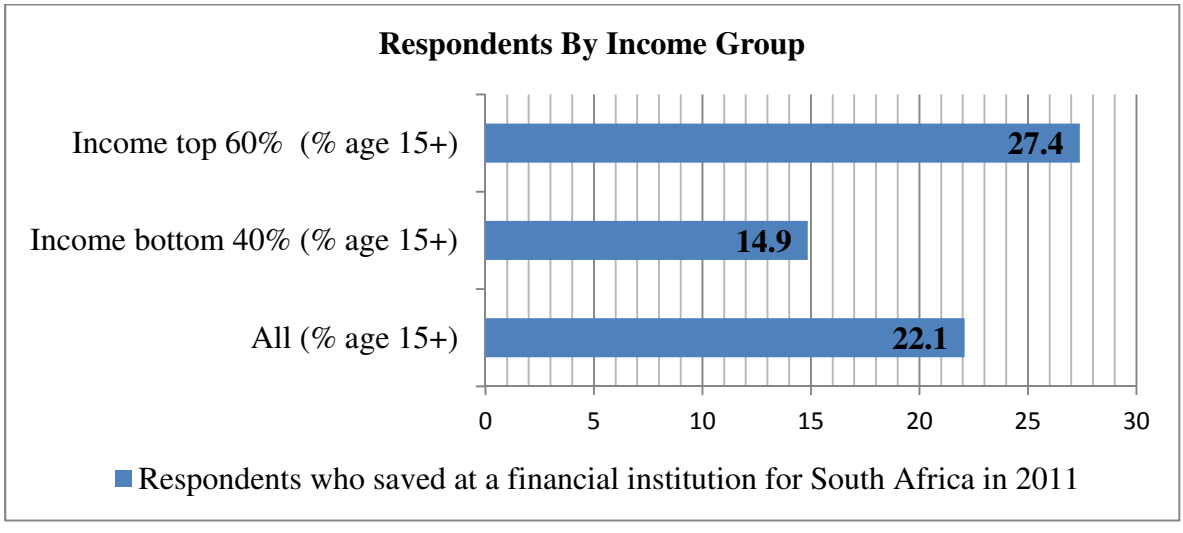
FIGURE 2.12: SAVED AT A FINANCIAL INSTITUTION IN THE PREVIOUS YEAR: BY GENDER



Source: Global Findex (2012).

Figure 2.12, indicates notable disparities in savings behaviour along gender lines. For instance, 23.7% of men reported having saved in the previous year while only 20.6% of women did, indicating a gender gap of 3.1 percentage points. According to Global Findex (2012), 22% of adults worldwide report having saved at formal financial institution in the past year while 14% did so in Sub-Saharan Africa (SSA). The largest contributor countries within the SSA region were Nigeria (24%), Kenya (23%) and South Africa (22%) placing SSA as the highest region excluding East Asia & Pacific with 28%. However South Africa is below the average of 24% and 45% for upper-middle income and high income countries respectively. When compared to BRICS countries namely, Brazil (10%), Russia (11%), India (12%), and China (32%), South Africa`s saving performance is better except with China.

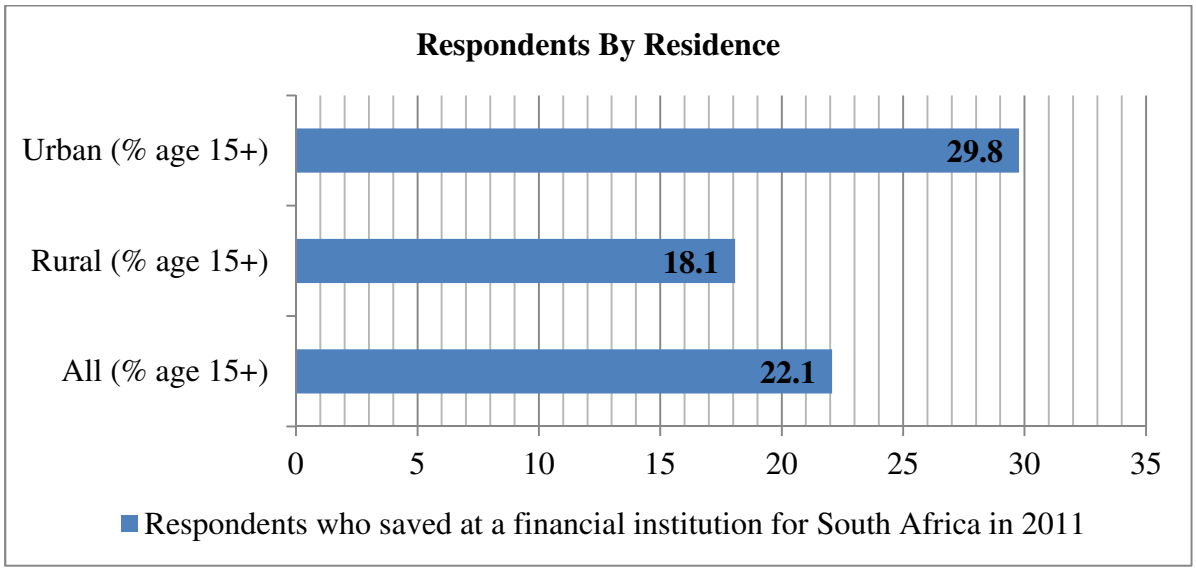
FIGURE 2.13: SAVED AT A FINANCIAL INSTITUTION IN THE PREVIOUS YEAR: BY INCOME DISTRIBUTION



Source: Global Findex (2012).

Analysis of household saving behaviour across income groups indicated that those in the top 60 percent income bracket are almost twice as likely as those in the bottom 40 percent income bracket to save at a formal institution. The gap averages about 12.5 percentage points. The 14.9% reported for the income-bottom 40 percent population group in South Africa is lower than the 17% reported for the income-bottom 40 percent population group for upper-middle income countries.

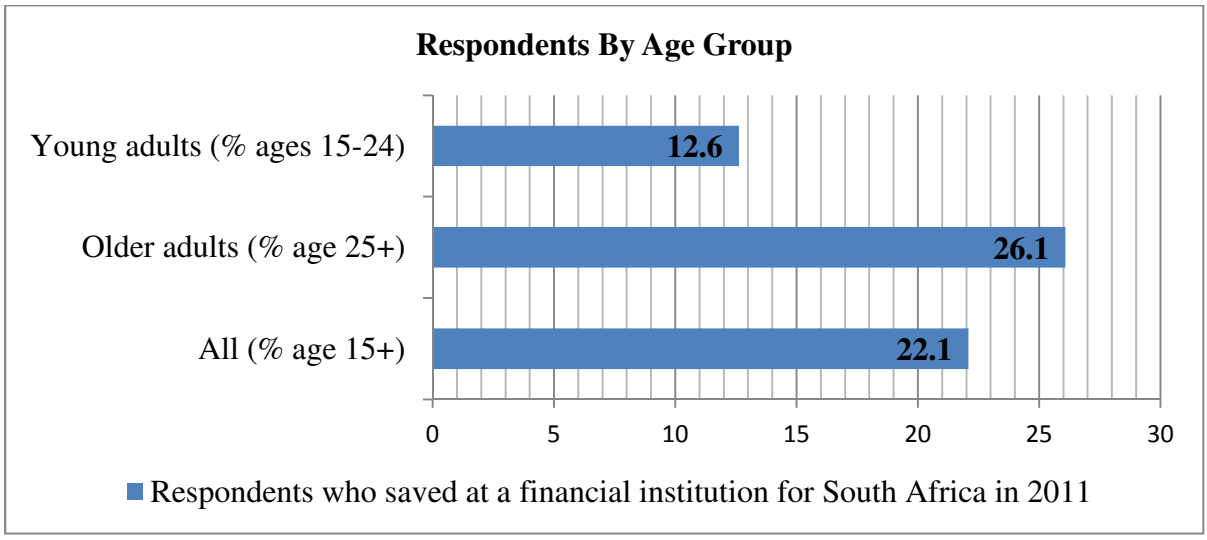
FIGURE 2.14: SAVED AT A FINANCIAL INSTITUTION IN THE PREVIOUS YEAR: BY LOCATION



Source: Global Findex (2012).

Wide disparities were also prevalent across rural and urban residents. For instance Figure 2.14 indicates that adults living in urban areas are more likely to save relative to those living in rural areas. The reported 18.1% share of South African adults living in rural areas who save was lower than the comparable 27% for upper middle-income countries. Demirguc-Kunt and Klapper, (2012) identified distance from a bank as a major barrier to financial access in rural areas.

FIGURE 2.15: SAVED AT A FINANCIAL INSTITUTION IN THE PREVIOUS YEAR: BY AGE



Source: Global Findex (2012).

Figure 2.15 portrays that the culture of saving is a major challenge among the young adult population of South Africa. 26.1% of the older adults, the working age population, who saved in the previous year were more than twice that of the young adults (12.6%). However, the share of South Africa`s working age population who report having saved in the previous year is lower than the upper-middle income countries (28%) and much lower than the average for high income countries (48%).

2.5.2.3 Credit

Access to credit is another vital pillar of financial services which plays a crucial role of smoothing consumption patterns and guarding against income or financial shocks allowing one to undertake investments and creating more income in the long-term. Basic financial products such as loans, credit cards and overdraft facilities can cushion one against unforeseen contingencies and so evens out consumption during economic hardships. Nieri (2007, p.118) states that the expression "credit exclusion" describes situations where individuals are unable to purchase from mainstream providers those services essential to achieve their credit needs for one reason or another. The reasons may be personal or economic or a previous record of bad debt. Kempson et al (2000) emphasises that the debate is more on access to affordable credit. He argues that one can be considered credit excluded if he/she has to incur charges that significantly exceed those in the conventional credit market. Due to the reason that a significant proportion of people prefer not to use credit despite the fact that they could gain access to it presents a much bigger challenge on the measurement of

credit exclusion than banking exclusion. Kempson et al (2000) suggests that by gauging the facilities that people have at their disposal regardless of whether they utilise them or not one can obtain a rough measure of the state of access to credit.

Collateral in the form of an asset is required to secure the loan. The types of assets generally used are real estate and some moveable or fixed properties. The asset which is offered for collateral can remain in the custody of the borrower or placed under the supervision of the third party or the lender. However, in the event of default by the borrower or failure to adhere to agreed financial obligations, the lender reserves the right to dispose the asset. Rodriguez-Meza (2006) states that the rights to earnings obtainable from the use of an asset can also serve as a form of collateral for the full outstanding amount.

Access to credit is normally limited by the restricted set of agreed assets that can be accepted by the bank as collateral. This is because the type and value of collateral used should be consistent with the amount of the funds being requested. Unfortunately, when poor households fail to secure the right kind of assets as collateral security they are denied their credit request. As a consequence, many are forced to resort to informal financial markets where they usually fall prey to high interest rate charges. Thus Rodriguez-Meza (2006) claims that widening the number of assets which can be accepted for collateral is therefore key to improving access to credit. In addition, those who already have access to credit would benefit in terms of improved terms and conditions in the form of longer terms to maturity, larger loan sizes, and lower interest rates.

Collard and Kempson (2005) noted that people who fail to obtain credit from banks or other conventional financial institutions will often resort to use sub-prime lenders where the terms and conditions may not be favourable and charges are costly. In the very worst case, the financially excluded usually turn to illegal and unscrupulous lenders, who charge exorbitant rates and use intimidation and violent strategies to recover their money when the borrower fails to honour their financial obligations in time.

2.5.2.4 Insurance

There are basically two types of insurance that are generally offered to consumers namely, short-term and long-term insurance. Examples of long-term insurance are health, disability and life insurance. In the event that the insured person dies, gets ill or becomes disabled,

these policies pay out a benefit against a claim. On the other hand, short term insurance is essentially for properties of individuals which include houses, cars, home contents, or other assets. In comparison to 17% and 3% of adults in developing economies and Sub-Saharan Africa who reported having paid for health insurance in the past 12 months, South Africa`s share was reported at 7% in 2011 (Global Findex, 2012). When compared to BRICS countries: Brazil (8%), Russia (7%), India (7%), and China (47%), South Africa is marginally lower than Brazil but substantially lower than China as indicated. In recent years, the president of South Africa expressed his commitment in implementing health insurance for all South Africans particularly the poor majority. However, it invoked a debate on how such a huge initiative would be funded sustainably.

The basic need to take on accountability for risk management of various forms compels one to need a range of affordable insurance products. In South Africa taking insurance on motor vehicles is a requirement while insuring the possessions of one's home is a personal responsibility. However no straightforward definition exists with regard to what insurance product types are considered important to enable one to identify those people who might be regarded as financially excluded. Nonetheless, a bank account is a necessary requirement when considering buying insurance. This is because premium payments for insurance services require payment from a bank account.

Insurance exclusion is somewhat a new concept and the formal sector in South Africa has not done much to integrate the low income to access affordable insurance services. As a result, a significant number of South Africans have remained without access to a suitable choice of formal insurance products and services. National Treasury (2011) reported that a 2009 study conducted by FinScope revealed that 15 percent of the coloureds and 23 percent of blacks in South Africa who were surveyed received cover against funeral costs by way of an informal burial society. According to FinScope (2010), 50 percent of adult South Africans were covered by insurance which was mainly funeral insurance. Their report also noted that the majority of such policies were sold by unscrupulous and unregistered providers. The Treasury policy document identified three areas of challenge that South Africa is currently facing regarding insurance: Firstly, the need to facilitate the taking-up of insurance policies by the disadvantaged and vulnerable groups; Secondly, a need to alter the composition of the insurance services to the poor and low-income people. It was identified that most of these

poor people were mainly protected against funeral expenses only, and; finally the need to ensure protection against unlicensed and unscrupulous operators.

The 2011 National Treasury insurance policy framework, was aimed at achieving the following objectives: (i) to expand access to a wide range of formal insurance products suited to the needs of the low-income households; (ii) to enable existing informal insurance providers to deliver formal insurance; (iii) to lower the barriers to entry in order to increase wider participation and promoting competition among the providers; (iv) to ensure protection of micro-insurance consumers; and (v) to facilitate efficient and effective supervision as well as enforcement.

2.5.3 Significance of Banking Services

National Treasury (2011) in its policy document, *A Safer Financial Sector to Serve South Africa Better*, highlighted the importance of the financial services sector in that it touches the life of every South African. It is recognised that lack of access to banking services not only constrains economic growth, but that it also traps people in poverty. At micro-level access to banking services enable people to make every day business, conduct transactions, invest, save and create wealth to achieve their future goals and provide for retirement, and to insure against unforeseen occurrences. At macro-level, the financial sector facilitates the attainment of growth and generation of employment.

The banking sector positively performs a vital role in the attainment of the developmental aspirations of a country. This critical sector functions as a catalyst for economic growth and creation of jobs, and so guarantees sustainability of economic development within a country. The financial sector in general and the banking sector in particular offer a base for its people to engage in business and exchange payments for services rendered. By performing this essential service the financial sector touches the lives of ordinary people around the globe.

Financial inclusion is a critical and fundamental pillar towards stimulating economic growth and development in South Africa. Hence the government has an obligation to ensure the facilitation, promotion, enhancement, accessibility and usage of suitable financial services and products for all South Africans particularly those who are currently excluded. However, during the apartheid years, a significant proportion of the population was excluded from banking services. Following the demise of apartheid regime, the new government committed

to ensuring financial inclusion and consumer protection, while maintaining strong policy on maintaining financial sector stability. Access to financial service has gained prominence among policymakers in developing countries in these recent years.

It is widely acknowledged that access to financial services is a contributing pillar in reducing poverty and encouraging social inclusion and that financial exclusion acts to retard efforts geared towards achieving these outcomes. Research evidence from reviewing literature highlights the importance of the role of a well-developed financial system for reducing poverty (Beck et al 2000; Beck et al 2004). Subsequent sections will endeavour to shed some light on why expanded access to financial services deserves so much attention especially in an economy like South Africa where issues such as crime, unemployment and education seem to be equally important. Porteous and Hazelhurst (2004) identified three critical areas as the key reasons why broader access to financial services matters for any economy. These include growth and job creation, poverty reduction and social exclusion.

2.5.3.1 Growth and Job Creation

There is a large body of literature and empirical studies that have sought to find the link between access to financial services and economic growth. There is however, inconclusiveness regarding the direction of causality. The question being: is it economic growth that leads to increased usage of financial services or is it financial services usage that generates economic growth. However, according to a study conducted by FinScope (2003), financial development leads to economic growth especially for developing countries. Porteous and Hazelhurst (2004) maintains that there are simple microeconomic chains of causality. Access to finance provides entrepreneurs with the necessary capital required to either start a business or develop the existing one thus creating employment and increasing national output. One of the key urgent social demands in South Africa is the need for housing finance. Easier access to finance, holding other things constant would increase demand for new houses which would stimulate the construction of new houses. This in turn increases consumption demand in areas of construction, furniture and other household durables. The previous example has demonstrated why expanded access to financial services is a crucial driver of domestic activity.

In recent years, SMMEs have become subjects of interest for policy makers as vehicles for promoting economic growth and employment in developing countries. As a result they have

become targets of policies due to their contribution to poverty reduction, employment generation, and private sector development. Making finance available for micro-enterprises is seen by many as one way of addressing the growing crisis of unemployment in South Africa. As a result, small businesses have gained high recognition from most governments as being integral to achieving high economic growth and employment. South Africa is currently grappling with the challenge of reducing unemployment currently recorded at 25.2 percent (representing a 0.2 percentage point rise compared to last year for the same period) as per the first quarter of 2012 (STATSSA, 2012).

A small business survey undertaken by FinScope (2010) has shown that approximately seventeen percent or one in six persons in South Africa of 16 years and older creates income through a small business and that 58 percent of small business owners are women. 22 percent of small business owners named the challenge of sourcing money as their major obstacle when starting up a business, while 14 percent identified cost of finance or accessing finance as the main obstacle to expanding their business. An interesting finding was that 76 percent of small business owners did not know of any establishment that offered help or advice to small businesses. Moreover, only 9 percent of small business owners knew of the role of banking institutions as a possible form of support, and only two percent actually sought financial assistance from banks.

This section has demonstrated that access to financial services by SMMEs and the public in general holds the potential to expand economic growth and thereby help reduce the current problem of unemployment in South Africa. As highlighted in previous sections there is need for the government to engage the private sector and key stakeholders in order to develop an enabling environment for the survival of SMMEs if the economy is to reap the benefits of economic growth and job creation.

2.5.3.2 Poverty Reduction

Access to financial services, although it is not a panacea has gained the attention of policy makers for its potentially significant role in reducing poverty. Fernando (2007) states that like the rich people, the poor can also benefit from accessing financial services in many ways. The poor households also need to obtain credit and insurance, to save, to invest, and to make electronic payments. Fernando (2007, p.2) states that the “poor need access to financial

services much more than the rich simply because the poor have little money". James et al (2005, p. 2) in their article *Rethinking Bank Regulation* states that:

> "Poorly functioning banks that simply funnel credit to connected parties and elites slow growth and exert a disproportionately negative influence on the poor and small businesses by depriving them of the capital they need to succeed. Unfortunately, billions of people live in countries with poorly functioning banks. Thus, banking policies matter because, banks influence the ability of people, rich and poor to improve their living standards."

It widely recognised that there are bidirectional causality relationships between financial inclusion and poverty. For example, financial exclusion can act as a primary source of poverty while poverty can cause reduced usage or demand of financial services. In both cases people need access appropriate and suitable financial services, which are best provided through banking institutions.

In 2000, many member countries of the United Nations adopted the Millennium Development Goals (MDGs) in which they committed themselves towards alleviating poverty by 2015. Poverty reduction is a key national goal. While empirical studies on the link between financial sector development and economic growth are well documented, little has been done regarding the nexus between access to financial services and poverty reduction. However, it is generally agreed that lack of access to finance keeps people trapped in poverty. A vast amount of literature has demonstrated that access to financial services helps reduce the overall levels of poverty. For instance, without access to capital or a bank loan it is difficult to create wealth, finance one`s education or improve one`s health status. Many researchers therefore claim that the delivery of these services to the poor is a powerful means of providing them with the chance to break away from poverty and improve their lives. Access to financial services, gives people the chance to improve their lives through creating jobs, smoothing income and consumption flows, enlarging and diversifying their businesses, and increasing their income and other benefits, such as their education and health care (Hao, 2005).

Financial services such as savings and insurance perform a vital role in reducing the impact of income shocks caused by illness or death thus reducing the vulnerability to poverty.

Ardington and Leibbrandt (2004, p.1) state that "Vulnerability is a cause of poverty and poverty is in turn a source of vulnerability." To achieve sustainable reduction in poverty, the authors argue that there is need for poor households to be capable of managing their risks effectively. Hence access to financial services helps one to manage risk and therefore reduce the impact of shocks. Lack of financial risk management assets which are available through being integrated into the financial system, makes it difficult for poorer households to cope during difficult times making them resort to desperate surviving ways such as removing their children from school, selling house equipment and borrowing at relatively higher interest rates thereby making them even more vulnerable to permanent poverty. In a country study of Indian villages, Jacoby and Skoufias (1997) observed that the tendency to remove children from school due to temporary economic shocks was more for households lacking financial access than those with better financial access. Hence, providing access to finance for the poor is a potential tool for economic development and poverty reduction. Thus, policies designed to lift the poor from poverty in South Africa should emphasise the importance of access to financial services as one way to equip the poor with the means to improve their lives. Hao (2005) cautioned that ensuring access should be done sustainably as he observed that varied poverty reduction approaches that many policy makers pursue fail to generate finance for the poor on a sustainable basis.

2.5.3.3 Social Exclusion

Access to financial services is seen as key in tackling wider social exclusion, as having a bank account can act as a gateway to many other services (Devlin, 2006). It is generally agreed that financial exclusion may either be a consequence or a cause of social exclusion, or both. In a broader sense, financial exclusion has been defined as practices or actions that prevent the poor and marginalised segments of our society from acquiring access to financial services. In a narrow sense financial exclusion entails the absence of access to particular services such as banking services, credit or insurance services or more precisely banking exclusion, credit exclusion or insurance exclusion respectively.

There are private and social costs of not having a bank account in a country where the majority of citizens are banked. Confronted with financial difficulties, the excluded individuals usually turn to illegal providers namely, loan sharks whose interest charges are very high. It is also reported that most of these unscrupulous lenders use threats and violence

if the borrower fails to honour their obligation on time. Balmer (2006) noted that this has further negative socio-psychological effects on such individuals causing even more severe health challenges of chronic depression and stress.

Given the link between social exclusion and financial exclusion, there are some societal ills that are a direct consequent of financial exclusion. Financial exclusion deprives people the opportunity to become an integral part of the community. Consequently, this hinders them from participating in the development of the country in which they live thus causing deeper poverty. As Hersi (2009, p.32) puts it "financial exclusion has a detrimental effect on how people run their lives and puts them in a very difficult position where they feel that they have lost control of their lives, causing anxiety and bringing about severe personal and community consequences."

Social exclusion has a number of undesirable consequences for the society. According to the World Bank (2006) social exclusion acts as a brake on development depriving people of the necessary means towards economic freedom. It is generally acknowledged that financial exclusion is at the centre of the problem of social exclusion that exists in society today. Financial exclusion highlights the failure of the mainstream financial providers in meeting people's demands for a variety of financial services that are suitable and adapted to their needs at a reasonable price. Lack of financial access therefore disconnects one from gaining access to the means of attaining basic needs of life such as decent housing, education or health care.

2.5.4 The consequences of "banking exclusion"

European Commission (2008) states that the intensity of the repercussions of financial exclusion is dependent on the dominating state of financial exclusion in a particular country. He argues that for example it is more difficult for one to depend on cash transactions within a country in which nearly everybody else is banked, than it is in a country where a considerable proportion of the population are unbanked. Furthermore, individuals without bank accounts encounter problems when handling cheques written to them by a third party. Many times they are required to pay a fee in order to cash in the cheque. This is the reason why in most developing countries there are informal numerous enterprises that offer cheque cashing services and other financial related services. To the majority of those excluded, such organisations are the only option available. As such Hersi (2009, p.28) states that the

unbanked pay a “huge premium for being excluded”. Moreover, it is also difficult to take formal employment when one lacks a bank account. This is because wages or salaries are paid out monthly on a regular basis by way of an electronic transfer to the employee’s bank. Moreover, purchasing bank services separately is generally more costly compared to accessing the same services within an established and continuous relationship with a bank. This explains why for non-customers of a bank accessing services such as cashing of cheques, bank transfers to third parties, payment of bills and so on is relatively expensive. Some services offer discounted rates for people making payments electronically namely, when one intends to purchase an air ticket. Hence, individuals without a credit or debit card forfeit the benefits of taking-up such regular discounts and promotional prices for goods and services bought in this way. In addition there are online products that one may require a credit card to make payment for which cash may not be readily acceptable.

One of the most serious financial repercussions for those without a transactional bank account arises when one cannot gain access to affordable products and services from banking institutions. Hence, they resort to illegal credit alternatives such as loan sharks where the interest charges are costly and the terms and conditions are not favourable. Such desperate alternatives are a real cause for concern in South Africa. These loan sharks will do everything in their capacity to recover their funds when borrowers default on their financial obligations. As a result there are reported cases where some have lost their lives to these unscrupulous lenders. Therefore the economic repercussions of seeking assistance via such organisations are severe and devastating for those excluded from the banking system.

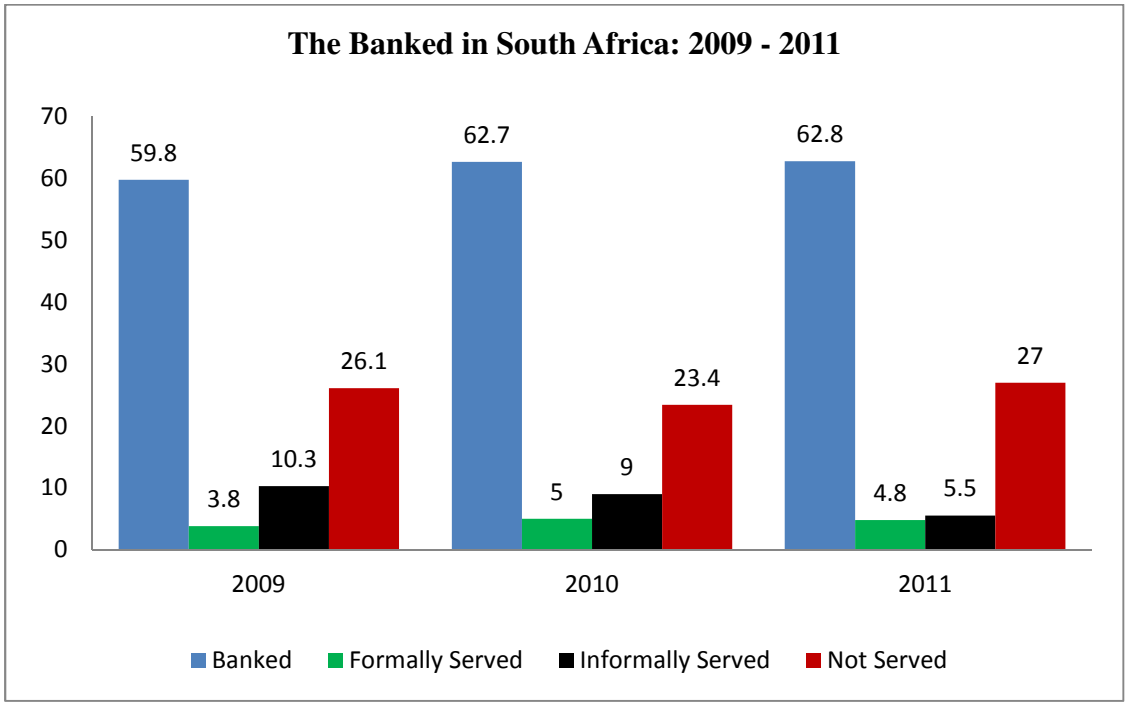
Finally the majority of the low-income households who have attempted to save their cash with informal markets have suffered substantial losses. Due to the inefficiency of intermediation in informal markets, borrowers face relatively high interest charges and many cannot access long-term loans. However, the persistence of these informal markets highlights the fact that there are serious supply-side limitations within the formal banking sector with regard to financial access.

2.5.5 Trends in South African banking

2.5.5.1 Access to Banking Services

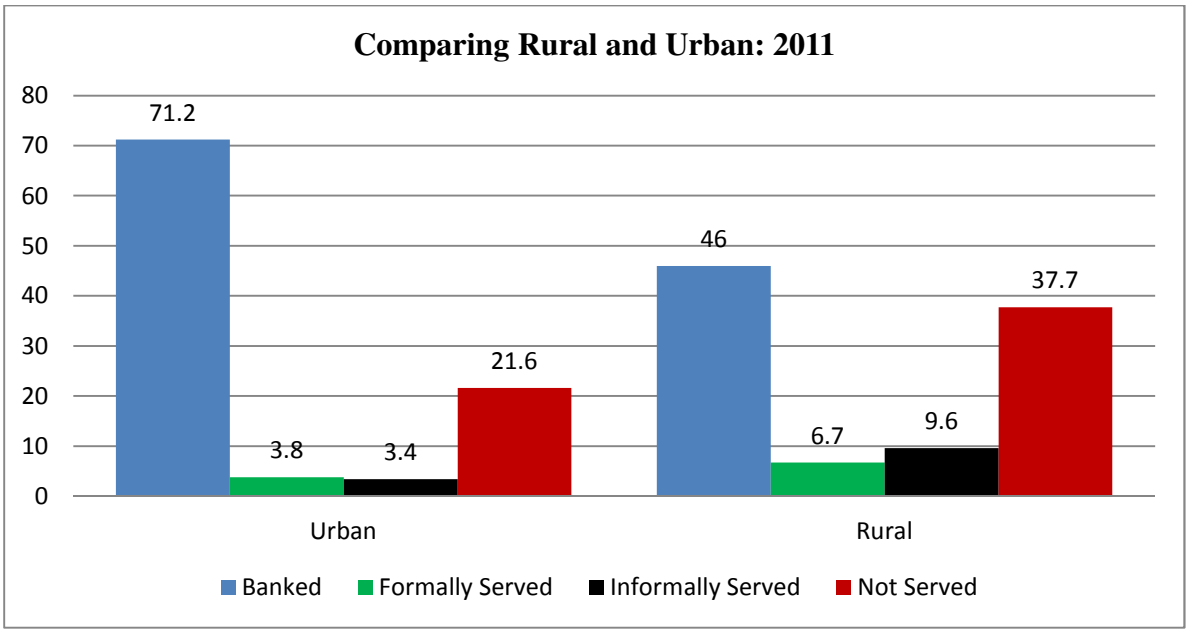
South Africa has a developed, fairly sophisticated, and well-regulated banking system which compares well with those of advanced economies. However, Okeahalam (2008, 1133) argues that South Africa falls behind many developed countries when evaluated on the basis of factors such as proximity to a bank branch, percentage of adult population with a bank account and access to bank services. According to FinScope (2011), 27 percent of South African adults are financially excluded, that is, they do not use both formal and informal services at all in their lives. This figure represented a significant 4 percent increase compared to 2010. 67.6 percent of South Africans are formally served and 62.8 percent have or use bank products and services and 4.8 percent have or use non-bank formal products and services but not any bank product. 1.6 million South Africans of 16 years and above rely solely on informal financial products or services, a decline of 1.4 million from 3 million recorded in 2010. Finally, the proportion that is not included into the financial system both informal and formal is 27 percent.

FIGURE 2.16: THE BANKED IN SOUTH AFRICA: 2009 - 2011



Source: Computed using data from FinScope (2011)

FIGURE 2.17: THE BANKED: COMPARISON OF RURAL AND URBAN: 2009 - 2011



Source: Computed using data from FinScope (2011)

Figure 2.17 shows a comparison of urban and rural access to financial services. The major difference between urban and rural state of financial inclusion is in the usage of bank products 71.2 percent of urban adults are banked compared to 46 percent of rural adults. This supports the earlier discussion that banks find the impetus to serve the up-market more profitable than serving the low-income clients. The assumption is that the urban market comprises the high-income clients while the rural market is predominantly made up of the low-income clients. FinScope (2011) highlighted the importance of the role which the informal sector plays in bridging the gap between the banked in urban and rural category. As seen in Figure 2.17 this role is considerably more pronounced in rural areas.

TABLE 2.4: INTERNATIONAL COMPARISON OF SELECTED FINANCIAL INCLUSION INDICATORS

	Australia	Brazil	India	Mexico	South Africa	United Kingdom	United States
Deposit value (% of GDP)	75.18	35.55	55.03	15.08	92.92	61.32	43.91
Loan value (% of GDP)	115.67	78.61	40.93	13.36	95.96	80.64	44.81
Bank branches per 100 000 adults	32	13	10	15	8	21	36
ATMs per 100 000	157	112	7	45	52	123	176
POS per 100 000	4 040	2 247	67	592	1 068	2 331	-
Value of SME loans (% of GDP)	15.33	3.77	4.34	-	10.71	-	4.93
1. Red shading indicates the lowest value 2. Green shading indicates the highest value							
Source: National Treasury Policy Document, 2011							

Source: various South African Reserve Bank publications

2.5.5.2 Profitability

Return on equity (ROE) and return on assets (ROA) are the two widely used performance indicators to evaluate profitability. ROA is found by expressing a bank's net income as a proportion of its total average assets in a given period. In essence it is an indication of how much profit a bank generates from every one rand in assets. High figures of ROA are desirable as long as the bank is not undertaking unjustified risk. On the other hand, ROE measures how much profit a bank generates with each rand of shareholder`s equity. It is therefore found by expressing a bank`s net income as a proportion of its average shareholders' equity. It also serves as an indication of how efficient a bank is in generating its profits. Banks that are highly geared (more debt relative to equity) tend to achieve high ROEs. As such as small equity base may lead to high ROE and constraining the capacity of a bank to borrow funds. For this reason Ikhide (2000) argues that a high ROE cannot always be taken to imply high efficiency. Nevertheless in comparison, ROE is more preferred to ROA.

TABLE 2.5: ROE & ROA: SOUTH AFRICA`S BANKING SECTOR

	2008	2009	2010	2011
Return on Equity (ROE)	20.65	15.88	14.64	16.39
Return on Assets (ROA)	1.15	0.94	0.97	1.15

Source: Bank Supervision Annual Report (2011) and Various Quarterly Bulletins SARB

Both the ROA and the ROE deteriorated in 2009, probably illustrating the impact of the financial crisis. The ROE continued to fall in 2010 and peaked in 2011. Following the drastic fall in 2009, ROA started recovering in each of the subsequent years as depicted in Table 2.5. According to Bank Supervision Department *Selected South African Banking Sector trends* (2012), the ROE and ROA was at 17.10 and 1.20 respectively indicating improvement in profitability in the banking industry.

2.5.5.3 Capital Adequacy

Capital adequacy is a concept based on managing or rearranging the existing capital structure in order to cushion the banking sector against potential losses as well as protecting depositors. Mishkin (2013) identified three reasons why banks need to hold optimum levels of capital. First and foremost, the right amount of bank capital helps to avoid the occurrence of a bank failure. A bank failure means that banks have failed to meet their financial obligation to pay their depositors and other creditors forcing the bank to shut down. Some analysts have argued that it was actually a "capital crunch that caused a credit crunch" in the 2007/2008 financial crisis. Mishkin (2013) in favour of this notion states that shortages of bank capital triggered at least in part the credit crunch that culminated in the global financial crisis. In 2010 in the wake of the sub-prime financial crisis, a global position under BASEL III was agreed upon to set standards and practices that would guarantee the preservation of adequate capital for the stability of the banking industry. BASA reported that South Africa began implementing the capital agreements along with other goals of the BASEL III accord in January of 2013.

Secondly, the return on equity is affected inter alia by the amount of capital the bank has. On the one hand, higher capital levels are necessary to safeguard the investment of owners by decreasing the potential for insolvency. On the other hand, holding more capital implies a

lower return on equity. The capital amount that a bank decides to hold is directly informed by the risk the bank is exposed to (KPMG, 1998). For this reason, Mishkin (2013) suggest an optimum trade-off between high capital sufficient to guarantee bank safety and a lower return that goes with high capital. However, during times of negative business outlook there is need for the bank to hold more capital to protect the holders of equity. Similarly, when bank managers are more optimistic they may decide to hold less capital to improve the return on equity.

Finally, regulatory authorities define the minimum capital requirements that all registered banks should maintain. In the case of South Africa, the tier 1 CAR minimum requirement rose from 11.8 in 2010 to 12.04 in 2011 as shown in Table 2.6. The table below highlights the capital position of the banking sector against the minimum requirements for the period 2008 to 2011.

TABLE 2.6: CAPITAL ADEQUACY RATIOS (CARS): SOUTH AFRICA`S BANKING SECTOR

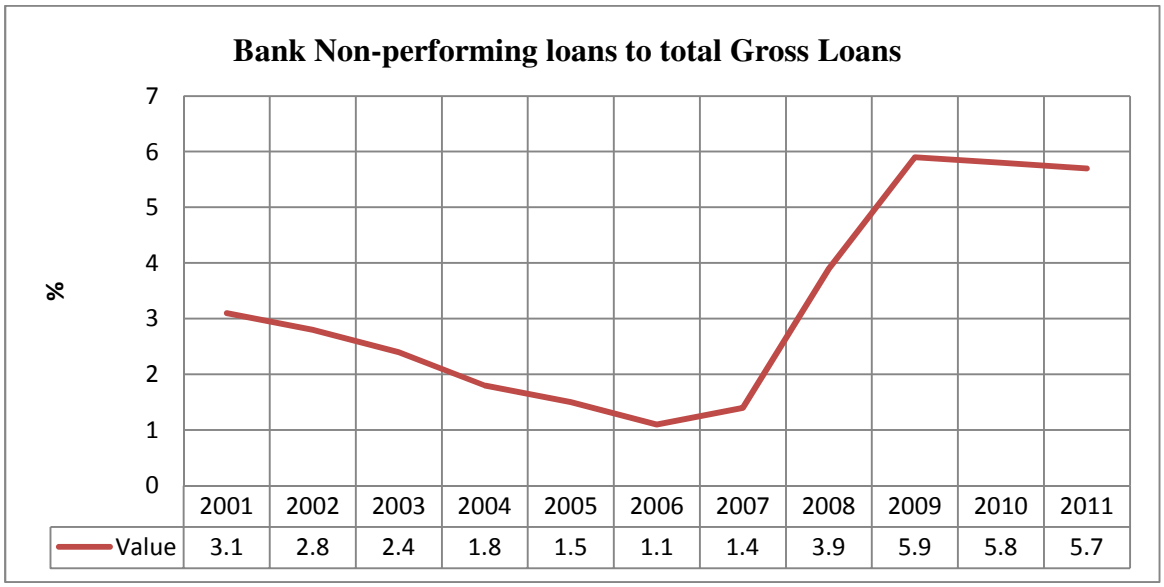
	2008	2009	2010	2011
Capital Adequate Ratio	13.01	14.12	14.88	14.91
Minimum required	10.22	10.97	11.80	12.04

Source: Bank Supervision Annual Report (2011) and Various Quarterly Bulletins SARB

Capital Adequacy Ratio (CAR) of the banking industry rose from 13.01 percent in 2008 to 14.91 percent in 2011, with banks sustaining their CARs above the stipulated requirement of at least 10 percent. The sector's Tier 1 capital-adequacy ratio was 11.8 percent as at the end of December 2010, compared with 10.97 percent in December 2009. As at the end of December 2011, the capital adequacy ratio (CAR) was 14.91 well above the statutory requirement of 12.04 representing an increase from 14.88 in December 2010 against the statutory requirement of 11.80. In a nutshell, the banking sector remained adequately capitalised.

2.5.5.4 Non-performing Loans

FIGURE 2.18: BANK NON-PERFORMING LOANS TO TOTAL GROSS LOANS

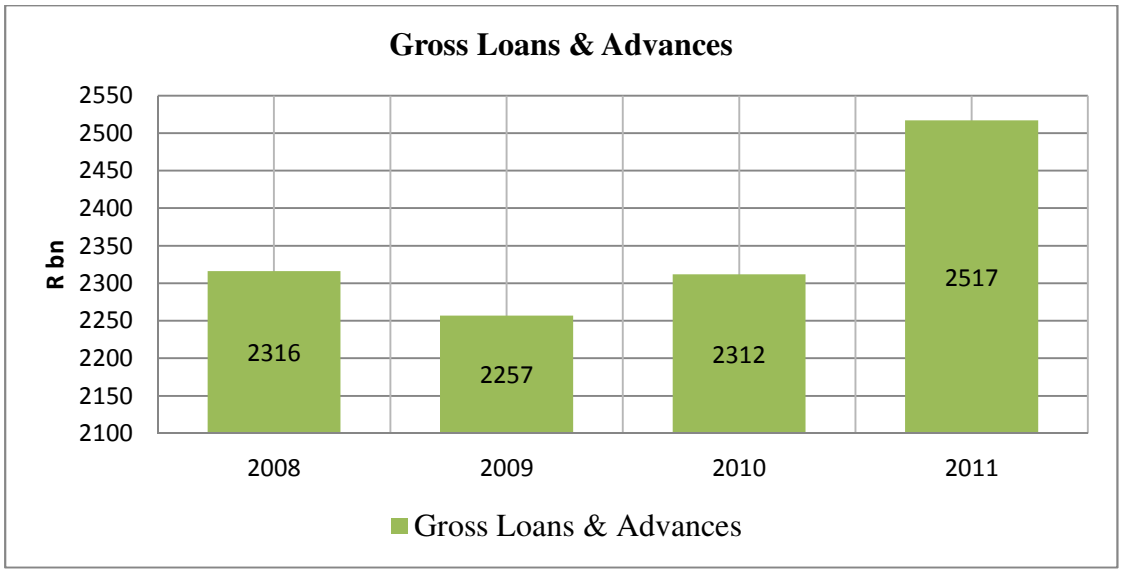


Source: Computed using data from Mundi Index

Another measure of bank performance that captures quality of bank loans is its level of non-performing loans (NPLs) or impaired loans. An increase in NPLs generally indicates inefficiency in lending. As the ratio of NPLs to gross loans increase, the provisions for loan losses also rise. NPLs are discussed further in subsequent chapters. KPMG (1998) states that as a rule of thumb, a level of NPLs approaching 7 percent is considered very high which may highlight very low prospects of success (KPMG, 1998).

The quality of bank loans, as indicated by the share of NPLs to gross loans decreased significantly between 2001 and 2006 as shown in Figure 2.18. However, the ratio began to worsen after 2007, rising to 3.9 percent in 2008 and 5.9 percent in 2009. Bank non-performing loans to total gross loans in South Africa were 5.7 percent as of 2011. Its highest value over the past 10 years was 5.9 percent in 2009 while its lowest value was 1.1 percent in 2006. From a lowest record of 1.1 in 2006, the NPL ratio rose sharply likely reflecting the onset of the global financial crisis. The highest record in 2009 may have been due to the global recession that placed a huge financial strain on consumers resulting in a large volume of defaulted loan obligations. However, following 2009 Figure 2.18 indicates that the value began decreasing in each of the subsequent years until 2011.

FIGURE 2.19: GROSS LOANS AND ADVANCES FOR SOUTH AFRICA`S BANKING SECTOR



Source: Computed using data from SARB: www.quantec.co.za

According to SARB Economic report (2011), bank`s loans and advances extended to the private sector first decelerated and then stagnated before progressing into a sustainable growth trend in the aftermath of South Africa's first recession in nearly two decades (see Figure 2.19). This improvement in the positive growth of gross loans and advances is attributed to the combined effect of improved interest rates which improved growth in real income, and the progressive easing of lending criteria of banks.

2.6 CONCLUSION

In this chapter we have given a general overview of South Africa`s banking structure in terms of the number of entities, shareholding, asset growth, private sector credit growth, market concentration and discussed the main policy initiatives that the government initiated since independence to expand banking services to all south Africans. The chapter has noted that the banking sector has remained well developed and well regulated and comparable to the banking industry of developed economies. It is generally viewed as highly sophisticated, with good technology, infrastructure and sufficient capital levels. The prudential regulatory system is believed to have cushioned the sector from the disastrous impact of the 2008 financial crisis. However, the concentration level of the banking industry emerged as a cause of concern with the big four representing over 84 percent of the total banking sector assets.

In this chapter, the main rationale of this thesis was argued: that due to the socio-economic profile of the low-income clientele, the banking institutions find the impetus to serve the up-market with greater profit potential than serving the down-market. It was noted that serving low-income clients resulted in a smaller contribution because the nature, structure and core business of commercial banks is not geared to serve this market. Moreover the product offering that most commercial banks provide fails to address the specific requirements or needs of the poor. This therefore is the source of argument that there is a possibility that an inverse relationship exist between enhancing bank efficiency and ensuring access to banking services. Hence this chapter has shown that, like the rich, the poor do have demand for banking services. They need financial services particularly banking services to enable them to smooth consumption, build and accumulate assets, manage risk and to take advantage of opportunities as they come.

Selected key performance indicators were also discussed to assess the current soundness of the banking system such as trends in profitability, capital levels, non-performing loans and banking access indicators. The chapter concluded that the banking industry is performing well overall but raised concern that there is still need to accelerate government efforts to ensure widespread access and usage of banking services by all South Africans. For example it has been noted that 62.8 percent of South African adults have or use bank products and services and that 71.2 percent of urban adults are banked compared to 46 percent of rural adults. This result confirmed the fact that the rural poor and the marginalised constitute the majority of South Africans who are unbanked.

The next chapter provides a theoretical framework and literature review on the conceptualisation of efficiency and productivity.

CHAPTER THREE

LITERATURE REVIEW

INTRODUCTION

This chapter provides a comprehensive review of both the theoretical literature and empirical studies underpinning banking productivity and efficiency. The theoretical section conducts a review of the literature on productivity analysis and efficiency measurement. The empirical section examines previous researches undertaken by distinguished researchers regarding banking efficiency and productivity and its macro-links with various aspects of the economy. In the wake of incidences of financial crises in recent years namely, the 2007/2008 US financial collapse and the most current 2012-2013 Cypriot banking crisis, the empirical section also discusses the underlying factors that culminated in the US sub-prime financial crisis.

THEORETICAL FRAMEWORK

3.1 CONCEPTUALISING EFFICIENCY

Efficiency measurement consist of two basic frontier approaches namely, the production frontier approach and the cost frontier approach. Thus efficiency can be measured by analysing the boundary or frontier of a cost or production function. The duality principle confirms this connection. For example, given that observed production cannot exceed its potential level or maximum possible, relative inefficiency would be captured by the amount by which observed production falls short of the production frontier. Likewise it's not feasible to achieve costs below the minimum or cost frontier. However observed cost can be equal or greater than the minimum cost frontier. Relative cost inefficiency would be measured by the amount of excess of observed costs above the minimum or cost frontier.

3.1.1 The Production function Approach

The conception of the production function as a frontier originally began with Farrell (1957). A production function is simply a process that involves the conversion of inputs into outputs. For instance, a bank is a financial intermediary institution that converts deposit funds into loans or interest income. Thus, a production function essentially represents efficient

transformation possibilities that satisfy a set of constraints. A production function exhibiting inefficiency can be stated in inequality form as:

$$y_i \leq f(X_i; \alpha) \qquad (1)$$

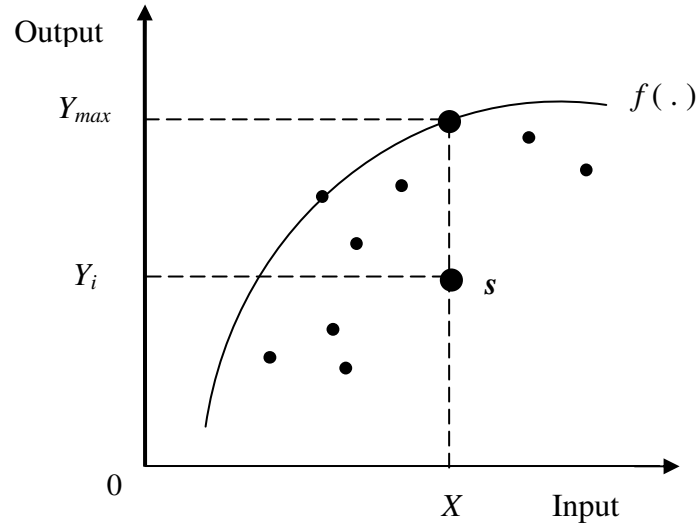
Where y_i denotes observed output at bank *i*, and X_i is a vector of inputs and α a vector of parameters which describe the transformation process. $f(.)$ is the production function and has the interpretation of the efficiency frontier or y_{max}. An inefficient bank would imply that potential output $(y_{\max})$ is greater than observed performance (y_i). Hence technical inefficiency entails that $(y_i - y_{max})$ is negative. The residual (ε_i) can be regarded as the variation between potential and observed performance and can be regarded as an indication of inefficiency.

$$\varepsilon_i = y_i - y_{max} \qquad (2)$$

The residual ε_i is strictly non-positive to guarantee that observed output is not greater than potential. In other words, $y_i > y_{\max}$ is not possible because it is impossible to achieve output greater than the maximum output implied by the production frontier. In Figure 3.1 below, decision making unit *i* (DMU_i[4]) is producing output (y_i) using input allocation *OX*. However, with input resources *OX* an efficient bank has the potential to produce maximum output $(y_{\max})$ as can been seen clearly on the diagram. The difference between actual and potential output, ε_i is negative therefore production at unit *i* is relatively inefficient. Ideally, the efficiency residual should be equal 0 for the production unit to be efficient since actual and potential outputs will be equal.

[4] Charnes, Cooper and Rhodes introduced the term "decision making units" (or DMU) which is now widely used in literature to denote an entity for transforming inputs into outputs.

FIGURE 3.1: EFFICIENCY AND THE PRODUCTION FRONTIER



Source: Ganley and Cubbin, 1992

3.1.2 The Cost function Approach

Ganley and Cubbin (1992) states that the duality principle between production and cost necessitates the existence of a cost function that relates to the production function specified in (1). A cost function connects the minimised total cost of a bank to its output and factor prices. Since it is possible that observed costs can exceed the minimum cost possible, the cost function may be written as an inequality:

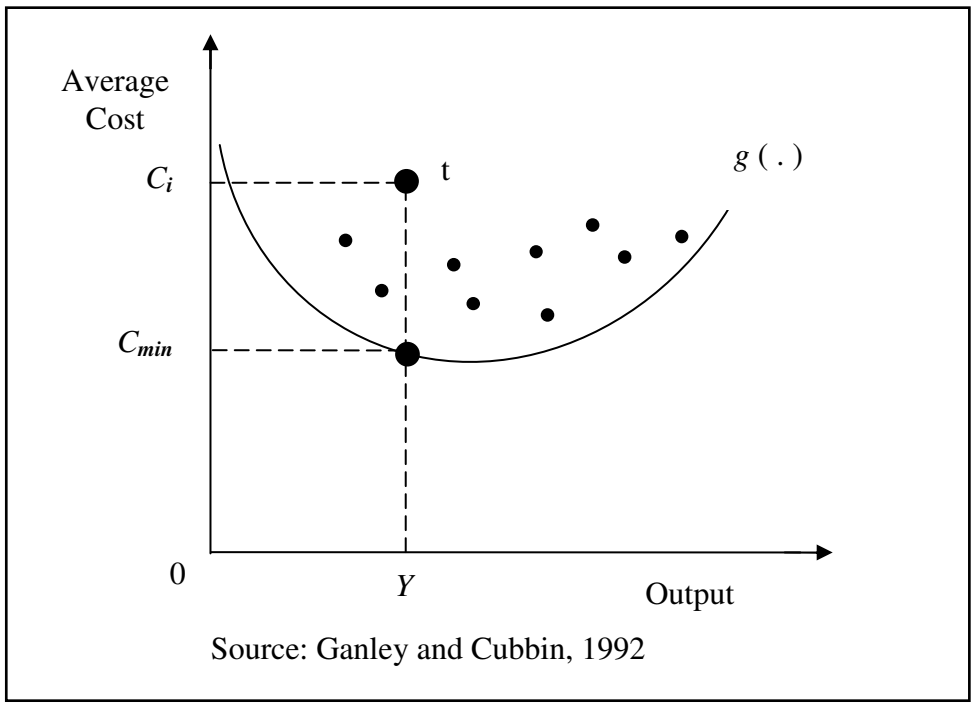
$$c_i \geq g(z_i; \beta). \qquad (3)$$

Where c_i represents average bank cost at bank *i*, z_i are contributing factors of costs, α is a vector of parameters and $g(.)$ defines the minimum costs $(c_{\min})$. Efficiency in ratio form is denoted by θ_i in (4) below as follows:

$$\theta_i = \frac{c_i}{g(z_i; \beta)} \qquad (4)$$

In the presence of inefficiency, observed costs are greater than the minimum costs and the efficiency residuals are positive. This means that the efficiency ratio is greater than unity. Therefore, the amount in excess of unity captures inefficiency. In the absence of inefficiency, observed costs and minimum costs are the same.

FIGURE 3.2: EFFICIENCY AND THE COST FRONTIER



Source: Ganley and Cubbin, 1992

Figure 3.2 depicts an analysis of efficiency using the cost frontier approach, where observed costs, c_i at unit i are greater than the minimum costs on the appropriate part of the frontier. Since frontier costs are the minimum feasible, observed costs cannot fall below minimum costs, i.e. $c_i \geq c_{\min}$.

3.2 CONCEPTS AND DEFINITIONS OF EFFICIENCY

3.2.1 Technical and Allocative Efficiency

Farrell (1957) who is considered to have pioneered and contributed a great deal to efficiency literature states that efficiency of a firm comprises two components allocative and technical efficiency. A firm is considered as technically efficient if the greatest output amount is achieved with the minimum input amount or if its utilisation of input resources generates maximum output. Allocative or price efficiency attempts to capture the degree to which a firm employs its input factors in their optimum proportion given input prices and the transformation technology (Coelli, 1996). It entails selecting among the different technically efficient input combinations that generates the highest achievable outputs. A combination of both technical and allocative efficiency captures total economic efficiency or overall efficiency.

3.2.2 Cost and Profit Efficiency

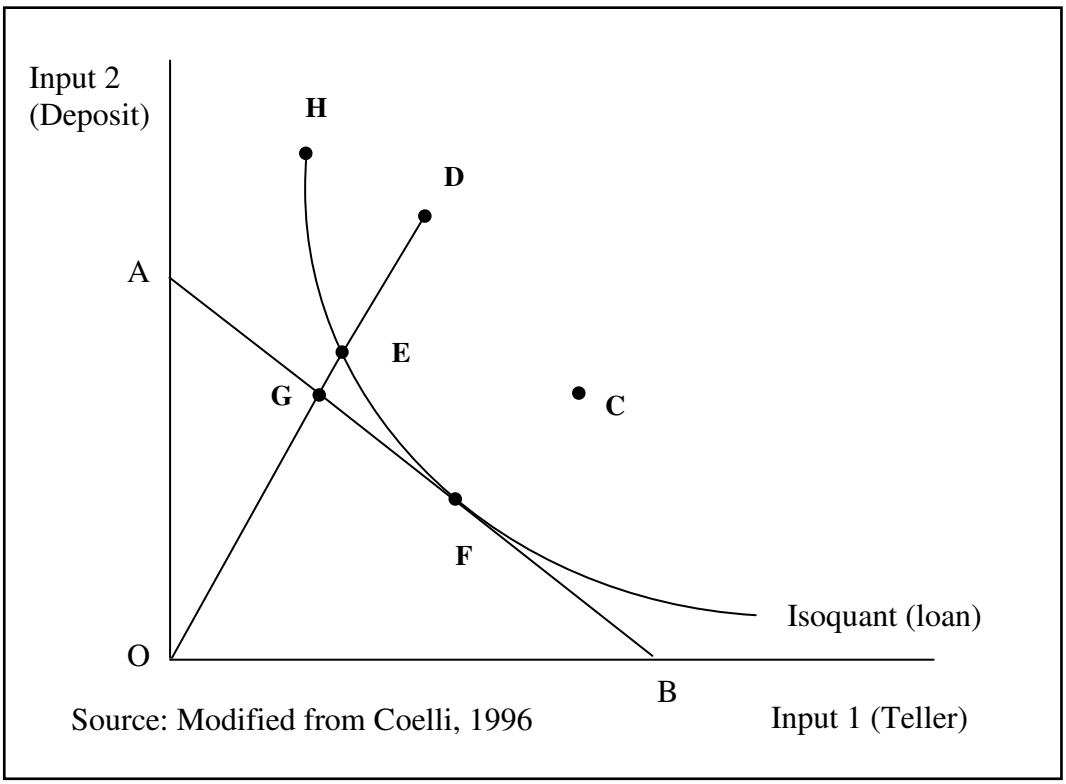
Depending on the researcher`s viewpoint and purpose of analysis, a choice can be made between profit and cost efficiency. Cost efficiency is a measure of the extent to which a bank's actual cost varies with that of the most efficient bank for generating an identical output vector under the same conditions of operation. Thus a bank is considered inefficient if it has more costs than the most efficient bank. Similarly, profit efficiency measures the ability of a bank in generating the maximum attainable profit given input prices and outputs. Thus a given bank is categorised as inefficient if its profits are less compared to the profits of the best-practice bank.

Again, depending on whether one has control over inputs or outputs, an output or input orientation can be specified. An input orientation measures input reductions that are required for a production entity to attain full efficiency without any output reductions. Input inefficiencies show the extent to which inputs must be decreased for the inefficient bank to lie on the efficiency boundary or frontier. An output orientation measures the expansion of output that is needed in order to achieve full efficiency without altering the inputs requirements. Thus output inefficiencies represent the required output increase necessary for the inefficient bank to become 100 percent efficient.

In Figure 3.3 below, a bank produces its output (loan or interest income) using a combination of two inputs (Labour and Deposits). A technically efficient bank is one that is located on the isoquant, that is, on the frontier such as **E, F** and **H**. Banks located at points **C** and **D** are technically inefficient. The measure of technical efficiency (TE) for a bank at point D is given as;

$$TE_D = \frac{OE}{OD}$$

FIGURE 3.3: FARRELL'S MEASUREMENT OF TECHNICAL AND ALLOCATIVE EFFICIENCY



Source: Modified from Coelli, 1996

This denotes the ratio of the least input requirement to actual input utilisation, given the input mix used by the bank at point **D**. The ratio **ED/OD** represents the percentage by which all inputs could be reduced without a reduction in output. If the bank at point **D** is to be efficient it has to relocate itself to point **E**. Technical efficiency assumes values ranging from 0 to 1. Given input prices, the isocost line **AB** represents the lowest cost of generating one unit of output. Allocative efficiency demands that production takes place at the point where the isoquant line is tangential to the isocost line. Given this definition, banks operating at points **E** and **H** are allocatively inefficient while technically efficient. The only bank that is exhibiting both allocative and technical efficiency is located at point D. The allocative efficiency of the bank operating at point **D** is given as:

$$AE_D = \frac{OG}{OE}$$

The ratio GE/OE denotes the percentage decrease in production costs that would be realised if production were to take place at the allocatively efficient point **F**. Farrell (Coelli, 1996) proposed that economic efficiency (EE) is measured as:

$$EE_D = \frac{OG}{OD}$$

The overall (economic) efficiency (*EE)* has the advantage that it easily decomposes into technical and allocative efficiencies.

$$\frac{OG}{OD} = \frac{OE}{OD} \times \frac{OG}{OE} \quad \text{That is, } EE = TE \times AE$$

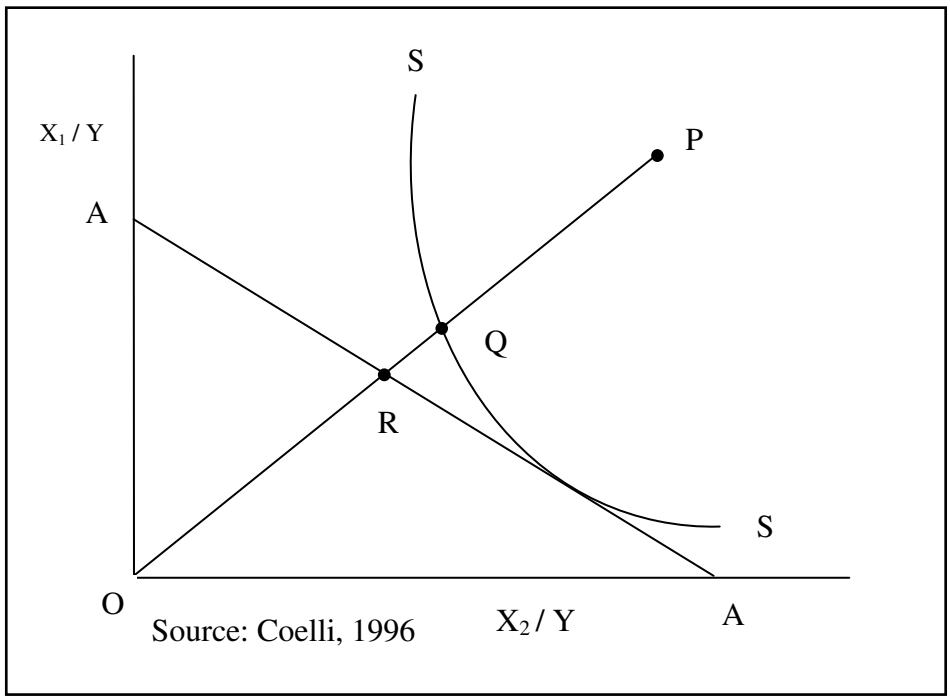
The measures obtained from Figure 3.3 represent input-oriented measures of efficiency. They are input-oriented as their focus is on the measurement of variations in input use between different banks for a standardised output.

3.2.3. Input-output efficiency measurement

3.2.3.1. Input-oriented measure

An input orientation approach measures the required input reductions that are needed for a production entity to achieve efficiency while holding output constant. In other words, input inefficiencies indicate the extent to which inputs must be decreased for the inefficient bank to lie on the efficiency frontier. Suppose a bank uses quantities, denoted by point P to generate a unit of output, the technical inefficiency of the bank can be stated by the ratio $\frac{QP}{OP}$ which represents the proportion by which all inputs could be reduced. Ideally the technical efficiency of the same bank operating at point P can be measured by the ratio: $TE = \frac{OQ}{OP}$ which is equal to $\left[1 - \frac{QP}{OP}\right]$. Efficiency values assume values between zero and unity (1). An efficient value of 1 indicates a fully technically efficient bank since it would be lying on the efficient isoquant SS.

FIGURE 3.4.: FARRELL'S EFFICIENCY MEASUREMENT (INPUT ORIENTATION)



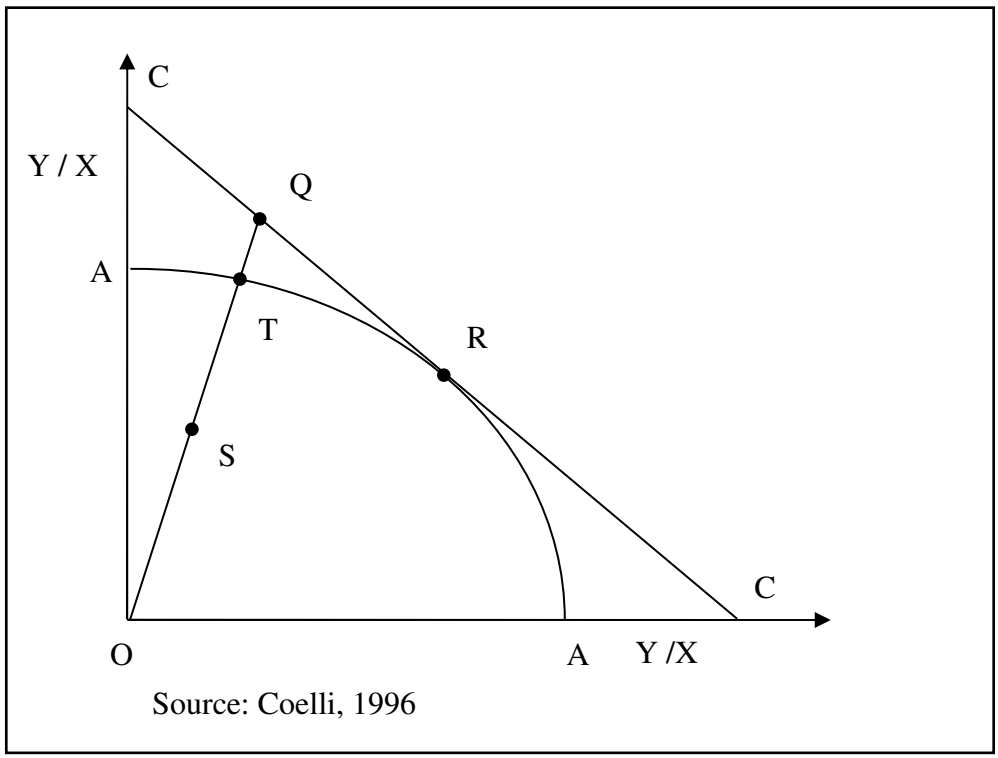
Source: Coelli, 1996

3.2.3.2. Output-oriented measure

An output orientation measures the needed increase in output required for an inefficient DMU to attain full efficiency without increasing any input usage. In Figure 3.5 below, the curve AA represent the maximum possible output attainable given the resources available. All points located inside the curve are technically inefficient relative to points on the frontier. For example, the distance ST represents technical inefficiency since it shows how far the point S is from achieving the maximum possible output with the given resources. At point S resources are being underutilised. Thus technical efficiency at point S is:

$$TE_S = \frac{OS}{OT}$$

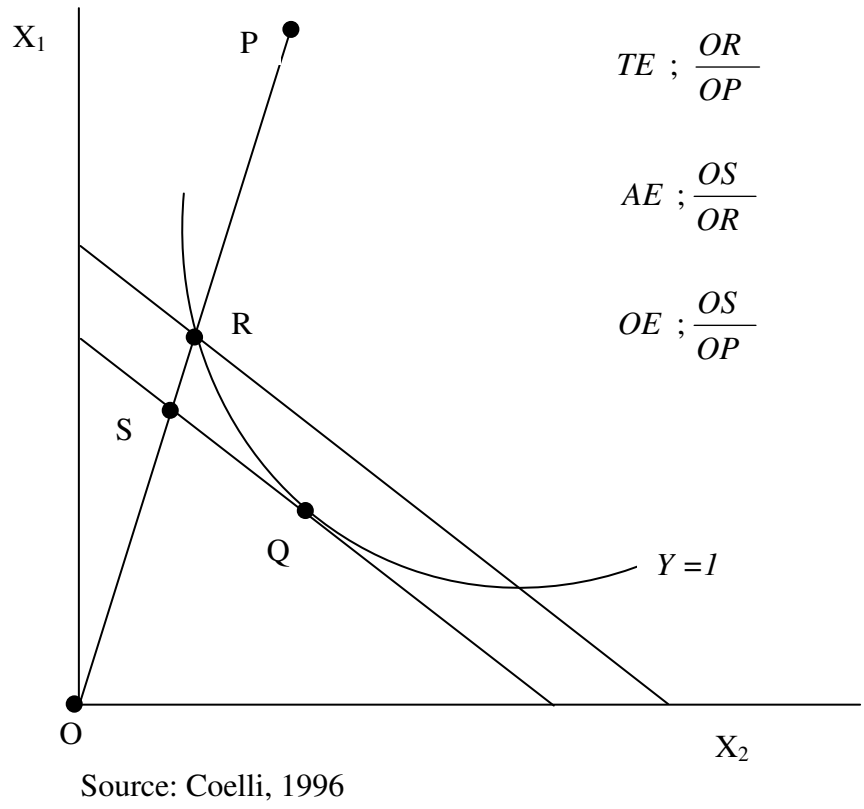
FIGURE 3.5: FARRELL'S EFFICIENCY MEASUREMENT (OUTPUT ORIENTATION)



Source: Coelli, 1996

Figure 3.6 depicts Farrell`s classic framework which enables the decomposition of overall efficiency into allocative and technical efficiency in which two inputs $\mathbf{X_1}$ and $\mathbf{X_2}$ are used to generate a single output **Y**. **Y=1** illustrates an efficient isoquant which shows all technically efficient combinations of X_1 and X_2 which generate Y. A DMU operating at Point P is inefficient since it is employing more input factors to produce Y. The degree of technical efficiency can be represented in ratio form as the proportion of optimum (best) and actual resource usage (OR/OP). Allocative efficiency can be analysed by taking into account the isocost line which represents the relative factor prices. All points lying on the isoquant Y=1 are technically efficient. However Q is the only point at which technical efficiency is achieved at the lowest cost. Hence a DMU operating at point Q is both allocatively and technically efficient. Allocative efficiency is stated as a ratio of minimum and actual cost (OS/OR) while overall efficiency *(OE)* is the product of technical and allocative efficiency.

FIGURE 3.6: CLASSIC FRAMEWORK OF EFFICIENCY BY FARRELL



Source: Coelli, 1996

3.3 EFFICIENCY MEASUREMENT

Measuring an organisation's efficiency is essentially about evaluating the relationship between the outputs it produces and the inputs it employs. Various techniques for evaluating efficiency of organisations or DMUs are classified into parametric and non-parametric. Efficiency can be estimated either as a *deterministic* frontier, or as a *stochastic* frontier. Under the deterministic method which is essentially a non-parametric approach, all variations from the efficiency frontier are attributed to inefficiency. The main non-parametric methods include Data Envelopment Analysis (DEA) and the Free Disposable Hull (FDH). On the contrary, parametric (stochastic) techniques assume that some deviations from the efficiency boundary are attributed partly to random elements or exogenous factors and partly to inefficiency. Mullineux and Murinde (2003) state that, the major drawback with parametric frontier approaches is that they demand identification of the specific functional form regarding either the production or cost function. As a result the accuracy of the generated efficiency measures depends on the precision of the specified functional form in approximating the correct underlying cost or production function. Three common parametric

methods include the Thick Frontier Approach (TFA), Distribution Free Approach (DFA) and the Stochastic Frontier Approach (SFA),

3.3.1 Parametric Methods

3.3.1.1 Stochastic Frontier Approach (SFA)

The parametric Stochastic Frontier Approach is founded on econometric techniques. Kablan (2010) states that by specifying a Cobb-Douglas, CES or trans-logarithmic function, the SFA essentially estimates the objective frontier function which may take the form of a cost or production function. Within this econometric approach any deviation from a specified technology is measured by a disturbance term which consists of two components, one accounting for noise or randomness and the other representing inefficiency. In inefficiency measurement the major challenge is to isolate actual inefficient behaviour from random elements that may influence performance. Hence the SFA recognises that any given firm`s cost may diverge from the boundary or minimum cost due to random variations or inefficiency. This property has made the parametric SFA to become more attractive compared to non-parametric approaches. According to Aigner et al (1977) the Stochastic Frontier Production function can be stated as:

$$y_i = x_i\beta + (v_i - \mu_i)$$

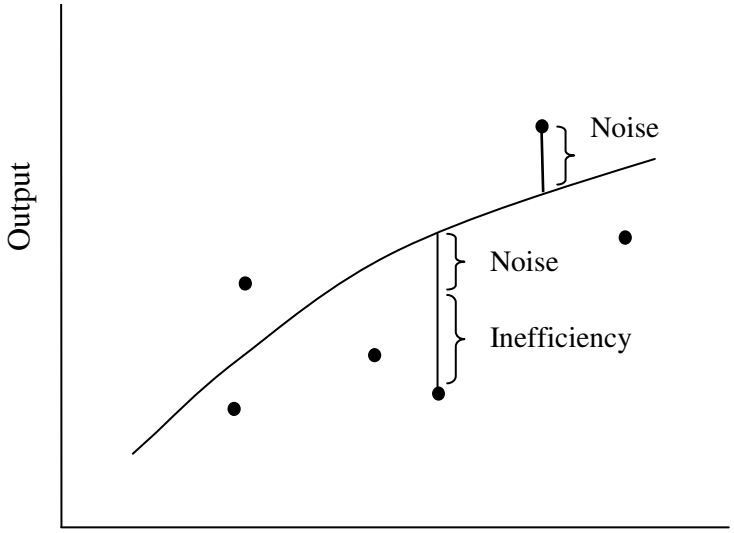
μ_i represent non-negative random variables and so account for inefficiency.
v_i represent random variables which are assumed to be *i.i.d*[5] N $(0, \delta^2)$ and independent of μ_i.
β *is* a vector of unknown parameters.
x_i is a vector of input quantities of unit *i*.
y_i is the average maximum output.

All elements beyond the control of the firm are captured by the disturbance or error term. These elements include unmanageable factors that affect the production function namely, diverse operating environments and econometric factors such as measurement errors or misspecification of the production function. The SFA requires the sample size to be sufficiently large to avoid problems of degrees of freedom. The distance between the observation and the estimated function captures inefficiency. However, the attractiveness of

[5] Independent and identically distributed.

this approach is its ability to isolate measurement errors, noise and external shocks outside the control of a production entity. The main disadvantage is that the reliability and precision of the obtained efficiency estimates depends on the accuracy of the selected functional form in representing the actual underlying production or cost function.

FIGURE 3.7: STOCHASTIC FRONTIER APPROACH



Source: Coelho, 2007

3.3.1.2 Distribution Free Approach (DFA)

The development of the DFA is attributed to Schmidt and Sickles (1984) and Berger (1993) and serves as an alternative to the conventional stochastic frontier technique especially when several years of data are available. While the SFA makes restrictive assumptions regarding the distribution of the error term, the Distribution Free Approach (DFA) makes no assumptions concerning the error term. Troutt et al (2005) state that the DFA is distribution free and completely data driven and demonstrate that such a model is capable of producing meaningful parameter estimates without making restrictive distributional assumptions. Mullineux and Murinde (2003) state that the v_{it} component of the composite error term discussed earlier under the parametric SFA framework is random and would be expected to average out to zero over time. Berger (1993) recommends a period of five years to be generally appropriate for allowing the errors to average out. Koutsomanoli-Filippaki (2007) cautions that if the selected period is too short, the random errors might not average out thereby over-estimating inefficiency. Again, if the selected period is too long, the bank`s

average efficiency may fluctuate over the period due to variations in environmental conditions rendering the derived efficiency measures meaningless. Assuming that these random errors indeed average out this process eventually produce average levels of efficiency across the sample of DMUs (Mullineux and Murinde, 2003). Therefore, these efficiency measures are generally standardized in comparison to the best-practice DMU in the sample.

3.3.1.3 The Thick Frontier Approach (TFA)

The Thick Frontier Approach (TFA) which is not commonly applied particularly in banking was developed by Berger and Humphrey (1991). Lang and Welzel (1998) claim that the TFA has the characteristics of both the non-parametric DEA and the parametric SFA. However, Bauer et al (1998) maintains that the TFA and the SFA adopts the same functional form for the cost function. Bauer et al (1998) states that the difference comes about in that the TFA applies regression analysis that is estimated using only those DMUs that form the lowest average cost quartile in each size category. These DMUs are considered to possess better than average efficiency measures and as such define a thick frontier of efficient DMUs. Likewise, DMUs in the highest average cost quartile are regarded as below average efficiency performers. The procedure under the TFA framework is that the DMUs to be evaluated are initially placed into quartiles according to their sizes. Their average cost over the time period is calculated. Then only those banks in each class size whose average cost is lowest define the data subset that is applied in the estimation of the thick frontier. The classification is done to ensure that an equivalent number of banks of all size categories are taken into account. The differences in error terms *within* the highest and lowest quartiles are taken to reflect random error. Whereas the predicted cost differences *between* the highest and lowest quartiles are assumed to capture inefficiencies and external differences in input prices and output quantities (Bauer et al, 1991).

Efficiency results generated by TFA have been regarded with suspicion since they are based on somewhat subjective assumptions. Bauer et al (1998) criticises the fact that the lowermost average cost quartile for all class sizes is taken to be a sufficient thick frontier of efficient DMUs. In addition, Bauer et al (1998) also point out that by estimating the difference between the highest and lowest quartile, the TFA only provides an indication of the broader level of total efficiency and not point efficiency estimates for each DMU. It is therefore essential to ascertain estimates of efficiency for each DMU for each period in order to enable

comparison with other frontier efficiency methods. Bauer et al (1998) suggests adjustments to be done to the TFA approach to facilitate this objective.

3.3.2 Non-Parametric Methods

3.3.2.1 Data Envelopment Analysis (DEA)

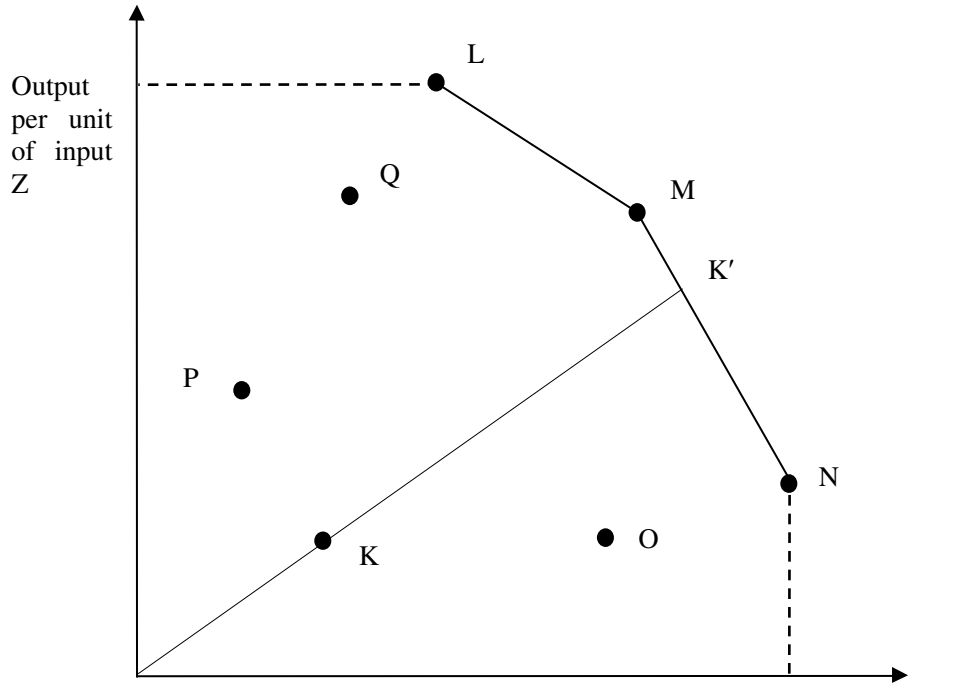
Data Envelopment Analysis is a non-parametric technique that assesses the performance of similar production or service entities known as decision making units (DMUs) which transform many inputs into many outputs (Cooper et al, 2011). The linear programming DEA technique was originally developed by Farrell (1957) and further modified by Charnes et al (1978) to estimate the efficiency of non-profit entities. The original intention was for application within the non-profit sector particularly government establishments such public departments, public hospitals and schools. However, since the DEA origination many improvements have been done to adapt it for various uses as well as application in profit-oriented businesses. In recent years, DEA has proved to be a popular tool of performance evaluation with wide application in banking, universities, hospitals and other service industries.

DEA is a non-stochastic technique that is used to provide an overall evaluation of technical and allocative[6] efficiency for a multiple-input-output firm (Coelli, 1996). The DEA model which was later developed by Charnes et al (1978) was input oriented and had constant returns to scale (CRS) technology. However, subsequent developments to the CRS model by Banker et al (1984) gave origin to a more representative variable returns to scale (VRS) model. Diewert and Parkan (1983, p.1) state that the DEA technique is considered *non-parametric* due to the fact that it makes no assumptions regarding the underlying technology for example that it "belongs to a certain class of specific functional form which depends on a finite number of parameters such as the well-known Cobb-Douglas functional form". Bowlin (1986) has argued that the non-parametric nature of DEA is an important property because the functional relationships underlying production of public entities may be abnormally complex and difficult to specify. Again, the DEA is classified as *non-stochastic* because it makes no restrictive assumptions regarding the probability distribution of errors. It constructs

[6] If price data on inputs and outputs is available, one can compute allocative efficiency measures using the DEA approach.

a piecewise "envelope" or boundary (frontier) such that all observed data points lay either on or below this boundary (Coelli, 1996). Therefore points lying on the boundary are regarded as fully efficient whereas points below the efficiency boundary are considered inefficient. The degree of input or output inefficiency is then captured by the vertical distance from the frontier. This is demonstrated by the illustration in Figure 3.8 below which depicts a DEA model with a solid line connecting the efficient DMUs *L, M* and *N* that represent achieved efficiency. The rest of the inefficient DMU such as *P, Q, K*, and *O* are then evaluated relative to the constructed efficiency frontier. For instance, DMU *K* is classified as inefficient and as such needs to shift to point *K′* on the efficiency boundary to be classified as efficient.

FIGURE 3.8: A DEA MODEL SHOWING AN EFFICIENCY FRONTIER



Source, Avkiran (1999)

It is essential to state that DEA measures relative and not absolute efficiency. DEA helps the researcher to determine the following:

i) The most productive DMUs which define best practice.
ii) The less-productive DMUs relative to the best-practice DMUs.
iii) The quantity of wastage by each of the less-productive DMUs
iv) The quantity of excess capacity or the capacity to increase outputs in less-productive entities while holding input resources constant.

The non-parametric DEA offers certain benefits over the parametric SFA: Firstly, It does not require the specification of the functional form that relates inputs and outputs as in standard estimations of regressions and does not also require the decision maker to express their own weighting structure for inputs and outputs. Secondly, the DEA approach is unit invariant. In normal circumstances the DEA relative efficiency coefficient is derived unaffected by units of measurements in the underlying data. According to Coelli (1996, p. 23) "changing the unit of measurement e.g. measuring quantity of labour in person hours instead of person years will not change the value of the efficiency measure." Thirdly, DEA approach can handle multiple outputs much better than parametric models of production, a useful property for analysing banking sector efficiency. Production within the banking industry is characterised by multiple inputs and multiple outputs. Finally, the DEA technique does not require price data for inputs and outputs which are difficult to obtain for most public institutions such as the banking service industry. The DEA approach requires only data for inputs and outputs. Despite the aforementioned advantages, DEA is not without its drawbacks. Many researchers have largely criticised the DEA technique for interpreting every deviation from the efficiency boundary as being inefficiency. By so doing DEA fails to account for randomness or exogenous factors that may affect organisational performance but which are not necessarily a reflection of inefficient behaviour. In recent years, two main methods that have grown popular in performance or efficiency evaluation are DEA and SFA. Table 3.1 contains a comparative analysis of DEA and SFA.

TABLE 3.1: A COMPARISON OF DEA AND SFA**.**

Category	DEA	SFA
Description	A non-stochastic method that constructs a nonparametric production frontier by fitting a piece-wise linear surface over the data points.	A parametric method that estimates a production frontier of the form: $y = f(x) + v - u$, where y is the output f(x) are input factors, v is a disturbance term that represents random shocks and u captures inefficiency. One can also estimate a cost frontier function.
Data Requirements	It only requires quantity data for both outputs and inputs for a sample of firms. If price data are available, one can use it to calculate allocative efficiency.	For a production function or distance function it requires quantity data on inputs and outputs for a sample of firms preferably over a number of periods. For a long-run cost frontier data on total costs, input prices, and output quantities is required. For a short-run cost frontier data on variable costs, variable input prices, and fixed input quantities and output quantities is needed.
Advantage	Identifies a set of peer DMUs (efficient with similar input and output mixes) for every inefficient DMU. Designed to evaluate DMUs with multiple inputs and multiple outputs. It does not require specification of a functional form for the production function. It does not complicate the analysis by making assumptions regarding the distributional properties of the inefficiency error term.	Attempts to account for noise. Environmental variables are easier to deal with. Standard statistical tests can be performed It's relatively easier to detect extreme data points or outliers.

Drawbacks	DEA can be affected by noise. Standard econometric tests are not applicable. It requires a sufficiently large sample size for valid and reliable estimates.	The breakdown of the error term into noise and efficiency may be affected by the specified distributional form.

The subsequent section discusses the non-parametric and deterministic Free Disposable Hull Analysis.

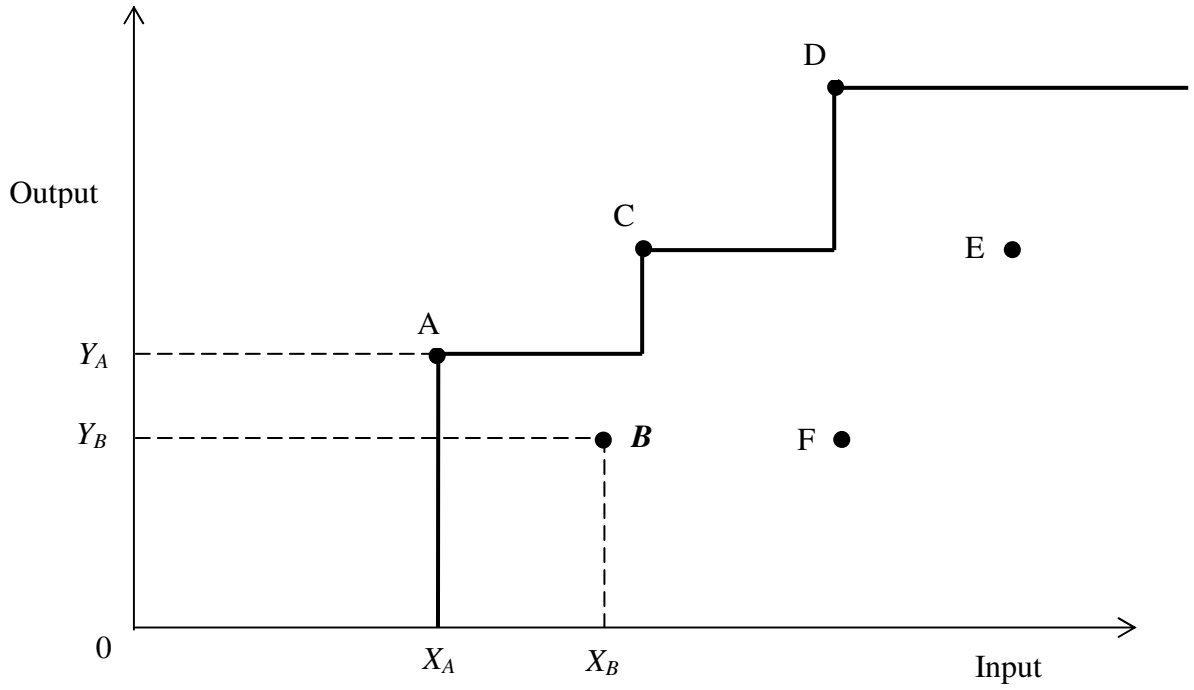
3.3.2.2 Free Disposable Hull Analysis (FDH)

The FDH was originally proposed by Deprins et al (1984) to serve as an alternative to the popular DEA approach. The FDH differs from the DEA in that it does not require convexity of the underlying technology, but maintains restriction of strong disposability of inputs and outputs (Walden and Tomberlin, 2010). Hence, the FDH approach can be considered to be a special case of the DEA approach with the least restrictions. Cherchye et al (2001) regards non-convexity as an attractive property of FDH on the basis that in efficiency analysis, one hardly finds any solid theoretical or empirical justification for proposing convex production sets. In addition, the convexity assumption has been criticised on the premise that the convex DEA model compares an inefficient DMU to an unobservable, non-existent and fictitious linear combination of efficient DMUs (Henderson, 2003). On the contrary, the FDH is not susceptible to this critique. The FDH identifies a single dominating (efficient) DMU for each inefficient DMU.

In the terminology of FDH analysis, a DMU is efficient if it is not "dominated by any other DMU" (De Sousa and Schwengber, 2005, p. 3). Hence under the FDH approach, the efficient frontier is defined by non-dominated firms. Considering an efficiency evaluation for a panel of banks, bank A is said to dominate bank B (*i*) if there is no input for bank A that exceeds the corresponding input for bank B or (*ii*) if there is no output for bank A that is less than the corresponding output for bank B and (*iii*) if there is at least one input for bank A that is less than the corresponding one for bank B and (*iv*) if there is at least one output for bank A that exceeds the output for bank B (Walden and Tomberlin, 2010). Therefore, for each inefficient DMU located within the interior of an FDH set, an existing reference DMU can be identified that proves superior in all input or output dimensions (Cherchye et al, 2001). Dlouhy (2009)

has shown that efficiency scores derived using the FDH approaches are relatively higher than the DEA-derived efficiency scores due to the relaxation of the convexity assumption of the FDH approach. Hence, Thrall (1999) states that DEA efficiency implies FDH efficiency but FDH efficiency does not imply DEA efficiency. Henderson (2003) argues that because of the convexity assumptions DEA calls inefficient too many DMUs. This approach is illustrated in Figure 3.9 using a single input single output production possibility frontier for simplicity.

FIGURE 3.9: FREE DISPOSABLE HULL (FDH) PRODUCTION POSSIBILITY



Source: Herrera and Pang, 2005

Suppose that Bank A and bank B require input X_A and X_B in order to generate output Y_A and Y_B in that order. The input efficiency of bank B would be given by the proportion $\frac{X_A}{X_B}$ while the corresponding output efficiency score would be defined as the proportion $\frac{Y_B}{Y_A}$. An efficiency score of unity means that the bank is operating on the efficiency boundary and therefore efficient. From the above illustration, it is clear that Bank A is dominating bank B, as bank A uses less input compared to the corresponding input factor for bank B ($X_A < X_B$) and bank A is producing more output compared to the corresponding output for bank B ($Y_A >$

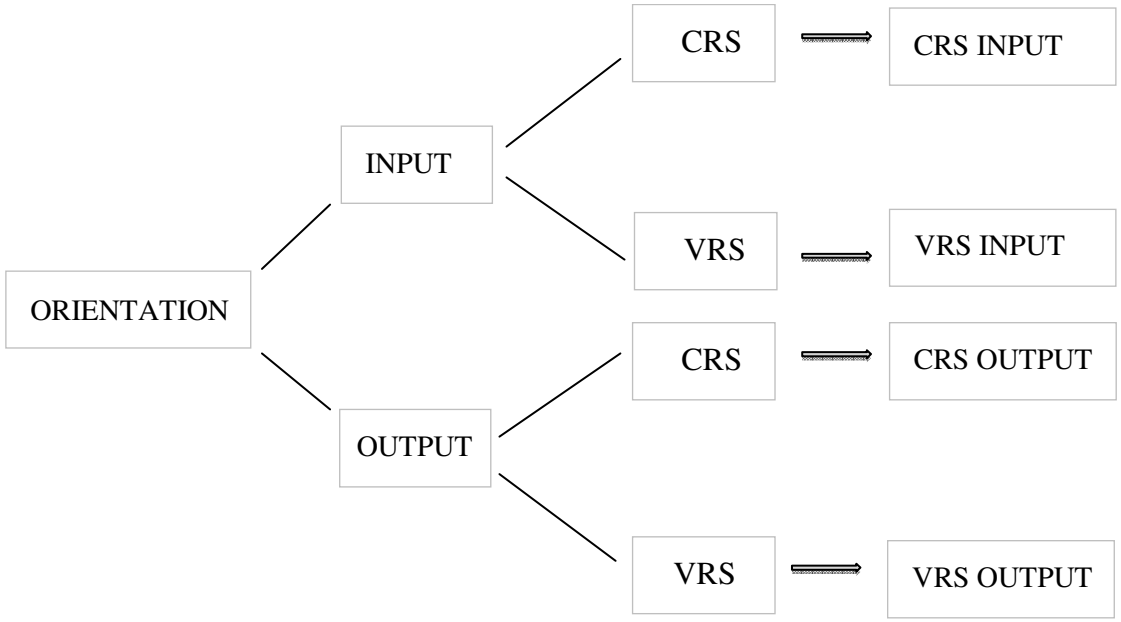
Y_B). Hence, bank A is superior in its utilisation of input and output production. Banks B is said to be non-dominating and therefore regarded as inefficient.

The major benefit of the FDH method is that for each inefficient DMU an existing best practice reference DMU is identified. However, the FDH is not without its disadvantages. Firstly, random or exogenous factors that might affect production are unaccounted for and therefore are included in the calculation of inefficiencies. Therefore similar the DEA approach, the deterministic FDH approach fails to separate deviations from the frontier into genuine inefficiency and random or exogenous factors. Secondly, the FDH has been criticised for allowing too many efficient DMUs or observations. Henderson (2003) states that a particular DMU with less amount of a certain input as well as output quantity that is significantly less relative to an efficient DMU may be considered efficient while the same DMU evaluated under the DEA framework would be deemed highly inefficient. Finally, it has been shown that this particular non-parametric FDH technique is sensitive to outliers thereby weakening its inferential power. In a nutshell, the distinction between DEA and FDH is that the FDH maintains free disposability while relaxes convexity, whereas the DEA assumes both the free disposability of resources and the convexity of the production set. Many empirical studies seem to favour the DEA approach compared to the FDH method. The literature behind the methodological framework underlying the DEA approach is detailed and presented in the subsequent section.

3.4 DEA MODELS

Two basic DEA type models may be adopted depending on the viewpoint of the researcher and the purpose of the study. These are the CRS model also known as the CCR model (after Charnes Cooper and Rhodes (1978)) and the VRS model also known as the BCC model (after Banker Charnes and Cooper (1984)). If the assumption is that economies to scale remain constant as DMU size increases then a CRS DEA model is appropriate compared to the VRS DEA model. Hence the CRS specification is only suitable when each DMU is operating at an optimum scale. However, this is not representative of reality. As a result the CRS model was further modified by Banker et al (1984) leading to a more flexible and realistic VRS model for use under variable returns to scale technologies. Coelli (1996) states that difference between CRS technical efficiency (TE) and VRS TE captures scale inefficiency. These model choices are portrayed in *Figure 3.10.*

FIGURE 3.10 BASIC DEA MODEL CLASSIFICATIONS



Source: Ozcan (2007)

In this chapter, we will also discuss two extensions to these basic models known as the Additive and Multiplicative model. While the Additive model permits both output and input orientations to be treated simultaneously in one model, a Multiplicative model replaces the usual linear piecewise frontier with a Cobb-Douglas piecewise or log-linear piecewise frontier in order to allow scale elasticities to be estimated.

3.4.1 The Basic CCR Model (1978)

3.4.1.1 The Dual Input Oriented CCR Model

The original CCR model by Farrell (1957) and Charnes et al (1978) was input-oriented and had constant returns to scale specification. In this analysis we will assume a sample of *n* DMUs using *m* diverse inputs to generate *s* diverse outputs. Furthermore, we assume that DMU_j employs quantity x_{ij} of input *i* and generates quantity y_{rj} of output *r* and that $y_{rj} \geq 0$ and $x_{ij} \geq 0$. Hence x_{ij} represents the actual quantity of input *i* employed by DMU_j while y_{rj} represents the actual quantity of output r generated by DMU_j. Within a multiple input multiple output context, efficiency is estimated by maximising weighted outputs to weighted inputs.

$$Efficiency = \frac{sum\ of\ \ weighted\ \ outputs}{sum\ of\ \ weighted\ \ inputs}$$

Symbolically, the above objective of the individual DMU to be evaluated can be expressed as a fractional programming (FP) maximising problem as formulation (**3.1**). Where *u* represents output weights and *v* represents input weights.

$$Max\ h_o(u,v) = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \quad \text{.....3.1}$$

Cooper et al (2011) states that the outputs to inputs ratio is used to evaluate relative efficiency of $DMU_j = DMU_o$ to be assessed compared to the ratios of all the $j = 1, 2, \ldots, n$ *DMUs*. In other words DMU_O is selected out of the j = 1..., n DMU`s for evaluation. Therefore, the fractional programming maximising problem for evaluating the relative efficiency of DMU_o under CRS is given by solving the following CCR model:

$$Max\ h_o(u,v) = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \quad \text{.....3.2}$$

$$\text{Subject to} \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \ for\ j = 1, \ldots, n$$

$$u_r, v_i \geq 0 \ for\ all\ i\ and\ r.$$

Where o is the DMU singled out for evaluation in the set of *j = 1,...,n DMU`s,*

y_{rj} (r = 1... s) – actual quantity of output r generated by bank *j*,

x_{ij} (i = 1…m) – actual quantity of input *i* generated by bank *j*,

u_r – weight assigned to output *r*,

v_i – weight assigned to input *i*.

Formulation **3.2** implies that the LP objective is to maximise output *y* while maintaining the condition that virtual output to virtual input ratio for all DMUs must equal to or less than one. The specified constraint indicates that each DMU in the sample is either on or below the efficiency boundary. Hence, the efficiency of all banks evaluated in this study have an upper bound of one. However, Cooper et al (2011) argue that the above formulation will yield an infinite number of solutions. Coelli (1996) also argues that if (*u**, *v**) is a solution, it follows that (α*u**, α*v**) is also a solution. The non-negativity condition for the *u* and *v* reflect the notion that all inputs and outputs have a non-zero value. Hence, in order to avoid this problem a constraint $\sum_{i-1}^{m} v_i x_{io} = 1$ is imposed. The resultant input oriented CCR model in the linear programming form becomes:

$$Maximize\ h_o = \sum_{r=1}^{s} \mu_r y_{ro} \quad \text{.....3.3}$$

$$Subject\ \ to \quad \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \upsilon_i x_{ij} \le 0, \qquad (j = 1, \cdots, n)$$

$$\sum_{i=1}^{m} \upsilon_i x_{io} = 1$$

$$\mu_r, \ge 0 \qquad (r = 1, \cdots, s),\ \upsilon_i, \ge 0 \qquad (i = 1, \cdots, m)$$

Coelli (1996) states that the notation changes in *u* and *v* in formulation **3.2** to μ and υ in formulation **3.3** represent the transformation referred to as the (dual) multiplier form of the LP problem. This first constraint ensures that the Pareto optimality conditions are satisfied since additional increases in this value can be achieved only if some of the output values y_{rj} are reduced or if some of the input values x_{ij} are increased. Cooper et al (2011, p.3) defines full efficiency as achievable by a particular DMU "if and only if none of its inputs or outputs can be improved without worsening some of its other inputs and outputs". This definition embraces the concept of Pareto and relative efficiency in the sense that a particular DMU is deemed fully efficient only if other performers' inputs and outputs are improvable without deteriorating some of its other inputs and outputs. It is important to recall that DEA measures relative efficiency as opposed to absolute efficiency. This is because each DMU is evaluated with reference to *j = 1,...,n DMUs.*

The second constraint shows that the weighted sum of inputs for a particular bank equals one. Weights u_r and v_i are considered as unknowns and their values are obtained in the LP solution. These weights are the variables of the problem and they determine which input a particular DMU is best in utilising or which output it is best in generating. The DMU is assigned higher rates to those inputs and output variables which it is more adept or best in utilising or in generating, and lower rates to others (Ganley and Cubbin, 1992). Sherman (1982) states that the values of the weights are chosen within the DEA model such that the DMU being evaluated achieves the highest efficiency ratio while maintaining the condition that an input-output ratio does not exceed unity for every DMU in the observation set. There are *n* DMU`s and hence the LP model must be run *n* times for every DMU.

The value h_0^* is an efficiency score which satisfies $0 \leq h_0^* \leq 1$ wherein $h_0^* = 1$ denotes 100 percent efficiency and $h_0^* < 1$ signifies the presence of inefficiency. The asterisk shows that the solution value from running the model is optimal. The deviation of the value of h_0^* from 1 $\left(1 - h^*\right)$ for each DMU captures the inefficiency score for each DMU.

3.4.1.2 The Primal Input Oriented CCR Model

Ganley and Cubbin (1992) state that, since the formulation in (**3.3**) is a dual (multiplier form) linear programming problem, it has a primal (envelopment) formulation. This primal formulation is constructed by minimising the amounts of *m* inputs required to meet specified levels of *s* outputs as follows:

Minimise: $$\theta - \varepsilon\left[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right] \quad \text{.....3.4}$$

Subject to: $$\theta x_{io} - S_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j, \qquad \left(i = 1, \ldots, m\right)$$

$$y_{ro} + S_r^+ = \sum_{j=1}^{n} y_{rj}\lambda_j, \quad \left(r = 1, \ldots, s\right)$$

$$\lambda_j \geq 0,\ j = 1 \ldots n \qquad \text{(Weights on branches)}$$

$$S_i^- \geq 0, i = 1 \ldots m \qquad \text{(Input slacks)}$$

$$S_r^+ \geq 0, r = 1 \ldots s \qquad \text{(Output slacks)}$$

In the above formulation, $\varepsilon > 0$ is a non-Archimedean component specified to be smaller than any positive real number. The notation λ_j denotes the output and inputs weights of other banks. The dual theorem of linear programming therefore implies that $h_0^* = \theta^*$. θ^* is the technical efficiency score to be computed for each DMU. Thus the solution to problem (**3.4**) is θ^*, the technical efficiency (TE) score. For a DMU_O to be defined as *fully* efficient (100%), it is a required condition to achieve both θ^*=1 and zero slacks $S_i^{-*} = S_r^{+*} = 0$ for all inputs and outputs. On the other hand, a DMU_O is said to be weakly efficient only if θ^*=1 and $S_i^{-*} \neq 0$ and/or $S_r^{+*} \neq 0$, for some inputs and outputs (Cooper et al, 2011). The * indicates optimal values of the variables. Hence both the primal and dual approaches provides an optimal solution for θ, $\theta^* = h^*$. This LP problem is applied *n* times once for each bank or DMU.

3.4.1.3 The Dual Output Oriented CCR Model

Cooper et al (2011) state that one can analyse efficiency from the output perspective where the objective is reoriented from maximising to minimising the virtual input to virtual output ratio where as usual $virtual\ input = \sum_{i=1}^{m} v_i x_{io}$ and $virtual\ output = \sum_{r=1}^{s} u_r y_{ro}$.

$$Minimise \quad \frac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{ro}} \qquad \text{.....3.5}$$

$$\text{Subject to} \quad \frac{\sum_{i=1}^{m} v_i x_{ij}}{\sum_{r=1}^{s} u_r y_{rj}} \geq 1 \quad for \ j = 1, \ldots, n$$

$$u_r, v_i \geq 0 \quad for\ all\ i\ and\ r.$$

As in the input-orientation context, the above output oriented formulation can be rewritten in the multiplier model form as:

$$Minimize \quad q = \sum_{i=1}^{m} v_i x_{io} \qquad \text{.....3.6}$$

$$\text{Subject to} \quad \sum_{i=1}^{m} \upsilon_i x_{ij} - \sum_{r=1}^{s} \mu_r y_{rj} \geq 0, \qquad (j = 1, \cdots, n)$$

$$\sum_{r=1}^{s} \mu_r y_{ro} = 1$$

$$\mu_r, \geq 0 \qquad (r = 1, \cdots, s), \ \upsilon_i, \geq 0 \qquad (i = 1, \cdots, m)$$

3.4.1.4 The Primal Output Oriented CCR Model

The corresponding primal LP form of **3.6** can be stated as:

Maximise: $$\phi + \varepsilon \left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+} \right] \quad \text{.....3.7}$$

Subject to: $$x_{io} - S_i^{-} = \sum_{j=1}^{n} x_{ij} \lambda_j, \qquad (i = 1, \ldots, m)$$

$$\phi y_{ro} + S_r^{+} = \sum_{j=1}^{n} y_{rj} \lambda_j, \quad (r = 1, \ldots, s)$$

$$\lambda_j \geq 0, \ j = 1 \ldots n$$

Similarly, a DMU_O is then defined to be fully efficient if and only if $\phi^* = 1$ and $S_i^{-*} = S_r^{+*} = 0$, for all i & r and weakly efficient if $\phi^* = 1$ and $S_i^{-*} \neq S_r^{+*} \neq 0$, for some i & r. Table 3.2 contains a summary comparison of both input and output orientations for the envelopment form and multiplier form of the CCR model. Cooper et al (2009) state, that the primal problem is sometimes referred to as the envelopment problem, while the dual problem is the multiplier problem. Ali and Seiford (1993) cautions that in order to avoid a mix-up of which one is dual or primal, the formulation involving virtual multipliers is always taken to be the multiplier (dual) form.

TABLE 3.2: A SUMMARY OF THE CCR DEA MODEL

INPUT-ORIENTATION	
Envelopment Model (Primal)	**Multiplier Model (Dual)**
Min $\theta - \varepsilon\left[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right]$	***Max*** $h_o = \sum_{r=1}^{s} \mu_r y_{ro}$
Subject to:	***Subject to:***
$\theta x_{io} - S_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j, \quad (i = 1, \ldots, m)$	$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} \upsilon_i x_{ij} \le 0,$
$y_{ro} + S_r^+ = \sum_{j=1}^{n} y_{rj}\lambda_j, \quad (r = 1, \ldots, s)$	$\sum_{i=1}^{m} \upsilon_i x_{io} = 1$
$\lambda_j \ge 0, \quad (j = 1 \ldots n)$	$\mu_r, \ge 0 \quad (r = 1, \cdots, s)$
	$\upsilon_i, \ge 0 \quad (i = 1, \cdots, m)$

OUTPUT-ORINTATION	
Envelopment Model (Primal)	**Multiplier Model (Dual)**
Max: $\phi + \varepsilon\left[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+\right]$	***Min*** $q = \sum_{i=1}^{m} \upsilon_i x_{io}$
Subject to:	***Subject to***:
$x_{io} - S_i^- = \sum_{j=1}^{n} x_{ij}\lambda_j, \quad (i = 1, \ldots, m)$	$\sum_{i=1}^{m} \upsilon_i x_{ij} - \sum_{r=1}^{s} \mu_r y_{rj} \ge 0,$
$\phi y_{ro} + S_r^+ = \sum_{j=1}^{n} y_{rj}\lambda_j, \quad (r = 1, \ldots, s)$	$\sum_{r=1}^{s} \mu_r y_{ro} = 1$
$\lambda_j \ge 0, \quad (j = 1 \ldots n)$	$\mu_r, \ge 0 \quad (r = 1, \cdots, s)$
	$\upsilon_i, \ge 0 \quad (i = 1, \cdots, m)$

Source: Cooper et al (2011)

3.5 EXTENSIONS TO THE BASIC CCR MODEL

3.5.1 The BCC Model (1984)

The CRS or CCR model discussed above was further developed by Banker et al (1984) to allow variable returns to scale that led to development of the VRS or BCC model. They demonstrated that the estimated measures obtained under both the CCR and BCC enabled the necessary decomposition of CCR efficiency into technical and scale efficiencies.

3.5.1.1 The Primal Input Oriented BCC Model

The major difference when comparing the CCR and the BCC model is embedded in the handling of returns to scale with the latter allowing for a more representative specification of variable returns. The BCC form is more flexible than the CCR version. The BCC model can be considered to be an extended version of the CCR model with an adjustment condition or additional constraint. Under the CRS (CCR) assumption, an increase in the amount of inputs consumed by a particular bank would lead to an equivalent increase in the amount of outputs it produces. In the case of a VRS technology, an increase in the amount of inputs consumed results in a more or less than equivalent increase in outputs produced. A DMU is designated CCR efficient if it attains both scale and technically efficiency. The same DMU under the BCC model only need to be technically efficient to be considered BCC efficient. This is because the BCC model measures pure technical efficiency.

Using the same setting as previously where we assumed n DMUs denoted by DMU_j $(j = 1, 2, \ldots, n\ DMUs)$ producing s different outputs, $y_{rj}\left(r = 1,2,\cdots,s\right)$ by consuming m different inputs, $x_{ij}\left(i = 1,2,\cdots,m\right)$. The efficiency of a particular bank denoted as DMU_O can be assessed using the BCC model of primal form as follows:

$$\textit{Minimise}: \quad \theta_o - \varepsilon\left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+}\right] \quad \text{.....3.8}$$

$$\textit{Subject to}: \quad \theta x_{io} - S_i^{-} = \sum_{j=1}^{n} x_{ij}\lambda_j, \qquad \left(i = 1, \ldots, m\right)$$

$$y_{ro} + S_r^{+} = \sum_{j=1}^{n} y_{rj}\lambda_j, \quad \left(r = 1, \ldots, s\right)$$

$$1 = \sum_{j}^{n} \lambda_j$$

$$\lambda_j \geq 0$$

$$S_i^- \geq 0,\ i = 1 \ldots m$$

$$S_r^+ \geq 0,\ r = 1 \ldots s$$

$\varepsilon > 0$ is a non-Archimedean component specified to be lesser than any positive real number. Zhu (2003) states that the model in **3.8** can be solved in a two-stage procedure. The first stage involves the process of achieving a maximal reduction in inputs via the optimal θ^* in **3.9** below as follows.

Minimise $\theta = \theta^*$ 3.9

Subject to: $\sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta x_{io} \qquad i = 1,2,\ldots,m,$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \leq y_{ro} \qquad i = 1,2,\ldots,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n$$

Under the input-oriented VRS model specified in **3.9**, the inputs are minimised while outputs are maintained at their present levels. In the second stage, we ignore the slacks and compute the value of θ^* first in **3.9** and then optimise the slacks by substituting θ^* in the LP problem stated in **3.10** below.

Maximise: $\left[\sum_{i=1}^{m} S_i^- + \sum_{r=1}^{s} S_r^+ \right]$ 3.10

Subject to: $\sum_{j=1}^{n} \lambda_j x_{ij} + S_i^- = \theta^* x_{io} \qquad i = 1,2,\ldots,m,$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - S_r^+ = y_{ro} \qquad i = 1,2,\ldots,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad\qquad j = 1,2,\ldots,n$$

A DMU is considered efficient only if $\theta^* = 1$ and $S_i^{-*} = S_r^{+*} = 0$ for all *i* and *r* and *weakly* efficient if $\theta^* = 1$ and $S_i^{-*} \neq 0$ and/or $S_r^{+*} \neq 0$ for some *i* and *r* (Zhu, 2003).

3.5.1.2 The Dual Input Oriented BCC Model

Banker et al (2004) show that the associated dual (multiplier) form of the BCC model specified in 4.8 can be stated as follows:

Maximise: $z = \sum_{r=1}^{s} u_r y_{ro} - u_o$, ...3.11

Subject to: $\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - u_o \leq 0, \quad j = 1.2,\ldots n$

$$\sum_{i=1}^{m} v_i x_{io} = 1,$$

$$u_r, v_i \geq \varepsilon, \quad u_o \text{ free in sign.}$$

All variables specified in **3.11** are restricted to be non-negative excluding u_o which can take any values zero, negative or positive. Hence, the obtained sign of u_o enables one to identify the nature of the returns to scale. Banker and Thrall (1992) identified the following three conditions to distinguish the different types/classification of returns to scale (RTS) within the BCC model in **3.11** as follows:

i) IRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* < 0)$ for all optimal solutions.

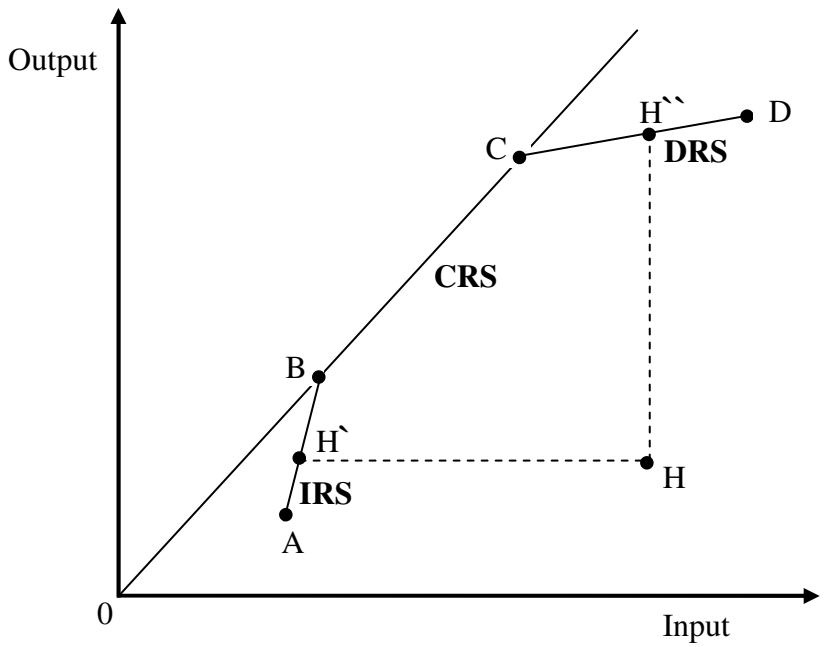
ii) DRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* > 0)$ for all optimal solutions.

iii) CRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* = 0)$ for at least one optimal solution.

Banker et al (2004) state that a BCC inefficient DMU can be projected onto the BCC frontier using optimal values from **3.8** via the following formulas sometimes known as Projection Formulas, where as usual, the symbol * represent optimal values:

$$\left\{\begin{array}{l}\hat{x}_{io}=\theta_o^* x_{io}-s_i^{-*}=\sum_{j=1}^{n} x_{ij}\lambda_j^*, \quad i=1.2,\ldots,m,\\ \hat{y}_{ro}=y_{ro}+s_r^{+*}=\sum_{j=1}^{n} y_{rj}\lambda_j^*, \quad i=1.2,\ldots,s,\end{array}\right\} \quad \text{.....3.12}$$

FIGURE 3.11 RETURNS TO SCALE IN DEA



Source: Banker et al, 2004

In an example provided by Banker et al (2004), five (5) DMUs *A, B, C, D*, and *H* are presented for evaluation as shown in Figure 3.11. The line OBC represents the CCR or CRS frontier. The lines AB, BC and CD characterise the BCC frontier and exhibits increasing, constant and decreasing returns to scale in that order. The portion between B and C is referred to as the region of most productive scale size (MPSS). Applying the projection formulas in **3.12** from the input orientation perspective, the BCC inefficient point *H* is projected onto the BCC efficiency frontier resulting in the point *H`* on line segment AB where there are increasing returns to scale (IRS). However, if the output-oriented BCC model is adopted the projection is onto *H``* on segment CD where decreasing returns to scale (DRS) prevail. As a result the input- and output oriented models may produce different outcomes in their analysis of returns to scale. In the above example, the input orientation BCC model evaluation of DMU *H* is showing IRS whereas applying the output orientation approach on

the same DMU indicates the presence of DRS. Banker et al (2004) state that this is because the input and output orientated models generate different projection points on the BCC frontier.

The illustration in Figure 3.11 also demonstrates the rationale behind why it is always advisable to maintain the interpretation that is relevant to the orientation adopted. For instance, if the BCC input model is adopted one can only give the input interpretation only and cannot interchange the interpretation as is the case with the CCR model. For instance, if the input oriented VRS model is used to evaluate *H*, then *H`* is the efficient target and the corresponding type of RTS for *H* is IRS. On the other hand, if the output VRS model is employed, *H``* becomes the efficient target and the classification of RTS for *H* will be DRS. Banker et al (2004) demonstrate that a DMU_O which is pronounced efficient under the CCR model will necessarily be designated efficient under the BCC model. However, they add that the reverse is not true. That is, a DMU_O can be considered efficient by the BCC model but not by the CCR model. The proof of the above theorem which was adapted from their study (Banker et al, 2004) is as follows:

$$\theta^*_{CCR} - \varepsilon\left[\sum_{i=1}^{m} S_i^{-*} + \sum_{r=1}^{s} S_r^{+*}\right] \leq \theta^*_{BCC} - \varepsilon\left[\sum_{i=1}^{m} S_i^{-*} + \sum_{r=1}^{s} S_r^{+*}\right],$$

Where the left-side and right-side expressions of the inequality represent optimal values of the CCR and BCC model respectively. Assuming that the CCR model defines DMU_O to be efficient which means that $\theta^*_{CCR} = 1$ and zero slacks. Then the resultant expression is as follows:

$$1 \leq \theta^*_{BCC} - \varepsilon\left[\sum_{i=1}^{m} S_i^{-*} + \sum_{r=1}^{s} S_r^{+*}\right]$$

The above expression shows that DMU_O will always be nominated to be efficient by the BCC model every time it is nominated to be efficient by the CCR model. Hence it is always the case that $\theta^*_{CCR} < \theta^*_{BCC}$ except when the point used to evaluate DMU_O is within MPSS region in which case the condition $\theta^*_{CCR} = \theta^*_{BCC}$ will be achieved (Banker et al, 2004).

3.5.1.3 The Output oriented BCC Model

The output-oriented BCC (VRS) model is basically stated as:

$$\textit{Maximise}: \quad \phi - \varepsilon\left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+}\right] \qquad \text{.....3.13}$$

$$\textit{Subject to}: \quad \sum_{j=1}^{n} \lambda_j x_{ij} + S_i^{-} = x_{io} \qquad i = 1,2,\ldots,m,$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - S_r^{+} = \phi y_{ro} \qquad i = 1,2,\ldots,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n$$

Similarly model **3.13** is calculated in a two stage procedure as follows (Zhu, 2003). Firstly, we find the optimal value of ϕ^* using the formulation in **3.13** while disregarding the slacks. We then optimise the slacks by fixing the ϕ^* in the following optimising problem.

$$\textit{Maximise}: \quad \left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+}\right] \qquad \text{.....3.14}$$

$$\textit{Subject to}: \quad \sum_{j=1}^{n} \lambda_j x_{ij} + S_i^{-} = x_{io} \qquad i = 1,2,\ldots,m,$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - S_r^{+} = \phi^* y_{ro} \qquad i = 1,2,\ldots,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0 \qquad j = 1,2,\ldots,n$$

A particular DMU will be evaluated as efficient if and only if $\theta^* = 1$ and $S_i^{-*} = S_r^{+*} = 0$ for all i and r and *weakly* efficient if $\theta^* = 1$ and $S_i^{-*} \neq 0$ and/or $S_r^{+*} \neq 0$ for some i and r. The subsequent sections will give a general framework of other various extensions to the basic CCR DEA model.

3.5.2 Additive DEA Models

While the previous DEA models, CCR and BCC versions requires one to differentiate between input-oriented and output-oriented models, the additive model treats both orientations simultaneously in one model. The additive model was developed by Charnes et al (1985) to simultaneously minimise inputs and maximise outputs. In this analysis, we shall denote the additive model, ADD_O.

3.5.2.1 The Primal Additive Model

Banker et al (2004) stated the basic additive model as follows:

$$\textit{Maximise}: \quad \left[\sum_{i=1}^{m} g_i^{-} S_i^{-} + \sum_{r=1}^{s} g_r^{+} S_r^{+}\right] \qquad \text{.....3.15}$$

$$\textit{Subject to}: \quad \sum_{j=1}^{n} \lambda_j x_{ij} + S_i^{-} = x_{io} \qquad i = 1,2,\ldots,m,$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - S_r^{+} = y_{ro} \qquad i = 1,2,\ldots,s,$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, S_i^{-}, S_r^{+} \geq 0 \qquad j = 1,2,\ldots,n.$$

Banker et al (2004) state that the additive model employs the goal vector approach wherein input and output slacks in the objective function (**3.15**) are assigned weights. This is done to guarantee that the units of measurement related to slack variables does not affect the optimal solution sets. A DMUo is rated ADD-efficient only if all slacks are zero. However, an ADD-inefficient DMU can be projected onto the efficient position by using the following projection formulae:

$$x_{io} - s_i^{-*} \Rightarrow \hat{x}_{io}$$
$$y_{ro} + s_r^{+*} \Rightarrow \hat{y}_{ro}$$

Where S_i^{-*} and S_r^{+*} are optimal input slacks and output slacks respectively obtained from the objective function **3.15**. $(\hat{x}_0; \hat{y}_0)$ are the coordinates for the point on the efficient frontier used

to evaluate DMUo. (Cooper et al, 2007) state that this constitute the first step in the analysis of returns to scale for additive models.

3.5.2.2 The Dual Additive Model

The associated dual problem to the primal formulation in **3.15** can be specified as follows:

$$\textit{Minimize} \quad \sum_{i=1}^{m} v_i x_{io} - \sum_{r=1}^{s} \mu_{ri} y_{ro} + u_o, \quad \text{.....3.16}$$

$$\textit{Subject to} \quad \sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} \mu_{ri} y_{rj} + u_o \geq 0, \qquad (j = 1, \cdots, n)$$

$$v_i \geq g_i^{-}, \quad u_r \geq g_r^{+}; \quad u_o \ \textit{free.}$$

Banker et al (2007) states that this minimising problem can be modified to yield a maximising problem as follows:

$$\textit{Maximize} \quad \hat{u}_o \quad \text{.....3.17}$$

$$\textit{Subject to} \quad \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - \hat{u}_o \leq 0, \qquad (j = 1, \cdots, n;\ j \neq 0)$$

$$\sum_{r=1}^{s} \mu_r \hat{y}_{ro} - \sum_{i=1}^{m} v_i \hat{x}_{io} - \hat{u}_o = 0,$$

$$u_r \geq g_r^{+}, \quad v_i \geq g_i^{-}, \quad \hat{u}_o \leq 0.$$

Cooper et al (2007) states that both the additive model and the BCC model have similar production possibilities. As such a DMU is considered ADD-efficient only if it is BCC-efficient, all slacks should be zero. This is depicted by the similarity of constraints imposed on the additive model`s objective function in **3.15**. Hence like the BCC model, the following types of RTS for the additive model are specified:

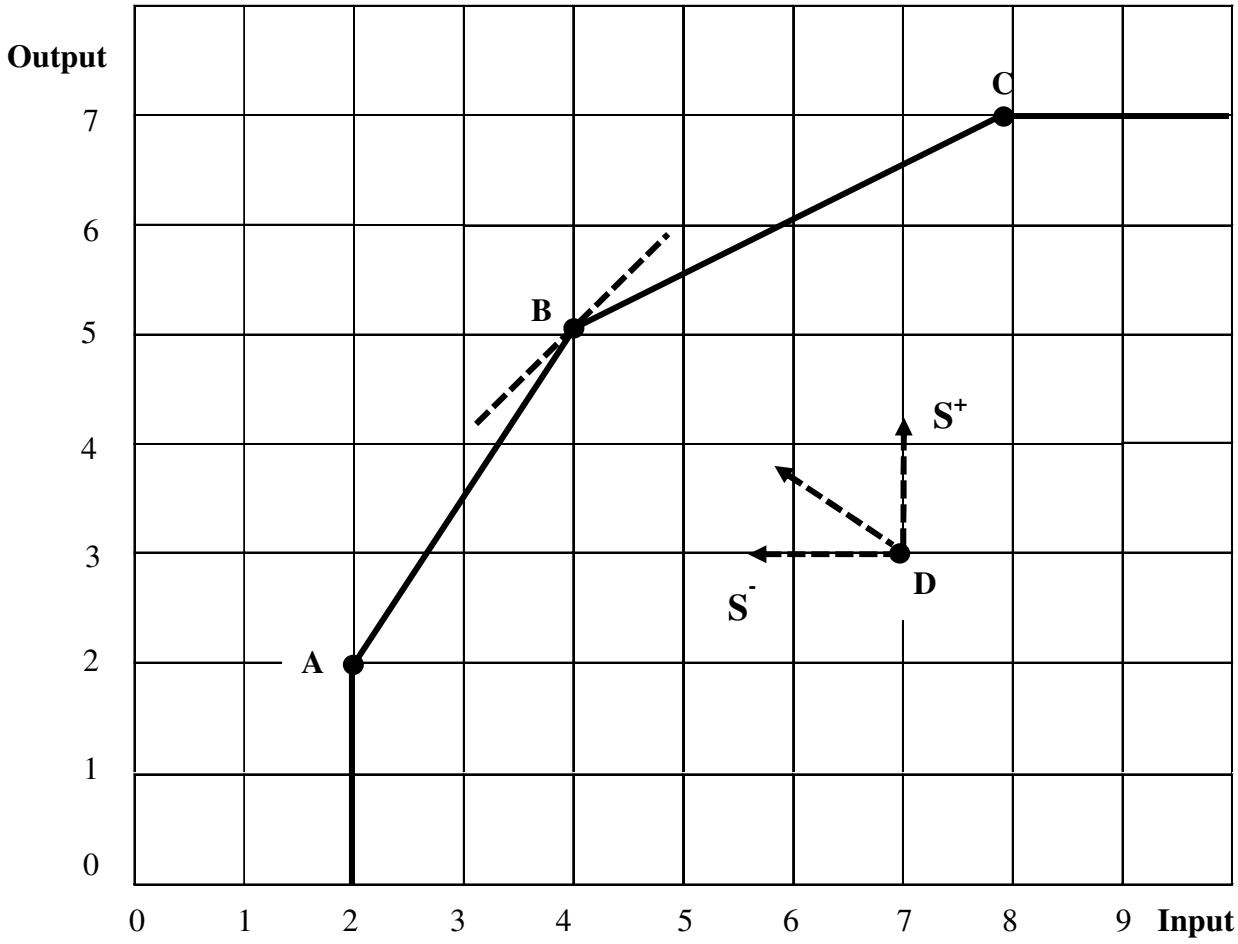
i) IRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* < 0)$ for all optimal solutions.

ii) DRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* > 0)$ for all optimal solutions.

iii) CRS dominate at $(\hat{x}_o, \hat{y}_o)$ only if $(u_o^* = 0)$ for at least one optimal solution.

We will conclude our discussion of the additive model with the aid of a diagram in Figure 3.12 wherein 4 DMUs *A* – *D* are shown. For simplicity a single input and single output setting is assumed and line segments *AB* and *BC* form the efficient frontier as shown in the Figure. DMU *D* is singled out for evaluation with other DMUs in the sample. The dotted arrows show the required input reduction (leftwards movement) and the required output augmentation (upward movement). Hence, the maximal value of S_i^- and S_r^+ is attained at point B as shown by the dotted line. Thus the additive model clearly demonstrates its ability to handle input excesses and output shortfalls simultaneously when projecting an inefficient point onto the efficient frontier.

FIGURE 3.12: THE ADDITIVE MODEL



Source: Cooper et al 2007

3.5.3 Multiplicative DEA Models

The multiplicative model which is accredited to Charnes et al (1982) incorporates the logarithmic values of output and input data [log(Y), log(X)]. Hence unlike other DEA models

which are usually piecewise-linear, the multiplicative model constructs an envelopment surface that is either log-linear or Cobb-Douglas (Bekiaris & Nakanishi, 2004). The piecewise log-linear model called the variant multiplicative model has CRS while the piecewise Cobb-Douglas model called the invariant multiplicative model has VRS. Charnes et al (1983) further developed the model to allow these models the property of non-dimensionality. This property implies that the generated efficiency measures should not be influenced by units of measurement used. Cooper et al (2011) state that efficiency frontiers of multiplicative models are not limited to be concave. They can be formulated to be concave in some regions and non-concave in others. However, despite its distinct properties, multiplicative models have not been widely applied in research compared to other models. A detailed analysis of this DEA type model is contained in efficiency studies by Charnes et al (1982, 1983) and Banker & Maindiratta (1986).

3.6 CONCEPTUALISING PRODUCTIVITY

Productivity is a simple indicator that describes the relationship between output and the inputs that generate that output and as such regarded as an important measure of economic performance. Total factor productivity (TFP) is defined as the ratio of total output to total inputs used in production. TFP is a broader concept compared to partial productivity given that it takes into consideration the specific combination of inputs used in the production process. On the other hand, partial productivity simply looks at the contribution of a single input factor to output and fails to take into account the substitutability among input factors. In the singe-output and single-input setting, productivity is simply the ratio of that bank`s output and input quantity. In real world cases where a bank produces several outputs using several inputs, TFP is measured as the ratio of aggregate output to aggregate input. Suppose that $x_{it} = (x_{1it},...,x_{Kit})'$ and $y_{it} = (y_{1it},...,y_{Jit})'$ denote the input and output quantity vectors of bank *i* in period *t*. Then the TFP of the bank is:

$$TFP_{it} = \frac{Y_{it}}{X_{it}} \tag{5}$$

Where $Y_{it} = Y(y_{it})$ is aggregate output and $X_{it} = X(x_{it})$ is aggregate input. *Y* (.) and *X* (.) are non-decreasing, non-negative, and linearly homogeneous aggregator functions. The related index number that measures the TFP of bank *i* in period *t* relative to bank *h* in period *s* is:

$$TFP_{hs,it} = \frac{TFP_{it}}{TFP_{hs}} = \frac{Y_{it} \big/ X_{it}}{Y_{hs} \big/ X_{hs}} = \frac{Y_{it} \big/ Y_{hs}}{X_{it} \big/ X_{hs}} = \frac{Y_{hs,it}}{X_{hs,it}} \qquad (6)$$

Where $Y_{it} = Y_{it} \big/ Y_{hs}$ is an output quantity index and $X_{it} = X_{it} \big/ X_{hs}$ is an input quantity index. Thus TFP growth can be expressed in terms of the ratio of output growth over input growth.

3.6.1 The Hicks-Moorsteen TFP Index

O`Donnell (2011) states that different aggregator functions give rise to different TFP indexes such as the Laspeyres, Paasche, Fisher, Lowe, Malmquist, Fare-Primont and the Hicks-Moorsteen. The Hicks-Moorsteen output and input aggregator functions are $Y(y) = [D_O(x_{hs}, y, s) D_O(x_{it}, y, t)]^{1/2}$ and $X(x) = [D_I(x, y_{hs}, s) D_I(x, y_{it}, t)]^{1/2}$ respectively. When these are substituted in equation (5) and (6) specified above, they give rise to the Hicks-Moorsteen TFP index:

$$TFP_{hs,it} = \left(\frac{D_O(x_{hs}, y_{it}, s)}{D_O(x_{hs}, y_{hs}, s)} \frac{D_I(x_{hs}, y_{hs}, s)}{D_I(x_{it}, y_{hs}, s)} \frac{D_O(x_{it}, y_{it}, t)}{D_O(x_{it}, y_{hs}, t)} \frac{D_I(x_{hs}, y_{it}, t)}{D_I(x_{it}, y_{it}, t)} \right)^{1/2} \qquad (7)$$

Where y_0 and x_0 are vectors of quantities and D_O (.) and D_I (.) are Shepherd`s (1953) output and input distance functions. Formulation (7) above was first proposed by Bjurek (1996) but is commonly known as the Hicks-Moorsteen index. O'Donnell (2011) states that it is called as such because it is the geometric average of two indexes that Diewert (1992) attributed to Hicks (1961) and Moorsteen (1961). O'Donnell (2011a) describes TFP indexes that can be expressed in terms of aggregate quantities as in equation (6) as being *multiplicatively-complete.* However, unlike the Hicks-Moorsten TFP Index, the popular Malmquist TFP index is not included among the class of multiplicatively complete TFP indexes. Thus O`Donnell (2010a, p.1) argues that for this reason it cannot be regarded as a valid measure of productivity change except under constant returns to scale technology - which he describes as "except in restrictive special cases". Briec and Kerstens (2011) in their paper entitled, "The Hicks–Moorsteen productivity index satisfies the determinateness axiom", proved that the Hicks–Moorsteen productivity index indeed satisfies the determinateness axiom under weak conditions on technology. However, on the contrary the popular Malmquist productivity index failed to satisfy the determinateness condition. In their conclusion the authors state: "We expect the Hicks–Moorsteen productivity index to gain in popularity in future empirical

work, especially when infeasible solutions are simply unacceptable" (Briec and Kerstens, 2011, p.10).

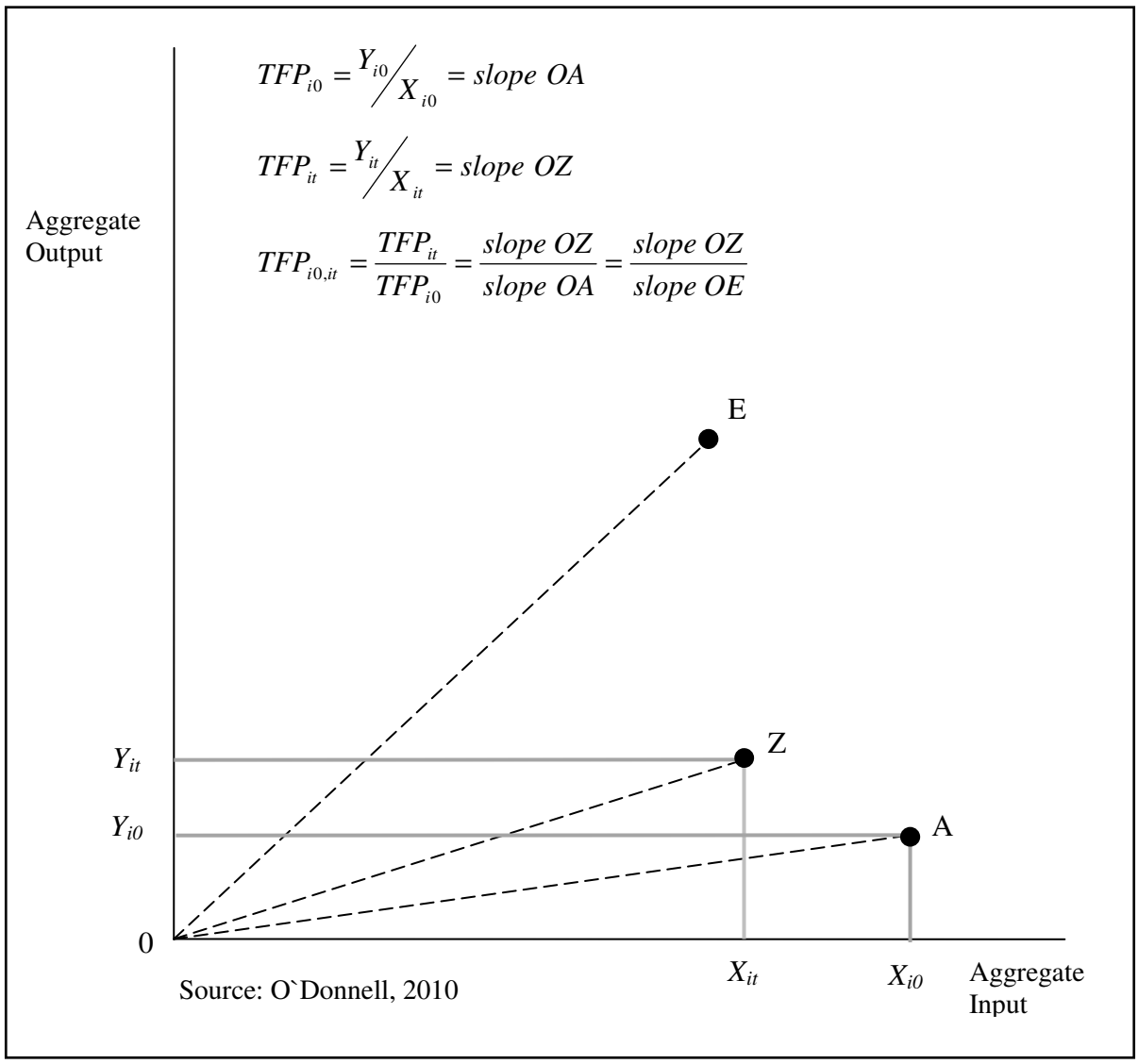
As alluded to in the introductory chapter, one of the contributions of this thesis is that it is a pioneering investigation within the South African banking industry to adopt the HMTFP index approach to decompose productivity into efficiency. Earlier studies have examined banking productivity by employing the popular yet imperfect Malmquist technique. Grifell-Tatje and Lovell (1995) demonstrated that the Malmquist is appropriate under the CRS specification and that with VRS the Malmquist productivity index fails to accurately measure productivity change. Coelli and Rao (2005) also criticised the Malmquist approach from the viewpoint that in the absence of CRS assumption, the outcome of running the Malmquist methodology may erroneously measure TFP changes arising from scale economies.

3.6.2 Components of Productivity Change

O'Donnell (2010c) shows that any multiplicatively-complete TFP index can be exhaustively separated into various measures of technical and efficiency change. Figure 3.13 below, portrays the concept in aggregate quantity space. The TFP of bank *i* in period 0 is shown by the gradient of the ray passing through the origin and point A, while TFP in period t is shown by the gradient of the ray passing through the origin and point Z, while TFP in period t is given by the slope of the ray passing the origin and point Z. It follows that the TFP index that measures the change in TFP between the two periods is:

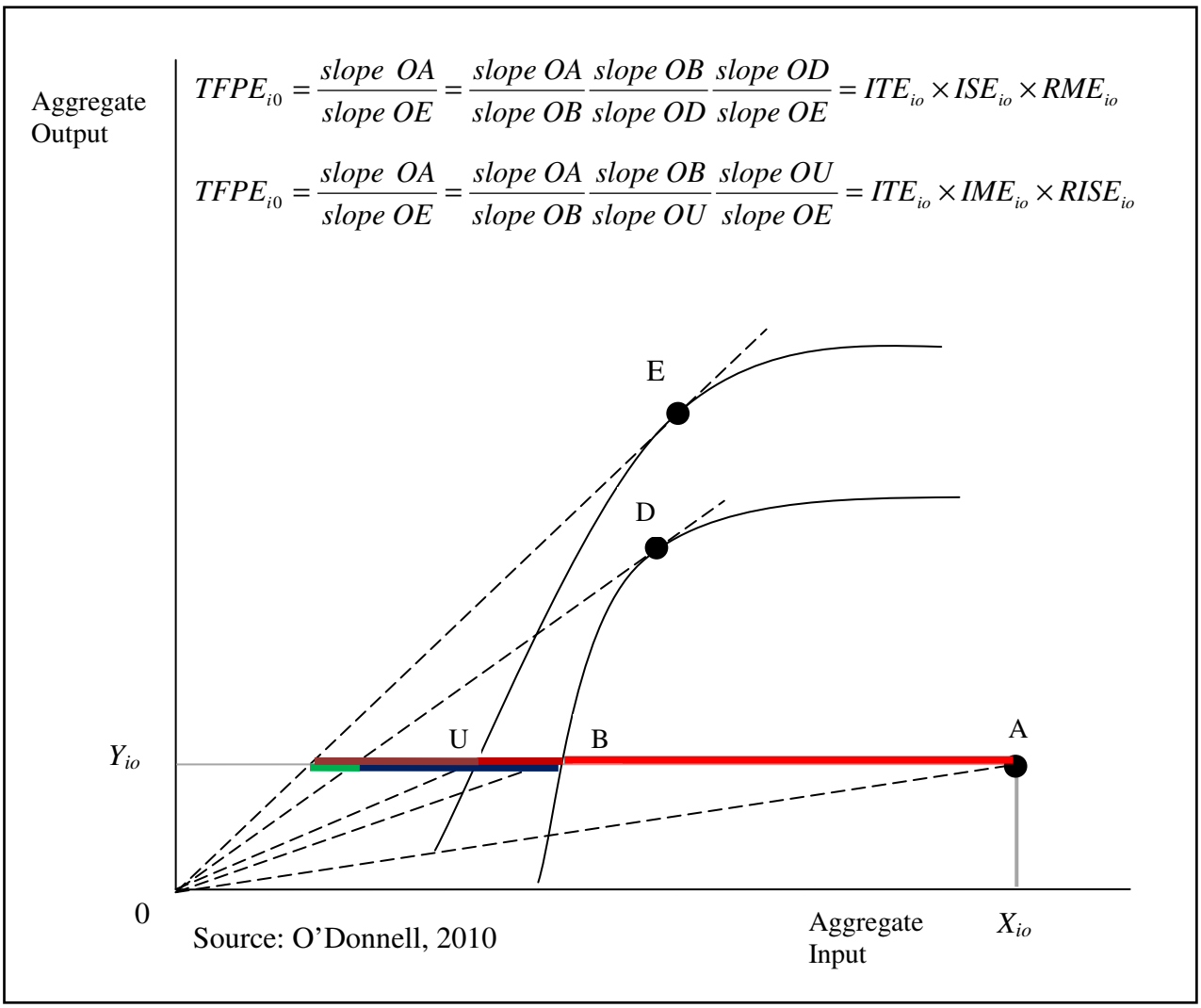
$$TFP_{i0,it} = slope\ OZ \Big/ slope\ OA.$$

FIGURE 3.13: MEASURING AND DECOMPOSING TFP CHANGE



O'Donnell (2011a, p.2) outlines that changes in productivity can be decomposed into measures of *technical change* which measures movements in the production frontier; *Technical efficiency change* which measures movements to or away from the frontier; *Scale efficiency* change which measures movements around the frontier surface to capture economies of scale and, *Mix efficiency* change which measures movements around the frontier surface to capture economies of scope.

FIGURE 3.14: TWO INPUT ORIENTED DECOMPOSITIONS OF TFP EFFICIENCY



O`Donnell (2010a, p.2-4) defined certain efficiency measures which feature prominently in input-orientated decompositions of TFP change as follows:

3.6.2.1 Input-oriented Technical Efficiency (ITE)

It measures the difference between observed TFP and the maximum TFP that is possible while holding constant the input mix, output mix and output level. ITE is illustrated in Figure 3.14 as the measure of the horizontal distance from point A to point B. It is equivalent to the measure of the difference in TFP at points A and B: $ITE_{io} = {slope\ OA}/{slope\ OB}$

3.6.2.2 Input-oriented Scale Efficiency (ISE)

It measures the difference between TFP at a technically efficient point and the maximum TFP that is possible while holding the input and output mixes constant but permitting the levels to

vary. Figure 3.14 represents this measure of efficiency as a movement from point B to point D: $ISE_{io} = {slope\ OB}/{slope\ OD}$. O'Donnell (2010a) refers to point D as the point of *mix-invariant optimal scale* (MIOS).

3.6.2.3 Residual Mix Efficiency (RME)

RME captures the difference between the possible maximum TFP on a mix-restricted frontier and the possible maximum TFP when input and output mixes (and levels) can vary. In Figure 3.14, it is depicted as a movement from point D to point E: $RME_{io} = {slope\ OD}/{slope\ OE}$

3.6.2.4 Input-oriented Mix Efficiency (IME)

It measures the difference between TFP at a technically efficient point on the mix-restricted frontier and the maximum TFP that is possible while holding the output level fixed. This measure of efficiency is portrayed in Figure 3.14 as movement from point B to point U: $IME_{io} = {slope\ OB}/{slope\ OU}$

3.6.2.5 Residual input-oriented Scale Efficiency (RISE)

RISE measures the difference between TFP at a technically and mix-efficient point and TFP at the point of maximum productivity. RISE is shown in figure 3.14 as a movement from point U to point E: $RISE_{io} = {slope\ OU}/{slope\ OE}$

3.6.2.6 TFP Efficiency (TFPE)

It measures the difference between observed TFP and the maximum TFP* possible using the available technology. Figure 3.14 illustrates this measure as a movement from point A to point E: $TFPE_{io} = {TFP_{I0}}/{TFP_0^*} = {slope\ OA}/{slope\ OE}$,

Where TFP_0^* denotes the maximum TFP possible using the technology available in period 0.

O`Donnell (2010a) states that Figure 3.14 shows just 2 of various pathways from point A to E and hence 2 of various decompositions of TFP efficiency which are:

$$TFPE_{i0} = \frac{TFP_{i0}}{TFP_0^*} = ITE_{io} \times ISE_{io} \times RME_{io}$$

$$TFPE_{i0} = \frac{TFP_{i0}}{TFP_0^*} = ITE_{io} \times IME_{io} \times RISE_{io}$$

Figure 3.15 and 3.16 below shows decompositions of total factor productivity into several efficiency measures from both orientations for a multiple input and multiple output firm. These efficiency measures are illustrated in Figure 3.15 and Figure 3.16. O`Donnell (2011) states TFP efficiency is a measure of overall productive performance, residual scale and mix efficiency are productive performance measures associated with economies of scale and scope.

$$TFPE_{it} = \frac{TFP_{it}}{TFP_t^*} \leq 1 \qquad \text{(TFP efficiency)} \qquad (8)$$

$$OTE_{it} = \frac{Y_{it} / X_{it}}{\overline{Y}_{it} / X_{it}} = \frac{Y_{it}}{\overline{Y}_{it}} = D_O\left(x_{it}, y_{it}, t\right) \leq 1 \quad \text{(Output oriented technical efficiency)} \qquad (9)$$

$$OSE_{it} = \frac{\overline{Y}_{it} / X_{it}}{\tilde{Y}_{it} / \tilde{X}_{it}} \leq 1 \qquad \text{(Output oriented scale efficiency)} \qquad (10)$$

$$OME_{it} = \frac{\overline{Y}_{it} / X_{it}}{\hat{Y}_{it} / X_{it}} = \frac{\overline{Y}_{it}}{\hat{Y}_{it}} \leq 1 \qquad \text{(Output oriented mix efficiency)} \qquad (11)$$

$$ROSE_{it} = \frac{\hat{Y}_{it} / X_{it}}{TFP_t^*} \leq 1 \qquad \text{(Residual output oriented scale efficiency)} \qquad (12)$$

Figure 3.15: Input-Oriented Measures of Efficiency for a Multiple Input Multiple Output Firm

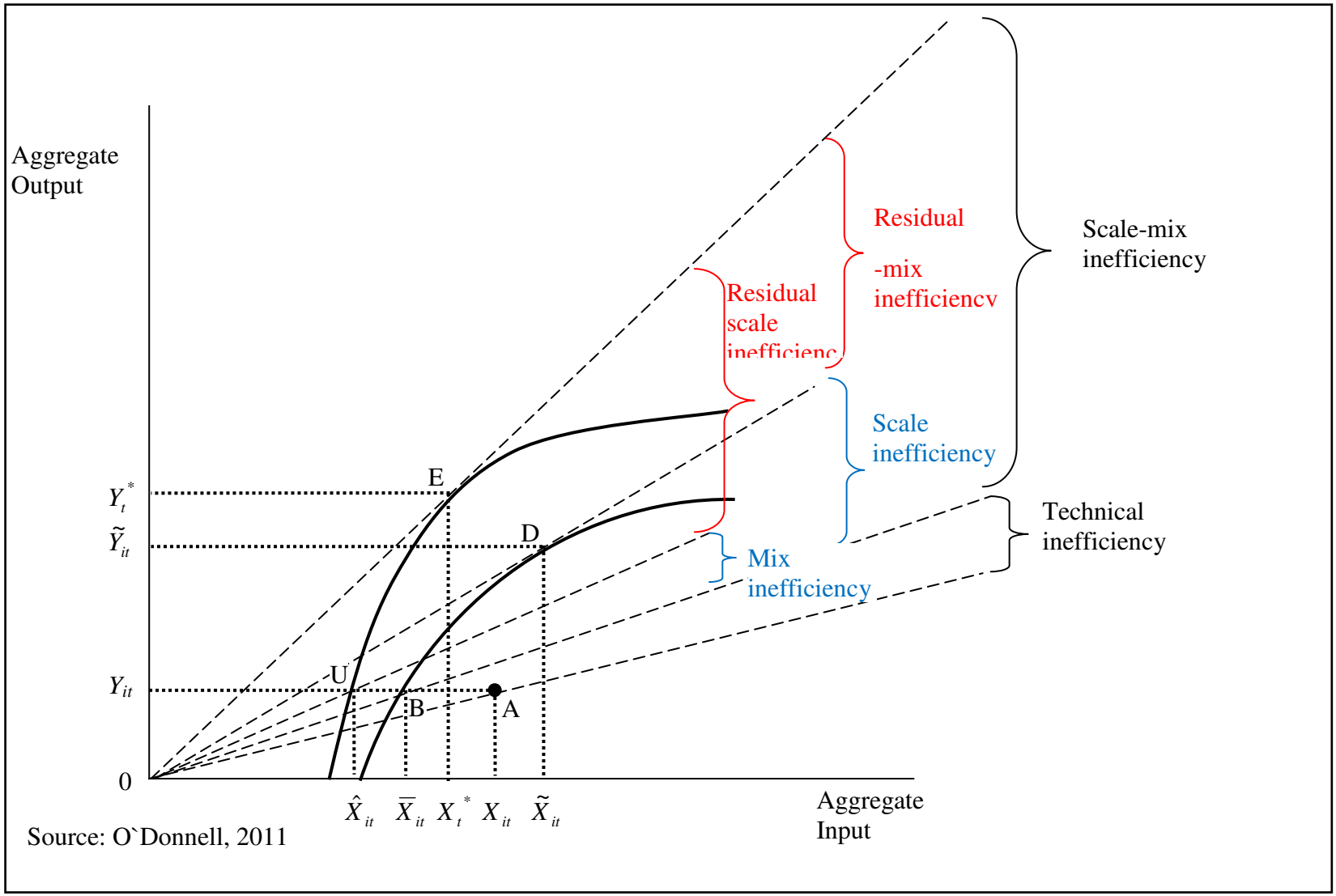
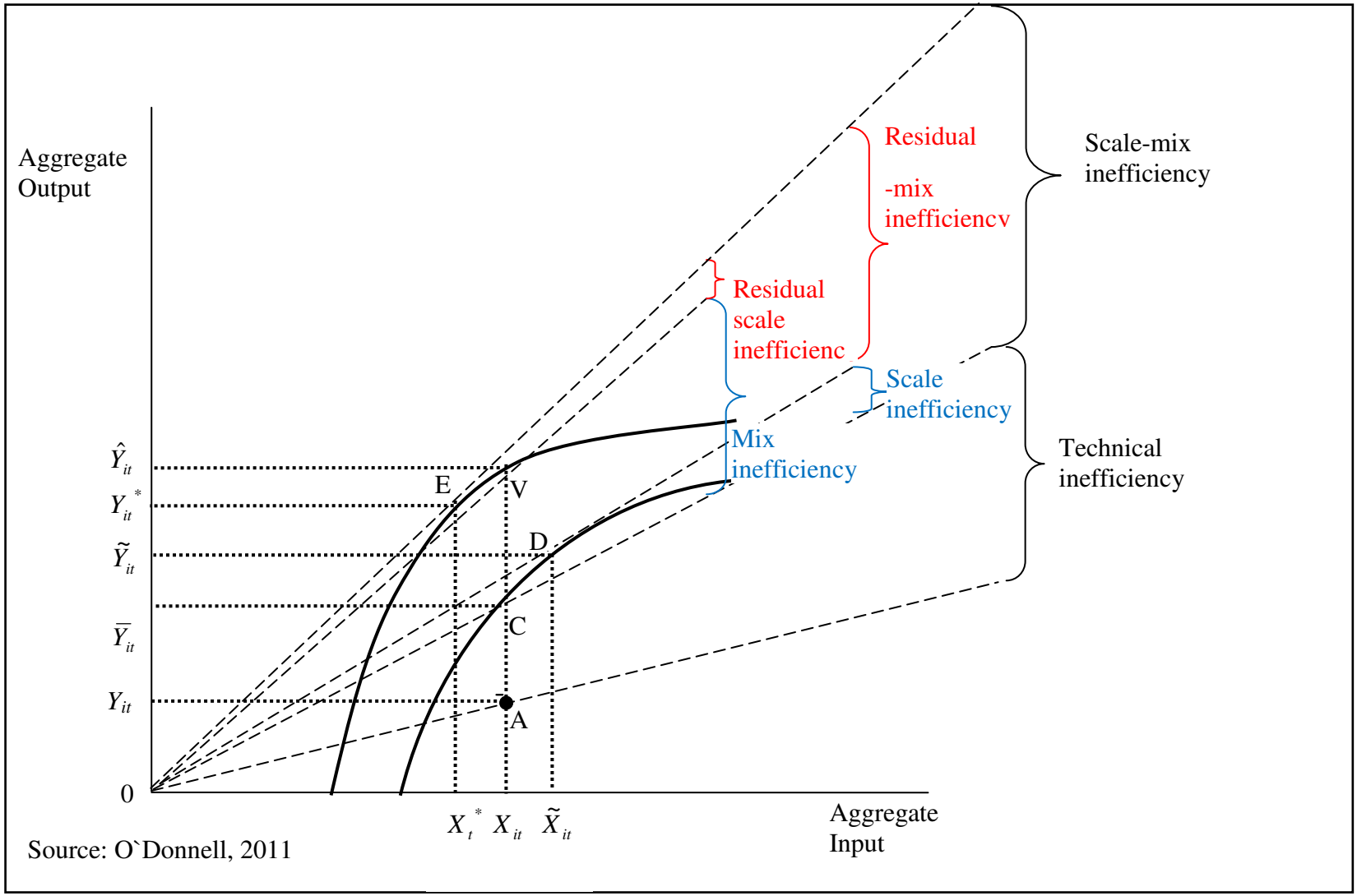


Figure 3.16: Output-Oriented Measures of Efficiency for a Multiple Input Multiple Output Firm



$$ITE_{it} = \frac{Y_{it} / X_{it}}{Y_{it} / \overline{X}_{it}} = \frac{\overline{X}_{it}}{X_{it}} = D_I\left(x_{it}, y_{it}, t\right)^{-1} \leq 1 \quad \text{(Input-oriented technical efficiency)} \tag{13}$$

$$ISE_{it} = \frac{Y_{it} / \overline{X}_{it}}{\tilde{Y}_{it} / \tilde{X}_{it}} \leq 1 \quad \text{(Input-oriented scale efficiency)} \tag{14}$$

$$IME_{it} = \frac{Y_{it} / \overline{X}_{it}}{Y_{it} / \hat{X}_{it}} = \frac{\hat{X}_{it}}{\overline{X}_{it}} \leq 1 \quad \text{(Input-oriented mix efficiency)} \tag{15}$$

$$RISE_{it} = \frac{Y_{it} / \hat{X}_{it}}{TFP_t^*} \leq 1 \quad \text{(Residual input oriented scale efficiency)} \tag{16}$$

$$RME_{it} = \frac{\tilde{Y}_{it} / \tilde{X}_{it}}{TFP_t^*} \leq 1 \quad \text{(Residual mix efficiency)} \tag{17}$$

Where TFP_t^* denotes the highest total factor productivity that is possible with the available technology in period t; $\overline{Y}_{it} \equiv Y_{it} D_O\left(x_{it}, y_{it}, t\right)^{-1}$ is the maximum aggregate output achievable when using x_{it} to generate a scalar multiple of y_{it}; $\overline{X}_{it} \equiv X_{it} D_I\left(x_{it}, y_{it}, t\right)^{-1}$ is the lowest aggregate input possible when consuming a scalar multiple of x_{it} to produce y_{it}; $\hat{Y}_{it}$ is the highest aggregate output achievable when using x_{it} to generate any output vector; $\hat{X}_{it}$ is the lowest aggregate input possible when using any input vector to generate y_{it}; $\tilde{Y}_{it}$ and $\tilde{X}_{it}$ are the aggregate output and input attained when TFP is maximised subject to the constraint that the output and input vectors are scalar multiples of y_{it} and x_{it} respectively (O`Donnell, 2011).

EMPIRICAL LITERATURE

Banking efficiency studies particularly in African countries are very scarce. The available empirical literature has been enriched with studies mainly from the developed world. In this section of the study, a brief empirical literature review relating to the economic performance of the banking sector is presented. This review is intended to show previous work in the area of banking efficiency and productivity in both developing and developed economies as well as South Africa. The key issues in this study are access to banking services, bank efficiency, the global financial crisis and the efficiency-employment nexus.

3.7 BANKING EFFICIENCY AND ACCESS TO BANKING SERVICES

In 2011, FinScope (2011) reported that 37.2 percent of South Africa`s adult population are unbanked while 27 percent are considered financially excluded - with no access whatsoever to either formal or informal financial services. Claessens and Laeven (2003) showed that technological advancement and reduction in production or distributional costs among commercial providers is not a guarantee of more or enhanced access. This is a very interesting proposition which forms the main objective of this thesis. The central objective in this study is to investigate the link between bank efficiency and access to banking services. The socio-economic profile of the low-income and poor households is such that it does not promote efficiency thereby suggesting the possibility of a trade-off between access and improvement in bank efficiency. It is essential at this moment to recall that allocative efficiency is concerned with the extent to which resources are devoted to activities with the greatest value. Presented with the relatively greater costs of information asymmetry normally related with serving low-income consumers, banks find it rational to serve high-income clients as the profit potential is greater. The question this thesis seeks to address is: does improving bank efficiency precludes the provision of banking services particularly to the low-income households.

According to Hawkins (2010), access to financial services equips consumers with the means to make payments, savings and make investments; manage risk; obtain credit and loans; and to make financial provision for retirement. She maintains that the concept of access is very complex involving availability, affordability and appropriateness which are indefinable therefore, subjective to measurement. While access to financial services is a broader concept, this thesis adopts the idea of access to banking services. This is because access to banking services as approximated by a bank account is a first necessary step in acquiring the ability to

save and make payments as well as opening the door to a host of other financial services. The World Bank (2009) study entitled *Banking the Poor* also measured access to banking services by considering the proportion of the population with a bank account per thousand adults and transactions offered at banks in each country. However, Hawkins (2010) also argues that the issue of bank account dormancy is important as it may reflect a lack of appropriateness of such bank services. For instance, if a considerable number of opened bank accounts are not being used, it may imply that bank fees and charges are prohibitively high. As such these services may not be affordable and therefore not appropriate for consumers. An IMF (2008) study noted that despite the significant increase in the population with access to a bank account, increasing access to affordable and better services remain a challenge. The major issues identified were: Access to bank accounts was categorised in terms of the salaried and the non-salaried; Provision of loans to small enterprises remained insufficient and; other banking services such as savings and insurance products trailed behind. The World Bank study (2006) on *Making Finance Work for Africa* identified information asymmetry and transaction costs as two major problems related to reaching the so-called difficult markets. Firstly, poor quality and scarcity of information about individual risks was identified as a greater barrier in assessing creditworthiness. Demirguc-kunt and Klapper, (2012) suggests relaxing documentation requirements as one of the ways to improve accessibility. Their study showed that in Sub-Saharan Africa reducing documentation requirements potentially increase the proportion of adults with an account by up to 23 percentage points. Secondly, their study noted that financial institutions find it difficult to serve the marginalised or distant customers as well as adapting product designs that suit these customer`s needs at an affordable cost. The study also noted that although banks may have the needed resources to set up new banking systems that can reach rural households, there is need to ensure that the unit cost of operating these systems remain sufficiently low. The World Bank (2006) also states that from the demand side, high minimum balances and monthly charges prohibits a great proportion of Africans from accessing formal financial services. In recent years, advances in technology and innovation such as cell phone banking and mobile branches have helped to reduce transaction costs at the same time improving access. This avenue holds greater potential in South Africa as the majority of people do have a cell phone.

Kablan (2010) adopted the stochastic frontier analysis (SFA) methodology to investigate the factors that impacts the degree of cost efficiency among sub-Saharan Africa (SSA) countries. First stage SFA results showed that SSA banks were generally cost-efficient with an

estimated efficiency score of 76%. An analysis by region showed that Southern Africa presented the highest efficiency score of 76% followed by Western Africa with 75% and 74% for the Eastern Africa region. In order to understand the determinant factors of efficiency, the following variables were considered: bank capitalisation, bank ownership, GDP per capita, the share of the population in rural areas, and non-performing loans. Of particular importance and related to the present study is the impact of growth of rural population on bank efficiency. Kablan (2010) argues that SSA banks tend to locate their branches in more economically developed regions at the expense of rural ones. Stating that banks with a greater proportion of rural population tends to be less cost-efficient because they cannot realise economies of scale. The second stage results indicated that indeed the variable representing the share of rural population had the expected negative sign with cost efficiency. This finding confirmed the notion that the larger share of the rural population limits bank efficiency by increasing costs because of lack of economies of scale for implementing bank technology (Kablan, 2010). Schoombee (2004) also states that commercial banks generally target the high-income clients and are reluctant to serve low-income consumers, SMEs and the poor mainly due to associated high costs.

In a provincial study of one large South African bank, Okeahalam (2008) assessed the marginal impact of expanding bank branches and the contents of a product suite on access to banking services. A sample of 7129 clients were analysed using the Maximum Likelihood method for the period 1999 – 2004. A product suite defined as "a composite of various products based on the customer profile" was considered to be determined by socio-economic factors such as age, gender, income and credit score (Okeahalam, 2008, p. 1133). Depending on product suite and the relative distance to the closest branch, three levels of access were defined namely, full, medium and low. Full access was defined as access to the total suite of products *and* the availability of a branch within 14.3 km of the residential address. Medium access was either access to total suite of products *or* access to a branch within 14.3 km. Low access was defined as limitations on access to the total variety of products within the product suite *or* branch distance exceeding 14.3 km. The results obtained showed that branch availability compared to product suite has a positive and greater marginal effect on low and medium access. Regarding full access, the study pointed out that the welfare increase of providing additional access to customers who at present have high access is negligible. The study concluded that improvement in the overall level of access is more pronounced via investment in bank branches than via an increase in the contents of the product suite. The

study also showed that an increase in bank branches is not as important as the geographical location of branches arguing that the clustering of branches in urban and larger cities is a form of managerial inefficiency and slack. In addition, their results also indicated that the client profile which the bank defines for each individual customer has an impact on the level of overall access. The reported evidence also confirmed that living in the townships (ghettos) has a negative and significant impact on access to branch banking services and on all levels of access. Finally, the income of clients was found to be positively related with the product suite and hence access.

This brief section has reviewed various views that are held concerning the subject of access to banking services. A comprehensive discussion of the wide-ranging concepts of access and the associated old and contemporary arguments were discussed in the preceding chapter. However, to date no known studies have actually examined the impact of bank efficiency on access, close South African studies include Hawkins (2010) and Okeahalam (2008). For instance Hawkins (2010) simply attempt to measure financial access while Okeahalam (2008) explores the impact of client profiles and bank branches on various levels of access to bank services using an econometric approach. This thesis therefore attempts to fill this gap in banking literature by broadening the scope of bank-specific regressor variables including inter alia total factor productivity efficiency on the level of access. The following section discusses the degree of competition and its implication on bank efficiency and access to bank services.

3.8 COMPETITION AND ACCESS TO BANKING SERVICES

Mlambo and Ncube (2011) states that competition reduces monopoly rents, prices and operational costs at which financial services are offered to the public, resulting in welfare gains for the public. This subject is critical in South African banking where the four largest banks in the country account for over 80% of total banking assets. A World Bank (2007) study also identified lack of competition in banking as one of the aspects that is related to low efficiency of commercial banks in Africa. The study noted that interest spreads, profits and overhead costs are high in African banking. Napier (2005a) observed that South African banks operate as a complex monopoly, with perceived high barriers to entry. Okeahalam (2000) argues that the presence of an oligopoly structure (imperfect competition) imply that the level of competition required to induce efficiency improvements may not exist. Mlambo

and Ncube (2009) also found that for the period 1999 – 2008, the structure of the South African banking industry was characterised by monopolistic competition.

A high level of market concentration raises the obvious question of whether efficiency and accessibility to banking services is being achieved in South Africa. In addition to ensuring improved efficiency in the production of financial services, quality of financial products and innovation, Claessens and Laeven (2003) added that competition matters for access to financial services by both firms and households. On the contrary, Petersen and Rajan (1995) hold the view that market power in banking may be to an extent beneficial for access to financing. Hence the view that competition is definitely good in banking may not have universal support. Edwards and Mishkin (1995) also concur, that while a competitive environment may induce efficiency, it may also increase risk taking activities through engaging in non-traditional banking activities as banks seek to maintain their market shares. In order to prevent the resultant instability in the banking sector, a solid regulatory and bank supervision framework is needed.

3.8.1 Studies on Competition and Banking Efficiency in South Africa

Okeahalam (2001) adopted the structure conduct performance (SCP) framework for the period 1997 to 1999 to assess the degree of concentration in South African banking sector. The SCP framework implies that a higher level of concentration leads to higher prices which in turn lead to higher profits. On the demand side, higher prices acts as a brake on efforts designed to increase access to banking services. The results from the model indicated that the South African banking sector was highly concentrated and characteristic of collusive oligopolistic behaviour. He noted that high concentration raises the likelihood of collusive oligopoly behaviour with the associated negative consequences for the consumer. Of particular interest to this study was the view that collusion leads to comfortable returns and reduces the incentive for banks to seek markets that are considered difficult[7] (Okeahalam, 2001). Okeahalam stated that the absence of service delivery is an aspect of welfare cost of high concentration. In his conclusion, the author urged the government to consider the Community Reinvestment Act which basically permits the government to establish particular prerequisites to be met by low-income groups for lending and to evaluate if banks are actually discriminatory.

[7] Rural and the poor or generally the unbanked.

Greenberg and Simbanegavi (2009) employed the Panzar-Rosse (1987) approach and the Bresnahan model (1982) to determine the level of competition in South Africa`s banking sector for the period 1998 to 2007. Depending on the market structures in which they operate, banks behave differently. They noted that there are efficiency implications if the banking sector is not competitive. The Panzar-Rosse model recognises banks as profit maximising entities that conducts business in a highly challenging market environment and so face conventional average and marginal costs. Mlambo and Ncube (2011) state that the H-statistic which is constructed from the reduced-form revenue function measures the elasticity of total revenues with regard to price changes of factor inputs. In order to account for possible heteroscedasticity, the authors conducted the Panzar-Rosse test on small, large and on all banks collectively. They found that with regard to interest income, the South African banking sector was highly competitive with an H-statistic of 0.7475 which was not significantly different from 1. The results also revealed that large banks with an H-statistic of 0.9672 operated under conditions consistent with perfect competition while small banks were characteristic of monopolistic competition exhibiting an H-statistic of 0.5795. They justified their interesting findings, arguing that small banks mainly operate in geographically segmented markets where they tend to serve niche markets allowing them to exercise market power (Greenberg and Simbanegavi, 2009). On the contrary, large banks in South Africa operated at a national and international scale where the competition is aggressive. Finally, results obtained from the Bresnahan model also supported their initial results that the South African banking sector was characterised by a high level of competition.

Another study of competition was conducted by Mlambo and Ncube (2011) who analysed the evolution of competition and efficiency of the banking sector in South Africa using firm-level data for the period 1999 -2008. The study investigated the relationship between market structure and efficiency. Their point of departure was that competition forces banks to reduce cost-inefficiencies and may result in welfare gains for the public. The data set consisted of a panel of 26 domestic and foreign banks. The methodology adopted was the variable returns to scale (VRS) input-oriented Data Envelopment Analysis (DEA). The intermediation approach was used to determine the choice of inputs and outputs. The input variables used were staff costs and deposits with loans and advances as the output variables. The DEA results showed that cost efficiency was 0.412 and that mean allocative efficiency was 0.609, while technical efficiency was 0.67. These results suggested that on average, inefficiency mainly resulted from wrong use or misallocation of inputs at existing input prices. The Panzar-Rosse results

revealed an H-statistic of 0.57 showing that the market structure of the banking industry was monopolistic competition. However, the Panzar-Rosse model was re-estimated with the DEA efficiency score included among the independent variables. Casu and Girardone (2006) justify the inclusion of DEA efficiency score on the grounds that it can be taken as a proxy for managerial ability. Managerial ability came out as an important factor for bank efficiency. The estimated H-statistic increased slightly to 0.575 from the previous 0.57. Their conclusion was that competition was a crucial determinant of bank efficiency and that the structure of the banking industry was characterised by monopolistic competition.

In addition to contributions made by [Okeahalam (2001), Greenberg and Simbanegavi (2009), Mlambo and Ncube (2011)] as highlighted in the literature review, another issue which this study addresses which has not been dealt with previously in the literature, is whether bank market concentration has an impact on access. A highly concentrated system generally invokes a concern regarding efficiency and the level of output production. Hence our study complements these studies and fills this existing gap by including a market structure variable to evaluate the impact of bank market concentration on the level of access to banking services in South Africa.

3.9 THE 2007/2008 SUB-PRIME FINANCIAL CRISIS

While a range of empirical studies in South Africa have sought to evaluate bank efficiency, none have examined the effects of the 2008 financial crisis on the banking sector efficiency and productivity using the methodological approach and procedures in this study. The financial crisis of the United States in 2007 and the recent 2012-2013 Cypriot banking crisis have generated worldwide concern on the financial sector of which the banking system forms the major component. These banking crisis episodes have shown how catastrophic problems in the financial sector can be for the entire economy and the global economy as a whole. The banking sector is crucial because it shelters the economy against instability and boost consumer confidence. However, ironically, the banking systems' fragility itself is a major source of instability. Mercan et al (2003) argues that the fragility of the fibre of banks and the fact that they are perceived to be institutions of confidence, threatens not just banking institutions but the entire economy. It is generally acknowledged that Africa escaped the contagion effects of the 2008 financial crisis due to the fact that its economies are hardly integrated in the global financial system. However, Mwega (2010) argues that Africa did not escape the impact of the crisis. According to South African Reserve Bank (SARB) (2011),

the devastating effect of the sub-prime financial crisis in South Africa was fairly cushioned by the solid macroeconomic policies, and prudent regulation in the domestic banking sector. However, as highlighted by Mnyande[8] (2012) in his key note address, South Africa is not immune to the crisis in the medium to long term. He stated that South Africa`s inter-linkages with the global economy means that its growth outcomes are highly dependent on those of the affected European economies. Hence, there is a need to investigate the magnitude of efficiency and productivity changes that were realised during the period of the crisis. This is important for three reasons. Firstly, South Africa is a developing or emerging economy with a highly developed banking system that compares well to that of advanced economies such as the US. Secondly, knowledge of the size of the impact of the crisis on the banking sector is crucial to inform policy makers to provide precise and targeted policy measures as well as whether to tighten bank supervision and regulation. Thirdly, pessimists have argued that it is too early to celebrate and that the worst case global challenges may have been merely delayed rather than avoided.

3.9.1 Mismanagement of Financial Innovation/Liberalisation in the Mortgage Market

A financial crisis is a great disturbance within financial markets which is marked by severe deteriorations in asset prices and collapsing of firms resulting in the general economic meltdown. There are a host of many factors direct and indirect which are believed to have culminated in the eruption of the 2008 financial crisis. However, mishandling of financial innovation within the sub-prime mortgage market, and a bubble burst within the housing sector are acknowledged to have been the underlying forces behind the crisis.

Mishkin (2010) states that the origins of a financial crisis starts when an economy innovates by introducing new types of loans or financial products known as financial innovation or when governments promotes financial liberalisation which is essentially the removal of limitations on financial institutions. It is important also to appreciate the benefits of both financial innovation and liberalisation in promoting financial development and the economy at large. Studies on the role of financial sector development and economic growth are well documented [Ahmad & Malik (2009); Honohan (2004); Khan & Senhadji (2000); Levine (1997)]. However, there is need to underscore that it is usually the mismanagement of innovation and liberalization that often precipitates danger as was the case with the sub-prime

[8] Chief Economist and Adviser to the Governor of Reserve Bank.

crisis. Before the crisis, it was only the credit worthy borrowers known as prime borrowers that could qualify for residential mortgages. However, due to improvements in technology and data mining techniques this led to the birth of a new class of risky residential mortgages (Mishkin, 2010). Moreover, these developments in computer technology facilitated a process known as securitization in which small loans are built into standard debt securities. Worth mentioning were a class of highly advanced financial products known as collateralized debt obligations (CDOs). These CDOs were paying out income streams from a collection of underlying assets and were designed to have certain risk characteristics that attracted investors of particular preferences. Mishkin (2010) states that the CDOs were so complex, that it was difficult to assess cash flows from the underlying assets and to identify the actual owners of these assets. These factors opened a door for subprime mortgage borrowers to access residential loans despite their poor credit records. According to The Economist (2007), by the end of 2006 20 percent of all new mortgages were sub-prime.

3.9.2 Agency Problems

In his speech Bernanke (2007) highlighted that the practices of some mortgage originators contributed to the problems that led to the financial crisis. The mortgage market was characterised by the principal-agent problem. For example, the mortgage brokers that initiated the loans did not make the required effort to ascertain whether the borrower was capable of paying off the loan. The reason being, that the mortgage broker's income was linked to the volume of mortgage deals sealed. As a result the brokers had little incentive to assess the capacity of the borrowers to fulfill their financial obligations. In some instances it is said that mortgage brokers encouraged subprime borrowers to take on loans they could not afford by falsifying information of mortgage applications in order to qualify them (Mishkin, 2010).

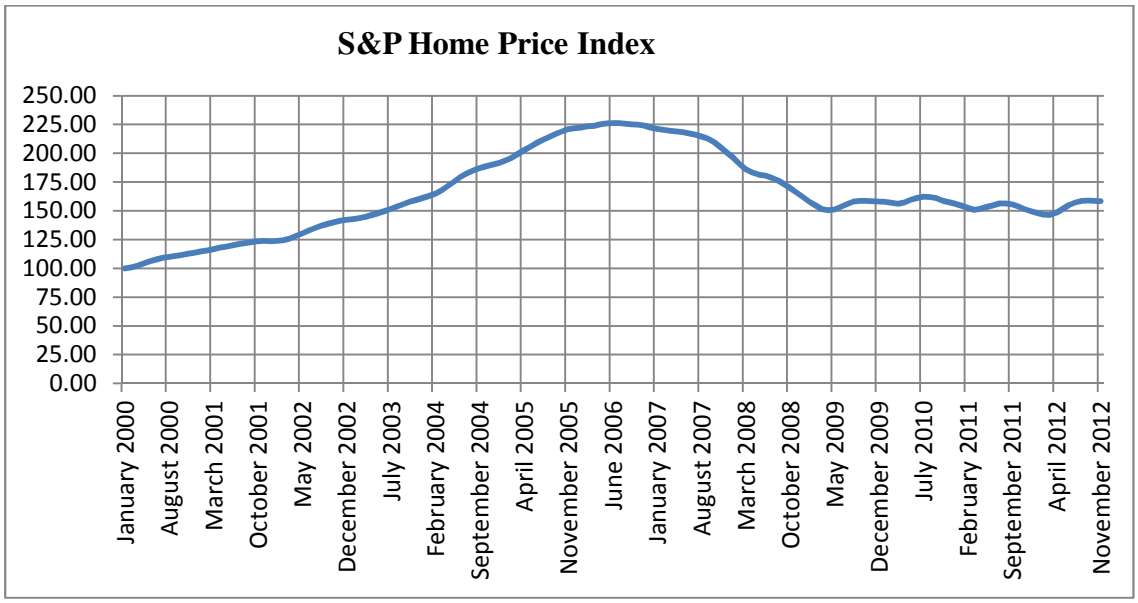
Precipitated by slow pace in the growth of mortgage origination and a strong demand for high-yield securities, the commercial and investment banks weakened underwriting standards. This was motivated by the fact that these institutions were gaining substantial income by endorsing mortgage backed securities and other products such as CDOs. This led to what Bernanke (2007, para.9) termed "*early payment defaults*" which essentially refers to defaults occurring within the few months of origination. Credit rating organisations also contributed to the eruption of the crisis. The credit ratings of the previously discussed CDOs and other securities were characterised by inconsistencies and irregularities. For example, the

credit rating agencies were also earning large incomes from rating debt securities of institutions that had an interest in seeing positive ratings. Hence the credit rating process became compromised and so fuelled the crisis.

3.9.3 Formation and Bursting of a Housing Bubble

The market for houses in the U.S performed a significant part in the outburst of the crisis. Lazarov (2009) states that the 2007 crisis that started in the U.S was precipitated by a combination of the credit boom as well as the formation and bursting of a bubble in the housing market. Policy makers in the U.S in their endeavour to promote home ownership weakened regulations that surrounded access to mortgages. In addition, huge inflows of credit from China, India and other countries into the US created surplus funds that led to competition by financial institutions for borrowers. This led lending institutions to eventually open their doors to sub-prime mortgage borrowers (Tienhaara, 2010). Subprime mortgages are loans that are made to borrowers who are regarded to have a high credit risk. These individuals often lack a strong credit history or have other undesirable characteristics that culminate to high probabilities of default. Inspired by financial liberalisation and democratisation of credit, a large number of these sub-prime borrowers were able to access mortgage loans and buy houses. Between the periods 1997 and 2006, the Standard and Poor (S&P) home price index reported American house prices to have risen by 124 percent (see Figure below). Mishkin (2010) states that these high prices implied that the appreciation of the value of houses could enable borrowers to refinance their houses with even larger loans. Even in the worst scenario, these sub-prime borrowers had an option of selling their houses to settle the loan. However, market participants did not anticipate any downward trend in house prices as a result the boom in house prices continued. After reaching their peak in 2006, the house price bubble eventually burst resulting in many subprime borrowers defaulting on their obligations as they realised that the value of their houses was now lower than the amount of their mortgages.

FIGURE 3.17: THE US HOME PRICE INDEX



Source: S&P Home Price Indices Economic Indicators, 2012

3.9.4 Studies in Developing and Developed Economies

Financial crises occur when an increase in asymmetric information triggered by a disruption in the financial system causes severe adverse selection and moral hazard problems that eventually cause financial markets to fail to execute their intermediary function efficiently (Mishkin, 2010). Therefore, if the banking system which plays the most important role within the financial system fails to facilitate the efficient flow of funds from savers to borrowers, then productive investment opportunities which are the output of banks are lost in the process. For instance during the crisis, the banking sector of most countries particularly advanced economies incurred additional cost of dealing with increasing non-performing loans and declining bank outputs. According to Banking Association South Africa (2010), the economic recession in 2009 which was a consequence of the global financial crisis affected consumer affordability and therefore spending patterns, resulting in consumers being reluctant to take on more debt. This resulted in the increase in non-performing loans which had a huge impact on the banks' loan books and so total assets and liabilities declined in 2009.

Although there is so much writing about the 2007/2008 global financial crisis and its impact in developing and developed countries in general, little has been done to investigate the efficiency and productivity of the banking system of these economies during the years of the

crisis. This study is a pioneering work in this regard. The subsequent sections review the few available studies that have attempted to analyse this subject matter in various countries.

In South Africa, Mabwe and Webb (2010) investigated the performance of the five largest commercial banks from 2005 – 2009 covering the period of the global financial crisis. Three aspects of bank performance namely, profitability, liquidity and credit quality were analysed using financial ratio analysis. The findings showed that overall performance improved considerably in 2005 and 2006. However, the impact of the global financial crisis was evident when the overall performance deteriorated in 2007 until 2009. Using the Student *t* test to test if there was any significant difference in profitability performance for the period 2005-2006 and the period 2008 – 2009, their results indicated that profitability deteriorated during the later period. However, they concluded that the South African banking system remained stable as there were adequately capitalised and profitable. Our study will complement Mabwe and Webb (2010) who analysed efficiency using financial ratios. Ncube (2009) states that despite the simplistic and straightforwardness of financial ratios, however, their exclusive use may not give enough insight on actual efficiencies. In addition, while this study applies the parametric Student *t* test, we will check for robustness of results by employing the non-parametric Wilcoxon matched-pairs signed ranks test in line with a study by Sufian (2009) who used the same technique to analyse the impact of the Malaysian financial crisis. Hence our study differs from that of Mabwe and Webb (2010), in terms of an alternative technique which preserves and complements the non-parametric context of the DEA methodology. The Wilcoxon matched-pairs signed ranks test is generally considered to be the non-parametric equivalent of the Student paired *t*-test. In a study of a sample of 269 banks, Wheelock and Wilson (1995) also applied the Wilcoxon matched signed ranks test to examine the significance of the differences of technical and allocative efficiency measures across different models. The models considered in the analysis included the production approach, the intermediation approach and various blends of both approaches. Their study showed that, on average, both technical and allocative efficiency was higher under the intermediation approach compared to the production approach.

Ngo and Nguyen (2012) undertook a study to evaluate the efficiency and TFP changes of Thai banks from the period 2007 – 2010. Their study employed the DEA and the Malmquist index approach on a panel of 27 major banks in Thailand. They found that the effect of the crisis was not immediate but that banks experienced a deterioration of TFP in 2010. For the

period 2007-2010, most of the efficiency changes were slightly above unity showing small efficiency improvements. However, pure efficiency change dropped from 100 percent in 2007 to 95.4 percent in 2008. On comparison, their study also showed that the effect of the financial crisis was not as severe as the regional Asian financial crisis of 1997. They also noted that local banks remained resilient and maintained their stability while foreign banks were more fluctuating with some improving and others getting worst. They highlighted that the crisis had little impact in Thailand on account of two reasons. Firstly, Thailand had experienced the 1997 Asian financial crisis before and as such it had prepared them well in terms of increased supervision, policies for risky management of banks and control of foreign capital inflows. Secondly, Thailand had low reliance on foreign sources of funding as well as its own exposure to foreign assets. Foreign banks in Thailand represent only 10 percent of the total banking sector assets.

In Kuwait, Abdmoulah and Laabas (2012) applied the SFA technique to investigate technical and allocative efficiency of the commercial banks from 1994-2009 covering the period of the subprime financial crisis. The results showed that Kuwaiti banks experienced two efficiency declines, one in 2000-2002 and the other serious downtrend since 2007. They noted that the decrease in efficiency in 2000-2002 was predominantly technical rather than allocative while in 2007 both technical and allocative efficiency deteriorated leading to greater loss in overall efficiency. They suggested that the decline in efficiency was likely caused by declining total banking income which was precipitated by lower investment activity. Moreover, they noted that Kuwati banks had no ability to make adjustments during the crisis in the short term. Hence it is important for policy makers, bank managers and Central bank supervision departments to ensure that banks are better prepared to respond and adjust in case of another financial crisis in the future.

Alzubaidi and Bougheas (2012) conducted a study to investigate the effect of the global financial crisis on the efficiency of European banks over the period 2005-2010. The non-parametric DEA methodology was employed to analyse bank efficiency for 225 sampled EU banks. They then compared the computed efficiency scores for the pre-crisis two-year period (2005-2006) with their corresponding scores for the post-crisis two year period (2009-2010). They found that the overall technical efficiency fell from a mean score of 68.4% to 56.1%, indicating a considerable decrease in efficiency of 12.3% implying that banks were generally unable to use their input resources efficiently after the crisis. They observed that the bigger

portion of this fall was due to input-orientated pure technical efficiency which fell by 8.9%. Scale efficiency decreased by 6.7% from 90.2% to 83.5% as a result of the crisis. Alzubaidi and Bougheas (2012) suggested that the decrease in banking efficiency may have been because interbank lending fell because of the uncertainty of investing in one another. Mckibbin and Stoeckel (2009) state that the risk premium on interbank borrowing increased sharply by 5%; loan loss provisions increased due to a rise in non-performing loans and the quality of assets held by banks decreased in general. The authors also noted that the effect of the crisis was not uniform across the EU area. They cited Belgium and Denmark as being worst affected by the crisis followed by institutions in Ireland, Greece, Finland and Netherlands. The authors also noted that the impact of the crisis was differentiated according to specialisation. For illustration, the largest drop was observed in the efficiency scores of commercial banks followed by savings banks and real-estate banks while cooperative banks were the least affected.

In Australia, Vu and Turnell (2011) conducted a study to evaluate bank profit and cost efficiencies as well as to assess the effect of the global financial crisis. The parametric stochastic frontier analysis (SFA) was employed over a 13 year period from 1997 until 2009. The studied panel consisting of 8 Australian banks was categorised into 4 major banks and 4 regional banks. A dummy variable was included in their model specification to take the value 1 for the period 2007-2009 and 0 otherwise. The other explanatory variables included loans to deposit ratio, total assets, equity capital to total assets ratio, non-interests expense to total assets ratio, net interest margin and return on equity. The obtained findings indicated that profit and cost efficiency scores were 87.76% and 87.45% respectively highlighting the capacity of the Australian banking industry to increase profits by 12.24% and to decrease costs by 12.55% in order to become efficient. During the global financial crisis, major banks and regional banks experienced a reduction of around 26% and 43.5% in profit efficiency respectively. With regard to cost efficiency, the decrease for major banks and regional banks was 5.03% and 5.7% respectively. The deterioration in efficiency was attributed to the fact that the financial crisis induced banks to stiffen their lending criteria in order to lessen their risk exposure. Consequently, this led domestic credit to expand at a comparatively lower rate of 4.5% in 2008 relative to a previous five year average rate of 14%. The global financial crisis dummy variable`s coefficient proved to be negative and significant under the profit efficiency model but insignificant under the cost efficiency model. The major banks were found to be relatively less cost efficient but better at generating profit in comparison to

regional banks. This implied that major banks were generally more inclined towards the maximisation of profit as opposed to minimisation of costs. Moreover, the results revealed that bank capitalisation and loans to deposits variables were major factors behind cost and profit efficiencies in Australia. Hence, it was suggested that bank management and policy makers could improve bank efficiency by ensuring that bank capitalisation is maintained at optimum levels that are consistent with capital adequacy requirements as well as ensuring better management of liquidity funding.

Another study of financial crises and bank efficiency was undertaken by Sufian (2009) when he adopted a two stage methodology framework to investigate the efficiency of Malaysian and Thailand banks during the 1997 Asian financial crisis. These two Asian countries were selected as there were the worst affected by the Asian financial crisis. The author first employed DEA under the VRS assumption and then a Tobit model for the period 1995-1999 to ascertain the relationship of DEA efficiency scores and the explanatory factors. Dummy variables were used to capture the pre-crisis (1992-1996), crisis period (1997) and the post crisis period (1998-2003). Three approaches to input and output variable definition were used in their study namely, the value added approach, the operating approach and the intermediation approach. The reason being that DEA results are sensitive to the nature of variables chosen. Under the intermediation approach, deposits along with labour and capital are regarded as inputs used in the production of loans and other securities. The value-added approach on the other hand considers both deposits and loans as outputs. Finally, the operating approach also known as the income approach regards total revenue i.e. interest and non-interest revenue as outputs while total expenses are considered inputs. He found that these different approaches produced deviating sets of efficiency scores. However, regardless of the approach employed, the results showed high levels of bank inefficiency especially the year following the crisis. This deterioration was particularly more evident when the intermediation approach was applied. He noted that most of bank inefficiency emanated from the fact that input resources were being underutilised. For Malaysian banks technical efficiency deteriorated sharply in the period 1998 with the intermediation approach whereas under the value added and operating approaches it only decreased in the period 1999. For Thailand, the results showed that TE decreased instantaneously due to the 1997 crisis under both the value added approach and the intermediation approach. Within the operating approach framework, Thailand banks experienced their first TE decline in the year 1999.

Their findings also showed that during the crisis period the Malaysian banking sector performed relatively better in terms of efficiency.

It is clear from the brief literature review that indeed not many studies have examined financial crises and their associated effects and in particular the impact of the recent global financial crisis on the efficiency and productivity of the banking sector of developing economies. To the best of the researchers' knowledge no study has examined the impact of the global financial crisis on banking sector efficiency in South Africa.

3.10 BANKING EFFICIENCY AND EMPLOYMENT

Unemployment is one of the crucial challenges facing South Africa with the official unemployment rate at 25.7% and nearly half of the young people unemployed. South Africa's unemployment rate remains far too high by local and international standards, and evidently contributes to much of the social fragmentation and anguish experienced in South Africa today. The 2011 Annual Budget highlighted unemployment as a problem which could have many adverse consequences if not curbed. Ikhide (2008), states that the basic benefit of enhanced efficiency is a reduction in spreads between lending and deposit rates thereby stimulating an increased demand for loans for industrial investment and greater mobilisation of savings through the banking system. This eventually increases investment, employment and production of output thus contributing to decreased unemployment in the economy. STATSSA (2010) Labour Force Survey established that the financial services industry in South Africa was the third largest employer in the country during the second quarter of 2010, representing about 13.1 % of total employment. Thus, there are both direct and indirect benefits of banking efficiency for employment. For these reasons, this study also seeks to shed some light on whether there exist a link between efficiency and employment. To the knowledge of the writer this relationship has not been explored in the banking literature in South Africa.

Many studies have looked at the linkages between bank performance and economic growth. In recent years, because of the financial crisis that eventually led to the current global financial recession, there has been a significant increase in unemployment in both developing and developed economies. As alluded to in the introductory chapter, one of the focus areas of this thesis is to investigate the relationship between bank efficiency and employment. The related research question is whether an increase in efficiency in banking has a negative

bearing on unemployment in South Africa. There are no papers in which authors have studied the connection between bank efficiency and employment.

3.11 STUDIES ON BANK EFFICIENCY IN SOUTH AFRICA

Okeahalam (2006) employed the Bayesian stochastic frontier approach to assess the production efficiency of 61 South African bank branches in the 9 provinces for the year 1999. The author found productive efficiency of banks to be 83.1% suggesting that on average banks could reduce their costs by 16.9% without altering their current output levels. Okeahalam also found that all branches were operating at increasing returns to scale and recommended levels of output to be increased either through regulatory reforms or competitive incentives. Hence, there is need by banks to expand their services to the unbanked population of South Africa as this will lead to efficiency improvement. The efficiency gains realised may also be transmitted to the general public in the form of reduced prices. While the study by Okeahalam (2006) analysed branches of one of the largest four banks in South Africa, this current study will widen the analysis to explore all the four largest banks including four others together representing over 90 percent of the entire banking sector. More over our analysis will extend beyond one period in order to analyse total factor productivity efficiency over time.

Van der Westhuizen (2008) employed DEA and Malmquist to evaluate the scale and technical efficiency and productivity changes of top four South African banks using balance sheet monthly data over a three year period. The results for scale and technical efficiency under both the input and output orientation were above 90 %. The findings also show that three banks operated on the increasing returns to scale region while one bank exhibited decreasing returns to scale. This presents an important implication to the present study because the presence of increasing returns scale implies that banks need to expand their services to the unbanked sectors or the unexplored markets. The obtained Malmquist productivity indices with variable returns to scale (VRS) specification for the four banks were as presented below. Where TFPCH, SECH, TECHCH and EFFCH denote total factor productivity change, scale efficiency change, technical change and efficiency change respectively.

Top Four Banks	TFPCH	SECH	TECHCH	EFFCH
A	1.016	1.000	1.016	1.000
B	1.005	0.999	1.006	0.999
C	1.020	1.001	1.019	1.001
D	1.004	1.001	1.006	1.002
Mean Score All four	**1.011**	**1.000**	**1.012**	**1.001**

The four banks were observed to have the potential to either decrease input consumption without decreasing outputs or increase the amount of outputs generated without requiring additional inputs. Moreover, no single bank had an optimal size as they were all operating either on the IRS portion or DRS region. However, empirical studies done by [Coelli and Rao (2005), Grifell-Tatje and Lovell (1995), Briec and Kerstens (2011),] have criticised the inadequacy of the Malmquist approach particularly on the basis that under the VRS assumption the outcome of calculating the Malmquist TFP index result in inaccurate measures of TFP changes. Hence this thesis fill this gap in literature by departing from the Malmquist productivity analysis approach used by Van der Westhuizen (2008) and instead apply the multiplicatively-complete Hicks-Moorsteen index approach.

Ncube (2009) conducted a study to analyse the cost and profit efficiency of banks in South Africa. Their study employed the parametric stochastic frontier approach to determine both cost and profit efficiency of four large and four small South African banks over the period of 2000-2005 classified according to the number of employees. The large banks whose number of employees exceeded 10 000 included ABSA, FirstRand Bank, Nedbank and Standard Bank. The small banks with employees less than 10 000 included African Bank, Capitec Bank, Investec Bank and Teba Bank. The average cost and profit efficiencies over the six periods were 92% and 55% respectively. Their study concluded that South African banks were relatively better at controlling cost than generating profit as indicated by the lower profit efficiency score and a higher cost efficiency score.

Kiyota (2009) conducted a study to determine the both profit and cost efficiency of commercial banks located in 29 Sub-Saharan African (SSA) countries for the period 2000 to 2007. A two-step methodology was adopted. Within the first step, the translog stochastic frontier approach was employed to calculate the profit and cost efficiency scores of the

sampled banks. In the second step, the Tobit model was utilised to examine the impact of environmental factors on efficiency of the banks. More important to the study was to establish if foreign SSA banks performed better compared to domestic SSA banks. Regarding profit efficiencies, their findings revealed that foreign banks were more efficient than domestic banks. Non-SSA foreign banks were also found to be relatively more cost efficient in comparison with SSA foreign banks and SSA domestic banks. The study also showed that medium size and large banks were cost efficient while smaller banks were profit efficient.

Oberholzer et al (2010) applied the DEA methodology to five of the largest banks in South Africa for the ten year period 1998 - 2007. The total assets of these five banks as at the end of 2007 represented 85.28% of the total banking assets in South Africa. Two DEA models were employed based on the definition of outputs included. Model 1 used only income statement data as outputs namely, the value of interest income and non-interest income. Under model 2, only balance sheet data outputs were considered, that is, the value of deposits, loans and equity. The empirical findings from DEA analysis revealed that the average technical efficiency of all the banks was 89.5% and 79% for Model 1 and Model 2 respectively. Therefore banks had the capacity to increase output by 10.5% and 21% without increasing their inputs, respectively. In addition, they found that the average allocative efficiency of all the banks were 98.5% and 89.3% respectively.

Van der Westhuizen (2010) undertook a study in South Africa to shed light on whether the income structure of a bank affects its efficiency. To address this research question, the DEA methodology was employed to evaluate four types of efficiency measures namely, technical (TE), allocative (AE), cost (CE) and scale efficiency (SE). A sample of four top South African banks was evaluated for the 10 year period from 1998 to 2007. Two models were constructed depending on whether the income was interest or non-interest. Interest income represents a bank`s traditional activities while non-interest income represents diversification from traditional activities. Hence, Model 1 and Model 2 were used to generate bank efficiency measures under the traditional and modern function of a bank. His findings indicated that banks showed improvements in average TE, CE and SE as they change from model 1 to model 2. The results obtained for all banks were as presented in the table below:

	TE	AE	CE	SE
Model 1 (Interest Income)	0.876	0.932	0.814	0.955
Model 2 (Non-interest income)	0.957	0.900	0.861	0.959

The results demonstrated clearly that as banks change to model 2 from model 1, they exhibited greater efficiency measures except for allocative efficiency. Hence, the author concluded that non-interest income contributed towards better technical efficiency, cost efficiency and scale efficiency. In the final analysis, it was concluded that banks should concentrate on non-interest income as a major source of income by diversifying their income from their traditional function to that of producing a wider range of services.

Maredza and Ikhide (2013) investigated the impact of the sub-prime financial crisis on the efficiency and productivity of the South African commercial banking sector using a two-stage methodology framework for the period 2000-2010. The Hicks-Moorsteen total factor productivity (TFP) index approach developed by O`Donnell (2010a) as opposed to the popular Malmquist TFP was utilised. Their first stage results showed that during the crisis period there was a noticeable but mild deviation of total factor productivity and efficiency measures. However, their second stage analysis using the censored Tobit model showed that the financial crisis was the main determinant of bank efficiency, indicating that total factor productivity efficiency was 16.96% lower during the crisis period compared to the pre-crisis period. Their study was the first in the South African banking sector to adopt the multiplicatively complete Hicks-Moorsteen approach. As stated before, this study adopts the same methodology to explore the link between bank efficiency and access to bank services.

3.12 STUDIES ON EFFICIENCY IN DEVELOPING ECONOMIES

Another study was conducted by Sathye (2003) to evaluate the productive efficiency of the Indian banking sector for the period 1997-1998. The DEA methodology was applied to three categories of banks namely, public, private and foreign. Of the 103 number of banks sampled, 27 were public owned, 34 privately owned and 42 foreign owned banks. In order to investigate how efficiency indicators change as inputs and outputs change, two models were constructed. Under model A, the input factors included both interest and non-interest expenses while output variables consisted of net interest and non-interest income. Within Model B inputs included size of staff complement and total deposits while outputs consisted of non-interest income and net loans. The findings indicated that under model A and B, the

mean efficiency scores of Indian banks were 0.83 and 0.62 respectively. They also reported that more foreign banks were in the highest efficiency quartile than private or public sector banks. As per model A, the public banks were more efficient relative to both foreign and private banks. As per model B, foreign banks were relatively more efficient in comparison to the other two categories while public banks performed better than private banks. Their findings also showed that in both models, the mean efficiency score for private sector banks was the lowest. They highlighted the reason that private banks in India, at the time, were in an expansion phase and that their higher fixed assets were yet to generate the commensurate returns.

Figueria et al (2006) undertook a study to investigate the impact of foreign ownership on bank performance and to establish if private banks performed relatively better than government owned banks. The study aimed to answer the question: does the type of ownership have any bearing on efficiency of African banks? Three methods were used to assess performance namely, bank performance ratios, DEA and the cost SFA. The results from each model were compared to check robustness. The study revealed that in Africa there was little evidence that privately-owned banks performed better than their state owned counterparts. But there was evidence that foreign owned banks were more efficient than domestically owned banks. Moreover, there was a suggestion that differences in performance may not have only been related to bank`s ownership but that they might have been associated with the national regulatory and economic environment in the countries in which banks operate.

In a study of Namibian commercial banks covering the period 1993 – 2006, Ikhide (2008) investigated operational efficiency using both the parametric cost SFA and financial ratios. Total loans and advances were considered as bank outputs whereas bank inputs consisted of deposits, capital and labour. The results showed that the Namibian banking sector was representative of economies of scale and that these economies could be realised by increasing the volume of operation. Again, evidence from the study established a low level of substitutability among input factors within the commercial banking industry.

In addition to the above study, Frimpong (2010) undertook a study to investigate the state of efficiency of Ghanaian banks in 2007 and to explore the linkage with profitability. The study employed the intermediation approach with deposits and total expenditures representing

inputs whereas outputs comprised advances and investments. The sample of banks used in the study consisted of eleven foreign banks, eight private local banks, and three government owned banks. In a two-step procedure, the DEA technique was employed to estimate efficiency scores. Their first stage findings reported an average TE score of 74% and that 18% of the banks were efficient while the rest (82%) had efficient scores ranging from 33% – 89%. The local privately owned banks proved to be the most efficient category with a score of 87%. Foreign owned banks were relatively more efficient with an average TE score of 72% compared to government-owned banks with an average 51%. In the second step, the efficiency-profitability matrix applied in the original work of Boussofiane et al (1991) was utilised to explore the relationship between efficiency and profitability. Four quadrants were identified namely, *star, sleeper, question* and *dog*. *Star* banks were those that achieved both superior TE and profitability; *sleeper* banks consisted of those banks that were highly profitable yet inefficient; *Question mark* banks were those lagging in profits by reason of their technical inefficiency; *Dog banks* consisted of those banks that operated at high efficiency but low profitability. Second stage analysis indicated that (6) 27%, (7) 32%, (7) 32% and (2) 9% of 22 banks included in the study were located within the star, dog, question and sleepers category respectively. The author concluded that the fact that 41% of the banks were located within the sleeper and question mark category was indicative of considerable inefficiencies existing in the Ghanaian banking industry.

Arjomandi and Valadkhani (2010) employed the Hicks-Moorsteen TFP (HMTFP) index approach to examine efficiency and TFP changes in the banking system of Iran for the period 2003 – 2008. The objective of their research was to investigate the effect of government`s regulations which were announced in 2005 on the performance of the banking industry. They found that although the banking system was inefficient, the technical efficiency level deteriorated significantly in 2006 following the introduction of regulation. Their study was the first to adopt the HMTFP approach to analyse productivity changes in banking. They employed the HMTFP approach over the Malmquist approach stating that with a VRS technology, the latter fail to estimate productivity changes accurately. This is consistent with developments by O`Donnell (2010) who showed that the Malmquist is an unreliable and biased method particularly when employed under a VRS specification.

Gitau and Gor (2011) conducted a study to measure productivity growth and its component for a panel of 34 Kenyan banks from 1999 – 2008. The DEA Malmquist TFP index was

employed in the study with a CRS specification. The obtained productivity index was also decomposed into technical efficiency and technological innovation. The input vector included loanable funds, capital and labour while the output vector consisted of customer net loans, other earning assets (such as loans by sectors and interbank funds) and investment securities (such as government bonds, treasury bills and other securities). Their results also showed that 26% of the sampled banks recorded productivity growth while the rest experienced a decline in productivity. The highest productivity growth recorded was 7.3% while the highest decline was 13.2%. Again, 65% of banks experienced efficiency growth while 35% experienced a decline. Only 18% of the banks had a technical efficiency growth while 82% experienced a decline. Their results were as presented in the table below:

Classification	**EFFCH**	**TECH**	**PECH**	**TFPCH**
Whole Sample	1.006	0.967	1.005	0.973
Small	0.999	1.059	0.999	1.058
Medium	1.002	1.004	1.001	1.006
Large	1.002	1.011	1.001	1.013
Foreign Owned	1.003	0.978	1.002	0.981
Locally Owned	1.008	1.030	1.001	1.039
Public Owned	1.000	0.935	1.000	0.935

As presented in their table of results, small banks had the highest productivity growth of 5.8%, large banks had 1.3% while medium banks had the lowest growth of 0.6%. Locally owned banks performed the best with a positive TFP change of 3.9% compared to both foreign and public owned banks that registered negative growth of 1.9 and 6.5 respectively.

In another study of Kenyan banks, Kamau (2011) investigated the state of bank productivity and intermediation efficiency for the period 1997 - 2009. The author adopted DEA to analyse intermediation efficiency and the Malmquist approach to estimate productivity growth. The intermediation approach was followed to define bank inputs and outputs to be used. The mean CRSTE and VRSTE for the entire period was 0.47 and 0.56 respectively. The scale efficiency of Kenyan banks was an average 0.84 implying that banks in Kenya had an inefficiency of 16% due to failure to operate on the optimal scale. Further examination of efficiency scores by size revealed that in general large banks were the most efficient. This was followed by medium banks which performed relatively better compared to small banks. Kamau's results were consistent in that she also found that large banks exhibited the best

performance relative to small and medium banks by a large margin in terms of technological innovation. Given that the composition of large banks was mainly foreign-owned, the author suggested that large banks had resources to spend in new technology which contributed to their increased efficiency. Moreover, the author highlighted that there was possibility of transfer of technology from their mother banks in developed economies. As a result, foreign owned banks proved to be the most efficient while domestic private banks were relatively better compared to domestic public. Foreign banks also showed the highest productivity index. In her conclusion the author recommended small and medium banks to create capacity for innovation in order to enhance their efficiency.

In India, Das and Kumbhakar (2012) analysed the impact of banking deregulation on efficiency and TFP change in the Indian banking industry. The investigation was carried out for a panel of 948 banks for a period of 1996 – 2005. In order to take account of quality differentials, their study used hedonic aggregator functions within the framework of input-distance functions. The cost efficiency of banks during the period of investigation averaged at around 66%. State-owned banks had the highest cost efficiency with a range of 65 – 75% compared to both private and foreign banks whose range was 60 – 70%. This result was consistent with the findings obtained by Sharma (2012) who found that state-owned banks performed comparably better than private banks. The authors stated that firstly it was because quality aspects of inputs and outputs were explicitly accounted for in their analysis. Secondly, they argued that state-owned banks have multiple goals that might have worked in their favour especially so when quality features are integrated. These goals included promoting employment of low-skilled works; allocating more loans to priority sectors; promoting job opportunities in rural areas by opening additional branches and embracing financial inclusion by bringing more people to formal banking. During the post-deregulation period, banks improved their efficiency from 61% in 1996 to 72% in 2005. The authors found that on average TFP growth was above 3.5% annually. In addition, technical efficiency was found to have performed a significant role in the achievement of TFP growth during the study period. In conclusion, the authors highlighted that there was need for focused attention in reducing transaction costs, improving credit delivery mechanism and extending banking facilities to a large segment of the unbanked population.

In another study of Indian banks, Sharma et al (2012) examined the determinants of efficiency and productivity as well as its relationship with bank specific factors. A two stage

methodological framework was applied to a sample of 64 banks for the period 2000 – 2010. The first stage DEA technique was employed to measure efficiency of the Indian banking sector. In the second phase, a censored Tobit model was then adopted to explore the relationship of bank-specific factors with bank efficiency. DEA results revealed that public banks had the highest efficiency scores compared to private and foreign banks. This finding was substantiated by the Tobit results that showed that state ownership of banks was positively and significantly associated with bank performance. Profitability was also found to be positive and significant implying that highly profitable banks were more efficient in their operations. Bank diversification practices were found to be negatively affecting the efficiency and performance of Indian banks. Finally, loan intensity, management quality, market share and bank size were found to be insignificant in influencing bank efficiency in India.

The reviewed empirical literature has therefore shown the various efficiency studies that have been conducted within the banking industry in various countries. Whereas these studies could serve to provide strategic lessons for the banking sector in general and to shed light to the present study, however, none has explored the main issues we have raised in this study. The present study is mainly important in order to inform formulation of appropriate policies for promoting financial inclusion. The section that follows discusses the potential that technology and innovation holds in addressing the key challenges related to access to banking services in South Africa.

3.13 TECHNOLOGY, INNOVATION AND ACCESS TO SERVICES

It has been argued that many of the products that are offered by commercial banks in Africa are poorly adapted to the low-income households` needs reflecting a need for innovation. On the other hand, some economists have argued that the cost of innovations can be so prohibitively high that the net benefit to banks may be marginal. However, a World Bank (2007) study showed that the major cost of these innovations are related to the initial set-up and mainly constitute fixed costs with very low marginal costs per new customer engaged offering greater potential for expanding access especially to those at the bottom of the pyramid. Lawson and Samson (2001, p.378) defines innovation "as the mechanism by which organizations develop new products, processes and systems that are required for adapting to changing markets, technologies and competition". However, Standard Chartered (2013, para.5) states that "Not all innovation involves technology - innovation can simply mean changing the way we do things so that more people can access banking services." They

maintain that mobile bank branches are an innovation in themselves with a focus to bring banking services to remote areas.

In this study we define technological innovation as comprising mechanisms through which technology can be enhanced to generate wealth and in turn contribute toward a better quality of life. Mobile phones are one of the contemporary and technology-based channels available with greater potential for making better the lives of the many "unbanked" South Africans. Ondiege (2010) argues in favour of technology-based banking services and explains that the cost of formal and traditional-based banking in Africa is very high and that even those customers with bank accounts often face high charges for moving their cash around due to high transactions costs. He states that since over 50% of the adult populations have access to mobile phones, this fact could be exploited to expand access to the rural and remote populations. Ondiege (2010) states that mobile banking (M-Banking) offers immediacy and efficiency. It makes basic banking services more accessible by minimising the time and distance to the nearest bank branches at the same time reducing a bank`s transactions costs. CGAP (2010) noted that of the developing countries for which data was available, South Africa had the highest share of the banked population who were accessing banking services via their cell phones. All large commercial banks in South Africa offer mobile banking as an additional channel for existing bank accounts. For example in 2005, MTN Banking, a joint venture of the cellular operator MTN and Standard Bank, developed a mobile phone-based transaction banking product which had specifically designed features for providing banking services to low-income customers (CGAP, 2010). A recent South African study by Musara and Fatoki (2010) showed that technological innovations indeed results in reductions in bank costs for customers and hence increased efficiency. However, their study also highlighted some related technology-based shortcomings such as fraud, technical complications in operating new technologies particularly for the illiterate and the need for personal or human contact. Moreover, some people maintain that M-Banking mainly benefits those in urban settings while posing some challenges to many who are located in rural remote areas.

Another avenue for innovation in recent years has been branchless banking which is the distribution of financial services outside traditional bank branches by making use of non-bank agents and information and communications technologies. Dias (2011) argues that innovation in the form of branchless banking extends opportunities to customers by leveraging existing infrastructure that already reaches unbanked people who otherwise would not be reached at a

profit if the services of traditional branches were used. Therefore technology and innovation have the capability of expanding the scale of access to banking services, reducing costs, reaching clients in remote and rural areas and enhancing quality and diversity of services.

This section has demonstrated great potential that advancement in technology and innovation holds in expanding access to banking services for South Africans in particular and for developing economies in general. It has shown how innovative banking may improve the level of access via enhanced technical efficiency particularly via its impact on low-cost service delivery. However, the initial development cost of these innovations and the risk that the scale of up-take will be too low are among the limitations that have retarded progress (World Bank, 2007).

3.14 CONCLUSION

This chapter reviewed various relevant empirical studies on the crucial subject of efficiency and productivity measurement, bank efficiency and access to banking services, the efficiency-employment nexus and financial crises. We followed key developments that have taken place in South African banking and in the analysis, existing research and methodological gaps were identified and are summarised as follows:

The present study attempts to fill the gap in the literature by departing from the traditional methodological approach of measuring and analysing efficiency and productivity. Analysis of efficiency and productivity have extensively applied the Malmquist TFP index approach which empirical evidence proved to be biased and inconsistent. Moreover, this study investigates the impact of the global financial crisis on TFP efficiency of the South African banking system using a two-step procedure and the non-parametric Wilcoxon matched-pairs signed ranks test. This approach represents a significant departure from the analysis of Mabwe and Webb (2010) who evaluated bank efficiency in South Africa using financial ratios and the parametric *t*-test to analyse the change in efficiency during the global financial crisis. Regarding access to banking services, this study complements the work done by Hawkins (2010) and Okeahalam (2008) by broadening the scope of bank-specific explanatory variables particularly total factor productivity efficiency. Finally, to date, no known studies have explored the link between banking sector efficiency and unemployment.

CHAPTER FOUR

METHODOLOGY

4.1 INTRODUCTION

This chapter discusses the first stage and second stage model specifications, data analysis and definition of variables to be used. In the first stage in which TFP is measured and subsequently decomposed into several efficiency indicators, the Hicks-Moorsteen Total Factor Productivity (HMTFP) DEA methodology will be utilised. As stated in the previous chapter, this particular approach was developed in order to redress the inadequacy of the popular Malmquist TFP index. The previous chapter has also discussed extensively the mathematical formulation of the DEA methodology, the basic constant returns to scale (CRS) model also known as the CCR[9] model and the variable returns to scale (VRS) model which is an extension of the CCR model. The VRS model is also known as the BCC[10] model. In order to actually measure and decompose productivity into various efficiency indicators, a computer program called DPIN which uses DEA linear programs for **d**ecomposing **p**roductivity **i**ndex **n**umbers is utilised.

4.2 MODEL SPECIFICATION

4.2.1 Choice of DEA Model for Empirical Analysis

The empirical model to be adopted in this study constitutes the BCC model. As previously stated, the assumption of the CRS DEA model is only suitable for situations where all banks are operating at their optimal scale. This assumption may not be realistic for banking institutions. The concept of efficiency adopted in this study is that of technical and cost efficiency and both the input and output orientation results will be reported. Cost efficiency is implied in the sense that our bank input variables constitute among others total operating costs. Here the notion of cost efficiency is inferred from technical efficiency which essentially measures a bank`s ability to produce a given set of outputs (total loans, interest income, non-interest income) using the least possible amount of inputs (inter alia operating cost). In view of the research objectives it is assumed that these efficiency concepts are appropriate and sufficient to shed light on this study. As hypothesised earlier the likely

[9] After its pioneers Charnes, Cooper and Rhodes (1978).
[10] After its pioneers Banker, Charnes and Cooper (1984).

channel through which gains in bank efficiency may feed through to improved access and employment is mainly via bank`s operating cost.

This study adopts a two stage methodology framework. Within the first stage, the HMTFP indices are generated and further decomposed into several types of output- and input-oriented efficiencies (O`Donnell, 2011) using the DEA program-based software called DPIN 3.0. These generated performance measures are analysed to determine if there has been a change in the efficiency and productivity of South African banking system during the period of the global financial crisis. In the second stage, the relationship between bank efficiency and access to banking services is assessed. The second stage analysis also seeks to shed light on the third sub-objective of this study namely; to establish the relationship between efficiency improvement and unemployment in South Africa.

4.2.2 Choice of DEA Model Orientation

Evaluated under an input orientation model, a bank is considered inefficient if it is possible to reduce any input without expanding any other input and without decreasing any output. Similarly a bank is deemed inefficient in an output oriented model if it is possible to expand any output without increasing any input usage or decreasing any other output. Therefore, inputs are considered to be controllable under input-orientation whereas outputs are deemed controllable in the output oriented model. Coelli (1996) suggests that a researcher should choose an orientation according to which of the two quantities of inputs and outputs the managers have control over. In the case of the CCR model, the input and output orientation measures of efficiency are always the same but yield different results under the BCC model. Therefore Martič et al (2009) state that with CCR models one can solve either the input oriented or output oriented model and give either of the two interpretations. On the contrary, if the BCC output model is adopted one can only give the output interpretation. In a bid to effect concrete policy positions and to deepen one's understanding of how banks utilise their resources to meet their goals, both orientations were adopted. The DPIN program for estimating TFP and decomposing TFP into various efficiency scores naturally reports both input-oriented and output-oriented measures.

4.3 DATA ANALYSIS AND DEFINITION OF VARIABLES

4.3.1 Sources of Data

To investigate the relationship between bank efficiency and access to banking services in South Africa, the researcher analysed a panel of eight South African banks classified as large (4 banks) and small (4 banks) for the period 2003 – 2011. Presently, there are 17 domestically owned banks that are registered with SARB, 14 branches of foreign banks and 3 mutual banks. However, the banking market is dominated by the four largest banks known as the "Big Four" which together contributed 84.1 percent to the balance-sheet size of the total banking sector at the end of December 2011 (SARB Annual Report, 2011). The big-four banks consist of Amalgamated Bank of South Africa (ABSA), FirstRand Bank (FRB), Nedbank, and Standard bank while the four small banks include African bank, Capitec, Sasfin, and Ubank (formerly known as Teba). The classification is based on the total assets of each bank as at 31 December 2011 (see Table 4.1 and Appendix 13).

TABLE 4.1: THE NUMBER AND CLASSIFICATION OF BANKS IN THE SAMPLE

LARGE BANKS	Total Assets (R Millions)	SMALL BANKS	Total Assets (R Millions)
STANDARD	889 250	AFRICAN BANK	49 236
ABSA	725 679	CAPITEC	22 230
FRB	665 525	UBANK	3 586
NEDBANK	585 033	SASFIN	2 767

Source: SARB Supervision Department, Annual Report, 2011

TABLE 4.2: LARGE BANKS IN TERMS OF OTHER VARIABLES

Year	Bank	Total Loans (Rm)	Interest Income (Rm)	Non-interest Income (Rm)	Customer Deposits (Rm)	Operating Expenses (Rm)	Number of Employees	Fixed Assets (Rm)
2011	**ABSA**	505462	49210	16514	431762	49517	35200	7268
2011	**FRB**	533347	41455	28578	472283	47875	36398	12026
2011	**NEDBANK**	490539	42880	15033	472740	44270	28494	6312
2011	**STANDARD**	561552	48196	18071	623295	50195	28422	8430

Source: Bankscope database: www.bvdinfo.com

TABLE 4.3: SMALL BANKS IN TERMS OF OTHER VARIABLES

Year	Bank	Total Loans (Rm)	Interest Income (Rm)	Non-interest Income (Rm)	Customer Deposits (Rm)	Operating Expenses (Rm)	Number of Employees	Fixed Assets (Rm)
2011	**AFRICAN**	41787	7647	7295	1666	7780	4978	852
2011	**CAPITEC**	18408	4347	2320	11660	3509	7194	543
2011	**UBANK**	1055.9	367.5	230	2882.3	479.2	793	47.8
2011	**SASFIN**	2931	434	425.4	1787.3	706.6	664	57.4

Source: Bankscope database: www.bvdinfo.com

Table 4.1 presents the classification of banks in each size category according to total assets while Table 4.2 and 4.3 shows other specific input and output variables for large and small banks respectively. Bank input and output data for the computation of bank efficiency and productivity measures within the first stage DEA analysis were obtained from Bankscope database (see appendices). Financial access survey data compiled by IMF (2012) for the period 2004 – 2011 was used to capture the level of access to banking services. Annual data used in the second stage analysis was obtained from three sources within the Quantec database covering the period 2004 - 2011. These sources include International Monetary Fund (IMF), Bureau for Economic Research (BER) and South African Reserve Bank (SARB).

4.3.2 Definition of Banking Input and Output Variables

An essential consideration when evaluating efficiency and productivity within the banking sector is the choice of inputs and outputs. When defining the inputs and outputs to be adopted two basic approaches are followed: the production approach and the intermediation approach also known as the asset approach. The *production* approach regards banks as firms that use inputs such as labour and capital to produce outputs such as deposits, loans and advances. On the other hand, the *intermediation* approach recognises the intermediary role of banks as accepting deposits to produce the outputs which are loans and advances. In this study, the intermediation approach is adopted. Previous studies that have used this approach inter alia include [Elyasiani and Mehdian (1990) and Berger and Humphrey (1997)] who argue that the intermediation approach is more desirable than the production approach as it is more inclusive of interest expenses which generally account for over fifty percent of total banking expenses. Elyasiani and Mehdian (1990, p. 543) also maintains that deposits are more suitably classified as inputs than as outputs since banks "buy rather than sell deposits".

Labour, total operating expenses, fixed assets, and total deposits are considered inputs while interest income, non-interest income and loans & advances are considered as output variables. Each variable has been chosen to reflect important characteristics of the main activities of commercial banks as indicated in South Africa`s banking industry and empirical literature. These variables and their descriptive statistics are presented in Table 4.4 and Table 4.5 respectively.

TABLE 4.4: VARIABLES FOR FIRST STAGE DEA ANALYSIS.

	DEFINITION
OUTPUTS	
Y_1	Interest Income
Y_2	Non-interest Income
Y_3	Gross Loans
INPUTS	
X_1	Total number of Employees
X_2	Total Operating Expenses (Total interest expenses plus total non-interest expenses)
X_3	Fixed Assets
X_4	Total Customer Deposits

TABLE 4.5: DESCRIPTIVE STATISTICS FOR FIRST STAGE VARIABLES

	Interest Income	**Non-interest Income**	**Gross Loans**	**Total Customer Deposits**	**Number of Employees**	**Fixed Assets**	**Operating Expenditure**
Mean	22008.13	8078.294	195072.4	179536.5	16581.35	2765.674	22907.53
Median	14957	6805	114158	78943.05	14148.5	1460.5	15763
Maximum	82797	37665	561552	623295	39738	12026	79746
Minimum	134.7	4.3	163.6	48.9	435	38.7	280.5
Std. Dev.	23148.47	8563.035	207069	195640.5	15145.26	3177.599	23683.81
Obs.	72	72	72	72	72	72	72
Cross-sections	8	8	8	8	8	8	8

4.4 FIRST STAGE ANALYSIS

As stated before, this study adopts a two-step methodology. In the first step, the efficiency measures are generated by the non-parametric program called DPIN 3.0 which uses the DEA methodology framework. The DEA models were extensively discussed in the preceding chapter. The generated measures are then analysed to determine if there was a significant change in the efficiency of South African banks during the period of the global financial crisis. Therefore, this section outlines the hypothesis testing procedure for carrying this investigation.

4.4.1 Testing Equality of Means

In order to investigate whether the difference in efficiency of banks in the pre-crisis period (2003 – 2006) was statistically and significantly different from the crisis period (2007 – 2009), the student`s *t*-test was employed. The test is performed to investigate the null hypothesis that the means of these two periods are identical against the alternative that they are different. Denoting the pre-crisis mean efficiency with μ_1 and the crisis period mean efficiency with μ_2, the test is set up as follows:

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

This equality of mean test follows a *t*-distribution:

$$t = \frac{\hat{\mu}_1 - \hat{\mu}_2}{se\ (\hat{\mu}_1 - \hat{\mu}_2)} = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{(Var(\hat{\mu}_1) + Var(\hat{\mu}_2) - 2Cov(\hat{\mu}_1, \hat{\mu}_2))}}$$

The calculated *t*-value is compared to the critical *t* value at the 0.05 level of significance. If the calculated *t*-value exceeds the critical *t*-value at 0.05 level of significance the null hypothesis is rejected. Alternatively if the *p*-value of the *t*-statistic is lower than 0.05 the null hypothesis can be rejected in favour of the alternative hypothesis.

Despite its simplicity, this approach essentially measures whether two independent means are equal therefore, in view of our objective is sufficient and appropriate to shed light to this study. This study has emulated Mabwe and Webb (2010) who applied the same student *t*–test to investigate the hypothesis that bank performance means (profitability, liquidity, and credit quality) for the period 2005 – 2006 was the same as for the period 2008 – 2009.

4.4.2 Wilcoxon Matched-pairs Signed Ranks Test

To check the robustness of the *t*-test findings, the non-parametric Wilcoxon signed ranks test for paired changes is performed to evaluate differences between the pre-crisis and the crisis period. Shaw et al (2000) state that the Wilcoxon test is a non-parametric equivalent of the parametric t-test. In this scenario the research emulates Sufian (2009) who applied both the student *t*-test and the Wilcoxon test to examine the difference in the banking sector efficiency levels between the pre- and post-1997 Asian financial crisis for Malaysia and Thailand. The Wilcoxon test is generally used when the normality assumption is not a necessary pre-condition. Shaw et al, 2000 argues that the normality assumption is avoided in the sense that Wilcoxon test is based on the rank order of the differences rather than the actual value of the differences.

In order to carry out the test we estimate the changes in the bank efficiency of all 8 sampled banks for the pre-crisis period (t = 2004 – 2006) and the crisis period (t = 2007 – 2009). We then calculate for each bank in the sample the difference between their average before-crisis score and their average crisis efficiency score. Ranks are then allocated to each absolute (ignoring the sign) difference with the smallest difference assigned a rank of 1, second smallest difference a rank of 2, third smallest difference ranked 3 and so on up to the 8^{th} rank, the number of banks in our sample. Then we separately find the total of the ranks of the positive differences and negative differences. Walker (2010) states that if the null hypothesis of equal means is correct, the total of the ranks of negative numbers and that of positive numbers should more or less be equal. Put differently the larger the difference between the two total values, the more likely that the null hypothesis will be rejected. The smallest of the two values is considered to be the Wilcoxon statistic and is then compared to the Wilcoxon critical value at the 0.05 level of significance using n = 8. The decision is to reject the null hypothesis that there is no difference between the pre-crisis and the crisis means if our calculated W-statistic is less than or equal to the appropriate critical value of 4[11] in this case. It is important to highlight that unlike most statistical tests the "Wilcoxon and Mann-Whitney tests are the only tests where a statistic is significant when it is below the critical value" (Walker, 2010, p. 237). The subsequent section discusses the second stage panel data models.

[11]For a level of significance of 0.05 and n = 8, the obtained Wilcoxon critical value from the Wilcoxon signed ranks table is 4.

4.5 SECOND STAGE ANALYSIS

Panel data models as opposed to time series models were used because of two important reasons. Firstly, panel data models and in particular the fixed effects model takes into account bank-specific heterogeneity (individuality). It is important to recall that one of the attractiveness of panel data analysis is that it enables one to analyse variables that vary over time but are constant across banks such GDP per capita, banking sector development, bank market concentration etc. Secondly, Gujarati (2004, p.637) states that by employing a combination of time series and cross sections, panel data provides "more informative data, more variability, less collinearity among variables, more degrees of freedom and more efficiency." Brooks (2008) also states that through pooling cross-sections and time-series data, one can increase the degrees of freedom, thereby improve the power of the test. Panel data models are employed to address the main objective namely, to study the relationship between access to banking services and bank efficiency. In order to examine the relationship between banking sector efficiency improvement and unemployment in South Africa the study also employs panel data techniques.

4.5.1 Panel Data Estimation Methods

Based on theoretical and empirical literature, various exogenous factors that are hypothesised to affect access to banking services and unemployment are analysed using panel data methods. There are three basic panel data estimation methods namely, Pooled OLS, Random Effects Model (REM) and Fixed effects Model (FEM).

4.5.1.1 Pooled Regression Model

The pooled regression model emphasises the joint estimation of coefficients using the ordinary least squares (OLS) thereby increasing degrees of freedom and decreasing standard errors of the coefficients (Baltagi, 2008). Its major weakness is that it pools all cross-sections and time series data and estimates a common regression disregarding possible bank-specific differences. To decide between a pooled regression model (restricted) and a fixed effects model (unrestricted), the F-test is performed. The F-test is constructed as follows:

$H_0 : \mu_1 = \mu_2 = ... = \mu_{N-1} = 0$

$H_A :$ Not all equal to 0

$$F\ statistic = \frac{(RSS - URSS)/(n-1)}{URSS/(nt-n-k)} \sim F_{(n-1),(nt-n-k)}$$

Where**:**

RSS - Residual sum of squares from the pooled model (restricted model)

URSS - Residual sum of squares from the FEM (unrestricted model)

n - Number of cross sections (banks)

t - Number of time periods

k - Number of explanatory variables

The F-test essentially involves a comparison of the residual sum of squares (RSS) of both models. This method is justified on the basis that OLS entails minimising the residual sum of squares. The rationale of the F-test is that if after imposing restrictions a much greater RSS is obtained then the restricted model is not supported by the data. Similarly, if the RSS decreases significantly following the restrictions then it is concluded that the restricted model is supported by the data. The null hypothesis of cross-section homogeneity is rejected if F-statistic > F-critical value at 0.05 level of significance.

4.5.1.2 Random Effects Model

In the random effects model (REM) or the error component model (ECM), the individual effects are treated as random draws from a larger population. Gujarati (2004, p.648) states that if the cross-section or individual-specific error component is zero there will be no difference between the pooled and the random effects model in which case one would simply run a pooled regression. Baltagi (2008, p.17) maintains that the random effects model is appropriate if draws are made randomly from a large population where N is generally large. In this study $N = T = 8$ is not sufficiently large to warrant the use of the Random effects model. As a result the REM versus FEM test is not performed and as such random effects are not accounted for in the specification.

4.5.1.3 Fixed Effects Model (FEM)

A Fixed Effects Model is a linear regression model in which the intercept terms vary over the individual units (banks). Gujarati and Porter (2009) states that these intercepts capture the different special features of each firm such as managerial style or the type of market each firm is serving. Griffiths et al (2008) also states that these fixed effects can be analysed to study the extent of firm heterogeneity and to examine any particular firms of interest. However, in the event that these individual effects are identical, one can then use a pooled least squares regression model. The fixed effects model can be estimated by employing the

"Within" Q estimation technique or the least square dummy variable (LSDV) technique. Both fixed effects modeling techniques namely, LSDV and "WITHIN"/Q estimation methods basically produces the same results. For making inferences, the LSDV is reported as it *estimates* as opposed to *calculating* the intercept coefficients and as such produces standard errors, *t*-statistic and *p*-values.

4.5.2 Modeling Access to Bank Services and Bank Efficiency

The main motivation driving this research has been to explore the nexus between gains in bank efficiency and access to bank services. In order to study this relationship the researcher estimated a translog function of the form:

$$\ln ACC_{it} = \alpha_{it} + \beta_1 \ln TFPE_{it} + \beta_2 \ln GDP_t + \beta_3 \ln BDEV_t + \beta_4 \ln BRAN_t + \beta_5 \ln HHI_t + \beta_6 \ln BC_{it} + \beta_7 \ln RUR_t + \mu_{it} \quad \text{.... [4.1]}$$

Where**:**

ACC_{it} - Number of deposit accounts with bank *i* in period *t*

$TFPE_{it}$ - Total factor productivity efficiency for bank *i* in period *t*

GDP_t - GDP per capita in period *t*

$BDEV_t$ - Domestic credit provided by banking sector (% of GDP) in period *t*

$BRAN_t$ - Number of bank branches per 100 000 adults in period *t*

HHI_t - Herfindahl Hicksman Index – Bank market concentration in period *t*

BC_{it} - Bank cost for bank *i* in period *t*

RUR_t - Rural population as a percentage of total population in period *t*

μ_{it} - Error term to capture other possible factors not specified.

4.5.2.1 Second stage variables and the expected priori

Access to Bank Services

The logarithm of the number of deposit accounts with a commercial bank was used as a proxy for access to bank services. This proxy was chosen in line with a World Bank (2009) study "Banking the Poor", which measured access to banking services for 54 countries by considering the percentage of the adult population with a bank account per thousand adults and transactions offered at banks in each of the countries. An IMF study conducted by Anayiotos and Toroyan (2009) also used the same indicator variable.

Total Factor Productivity Efficiency (TFPE)

The logarithm of TFPE was used. This efficiency variable which constitutes the main focus of this study was generated using the Hicks-Moorsteen total factor productivity (TFP) index approach. It is important to recall that in this study the TFPE score for each bank is generated by evaluating among other inputs, the operational costs of each bank relative to the minimum cost of a fully efficient bank that is required to produce a defined set of outputs. It is mainly via this cost channel that changes in TFPE are expected to be transmitted through to access. Schoombee (2004) states that commercial banks generally do not serve the poor mainly due to the associated high costs involved. As such one is inclined to expect a negative relationship between this measure of efficiency and access. The underlying argument is that banks' appetite to achieve and maintain good scores on TFPE and cost efficiency, hold the potential to reduce access for consumers particularly the low-income groups. Moreover, given the low economic profile and social development of the poor, the impetus to "move up" the market is more rewarding than "moving down" the market. Alternatively, we are also conscious of the fact that the benefits of a decrease in operational cost (increased TFPE/cost efficiency) could be passed on to consumers in the form of better service charges thereby increasing access. One World Bank (2006) study reported that high minimum balances and monthly charges prevent a large proportion of the African population from accessing bank services. Hence, it is a possible outcome that gains in TFPE (via the cost channel) may feed through to improved access exhibiting a positive relationship between the two variables. However, we are more convinced of the former argument as opposed to the latter.

Gross Domestic Product per capita (GDP)

Here the logarithm of GDP is used as a composite proxy to capture the impact of changes in macroeconomic conditions on efficiency and access to bank services. This is done in an attempt to answer the question: Does the macroeconomic conditions in a country matter for bank performance. In its bank supervision report the SARB (2009, p.38) noted that "while many of the risks are at bank level, the macroeconomic environment influences them". We do acknowledge the commonly held direction of causality that a stable and vibrant banking system is a vital engine for growth. However, it may be argued changes in macroeconomic conditions affect a bank`s financial health and its performance. Any given banking system is relatively more likely to face challenges in an economy whose macroeconomic fundamentals are poor. This is because during turbulent economic times the proportion of NPL to gross loans rises, while profitability generally declines impacting negatively on the efficiency of a

bank. The South African Reserve Bank (SARB, 2011), reported that the devastating effect of the sub-prime financial crisis in South Africa was partly cushioned by the countries' solid macroeconomic fundamentals. Hence, changes in the macroeconomic conditions of an economy have implications on efficiency and access.

Banking Sector Development (BDEV)

The traditional measure of banking sector development in line with previous studies is the proportion of domestic credit provided by the banks as a proportion of GDP. Well-developed banking institutions improve the intermediation process and are expected to contribute positively towards efficiency and access. Hence this variable is expected to impart a positive effect on access. Logarithm values were generated for the variable and used in the estimation.

Number of Bank Branches (BRAN)

BRAN represents the logarithm of the number of bank branches per 100 000 adults. We expect a positive relationship between branch network and access indicating that an increase in the number of bank branches should provide the would-be clients with more geographical choices at which to conduct transactions. However, Okeahalam (2008, p.1133) argues that an increase in the number of branches may not yield the improved access unless these branches are located at points of actual demand.

Herfindahl Hicksman Index (HHI)

Economic theory prescribes that competition is generally good for efficiency with welfare gains for the public. Competition within banking can take the form of price-competition where banks will be pushed to lower their lending rates in order to keep their market share. Therefore, this form of competition appears to hold great potential for expanding access to the unbanked via the cost efficiency channel. Moreover, Duygun et al (2013) maintains that banks may also engage in non-price competition by offering a wider range of products and services. The logarithm of HHI was included in this study to determine whether a concentrated banking system contributes towards increased access. It is generally acknowledged that banks in highly concentrated markets avoid competing among themselves and so refrain from reducing their interest spreads resulting in inefficiency. The general expectation is that high levels of concentration are negatively correlated with competition. However pessimists have argued that the level of market concentration does not matter because the nature of competition within banking is rather product specific. However, this

study is more inclined towards the former argument. Okeahalam (1998) found evidence that the banking sector of the common monetary area[12] of Southern Africa was highly concentrated and that this high concentration level reduced the incentive for banks to improve efficiency.

Bank Cost (BC)

A bank`s cost is expected to have a negative impact on accessibility. We argue that in order for banks to maintain their profit margins they are more inclined to pass on the increased cost to the consumers with an obvious deterring effect on potential clients. The aggregate of total interest expenses and total non-interest expenses constitute a bank`s operating expenses. The logarithm of a bank`s cost was also taken as a proxy for bank`s service charges.

Rural Population (RUR)

In order to ascertain the impact of the growth of rural population on access to bank services the logarithm of the proportion of the rural population variable was included in line with the study of Kablan (2010). He argues that banks generally tend to locate their branches in more economically developed regions at the expense of rural ones. The inclusion of this variable is critical in order to inform policy formulation to establish if banks are indeed generally biased against providing their services to the rural population of the economy. A negative relationship between rural population and access is expected.

The discussed model is estimated using the appropriate panel data technique following the F-test of determining the validity of fixed effects. The relevant estimation technique is applied and results are presented and discussed in chapter six. The following section presents the model and the definition of variables for analysing the efficiency-unemployment nexus.

4.5.3 Modeling Unemployment and Bank Efficiency

To investigate the relationship between bank efficiency and unemployment in South Africa, we estimated the following translog function:

$$LnUNEM_t = \alpha_{it} + \beta_1 LnTFPE_{it} + \beta_2 LnSPR_{it} + +\beta_3 LnCIR_{it} + \beta_4 LnBDEV_t + \beta_4 LnGDPG_t + \mu_{it} \quad [\mathbf{4.2}]$$

.....

[12] South Africa dominates the Common Monetary Area of Southern Africa. Countries within the CMA include South Africa, Namibia, Lesotho and Swaziland.

Where**:**

$UNEM_t$ - Unemployment rate in period *t*

$TFPE_{it}$ - Total factor productivity efficiency for bank *i* in period *t*

SPR_{it} - Interest spread for bank *i* in period *t*

CIR_{it} - Cost to income ratio for bank *i* in period *t*

$BDEV_t$ - Domestic credit provided by banking sector (% of GDP) in period *t*

$GDPG_t$ - GDP growth in period *t*

μ_{it} - Error term to capture other possible factors not specified.

4.5.3.1 Second stage variables and the expected priori

Unemployment

The logarithm of unemployment rate in South Africa was used to capture the level of unemployment. This variable is hypothesised to be influenced by the right-hand variables such as banking sector efficiency, bank interest spreads, cost-to-income, banking sector development, and GDP growth. The government consumption variable was not specified in the model due to high collinearity with the GDP variable.

Total Factor Productivity Efficiency (TFPE)

The logarithm of TFP efficiency was used. This efficiency variable which constitutes the main focus of this paper is generated in the first stage using the Hicks-Moorsteen total factor productivity (TFP) index approach. We expect the effect of bank sector efficiency on unemployment to be negative. This follows from the argument that enhanced efficiency implies achieving more with less input resources. These gains are then passed on to recipients of banking services through affordable services, easy access to banking services, increased loan demands thereby increasing investment, economic activity and employment in the process.

Interest Rate Spread (SPR)

The logarithm of interest rate spread was included to capture the role and importance of intermediation efficiency on unemployment in South Africa. Interest spread is the difference between the lending rate and the deposit rate. We expect that bank efficiency gains in the

form of narrowing of interest spreads should stimulate an increased demand for loans for investment and greater mobilisation of savings. Ikhide (2008) argues that wide spreads affect banks' basic function of intermediation hence it distorts prices thereby slowing down the role of the banking system in contributing to economic growth. Thus, the variable is expected to exhibit a negative relationship with TFPE. In line with many empirical studies we used the net interest margin (NIM) variable as a proxy variable.

Cost-to-Income (CIR)

A traditional measure of cost efficiency used in empirical studies is the cost-to-income ratio. This efficiency indicator expresses a banks` total operating cost (non-interest expenses) as a proportion of its total operating income. An increase in this ratio is interpreted as cost inefficiency and is expected to be positively related with unemployment. Thus, a positive relationship between cost-to-income ratio and unemployment is expected.

Banking Sector Development (BDEV)

A measure of banking sector development used in empirical studies is the proportion of domestic credit provided by the banks as a ratio of GDP. An increase in the volume of credit extended by the banking sector is expected to increase economic activity therefore reducing unemployment. This variable is therefore, expected to exert a negative impact on unemployment.

Gross Domestic Product Growth (GDPG)

In line with economic theory, job opportunities emanate from a growing economy. We therefore expect real GDP to have a negative impact on unemployment. However, in recent years a reduction in job opportunities has been observed in the formal sector despite positive GDP growth rates highlighting a case of jobless growth.

4.6 CONCLUSION

This chapter has explored the methodological framework of the study: various DEA models, model specification, data analysis and definition of variables. It has set forth an analytical framework in which the impact of bank efficiency on access and unemployment will be investigated. A two-step methodology framework is applied to a sample of 8 banks for the period 2004 – 2011.

In the first step, the efficiency measures are generated using the non-parametric DEA methodology framework. The study selected the variable returns to scale BCC model over the constant returns to scale CCR model. The chapter highlighted that the CCR model is only suitable for situations where all the banks are operating at an optimal scale. This assumption is not plausible within banking particularly for a panel of large and small banks. The concept of efficiency chosen in this study is technical efficiency and cost efficiency whereas both the input and output orientation results will be reported. The intermediation approach as opposed to the production approach was followed in the definition of the bank inputs and outputs. Labour, total operating expenses, fixed assets, and total deposits were considered inputs while interest income, non-interest income and loans & advances were defined as output variables. Each variable was carefully selected to reflect important characteristics of the main activities of commercial banks as indicated in South Africa`s banking industry and empirical literature. Next, the efficiency measures from the DEA output are then analysed to determine if there was a significant change in the efficiency of South African banks as a result of the global financial crisis. To carry out this task the researcher discussed two tests namely, the parametric *t*-test and the non-parametric Wilcoxon matched pairs signed ranks test.

In the second step, the generated efficiency scores are used with other variables as regressors within the access-efficiency model and the unemployment-efficiency model. In each case, various other factors as informed by theoretical literature and empirical research were discussed as well as their expected impact on the regressand. Three panel data estimation techniques namely, the pooled regression model, fixed effects model and the random effects model were also discussed. To aid the selection of the most appropriate estimation technique, the procedure behind the F-test was outlined. The succeeding chapter runs the DEA models discussed in this chapter using the *DPIN version 3.0* which uses the Data Envelopment Analysis (DEA) programs to generate and decompose the HMTFP indices into several efficiency measures. The analysis software package used in the second stage is *EViews 8*.

CHAPTER FIVE

ESTIMATION AND ANALYSIS OF BANK EFFICIENCY

5.1 MEASURING EFFICIENCY USING FINANCIAL RATIOS

While this study adopts the non-parametric approach as the main first-stage methodology to measure bank efficiency, we present accounting ratios in order to complement one`s understanding of bank efficiency. In this section, only specific ratios that relate to the selected bank output and input variables as discussed in the previous chapter are analysed.

TABLE 5.1: MEASURING EFFICIENCY USING FINANCIAL RATIOS

SOUTH AFRICAN BANKING SECTOR TRENDS		
YEAR	**Cost To Income**	**Net Interest Income To Interest-Earning Assets**
2004	63.9	3.00
2005	66.3	2.70
2006	58.8	3.10
2007	56.9	3.30
2008	49.00	3.24
2009	51.13	3.10
2010	56.43	3.13
2011	55.25	3.38

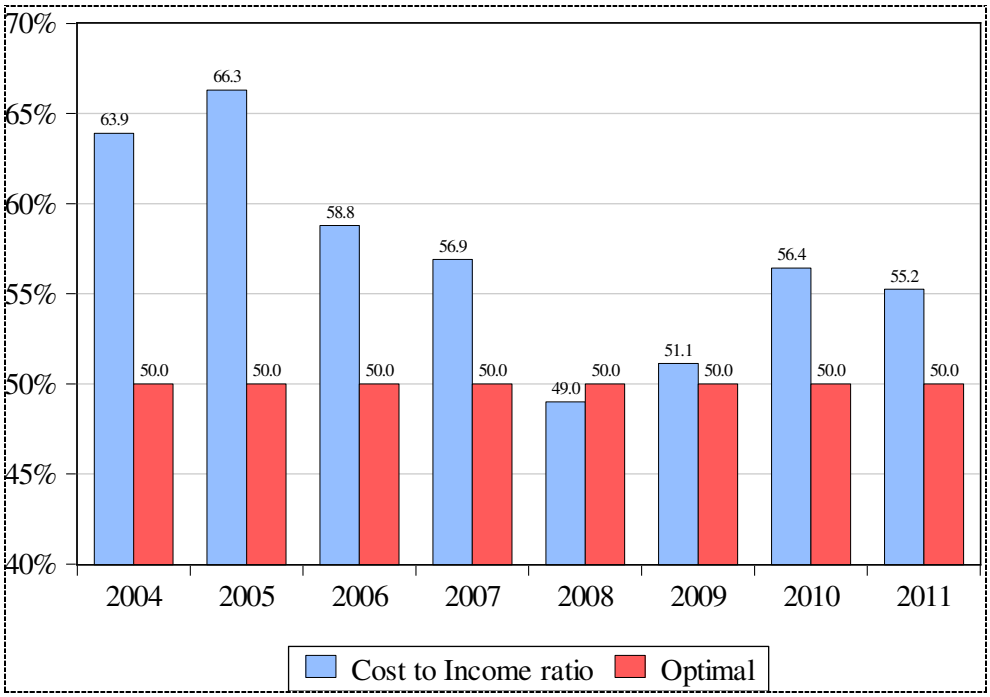
Source: SARB various publications.

Conventionally, financial ratios have always been used in banking to measure efficiency and other performance indicators. However, Ikhide (2000) states that despite their extensive use, caution should be exercised when making inter-bank or inter-country comparisons because of the possible differences in business mix, capital structure and accounting standards and practices. The researcher selected banks that are relatively homogeneous in their business operations and specialisation. The assumption is that since these banks operate in the same policy environment they comply with standard reporting practices as well as capital adequacy ratio requirements. Capital adequacy is a concept based on managing or rearranging the existing capital structure in order to protect depositors and to cushion the banking sector against potential losses. The SARB Supervision department currently uses accounting ratios to measure various aspects of bank performance. In particular, the SARB uses cost-to-income

ratio (CIR) to measure cost efficiency. In addition to CIR, this section focuses on non-interest income to gross revenue, net interest margin (NIM), and impaired loans (NPLs) to gross loans.

5.1.1 Cost-to-Income Ratio

FIGURE 5.1: COST EFFICIENCY USING COST TO INCOME RATIO



Source: Computed using data from SARB

Traditionally, bank efficiency has always been measured using cost-to-income ratio (CIR). This efficiency indicator expresses a banks` total operating cost (non-interest expenses) as a proportion of its total operating income. The ratio measures how much it costs the bank to generate a rand`s worth of income. A bank is regarded as relatively efficient if it incurs less cost to generate R1 of revenue compared to an inefficient one. An increase in the cost-to-income ratio is not desirable since it implies that either costs are increasing or revenue is decreasing. The lower the ratio the better thus banks should strive to drive the ratio down as much as possible. The cost-to-income ratio is calculated as follows:

$$CIR = \frac{Non - interest\ expenses}{Net\ interest\ income + Non - interest\ income}$$

While there is no ideal or optimum benchmark regarding CIR among banks, an informal investigation conducted by SARB (2006) into the CIRs reported by the largest and most efficient international banks indicated that these efficient banks maintained CIRs of approximately 50 percent. A cost-to-income ratio of 50 percent indicates that it cost the bank 50c to generate a R1 of income. Figure 5.1 indicates that since 2005, the ratio has been decreasing implying significant improvements until 2008 when it recorded the lowest ratio (most efficient) below the 50 percent optimal mark. The ratio started increasing in 2009 until 2010, perhaps highlighting the adverse impact of the 2008 financial global crisis. At the end of 2011, the CIR amounted to 55.25, an improvement from the previous year. Appendix 6 shows that African bank has been the most performing in the sample in decreasing its cost-to-income ratio. Compared to other banks, the African bank reported measures that were significantly below 50 percent in each year of the study period with an average of 39.50 percent. African bank is incurring the least cost (R39.50) to generate R1 of revenue and has kept its cost under control. Standard bank was in the second place followed by FRB and Capitec with a ratio amounting to 52.37, 53.82 and 54.85 respectively. On the contrary, Teba bank was the least performing with ratios above 70 percent in each period and recording an average of 83.61 percent for the period 2005 – 2011. This was followed by Sasfin, Absa and Nedbank with an average ratio of 63.30, 57.99 and 57.79 in that order.

The cost-to-income ratio (CIR) analysis for the period 2005 – 2011 indicated an average lower ratio of 0.55 for large banks compared to 0.60 among small banks. This is in line with Casu et al (2006) who found that generally CIR is higher for small banks compared to large banks. Vittas (1991) also found that retail banks are generally associated with higher operating costs as opposed to wholesale corporate banking. Furthermore, empirical studies [KPMG, 1998; Hess and Francis, 2004] have reported evidence that banks that obtain a substantial share of their funding from retail customers tend to have higher CIR than those that rely on wholesale funding. This high CIR among retail banks reflects the high cost involved in serving retail accounts and maintaining branches (KPMG, 1998). However, South Africa`s CIR within the banking sector is in line with that of other emerging markets and advanced economies. Table 5.2 provides an international comparison of South Africa`s performance.

TABLE 5.2: CAPITAL TO INCOME RATIO – AN INTERNATIONAL COMPARISON

COUNTRY	(%)
SOUTH AFRICA	**57**
AUSTRALIA	44
BELGIUM	58
BRAZIL*	57
CANADA	60
CHINA*	38
FRANCE	71
GERMANY	85
GREECE	73
INDIA*	44
ITALY	73
JAPAN	62
MALAYSIA	41
MEXICO	64
RUSSIA*	90
SWEDEN	58
UK	62
USA	61
Red shading indicates better performance than South Africa. *Emerging BRICS countries are highlighted by an asterisks.	

Source: Federal Reserve Bank of St. Louis[13]

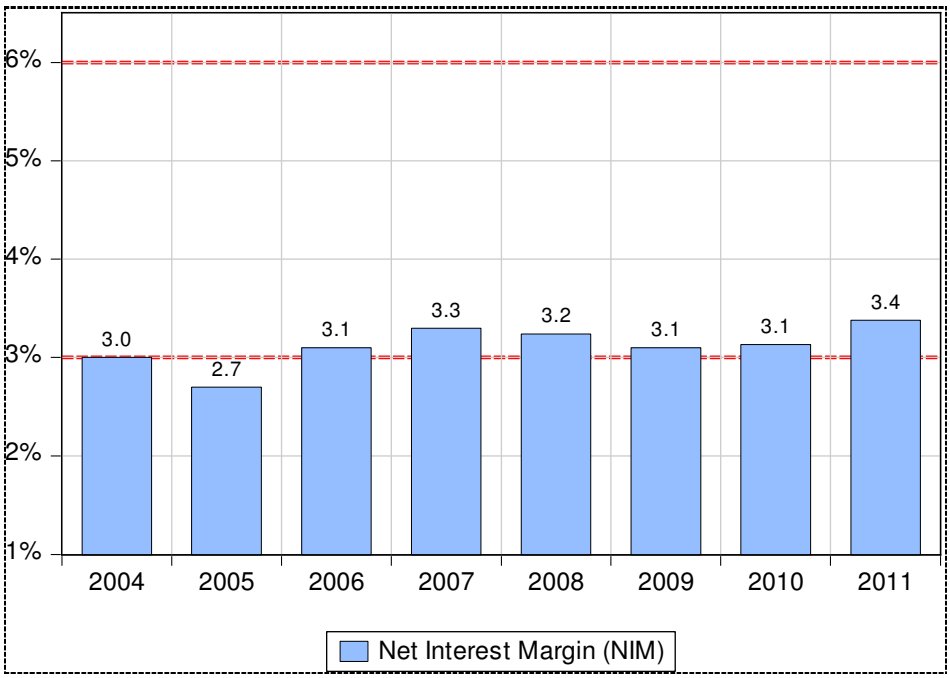
5.2.2 Net interest Income to Average Interest-bearing Assets

Net interest margin (NIM) is the percentage of net interest income to average interesting-earning assets calculated as: $NIM = \frac{Interest\ Income - Interest\ \ Expense}{Interest\ Bearing\ Assets}$.

KPMG (1998) study noted that a NIM of below 3 percent is generally considered low while an excess of 6 percent is regarded as high. In general, a wide net interest margin above 3 percent indicates the ability of a bank to manage well its interest bearing assets and liabilities. However, a very large wide net interest margin exceeding 6 percent may indicate some degree of inefficiency as this might mean excessive returns to investors at the expense of borrowers and depositors (KPMG, 1998). Moreover, such wide interest spreads have important implications for financial intermediation.

[13] Computed using bank-by-bank data from Bankscope

FIGURE 5.2: NET INTEREST MARGIN



Source: Computed using data from SARB: www.quantec.co.za

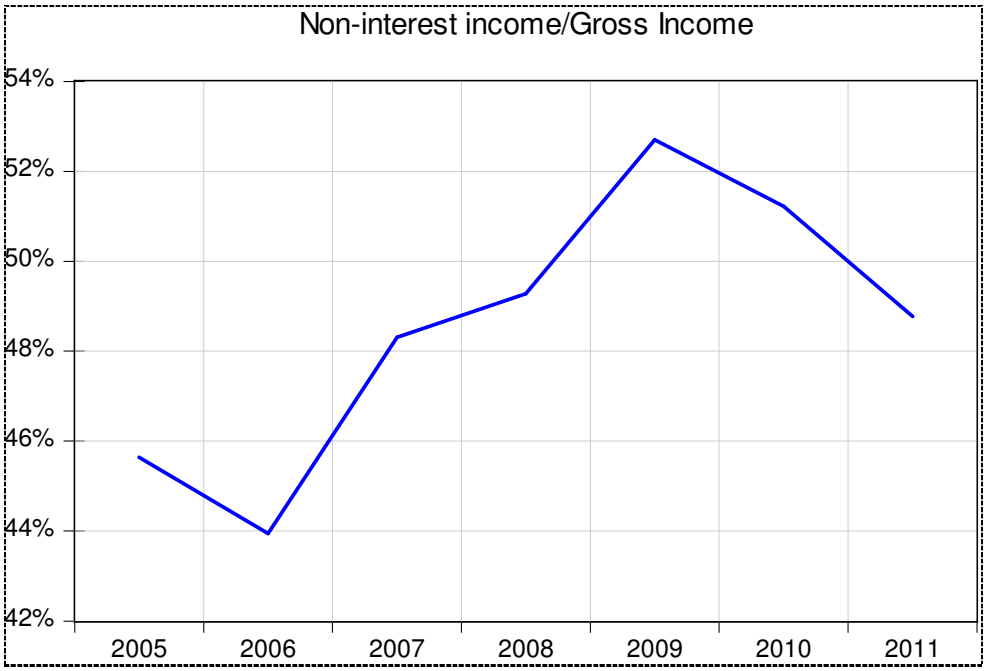
Evidence in Figure 5.2 indicates that with the exception of 2005 when the lowest NIM of 2.7 percent was recorded the banking sector managed to keep its NIM above the minimum 3 percent. However, there is a need to highlight that there are marked disparities between large and small banks in the research sample with regard to NIM. For example, average NIM for large banks was 3 percent compared to 23 percent for small banks. KPMG (1998) maintains that large banks tend to have lower margins compared to small banks. This study postulates that this is indicative of the fact that generally a greater proportion of total income for large banks is attributable to non-interest income compared to small banks. Hence, relative to small banks whose business models are predominantly retail-oriented, large banks tend to realise less interest income resulting in lower interest margin. Large banks predominantly engage in non-interest activities such as bond trading, currency, asset management services, and other fee based financial services. However, we acknowledge that in South Africa some small banks handle corporate transactions but not in the same scale as large wholesale corporate banks. Thus the hypothesis that at least part of the wide disparity in net interest margin between small and large banks may be attributed to the difference in the composition of income appear plausible.

5.2.3 Non-Interest Income to Gross Revenues

Non-interest income to gross revenue measures the proportion of a bank's total income that has been generated by non-traditional or non-lending activities. It indicates that a bank is not dependant on its lending activities to generate income and therefore regarded as a measure of diversification. Ikhide (2000) found a negative relationship between non-interest income and interest spreads in South Africa. Interest spread measures the difference between deposit and lending rates. The narrower the interest spread the more efficient the bank. He argued that the major sources of non-interest income in South Africa and Botswana as opposed to other developing economies are security related transactions and exchange earnings resulting in narrowing of interest spreads. Kohler (2013) in his study recommends retail-oriented banks to increase their share of non-interest income to become stable. He argues that this enables them to diversify their income and so become more resilient to economic conditions. It is generally argued that relative to interest income, non-interest revenue is less dependent on the economic conditions and as such reduces the cyclical variations in bank revenue (Stiroh, 2002). The impact on bank stability emanates from fee and commission income which is generally considered to be less volatile. In developing economies, a large proportion of non-interest income comes from bank fees/charges and commissions accounting for more than 70 percent of total non-interest income. Appendix 9 indicates that the share of net fees and commissions to total non-interest income averaged 75 percent in our sample of banks. Fee income and commissions could be generated from increasing the volume of mortgage loans, securitisation, consumer credit, foreign exchange activities and expanding bank services to the unbanked. However, in the case of reaching the previously unbanked, the cost of extending bank services should not exceed the revenue to be generated. Banks could also increase non-interest income through increases in fees on loan and deposit accounts. Ikhide (2000) states that, this later form of increase in non-interest income may not result in narrowing of interest spreads. However, if a particular increase in bank fees is perceived by the general public to be unjustifiable this could result in consumers substituting with cheaper banks with the obvious negative impact on financial access for potential consumers. Lozano-Vivas and Pasiouras (2008) conducted a study on a panel of 87 countries to determine bank efficiency with and without non-traditional activities. Their results revealed that when non-interest income was included as a bank output both profit and cost efficiency scores improved significantly relative to the traditional-interest income model. In a more recent study, Van der Westhuizen (2010) undertook a study in South Africa to investigate whether a shift in the composition of a bank`s income from interest income to non-interest income as the main

source of income had any positive impact on bank efficiency. His results showed that indeed a shift towards non-interest income as the main source of income contributed to the improvement of technical, cost and scale efficiency. Historically, non-interest income represents a small portion of a bank's total revenue. However, recently this ratio has been on the increase. In our sample, the proportion of non-interest income to gross income range between 41 to 54 percent and 31 to 61 percent among large and small banks respectively. With an average non-interest income ratio of 69.4, Sasfin bank outperformed the rest of the banks followed by FRB and African bank with averages of 57.4 and 51.4 percent respectively.

FIGURE 5.3: NON-INTEREST INCOME TO GROSS INCOME



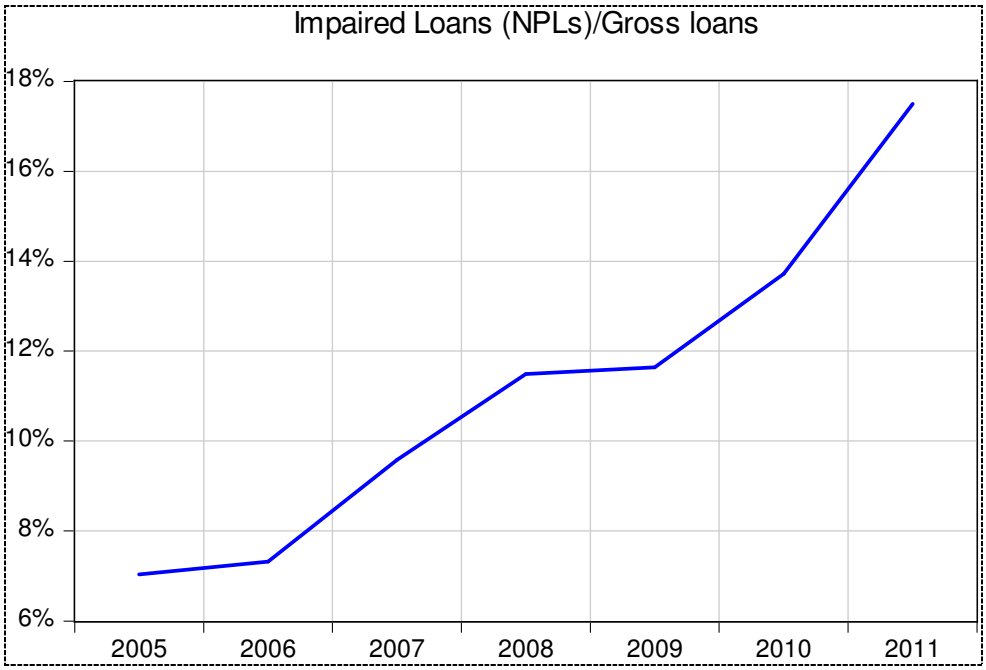
Source: Computed using data from Bankscope: www.quantec.co.za

5.2.4 Impaired Loans (NPLs) to Gross Loans

The proportion of non-performing loans (NPLs) or impaired loans to gross loans captures the quality of loans. An increase in this ratio indicates deteriorating quality of loans and therefore, reflects on the inefficiency of the management. By employing the Granger causality test, Berger and DeYoung (1997) found that high levels of non-performing loans Granger-cause decreases in cost efficiency due to extra cost of administration related to these problem loans. Moreover, low bank cost efficiency was found to Granger cause increases in NPLs implying that inefficient management worsens the problem of NPLs. Berger and DeYoung (1997, p.853) identified some of the extra operating costs that are related to NPLs

to include among others: (i) monitoring of non-performing borrowers and the value of their collateral (ii) cost of negotiating alternative remedial arrangements (iii) cost of repossessing, keeping and eventually disposing of collateral in the event of default (iv) extra cost of preserving the bank`s safety and soundness record to market participants and bank supervisors and (iv) the digression of senior management focus away from handling other operations problems.

FIGURE 5.4: IMPAIRED LOANS TO GROSS LOANS



Source: Computed using data from Bankscope: www.quantec.co.za

Figure 5.3 portrays the behavioral pattern of NPLs over the seven year period from 2005 – 2011. The results indicate that since 2005, NPLs have been increasing until 2008 and 2009 when the ratio remained constant averaging 11 percent. Thereafter, the share of impaired loans rose sharply at an increasing rate possibly highlighting the impact of the economic recession that followed the financial crisis. The bank that reduced its NPLs significantly was FRB with an average of 3.54 followed by Standard bank with an average of 4.02 percent. The worst performance was from Teba and African bank with averages of 30.24 and 28.98 respectively. For further analysis, all the discussed ratios for each bank are in Appendix 6 through Appendix 10.

The relative high level of NPLs among small banks has been blamed on the growth of unsecured lending in the retail market. Unsecured lending has been on the increase in recent years with smaller banks such as African bank and Capitec dominating the space. According to Kagiso Asset Management (2013) the asset quality of African bank which is the largest provider of unsecured credit in South Africa has been deteriorating as indicated by the increase in its NPLs and bad debts. However, Moody`s Investors Services (MIS) (2013) reported that the impact of NPLs among large commercial banks is limited as unsecured lending constitutes a small proportion of their loan portfolios. The low rate of NPLs among large banks has also been attributed to the view that large banks are relatively more effective in screening their loan customers compared to smaller banks. Generally, the increase in NPL has been a consequence of the economic recession brought about by the financial crisis. SARB (2009) reported that the growth in impaired loans was a result inter alia of the weakening economy, rising interest rates and the increasing household debt.

5.2 MEASURING EFFICIENCY USING THE NON-PARAMETRIC APPROACH

This section presents and analyses the findings of estimating and decomposing the TFP indices of a panel of eight South African banks for the period 2003 – 2011. The concept of efficiency applied in this study is that of technical and cost efficiency whereas both the input and output orientation results will be reported. The empirical model adopted in this study is that of Banker Charnes and Cooper (BCC model, 1984). As previously stated, the major distinction between the CCR model and the BCC model is the handling of returns to scale with the BCC model allowing for a more realistic concept of variable returns to scale. The assumption of the CRS DEA model is only suitable for situations where all the banks are operating at an optimal scale. Since banking in South Africa is heavily concentrated with the four largest banks the CRS assumption may not be realistic for a panel of large and small banks. Hence, the presentation of the first-stage results is guided by the methodological framework developed in the previous chapter. The computation and decomposition of TFP scores was undertaken using *DPIN 3.0* software package developed by Donnell (2011) that uses DEA programs.

5.2.1 Empirical DEA Results

Since DEA is affected by extreme values, the sample was categorised into four large and four small banks. The DPIN program was then run for each size category. The sample in this study is a fair representation of the entire banking sector and as such the results of this study can be interpreted as being representative of the total banking sector. This is particularly so given that the sample includes the four largest banks (big-four) which CGAP (2011) reported as accounting for over 90 percent of retail banking in South Africa. It is important to note that efficiency scores ranges from zero to one. A score below unity represents an inefficient bank which is located below the production frontier while a score of one implies that the bank is fully efficient and lies on the frontier of the production technology. Regarding productivity, a value greater than one indicates positive TFP growth from period *t* to period *t+1*.

Running the DPIN program with VRS assumptions (BCC model), all South African banks in the sample exhibited an overall TFP and TFPE score of 1.35 and 0.59 respectively during the study period. Table 5.3 presents the total factor productivity (TFP) averages and total factor productivity efficiency (TFPE) averages for each size category and for all banks in each year of the study. Small banks appear to be more *productive* than large banks while large banks proved to be more *efficient* compared to small banks (see Figure 5.5 and Figure 5.6). For instance, the average TFP score for the period 2003 – 2011 was 1.31 and 1.40 for large and small banks respectively. The results indicate that small banks had the higher TFP index (1.40) than large banks (1.31). However, both categories had means above unity suggesting improvement in the period. The highest TFP index for the entire period of 1.55 was recorded in 2008 before a drastic decline in 2009 to 1.33. This coincided with the period of the onset of the intensifying 2008/2009 global financial crisis.

TABLE 5.3: SUMMARY STATISTICS OF TFP & TFP EFFICIENCY, 2003 – 2011.

		LARGE BANKS		SMALL BANKS		ALL BANKS	
2003		**TFP**	**TFPE**	**TFP**	**TFPE**	**TFP**	**TFPE**
	MEAN	1.000	1.000	1.000	0.902	1.000	0.951
	MAX	1.000	1.000	1.000	1.000	1.000	1.000
	MIN	1.000	1.000	1.000	0.790	1.000	0.895
	STDEV	0.000	0.000	0.000	0.114	0.000	0.057
2004							
	MEAN	1.277	0.727	1.465	0.387	1.371	0.557
	MAX	1.356	0.883	2.318	0.798	1.837	0.841
	MIN	1.207	0.509	0.815	0.034	1.011	0.272
	STDEV	0.061	0.186	0.625	0.352	0.343	0.269
2005							
	MEAN	1.531	0.471	1.387	0.366	1.459	0.419
	MAX	1.966	0.804	1.770	0.902	1.868	0.853
	MIN	1.198	0.242	1.009	0.136	1.104	0.189
	STDEV	0.351	0.244	0.328	0.359	0.340	0.302
2006							
	MEAN	1.391	0.737	1.554	0.326	1.473	0.532
	MAX	1.889	0.873	2.372	0.704	2.131	0.789
	MIN	1.102	0.578	1.223	0.070	1.163	0.324
	STDEV	0.358	0.122	0.552	0.310	0.455	0.216
2007							
	MEAN	1.346	0.734	1.277	0.468	1.312	0.601
	MAX	1.402	0.777	1.605	0.705	1.504	0.741
	MIN	1.297	0.662	0.963	0.092	1.130	0.377
	STDEV	0.047	0.053	0.264	0.277	0.156	0.165
2008							
	MEAN	1.362	0.607	1.745	0.604	1.554	0.606
	MAX	1.504	0.696	3.205	1.000	2.355	0.848
	MIN	1.234	0.528	1.014	0.271	1.124	0.400
	STDEV	0.111	0.069	0.990	0.347	0.551	0.208
2009							
	MEAN	1.376	0.512	1.280	0.526	1.328	0.519
	MAX	1.540	0.650	1.610	0.867	1.575	0.7585
	MIN	1.247	0.431	1.089	0.156	1.168	0.2935
	STDEV	0.128	0.095	0.247	0.346	0.1875	0.2205
2010							
	MEAN	1.311	0.601	1.510	0.436	1.4105	0.5185
	MAX	1.615	0.774	1.897	0.808	1.756	0.791
	MIN	1.087	0.469	1.098	0.154	1.0925	0.3115
	STDEV	0.220	0.138	0.374	0.323	0.297	0.2305
2011							
	MEAN	1.205	0.686	1.367	0.565	1.286	0.6255
	MAX	1.252	0.910	1.751	1.000	1.5015	0.955
	MIN	1.160	0.570	1.104	0.222	1.132	0.396
	STDEV	0.038	0.160	0.276	0.368	0.157	0.264
OVERALL	***MEAN***	**1.31**	**0.68**	**1.40**	**0.51**	**1.35**	**0.59**
OVERALL	***STDEV***	**0.22**	**0.19**	**0.47**	**0.33**	**0.34**	**0.25**

An examination of the results confirmed large banks to exhibit a higher TFP efficiency score of 68 (0.68) percent compared to small banks with an average score of 51 (0.51) percent. TFPE score represents our main performance indicator of particular interest in this study. TFPE is a measure of overall productivity performance. It actually measures the difference between observed (actual) TFP and the maximum TFP* attainable using the available technology. This measure was an average 59 percent for all the banks which means that for the nine year period 2003 – 2011 banks fell short by 41 percent to realise the maximum productivity that was achievable with their technology. Another way of putting it is that all banks needed 59 percent of the resources actually consumed in generating banking output. The standard deviation figures suggest that dispersion or variability of both performance indicators is wider for small banks compared to large banks. For instance, large banks had a standard deviation of 0.146 and 0.113 for TFP and TFPE respectively compared to 0.406 and 0.311 for TFP and TFPE for small banks respectively. These highly dispersed scores show more scope for improvement particularly among the small banks.

We posit that this disparity in efficiency is due to the fact that large banks and small banks operate different business models and emphasise different focus areas. For instance the business model of retail banking is mainly associated with small banks while large banks mostly operate in the wholesale corporate market. Akhigbe and McNulty (2005) argues that the business model of small banks generally require relatively high cost whereas larger banks preserve low costs. Studies [Vittas, 1991; Casu et al, 2006] in banking literature have confirmed that in general the cost to income ratio is relatively higher for a small bank compared to a larger bank. Okeahalam (2001) noted that for corporate banking interest income and fee income spreads are competitively determined resulting in higher levels of efficiency within this business-type bank model. Moreover, Berger (1995) suggests that the large market share usually associated with large banks maybe the result of better efficiency and lower costs.

The reported TFP efficiency score for large banks is consistent with the previous study by Mlambo and Ncube (2011) who found average technical efficiency to be 67 percent for the period 1999 – 2008. However, the current findings are much lower than those obtained by Okeahalam (2006) who found an efficiency score of 83.1 percent for 61 bank branches of one large South African bank. It is essential to indicate that these results are not easily comparable to previous studies due to the novelty of the issues being addressed in this study.

The estimation technique, the varied sets of inputs and output variables, the sample size, the duration and actual coverage of the time frame all add to the complexity of carrying such an exercise.

FIGURE 5.5: TFP FOR LARGE BANKS VS. SMALL BANKS

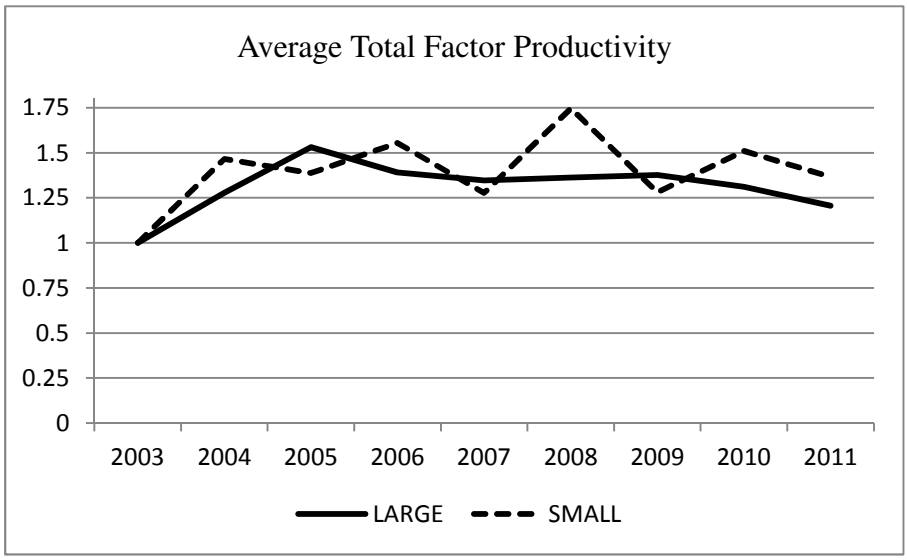
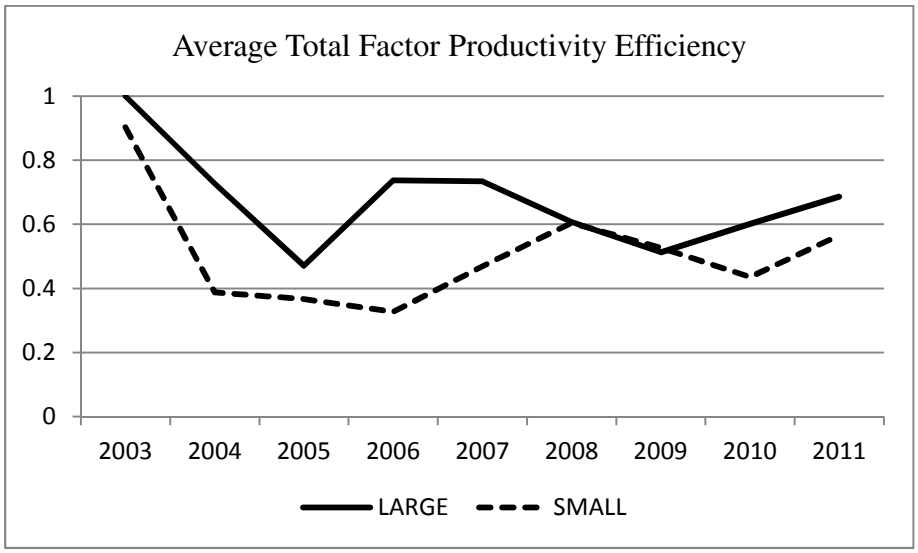


FIGURE 5.6: TFP EFFICIENCY FOR LARGE BANKS VS. SMALL BANKS



Appendix 2 and Appendix 3 present the estimates of *levels* of productivity and the various input-oriented and output-oriented efficiency scores for all the banks for the entire period. An estimated TFP average score of 1.35 for all banks indicated a positive growth of 35 percent over the 9 year period of evaluation. An average output-oriented technical efficiency (OTE) score for large banks of unity implied that large banks were fully efficient in their production of banking output. Correspondingly, an input-oriented technical efficiency (ITE) for small

banks of 0.992 implies that with the endowment of inputs that were available, each bank on average had the potential to expand its output by 0.088 percent. Similarly the obtained ITE estimate of 0.998 meant that small banks had the potential to decrease their inputs by 0.002 percent without altering their output. In other words, this represented input wastage of 0.002 percent. Again, large banks were fully efficient in their utilisation of input factors.

FIGURE 5.7: AVERAGE TFP BY PERIOD

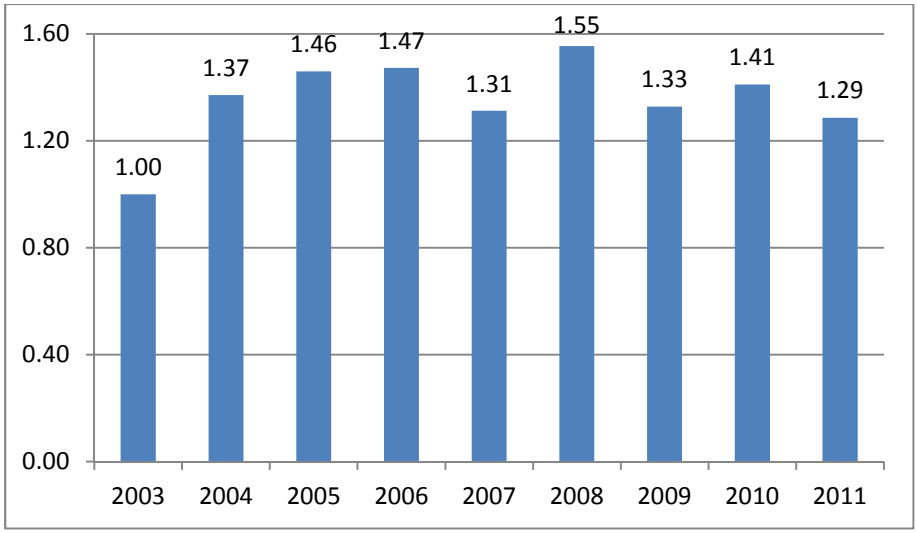


FIGURE 5.8: AVERAGE TFP EFFICIENCY BY PERIOD

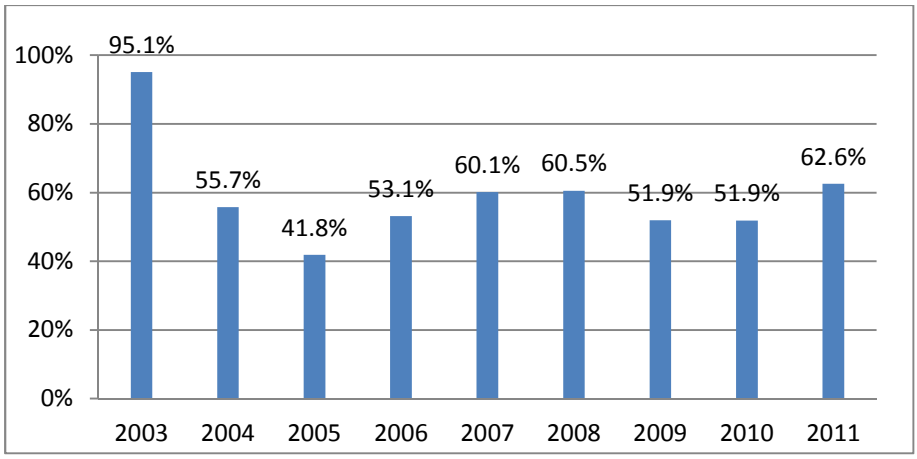
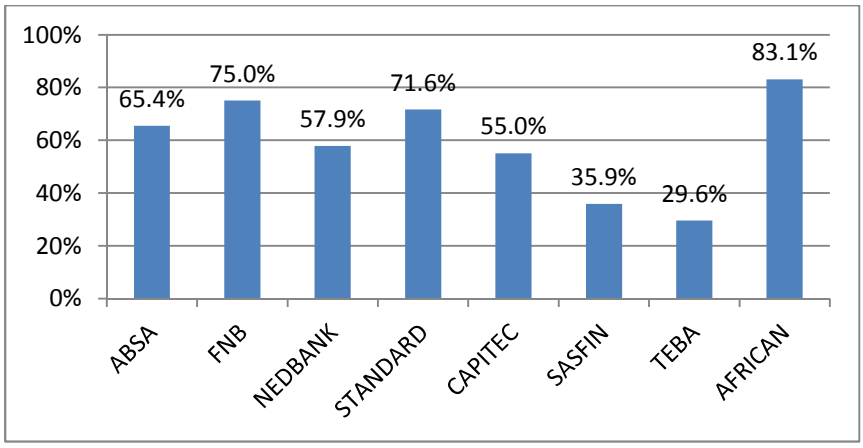


Figure 5.7 and Figure 5.8 results indicate a noticeable change in the scores for the period 2008 – 2009. This period coincides with the worst performance noted during the 2008-2009 when economies were at the height of the financial global crisis. The worst performance is evident across all the different efficiency and productivity scores especially for the year 2009. For example, TFP deteriorated from 1.55 in 2008 to 1.33 in 2009 while TFPE fell from 0.61 in 2008 to 0.52 in 2009.

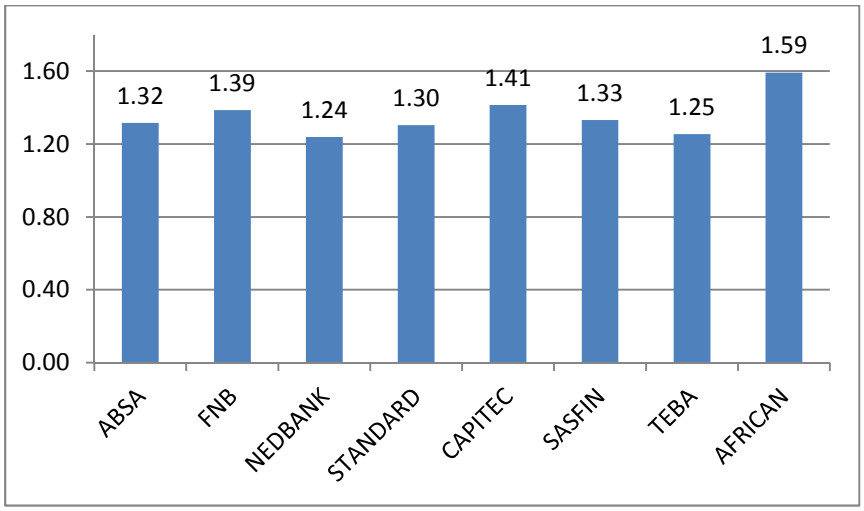
FIGURE 5.9: AVERAGE TFP EFFICIENCY BY BANK FOR THE ENTIRE PERIOD



All large banks exhibited TFPE scores above their mean score of 0.68 while two banks (50 percent) of a sample of small banks scored below their 0.51 mean score. Figure 5.9 reports African bank to be the most efficient bank of all the banks that were included in the study followed by FRB. Capitec was the second-most efficient among small banks with Sasfin and Teba being the least with TFPE scores below their category mean. FRB was the most efficient among large banks followed by Standard and Absa in that order. Nedbank was the least efficient bank.

Figure 5.10 below depicts that the bank that improved its productivity the most was African bank where the average score for the period was 1.59. Overall, the results show an improvement in the average TFP for all banks. All banks except Nedbank and Standard bank had TFP scores of less than their group average of 1.31. Similarly African bank and Capitec bank were above their average of 1.40 while Sasfin and Teba scored below average. As stated earlier small banks were relatively the most productive. Overall, African bank represented the best-practice bank for all banks to emulate regarding both TFP and TFPE.

FIGURE 5.10: AVERAGE TFP BY BANK FOR THE ENTIRE PERIOD



Inefficiencies can also occur as a result of resources being employed in the wrong mix. For instance, an input-oriented mix efficiency (IME) score of 1 means that the combination (mix) of inputs being used is in the most efficient state for producing the current outputs level. Hence, the mean score obtained of IME of 99.5 for small banks represents a marginal inefficiency of 0.50 percent that was due to input resources being employed in the wrong mix. However, large banks exhibited an IME score of 1 implying that their banking output was attained with the input mix that was most efficient. The study also identified evidence of combined scale and mix inefficiencies arising from banks not operating on their most productive scale and input/output mix. For example, the average input-oriented scale-mix efficiency (ISME) and output-oriented scale-mix efficiency (OSME) for large banks was 67.5 percent and 67.5 percent respectively. This means that banks exhibited inefficiency of 32.5 percent and 32.5 percent due to input utilisation and output production that was associated with both economies of scale and economies of scope respectively. Scale efficiency is achieved when a bank operates on the minimum point of its average cost curve, while economies of scope are achieved when the cost of jointly producing a range of outputs is less than the cost of producing them independently. Similarly, the average input-oriented scale-mix efficiency (ISME) and output-oriented scale-mix efficiency (OSME) for small banks was 51 percent and 51.6 percent respectively.

A change in total factor productivity (ΔTFP) which captures a contraction or expansion of the production possibilities frontier on average improved marginally for large banks and considerably for small banks. As summarised in Appendix 3 and Appendix 4 this

performance indicator was 1.004 and 1.16 for large and small banks respectively. Overall, the change in TFP (ΔTFP) for all banks for the entire period was 1.08. This result is also consistent with Van der Westhuizen`s (2013) study that revealed progress in TFP change for the big-four banks of 6 percent (1.06). Van der Westhuizen (2013) investigated the level of efficiency in the banking sector using the Malmquist TFP approach to analyse the four largest South African banks for the period 1994 to 2010.

Overall, a technological change (ΔTECH) estimate of 1.07 for all banks indicated technological progress of 7 percent in the South African banking industry during the period. Similar results were found by Van der Westhuizen (2013). The author reported that despite technological regress with one bank, the average for all banks was 1.06 which is very close to the current study. This study also found technological regress with two large banks bringing the overall score for the group to a technological regress of 0.5 percent. Hence, the reported overall technological progress of 6 percent (1.06) was predominantly driven by small banks that recorded 14 percent technological progress (1.14) offsetting the technological regress that emanated from large banks of 0.5 percent (0.995). The rest of the findings are in Appendix 2 through Appendix 4. The TFPE scores for each bank in the sample are used in the second stage analysis among other determinant factors of accessibility to bank services.

5.3 EMPIRICAL INVESTIGATION OF EQUALITY OF MEANS

5.3.1 The Student *t* – Test

Apart from estimating and decomposing TFP indices it was imperative to determine if there was a significant change in the efficiency of South African banking system as a consequence of the global financial crisis. Thus far, evidence deduced from analysing the performance indices in the previous section revealed that most of the efficiency measures clearly deteriorated during the period 2008-2009. In this section, it is empirically established if this deterioration was statistically significant by employing the student`s *t*-test.

TABLE 5.4: EQUALITY OF PRE-CRISIS AND CRISIS MEANS TEST

Equality of Means Test	PRE-CRISIS MEAN (μ_1) (2003 – 2006)	CRISIS MEAN (μ_2) (2007 – 2009)	MEAN DIFFERENCE $(\mu_1 - \mu_2)$	*t-TEST* *P*-value	CONCLUSION
TFPE $H_0 : \mu_1 - \mu_2 = 0$ $H_A : \mu_1 - \mu_2 \neq 0$	0.615	0.575	0.04	0.62	We fail to reject H_o and conclude that the means of the two periods are the same.
CIR $H_0 : \mu_1 - \mu_2 = 0$ $H_A : \mu_1 - \mu_2 \neq 0$	57.6	57.3	0.3	0.95	We fail to reject H_o and conclude that the means of the two periods are the same.

P-values: * / [**]/ (***) denotes significance at 10%, / [5%] / (1%) level of significance respectively.

Table 5.4 shows that the mean TFPE before the crisis was higher than during the crisis period. On the contrary, in terms of CIR, the banking sector performed better during the crisis compared to the pre-crisis period. This is in line with results obtained by Mabwe and Webb (2010) who found that CIR actually improved during the crisis compared to pre-crisis period. However, in both tests regarding TFPE and CIR, the *p*-values are greater than 0.05 implying that the differences in the banking sector performance between the pre-crisis and crisis-period are statistically insignificant. Therefore, one can conclude that both TFPE and CIR indicators were not significantly affected by the global financial crisis. These findings are consistent with the views of the Bank Supervision Annual Report (2009, p.4) that broadly described the financial sector as "remaining vigilant" despite the difficult circumstances that came with the crisis. This vigilance was attributed to effective supervision and regulation of the banking sector.

5.3.2 The Wilcoxon Matched-Pairs Ranks Test

Table 5.5 below presents the results of performing the Wilcoxon test. As stated earlier in the previous methodology chapter, the lesser of the two sums of the ranks of the positive and negative differences becomes the Wilcoxon statistic. According to evidence in Table 5.5 the

sum of the ranks of the positive differences $\left(\sum \boldsymbol{R}^{+} = 16\right)$ is the lesser of the two and therefore consitute our Wilcoxon test statistic of 16 to be compared with a Wilcoxon critical value of 4 $\left(\boldsymbol{W}_{0.05,8} = 4\right)$. Contrary to standard statistical inference the Wilcoxon test is constructed in such a way that if the W-statistic is less or equal to the critical value one rejects the null hypothesis and conclude that there is no difference between the before and the after measures. The results obtained for all the eight banks in the sample indicates that *W-statistic* = 16 > *W-critical* = 4 which leads to non-rejection of H_0. We therefore conclude that there was no difference between the pre-crisis total factor productivity efficiency (TFPE) and the crisis TFPE. Hence the outcome of the Wilcoxon test is consistent with the previous t-test results.

TABLE 5.5: WILCOXON SIGNED RANK TEST RESULTS, ALL BANKS

ALL BANKS	**ABSA**	**FRB**	**NED**	**STAN**	**CAP**	**SAS**	**TEBA**	**AFR**
PRE	0.535	0.818	0.480	0.747	0.107	0.251	0.281	0.801
CRISIS	0.648	0.640	0.561	0.621	0.766	0.367	0.173	0.826
CHANGE	-0.113	0.178	-0.082	0.126	-0.659	-0.116	0.108	-0.024
RANK	**(-) 4**	**7**	**(-) 2**	**6**	**(-) 8**	**(-) 5**	**3**	**(-) 1**
$\boldsymbol{W} = \sum \boldsymbol{R}^{+} = 16$								
$\sum \boldsymbol{R}^{-} = 20$								
The Wilcoxon statistic in this case is the sum of the ranks of positive differences, $\boldsymbol{W} = 16$. The two-tailed Wilcoxon critical value for $n = 8$ at 0.05 level of significance is **4**.								

The Wilcoxon test was then applied to large and small banks separately particularly to examine whether the four largest South African banks which are relatively integrated into the global financial system were affected by the global financial crisis. The results are presented in Table 5.6 and Table 5.7 for large and small banks respectively. In each of the two tests the null hypothesis of equality could not be rejected as the W-statistic exceeded the W-critical value.

TABLE 5.6: WILCOXON SIGNED RANK TEST RESULTS, LARGE BANKS

LARGE BANKS	ABSA	FRB	NEDBANK	STANDARD
PRE	0.535	0.818	0.480	0.747
CRISIS	0.648	0.640	0.561	0.621
CHANGE	-0.113	0.178	-0.082	0.126
RANK	**(-) 3**	**4**	**(-) 1**	**2**
$W = \sum R^+ = 6$				
$\sum R^- = 4$				
The Wilcoxon statistic in this case is the sum of the ranks of negative differences, $W = 4$. The two-tailed Wilcoxon critical value for $n = 4$ at 0.05 level of significance is **0**.				

TABLE 5.7: WILCOXON SIGNED RANK TEST RESULTS, SMALL BANKS

SMALL BANKS	CAPITEC	SASFIN	TEBA	AFRICAN
PRE	0.107	0.251	0.281	0.801
CRISIS	0.766	0.367	0.173	0.826
CHANGE	-0.659	-0.116	0.108	-0.024
RANK	**(-) 4**	**(-) 3**	**2**	**(-) 1**
$W = \sum R^+ = 2$				
$\sum R^- = 8$				
The Wilcoxon statistic in this case is the sum of the ranks of positive differences, $W = 7$. The two-tailed Wilcoxon critical value for $n = 4$ at 0.05 level of significance is **0**.				

Table 5.8 and Table 5.9 provide a general comparison of South Africa with other emerging markets on selected banking indicators of profitability and capital levels. This comparative analysis is intended to partly shed some light as to why the South African banking industry despite the presence of foreign banks relatively managed to come out of the crisis unscathed. Contrary to experience in many emerging and developed countries the SARB Supervision report (2009) noted that South African banks did not require any form of liquidity support from either the government or the Reserve bank. In particular, the report identified adequate capital levels, low leverage ratio, and limited exposure to foreign assets and funding among the main factors contributing towards how the South African banking sector weathered the financial crisis.

TABLE 5.8: CAPITAL ADEQUACY RATIO – A COMPARISON WITH EMERGING MARKETS.

	Capital Adequacy Ratio %		
	2007	***2008***	***2009***
SOUTH AFRICA	**12.8**	**13**	**14.1**
BRAZIL*	18.7	18.3	18.8
CHILE	12.2	12.5	14.3
CHINA*	8.4	12	11.4
CZECK REPLUBLIC	11.6	11.6	14
GREECE	11.2	9.4	11.7
HUNGARY	10.4	11.2	12.9
INDIA*	12.3	13	13.2
KOREA	12.3	12.3	14.4
LITHUANIA	10.9	12.9	14.2
PAKISTAN	12.3	12.3	14.1
PERU	12.1	11.9	13.5
POLAND	12	11.2	13.3
RUSSIA*	15.5	16.8	20.9
Red shading indicates better performance than South Africa. *Emerging BRICS countries are highlighted by an asterisks.			

Source: IMF Global Financial Stability Report (2010)

TABLE 5.9: PROFITABILITY INDICATORS – A COMPARISON WITH EMERGING MARKETS.

	Return on Equity (ROE) %			Return on Assets (ROA) %		
	2007	*2008*	*2009*	*2007*	*2008*	*2009*
SOUTH AFRICA	**18.1**	**28.7**	**15.9**	**1.4**	**2.1**	**0.9**
BRAZIL*	28.8	15.3	20.4	2.9	1.5	1.9
CHILE	16.2	15.2	18	1.1	1.2	1.2
CHINA*	16.7	17.1	15.1	0.9	1	0.8
EGYPT	15.6	14.1	13	0.9	0.8	0.8
GREECE[14]	14.8	3.2	-1.5	1	0.2	-0.1
HUNGARY	18.4	11.6	9.8	1.2	0.8	0.7
INDIA*	13.2	12.5	12.3	0.9	1	1
KOREA	14.6	7.2	5.8	1.1	0.5	0.4
LATVIA	24.3	4.6	-41.6	2	0.3	-3.5
PAKISTAN	15.4	7.8	8.6	1.5	0.8	0.9
PHILIPPINES	8.7	6.9	10.8	1	0.8	1.2
POLAND	24.9	20.5	10.7	1.9	1.5	0.8
ROMANIA	11.5	17	2.7	1.3	1.6	0.2
RUSSIA*	22.7	13.3	4.9	3	1.8	0.7
THAILAND	1.2	10.3	9.5	1	1	1.1

Red shading is an indication of better performance than South Africa.
*Emerging BRICS countries are highlighted by an asterisks.

Source: IMF Global Financial Stability Report (2010)

5.4 CONCLUSION

This chapter evaluated the total factor productivity and efficiency of the eight South African commercial banks categorised according to their relative sizes for the period 2004 – 2011. Two methodological approaches were employed to measure efficiency namely, financial ratio analysis and the non-parametric approach. While the non-parametric approach remained the main methodology of analysis, financial ratios were used to complement and to deepen our understanding of bank efficiency. Only particular ratios that relate to the chosen banking outputs and inputs were selected for analysis. These selected ratios were analysed with the hope that they would shed an even greater light for future analysis on the potential feedback channels through which changes in efficiency may impact on access to bank services and

[14] In 2013 Greece was downgraded by MSCI from being a developed country to emerging market.

unemployment. Analysed ratios included cost-to-income ratio a proxy for cost efficiency, net interest margin a proxy for intermediation efficiency, non-interest income to gross revenue a proxy for risk diversification, and NPLs to gross loans a proxy for quality of services. The measurement and the decomposition of total factor productivity into several measures of efficiency was performed using the program *DPIN 3.0* developed by O'Donnell (2011). The BCC model was specified and the DEA program was run with the variable returns to scale (VRS) assumption.

Total factor productivity efficiency (TFPE) constitutes the main variable of interest in this study for exploitation in the second stage analysis. The findings revealed that large banks performed relatively better in terms of TFPE. All large banks recorded above their group average TFPE score of 0.68 while 50 percent of small banks performed below their group average of 0.51. The researcher noted that the variation or dispersion was much wider for small banks than for large banks highlighting that there is more scope for improvement particularly for small banks. The study postulated that the variance in efficiency measures between small and large banks could reflect the difference in their business models. Akhigbe and McNulty (2005, p.298) argues that the business model of the small bank is designed to offer "customized and personalized service but at high cost, while larger banks aim to deliver relatively uniform financial services to large groups of customers at lower cost". Moreover, larger banks are more likely to reap economies of scale due to their size resulting in greater efficiency. Lastly we argue that large banks in South Africa operate a mixed business model where they offer both retail and wholesale corporate banking. As such the blend enables them to diversify their business operations allowing them to deliver better efficiency than small retail-oriented banks. However, we acknowledge that in South Africa some small banks handle corporate transactions but not in the same scale as large wholesale corporate banks.

Finally, we performed the Student's *t* test and the Wilcoxon Matched-pairs Signed Ranks test to determine if there was a significant change in the efficiency of the banking sector as a result of the global financial crisis. The results for each performed test revealed the difference in both the mean TFPE measures and cost-to-income ratio indicators between the pre-crisis period and crisis-period to be statistically insignificant. This result was found to be consistent with the views of the SARB report (2009, p.4) that indicated that the banking sector "remained liquid and well capitalised".

CHAPTER SIX

EFFICIENCY AND ACCESS TO BANK SERVICES IN SOUTH AFRICA

6.1 INTRODUCTION

One of the main objectives of this study is to investigate the nature of the relationship between bank efficiency gains and access to bank services in South Africa. According to a Paulson and McAndrews (1998), the provision of banking services for clients of low income status is difficult to do profitably. Therefore given the relatively higher costs normally related with serving the low-income clientele, banks find it rational to serve the high-income clients as the profit potential is greater. A dilemma exists when the same banks must expand services to low income groups which constitute a major segment of the population. According to World Bank (2013) indicators the percentage of South Africa`s population living on less than $2.00[15] a day in 2009 was 31.3. One aspect of efficiency known as allocative efficiency prescribes that resources should be employed to the activities with the greatest expected value. According to Okeahalam (2006), banks argue that transactions done by the low-income clients are small in terms of returns and yet high in volume driving up the average cost of production. Therefore the expected returns and profitability to be realised from providing services to poor rural and remote clients is low. We therefore argue in this study that bank`s appetite to achieve good scores on technical, cost or profit efficiency appear to hold the potential of reducing access to services for consumers particularly the poor majority. In a nutshell this study is an attempt to shed light on whether the attainment of greater efficiency by banks results in enhanced or reduced access for consumers.

6.2 PRESENTATION OF EMPIRICAL RESULTS

6.2.1 Stationarity Test Results

This section provides empirical results on the impact of bank efficiency on access to bank services using the appropriate panel data estimation method. We first perform unit root tests on the data series to establish whether the series are stationary or not. A crucial preliminary step in the process of building a robust econometric model is to understand the time series properties and characteristics of the data involved. It is therefore crucial to test for stationarity

[15] Using 2005 International Prices

of each panel series to be used in the estimation. A series is said to be stationary if it's mean, variance and covariance structure do not change over time. Disregarding the problem of non-stationarity (unit root) when it is actually present leads to spurious or nonsensical results. Several panel data unit root tests were performed on the data and the results are presented in Table 6.1. The IPS unit root which is accredited to Im, Pesaran and Shin (2003) was more preferred to those proposed by Levin & Lin (1992) and Madalla & Wu (1999). De Wet and Van Eyden (2005) states that the IPS test preserves small sample properties and that it is generally more intuitive in its structure than the Levin and Lin (1992) test. However, all the tests produced stationary variables in levels (see Table 6.1).

TABLE 6.1: PANEL DATA UNIT ROOT TESTS

Tests	***Unit root process (Common / Individual)***	***Statistic***	***P-value***
Null Hypothesis	Each individual series contains a unit root		
Levin, Lin, Chu (LLC)	Common	*-10.4511*	0.0000***
Im, Pesaran & Shin (IPS)	Individual	*-2.70597*	0.0034***
ADF - Fisher Chi-square	Individual	*138.358*	0.0030***
PP - Fisher Chi-square	Individual	*151.575*	0.0003***

* / [**]/ (***) denotes significance at 10%, / [5%] / (1%) level of significance respectively.

6.2.2 Diagnostic Tests: Pooled vs Fixed Effects

Having performed the necessary unit root test and confirmed that there is no unit root present in the panel data, we proceed to the next step. The next step is to choose the appropriate estimation method. As detailed in the methodology section there are three panel data estimation methods namely, pooled regression model, fixed effects model (FEM) and random effects model (REM). Both the FEM and REM takes into account the bank-specific features while the pooled OLS model pools all cross sections together and estimate a common regression model disregarding the heterogeneity or distinctiveness of the cross-sections. The random effects model is appropriate if draws are made randomly from a large population where N is generally large relative to T (Baltagi, 2008, p.17). However our panel sample $N = 8 = T$ is not sufficiently large to permit the use of the random effects model. As such the pooled model versus random effects model test was not carried out.

To decide between a pooled regression model (restricted) and a fixed effects model (unrestricted), the F-test of pooled regression model (restricted) versus individual fixed effects model (unrestricted) was constructed and confirmed the fixed effects model as our robust and representative model. Under the null hypothesis that cross-sections are homogeneous, the F-test constructed was as follows:

$$F\ statistic = \frac{(RSS - URSS)/(n-1)}{URSS/(nt-n-k)} \sim F_{(n-1),(nt-n-k)}$$

$$F\ statistic = \frac{(0.000182 - 0.0000683)/7}{0.0000683/41} = 9.75047 \sim F_{7,41,0.05} = 2.2429$$

The decision is to reject the null hypothesis if the F-statistic exceeds the F-critical value. Since F-statistic = 9.75 > F-critical = 2.24 we reject the null hypothesis that our sample of banks are homogeneous implying that the FEM allowing for bank-specific effects is a better model specification. EViews redundant fixed effects likelihood ratio test confirmed the same results and conclusion (see Table 6.2). Again EViews external programs for pooled versus individual fixed effects test were utilised to check consistency of our conclusion. Both the F-test and χ^2 (Chi-Square) tests results confirmed that the FEM was the appropriate model that fit our data.

TABLE 6.2: DIAGNOSTICS TESTS

Test	Test Statistic	Critical Value / P-values	*Inference*
Redundant Fixed Effects Test H_0 : Cross-sections are homogenous H_1 : Cross-sections are heterogeneous	F = 9.747 χ^2 = 54.871	P-value = 0.000 P-value = 0.000	*We reject H_0 and conclude that bank specific features should be accounted for using the FEM.*
Pooled vs. Fixed Effects: $H_0 : \mu_1 = \mu_2 = ... = \mu_{N-1} = 0$ H_A :*Not all equal to 0*	F = 4.3923	$F_{7,41,0.05} = 2.2429$	*We therefore reject H_0 implying that the FEM is a better model to allow for bank heterogeneity.*
Test for Serial Correlation: $H_0 : \rho = 0$ (no serial correlation) $H_A : \rho \neq 0$ (serial correlation)	DW = 2.094	No correlation if: $D_U < D_W < 4 - D_U$ $1.8851 < D_w < 2.1149$	*We conclude that the iterative procedure did remove negative serial correlation*

The estimated correlation parameter ρ = - 0.458918 was used to correct serial correlation through the generalized least squares (GLS) estimation method. We concluded that the iterative procedure did remove serial correlation based on the improvement noted of a Durbin-Watson from 2.99 to 2.09.

Heteroscedasticity: $H_0 : \sigma_i^2 = \sigma$ (homoscedastic errors) H_A :*Not equal for all i* (heteroscedastic errors)	White diagonal standard errors and covariances were used to correct the problem of heteroscedasticity.

6.2.3 Heteroscedasticity – Remedial Measures

In line with the suggestion by Baltagi (2008), the White-diagonal standard errors and covariances were used to correct for the presence of possible heteroscedasticity. The problem of heteroscedasticity occurs often with cross sectional data as opposed to time series data. Baltagi (2008) argues that in panel data analysis, the assumption of homoscedasticity may not be plausible due to the different variation in sizes of the cross sections. Three White coefficient covariance methods are considered namely, White-cross-section, White-period

and White-diagonal. The rule of thumb is to use the White diagonal method if our values for N and T satisfies $\frac{1}{2}N < T < 2N$. Recall that $N = T = 8$.

Classical linear regression modelling requires that the variance of the residuals be homoscedastic or constant irrespective of the values of the independent variables. Heteroscedasticity therefore arises when the variance or spread of the residuals is not constant. The need for equal spread of residuals comes about because in the estimation of a regression function, OLS assigns equal weight (importance) to every observation when minimising the residual sum of squares (RSS). Ideally in order to accurately estimate a regression function, observations that are closer to their respective mean values should be given more weight relative to those that are scattered about (Gujarati and Porter, 2010). Baltagi (2008, p.87) states that assuming homoscedastic residuals when heteroscedasticity is present will still yield unbiased and consistent estimates of regression parameters but the estimates will no longer be efficient. The author further argues that the standard errors of these estimates will be biased and therefore suggests the use of robust standard errors to correct for possible presence of heteroscedasticity. Hence if one persist and use OLS, the *t* and *F* tests obtained will be distorted leading to wrong inference.

6.2.4 Serial Correlation – Remedial Measures

We also tested for the problem of serial correlation in the model. One important assumption underlying classical linear regression is that the residuals be uncorrelated or be independent of one another. The problem of serial correlation is common with time series data and as such also require attention in panel data estimation. Gujarati (2004) demonstrates that if the assumption of no serial correlation is violated the OLS estimators although unbiased and consistent are no longer efficient. In other words the variance will no longer be minimum and so the standard errors of the derived estimates will be biased. As a result, the standard t , F and χ^2 tests will not be validly performed. The original FEM showed evidence of negative correlation with a D-W statistic of 2.98 as shown in column 3 of Table 6.3. The necessary corrective procedures were performed to transform the original data using the Generalised Least Squares (GLS) technique. Studenmund (2011) defines the GLS as a method of eliminating pure first-order correlation from an equation thereby restoring the minimum variance property to its estimation. Therefore GLS is simply "OLS applied to the transformed model that satisfies the classical assumptions" (Gujarati and Porter, 2010, p. 442).

The first step in GLS estimation is to run the regression that contains first-order serial correlation: $Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it}$ where $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + u_{it}$ **.....6.1**

Here, *Y* is the dependent variable access and *X* denotes a vector of all included explanatory variables while ρ is the serial correlation coefficient. As indicated, equation 6.1 contains a composite error term that is made up of the serially correlated component $\rho\varepsilon_{i,t-1}$ and the classical (not serially correlated) error term u_{it}. To remove the serially correlated component from equation 6.1, we transform it using the GLS technique as follows:

$$\left(Y_{it} - \rho Y_{i,t-1}\right) = \beta_0\left(1-\rho\right) + \beta_1\left(X_{it} - \rho X_{i,t-1}\right) + u_{it} \quad \textbf{.....6.2}$$

This transformation [equation 6.2] now represents an equation that no longer contains the serially correlated component and can be re-stated as:

$$Y_{it}^{*} = \beta_0^{*} + \beta_1^{*} X_{it}^{*} + u_{it} \quad \textbf{.....6.3}$$

where $Y_{it}^{*} = Y_{it} - \rho Y_{i,t-1}$, $X_{it}^{*} = Y_{it} - \rho X_{i,t-1}$, $\beta_0^{*} = \beta_0 - \rho\beta_0$

In empirical research, the challenge is that the correlation parameter ρ that is used in the GLS estimation to correct serial correlation is not known and have to be estimated using inter alia the Cochrane-Orcutt iterative procedure or the Durbin two-step procedure. This study applied the popular Cochrane-Orcutt (C-O) iterative procedure. Gujarati and Porter (2010) argue in favour of the C-O iterative procedure citing one of its advantages as its capacity to estimate higher order autoregressive systems. Returning to our access-efficiency regression, we used the C-O iterative procedure by running this regression model containing serial correlation:

$$ln\,ACC_{it} = \alpha_{it} + \beta_1\,lnTFPE_{it} + \beta_2\,ln\,BC_{it} + \beta_3\,ln\,BDEV_t + \beta_4\,ln\,BRAN_t + \beta_5\,ln\,HHI_t + \beta_6\,ln\,GDP_t + \beta_7\,ln\,RUR_t + \mu_{it}$$

..... **6.4**

To estimate the correlation parameter $\hat{\rho}$, the residuals ,*v,* generated in equation 6.4 were then regressed on their one period lagged values according to the following specification:

$$v_{it} = \rho v_{it-1} + u_{it} \quad \textbf{.....6.5}$$

Running the regression 6.5 yielded $\hat{\rho} = -0.458918$. This estimated correlation parameter was then used to transform our panel series as follows: $Y_{it}^{*} = Y_{it} - \left(-0.458918\right)Y_{i,t-1}$, and $X_{it}^{*} = Y_{it} - \left(-0.458918\right)X_{i,t-1}$. Equation 6.4 was then re-estimated but with the transformed variables. These steps are repeated until further iterations remove the problem of serial correlation. Brooks (2008) states that in practice a small number of iterations usually less than 5 is generally sufficient. However, after performing the C-O procedure once and

carrying out the GLS estimation, the resultant model was subsequently tested to establish if the problem was removed. We did not expect the problem of serial correlation to be persistent due to the short time dimension of our panel. As state earlier, autocorrelation or serial correlation is mainly a problem with time series data. Based on the noted improvement of a change in the D-W statistic from 2.99 to 2.09, we then concluded that the GLS remedial procedure did remove serial correlation. The final GLS-FEM results of estimating equation **6.4** are tabulated in Table 6.3 column 4.

TABLE 6.3: ACCESS TO BANKING SERVICES IN SOUTH AFRICA (2004 – 2011)

	Pooled OLS Model	**Fixed Effects Model 1**	**Fixed Effects (GLS) Model 2**
Dependant Variable	***Deposit accounts with commercial banks***		
Constant	205.7766*** *(0.0000)*	213.7431*** *(0.0274)*	**185.1434*** *(0.0274)***
Efficiency	0.00159 *(0.4309)*	0.00174 *(0.3297)*	**0.00125** *(0.0193)***
GDP per capita	1.9769*** *(0.0000)*	1.9478*** *(0.0000)*	**2.4939*** *(0.0000)***
BDEV	0.8447*** *(0.0000)*	0.8153*** *(0.0000)*	**1.0708*** *(0.0000)***
Branch Network	0.8691*** *(0.0000)*	0.8632*** *(0.0000)*	**0.7499*** *(0.0000)***
Herfindahl Hicksman Index	- 4.7433*** *(0.0000)*	- 4.9513*** *(0.0000)*	**- 4.0866*** *(0.0000)***
Bank cost	- 0.0148*** *(0.0000)*	- 0.0184*** *(0.0000)*	**- 0.00952*** *(0.0000)***
Growth of Rural Population	- 51.0804*** *(0.0000)*	- 52.9247*** *(0.0000)*	**- 33.1733*** *(0.0000)***
Adjusted R^2	**99.91**	**99.94**	**99.99**
D-Watson statistic	**2.049**	**2.987**	**2.094**

P-values are reported in parentheses:

* / [**]/ (***) denotes significance at 10%, / [5%] / (1%) level of significance respectively.

6.2.5 Discussion of Results

Total Factor Productivity Efficiency

Of paramount value to the present study was to establish the significance and magnitude of the impact of bank efficiency gains on access to banking services. A pofsitive and significant coefficient found implies that access to bank services is influenced by the efficiency with which banks operate. On average, a 10 percent increase in bank efficiency improves access to bank services by 0.013 percent. Although the size of the marginal effect of bank efficiency appear to be small we argue that these effects are economically significant. For illustration using 2009 South African bank access data, a 0.013 percent increase in access represents 12 deposit accounts per 1000 adults as a result of a 10 percent improvement in banking sector efficiency. Similarly, using 2012 latest IMF (2013) data of 1373 deposit accounts per 1000 adults for South Africa, a banking sector efficiency gain of 10 percent would contribute 18 deposit accounts per 1000 adults. Hence in the light of a World contribution of 65 deposit accounts per 1000 adults recorded in 2009, we argue that a 0.013 percent contribution of 18 accounts resulting from a 10 percent improvement in bank efficiency is economically significant particularly so from the standpoint of a single country.

The FEM component showed the individual intercepts for large banks to be small compared to those for smaller banks. Gujarati and Porter (2009) states that these intercepts within a fixed effects model captures the different special features of each firm such as managerial style or the type of market each firm is serving. Therefore bank specific features appear to be important in explaining access to services. Our results imply that small banks generally contribute more towards access. In recent times small banks particularly Capitec and African bank have been very innovative and instrumental in reaching the unbanked. However, we expected large banks which are relatively more efficient to have contributed more towards reaching the unbanked.

This reported positive impact of bank efficiency runs contrary to our earlier expectation of an inverse relationship between bank efficiency and access. Our earlier point of departure was that banks in their pursuit for better efficiency scores neglect serving the low-income or low-profile households. The reason being that banks find it rational to serve the high-income or high-profile households as the profit potential are greater thereby suggesting the possibility of a trade-off. The average cost of extending services to the poor and rural households is usually

high and the expected level of profit is generally low. Given that a considerable proportion of South Africans constitute the low-income and poor people we were more inclined to suspect a negative association with access. We however underscore in the light of our results the need for banks to remain efficient in order to augment efforts towards improving accessibility to bank services for the unbanked South African people. Nonetheless, we argue that attempts by banks to remain profitable will not necessarily affect outreach. In recent years the success stories of Capitec and African bank the largest providers of unsecured lending have challenged the existence of a trade-off between sustainability (profitability) and outreach. Hence banks have the potential to extend their services to low-income and to the poor and still remain sustainable.

GDP per capita

The coefficient on GDP was significant and positive as expected, indicating that on average increases in output/income are associated with higher level of access to banking services. Income facilitates easy access to bank services by enabling households to afford the related cost of accessing bank services. Our results are consistent with those obtained by Okeahalam (2008) who reached the conclusion that income has an impact on the overall level of access to bank services. The author found higher GDP per head and bank efficiency to be positively related with access to deposits. A positive and significant coefficient of 2.49 implies that a 1 percent improvement in GDP per capita results in a more than proportionate increase in the access to bank services of 2.49 percent. The level of access to bank services is therefore highly income elastic.

Banking sector development

The variable BDEV which is a proxy for banking sector development is also reported as statistically significant and bears the expected sign. This is consistent with our earlier hypothesis that developments within the banking sector as approximated by an increase in the volume of credit to the domestic economy has a positive impact on access to bank services.

Number of Branches

BRAN which is a measure of the number of bank branches per 100 000 is reported positive and significant confirming a positive link with access to bank services as previously predicted. It is argued that more branch penetration provides more scope for choice and is likely to encourage access. Hence branch intensity is an important factor that influences

access to bank services. This is in line with the result reported by Ikhide (1996) who showed that increased availability of bank branches can increase savings by improving the banking habit. However we highlight the need to evenly expand these branches throughout the country particularly in the marginalised areas. Okeahalam (2008) argues that many bank branches are clustered in urban areas while the majority South African live in townships and rural areas.

Herfindahl Hicksman Index

The coefficient on banking market concentration was significant and negative in line with economic theory implying that lack of competition as indicated by an increase in bank market concentration (HHI) reduces the level of access by 4.4 percent. Generally, our finding lends support to previous studies suggesting that increased competition results in lower prices and hence increased access. Beck et al (2003) in their study of the impact of bank competition on firm`s access to credit found bank market concentration to be negatively related with access to credit particularly for developing countries. However, the authors found that this relationship turned insignificant for developed countries. Our findings underline the importance of reducing the entry barriers to encourage wider participation thereby promoting competition among bank service providers. Simpasa (2013) argue that banking sector competitiveness reduces the cost of financial intermediation and enhances the delivery of quality services. Bank market concentration appears to affect the level of access via the efficiency channel. Our point of departure is that an increase in the concentration of the banking market increases the banks` market power and hence less exposure to competition. Lack of competition signals a lack of incentives for achieving more output with minimum input usage.

Bank cost

A negative and significant bank cost variable confirmed our earlier expectation of a negative impact on access. Clients opt for banks that will provide them with the highest service satisfaction at lowest cost. From the supply side, banks attempt to minimise their operating cost while maximising profitability. A bank that operates with high cost is likely to pass on the cost to its clients to maintain its profit margins. Hence, the overall effect of high bank operating cost is to reduce the level of banking access. Our findings show that on average a 10 percent increase in operating costs reduces access by 0.09 percent. Using the latest 2012

South African data this represents a decline of 92 deposit accounts per 1000 adults resulting from a 10 percent increase in bank operating cost.

Growth of rural population

A significant and negative coefficient confirmed the notion that a larger share of the rural population increases bank inefficiency by increasing costs thereby reducing access to banking services. A negative coefficient of 33.2 implies that a 1 percent general increase in the rural population results in a more than equivalent decrease in access to services of 33.2 percent. This finding is in line with that of Kablan (2010) whose paper investigated the determinant factors of cost efficiency in sub-Saharan Africa (SSA). The variable representing the proportion of population living in the rural areas produced a negative sign with cost efficiency. Kablan (2010) argues that SSA banks tend to locate their branches in more economically developed regions at the expense of rural ones and citing that banks with a high market share of rural population tends to be less cost-efficient because they cannot realise economies of scale. From this background, we speculate that banks are somewhat biased against providing services to the low-income rural population. Hence we suggest that investment in rural infrastructure could help in this regard. It is worth mentioning that of the various factors included in our model, the growth of rural population variable had the greatest marginal impact on access. Hence overall, this reveals that improvements in the level of access will be more pronounced via infrastructure investment in the rural areas.

6.3 CONCLUSION

In this chapter we have analysed the nexus between gains in bank efficiency and access to banking services. In order to study this relationship, we applied panel data techniques, in particular the fixed effects model. Our findings confirmed a positive and significant relationship between gains in bank efficiency and access. Evidence obtained from analysing bank fixed effects pointed to the fact that large banks were not as proactive as small banks in promoting the goal of expanding access. We expected large banks which are relatively more efficient to have contributed more towards reaching the unbanked. Instead, our results imply that small banks generally contribute more towards access compared to large banks. In recent times small banks particularly Capitec and African bank in South Africa have been very innovative and instrumental in reaching the unbanked. We also found bank branch penetration to have a positive influence on access through providing an expanded range of choice to clients. Among the chosen factors included in our model, the rural-population variable was found to have the largest but negative relationship with access. Hence, we confirmed the notion that serving a larger share of the rural population increases bank inefficiency by increasing costs hence the trade-off relationship. In light of this outcome, we propose that the government support investment in rural infrastructure as one way to broaden access on a larger scale. Lastly, the study reported that a highly concentrated bank market and high bank operating costs have a detrimental impact on access to bank services.

CHAPTER SEVEN

EFFICIENCY AND UNEMPLOYMENT IN SOUTH AFRICA

7.1 INTRODUCTION

Unemployment is one of the crucial challenges facing South Africa with the official unemployment rate currently at 25.6 percent (STATSSA, 2013). The 2011 annual budget highlighted unemployment as a problem which could have many adverse consequences if not curbed. National Treasury (2011, p.17) identified enhancing "...efficiency across the economy" as one important job driver for accelerating growth and employment creation in achieving the targets of the New Growth Path. In a recent budget speech statement, Mr Gordhan (2013, p.7), the Minister of Finance, also highlighted the need "to raise productivity, and diversify the economy, to create jobs and raise living standards." In this chapter, the researcher unpacks the relationship between banking sector efficiency improvement and unemployment and provides strategic and forward-looking lessons for the banking sector. This is crucial in order to inform formulation of suitable banking sector policies that are supportive of employment creation in line with the targets of South Africa`s New Growth Path.

STATSSA (2013) quarterly labour force survey reported that both formal and informal sector employment increased by 109 000 and 30 000 persons respectively in the second quarter compared to the first quarter of 2013. The formal sector employment increase of 109 000 was largely driven by the finance industry that contributed 47 000 persons representing 43 percent of employment growth during the period. The Banking Association South Africa (BASA, 2012) reported that the financial sector assets totalled over R6 trillion in 2012 with the banking sector accounting for over 50 percent. As such, banks are the most dominant institutions within the finance industry. The banking sector acts as a catalyst for economic growth and as such holds great potential for employment creation through various transmission mechanisms. Therefore, an unstudied but important issue is whether changes in banking sector efficiency are reflected in national employment. In other words, does the enhancement of bank efficiency confer positive gains as far as South Africa`s employment generation is concerned? This chapter seeks to establish the direction and magnitude of the impact of banking sector efficiency on unemployment in South Africa.

The chapter is structured as follows: The following section, presents results of unit root or stationarity test and panel data estimation. The subsequent section discusses the results. Finally, the last section provides a conclusion to the chapter.

7.2 PRESENTATION OF EMPIRICAL RESULTS

7.2.1 Stationarity Test Results

In keeping with standard procedure for ensuring robust econometric model construction, the panel series was tested for the problem of unit root using the panel data techniques applied in the preceding chapter. The previous chapter elaborated what stationarity is and its significance in estimation. The results of the four unit root tests, LLC, IPS, ADF-Fisher and PP-Fisher are summarised in Table 7.1 below. The results suggest that the null hypothesis of the presence of unit root in levels was rejected denoting that all variables are stationary in levels.

TABLE 7.1: PANEL DATA UNIT ROOT TESTS

Tests	***Statistic***	***P-value***
Null Hypothesis	**Each individual series contains a unit root**	
Im, Pesaran & Shin (IPS)	-5.59811	*0.0000***
ADF - Fisher Chi-square	197.312	*0.0000***
PP - Fisher Chi-square	132.718	*0.0078***
Null Hypothesis	**Assumes common unit root process**	
Levin, Lin & Chu (LLC)	-17.8576	*0.0000***

* / [**]/ (***) denotes significance at 10%, / [5%] / (1%) level of significance

7.2.2 Diagnostic Tests: Pooled vs Fixed Effects

After performing stationarity test and having established that there is no unit root present in the panel data, we continued to the next step. The same procedure as in the preceding chapter was followed in order to determine the choice of the appropriate panel data technique to be used. As mentioned earlier in the initial analysis of the main objective, the panel sample $N = T = 8$ was not sufficiently large to permit the use of the random effects model. Baltagi (2008, p.17) states that the random effects model is suitable when the sample population is generally large particularly where N is greater than T. As such the pooled model versus random effects model test was not carried out. Hence the F-test of pooled regression model versus fixed

effects model was constructed and confirmed the pooled model as the representative model. Under the null hypothesis that cross-sections are homogeneous, the F-test constructed was as follows:

$$F\ statistic = \frac{(RSS - URSS)/(n-1)}{URSS/(nt-n-k)} \sim F_{(n-1),(nt-n-k)}$$

$$F\ statistic = \frac{(0.013546 - 0.013466)/7}{0.013466/51} = 0.0433 \sim F_{7,51,0.05} = 2.195$$

Since F-statistic = 0.0433 < F-critical = 2.195 we failed to reject the null hypothesis that our sample of banks are homogeneous indicating that individual effects are not valid and should not be accounted for. The original pooled regression results showed evidence of positive correlation with a Durbin-Watson statistic of 1.26 as shown in column 2 of Table 7.2. However, the necessary corrective procedures were performed to transform the original data using the Generalised Least Squares (GLS) technique. The cross-section Seemingly Unrelated Regression (SUR) weights were chosen to correct for heteroscedasticity and contemporaneous correlations (Zellner, 1962). The final pooled GLS model results are tabulated in Table 7.2 column 3. The pooled OLS results were reported for comparative analysis in column 2 of Table 7.2. These results indicated significant improvement of the D-Watson statistic from 1.26 to 2.07. We estimated the following translog function:

$$LnUNEM_t = \alpha_{it} + \beta_1 LnTFPE_{it} + \beta_2 LnSPR_{it} + +\beta_3 LnCIR_{it} + \beta_4 LnBDEV_t + \beta_4 LnGDPG_t + \mu_{it}$$

..... [**7.1**]

Where**:**

$UNEM_t$	-	Unemployment rate in period *t*
$TFPE_{it}$	-	Total factor productivity efficiency for bank *i* in period *t*
SPR_{it}	-	Interest spread (intermediation efficiency) for bank *i* in period *t*
CIR_{it}	-	Cost to income ratio (cost efficiency) for bank *i* in period *t*
$BDEV_t$	-	Domestic credit provided by banking sector (% of GDP) in period *t*
$GDPG_t$	-	GDP growth in period *t*
μ_{it}	-	Error term to capture other possible factors not specified.

TABLE 7.2: BANKING SECTOR EFFICIENCY AND UNEMPLOYMENT IN SOUTH AFRICA

	Pooled OLS Model	**Pooled GLS Model**
Dependant Variable	**Unemployment rate**	
Constant	2.3134*** *(0.0000)*	2.4682*** *(0.0000)*
TFPE	-0.0233 *(0.7348)*	-0.0227*** *(0.0000)*
SPR	0.0108 *(0.8117)*	0.0320** *(0.0443)*
CIR	0.0054 *(0.7658)*	0.0032 *(0.0000)*
BDEV	-0.4169*** *(0.000)*	-0.4874*** *(0.0000)*
GDPG	-0.0020 *(0.1333)*	-0.0027*** *(0.0000)*
Adjusted R^2	**0.306**	**0.902**
D-W statistic	**1.26**	**2.07**
F-statistic	**6.55**	**116.62**

P-values are reported in parentheses:

* / [**]/ (***) denotes significance at 10%, / [5%] / (1%) level of significance respectively.

7.2.3 Discussion of Results

This section discusses the findings based on the final pooled GLS model results that are reported in Table 7.2 (column 3). All explanatory variables included in the model are highly significant. The F-statistic also indicated that the included explanatory variables are jointly significant in explaining the dependent variable which is unemployment. The reported adjusted R^2 value of 0.902 for the pooled GLS model implies that our model explains 90.2 percent of the variation in unemployment.

Of paramount value to the present study is establishment of the significance and magnitude of the impact of gains in banking sector efficiency on unemployment. Table 7.2 shows that the coefficient on total factor productivity efficiency (TFPE) was significant at 5 percent level with a negative sign in line with our earlier hypothesis. A negative sign implies that positive gains in banking sector efficiency are associated with a reduction in unemployment in the economy. A significant coefficient found means that unemployment is indeed influenced by

the efficiency with which banks operate. However, the size of the incremental effect appears to be marginal. On average, a 1 percent increase in bank efficiency stimulates a decrease in unemployment by 0.023 percent. However, the size of the marginal effect is economically significant if evaluated from the background that the finance industry on average contributes 4 percent to employment every year. This reported negative impact of bank efficiency on unemployment is in line with our earlier expectation. We therefore, underscore the need for the banking sector to maintain high efficiency in order to augment efforts to achieve the objectives of the New Growth Path aimed at creating five million jobs in South Africa by 2020. We also advocate for banking sector policies and incentives that are directed at enhancing the efficiency of the banking sector.

A positive and significant interest spread variable confirmed our earlier expectation of the adverse impact of wide spreads on unemployment. This variable was included to capture bank intermediation efficiency. Ikhide (2008) argues that historically, government regulations and policies were held responsible for wide spreads. He however, argues that with the advent of financial liberalisation in many African countries, efficiency studies in banking have attributed wide interest spreads to be a consequence of inefficiency. The findings show that, a 1 percent increase in interest spreads exacerbates unemployment by 0.011 percent. All other things being equal, we expect narrow interest spreads to encourage greater mobilisation of savings and to stimulate increased loan demands for business investment and household consumption generating job opportunities in the process. We therefore, confirm the intermediation efficiency channel to be one conduit through which the banking sector contributes towards reducing unemployment in the economy.

Cost-to-income ratio, a proxy for cost efficiency was found to be significant and carrying a positive sign in line with our earlier hypothesis. This was expected given that deterioration in cost efficiency as indicated by an increase in cost relative to income generally increases the cost of doing business and so increases unemployment. A bank that operates with high cost is likely to pass on the cost to its clients to maintain its profit margins. Hence, the overall effect of cost inefficiency is to worsen the level of unemployment. Our findings show that on average a 10 percent increase in cost inefficiency (cost-to-income ratio) increases unemployment by 0.054 percent. These results confirm the cost channel and the interest spread channel as some of the potential channels through which gains in bank efficiency transmit through to reduced unemployment.

BDEV, a variable that represents the logarithm of the volume of domestic credit provided by the banking sector as a percentage of GDP was negative and significant at 1 percent level of significance as expected. This means that the volume of domestic credit to the economy exert a negative influence on unemployment. This variable which is a traditional proxy for banking sector development, had the largest marginal effect on unemployment. For example a 1 percent increase in the volume of domestic credit provided by the banking sector contributed a 0.49 percent decrease in unemployment. Hence, both financial deregulation and liberalisation presuming they result in increased credit/GDP ratio have a significantly large impact on efforts geared towards reducing unemployment in South Africa.

Lastly, Table 7.2 shows that economic growth is statistically significant at 1 percent level of significance and carrying a negative sign. The estimated elasticity coefficient indicates that following a GDP growth of 1 percent unemployment decreases by 0.0027 percent. This economic growth variable exerted the weakest influence on unemployment relative to all variables captured by the model. In recent years, South Africa`s economy has presented a classic case of job-less growth. The size of the marginal impact of GDP growth possibly highlights the fact that in the past years economic growth in South Africa has not been accompanied with increased job opportunities particularly in the formal sector.

7.3 CONCLUSION

This chapter analysed the relationship between bank efficiency and unemployment in South Africa using a panel of 8 banks for the period 2004 – 2011. In safe keeping with robust econometric model estimation procedure, it tested the panel series for stationarity using four methods namely, IPS, LLC, ADF-Fisher and PP-Fisher. All the tests produced stationary variables in levels. Panel data techniques were then applied, in particular the pooled GLS model. Contrary to our expectations, the F-test null hypothesis of cross-section homogeneity was not rejected implying that individual effects could not be accounted for. However, the original pooled regression model showed evidence of positive correlation and the necessary corrective procedures were performed to transform the original data using the Generalized Least Squares (GLS) technique. The cross-section SUR weights were then chosen to correct for heteroscedasticity and contemporaneous correlations that are common with panel series.

Our final pooled GLS regression results confirmed a negative and significant relationship between bank efficiency and unemployment. Hence, there is indeed evidence that improvement in the banking sector efficiency transmit through to reduced unemployment in the economy. We also found both interest spreads (a proxy for intermediation) and cost-to-income ratio (a proxy for cost efficiency) to be positively related with unemployment. An increase in each of these two factors interpreted as inefficiency was found to exert an increasing effect on unemployment. Therefore, we have confirmed the bank cost channel and the interest spread channel as potential channels through which bank efficiency gains transmit through to reduced unemployment. Furthermore, economic growth and the volume of domestic credit (% of GDP) provided by the banking sector a proxy for banking sector were found to be negatively related with unemployment. The banking-sector-development variable had the largest marginal effect on unemployment. In view of this outcome we recommended financial deregulation and liberalisation presuming these policies result in increased credit/GDP ratio. Lastly, we found the positive marginal impact of economic growth to be negligible relative to all variables captured by the model. To the best of our knowledge this is the first study to explore the banking sector efficiency-employment nexus in South Africa.

CHAPTER EIGHT

SUMMARY, CONCLUSION AND POLICY RECOMMENDATIONS

8.1 INTRODUCTION

This final chapter provides a concluding overview of the whole study on bank efficiency and its relationship with access to banking services and unemployment. The primary objective of this study has been to investigate the nature of the relationship between bank efficiency gains and access to banking services in South Africa. Secondly, it was imperative to determine if there was a significant change in the total factor productivity efficiency of South African banking system during the period of the global financial crisis. Thirdly, the study also sought to establish the link between banking sector efficiency and unemployment in South Africa.

We applied a two-stage methodology framework to a sample of eight banks for the period 2004 – 2011. In line with most previous studies, we followed the more preferable intermediation approach of defining bank inputs and outputs. Using a more flexible and realistic BCC model with variable returns to scale specification, we generated the total factor productivity efficiency scores within the first stage by applying the Hicks-Moorsteen total factor productivity (HMTFP) index approach. We used the HMTFP approach in an attempt to redress the inadequacy of the popular Malmquist TFP index which recent developments by O'Donnell (2010) has shown to be biased and inconsistent when applied to varying returns to scale technologies. We then performed the Student's *t* test and the Wilcoxon matched-pairs signed ranks test on the generated scores to determine if there was a significant change in the efficiency of the banking sector as a result of the global financial crisis. Finally, within the second stage, we then used appropriate panel data techniques to investigate the link between banking sector efficiency and access to banking services as well as unemployment.

8.2 KEY FINDINGS

First-Stage Analysis

This study investigated efficiency of the banking sector in South Africa for the period 2004 – 2011. The data collected from Bankscope database was analysed using *DPIN 3.0*, a program that uses the DEA methodological framework. The BCC model was used in the estimation

and hence the variable returns to scale assumption was adopted. We followed the intermediation approach as opposed to the production approach when we defined the bank inputs and outputs. The intermediation approach recognises the intermediary function of banks as accepting public deposits to produce various outputs. Interest income, non-interest income and loans were considered outputs while the number of employees, customer deposits, fixed assets, and total operating expenses were considered as inputs. The bank outputs were carefully selected to reflect both the traditional (interest income) and non-traditional activities (non-interest income) of modern banking. The major findings in this study are as follows:

The evidence in this study revealed that for the period 2004 – 2011 banks recorded a TFP and TFPE score of 1.35 and 0.59 respectively. A further comparison of performance within the two categories revealed that large banks were better performing than small banks in terms of TFPE while small banks performed relatively batter in terms of TFP. Large banks exhibited a higher TFP efficiency score of 68 (0.68) percent compared to small banks with an average score of 51 (0.51) percent. All the large banks exhibited TFPE scores above their group mean score of 0.68 while two small banks (50 percent) scored below their 0.51 group mean score. The average TFP score was 1.31 and 1.40 for large and small banks respectively. However, both categories had means above unity suggesting that there was an overall improvement during the period. 50 percent of banks in each size category had TFP scores of less than their group average of 1.31 and 1.40 for large and small banks respectively. An examination of standard deviation figures suggested that variability of both performance indicators was wider for small banks compared to large banks. This highlighted more room for improvement particularly among small banks. Again, this wide variability and the overall poor performance of small banks could be indicative of the presence of economies of scale within the retail market or that small banks are operating at the falling portion of their average cost curves.

Apart from estimating and decomposing TFP indices we investigated if there was a statistically significant change in the efficiency and productivity of South African banking system as a consequence of the global financial crisis. The evidence in this study showed that most of the efficiency and productivity measures decreased during the period 2008-2009, the period that corresponds to the financial crisis. The highest TFP index for the entire period of 1.55 was recorded in 2008 before a drastic decline in 2009 to 1.33. Similarly in 2008, the banking sector TFPE score was 60.5 percent before deteriorating to 51.9 percent in both the

years 2009 and 2010. However, based on the parametric student *t*-test and the non-parametric Wilcoxon signed ranks test, we found the difference in the mean productivity efficiency measures between the pre-crisis period and crisis-period to be statistically insignificant. It is important to note this seemingly contradictory finding. While the first analysis merely confirmed that indeed efficiency measures decreased during the period of the crisis. The second investigation was to quantify the magnitude of the change and establish if this change represented a significant (statistically noteworthy) departure from average values. The fact that both tests unanimously concluded that there was no significant difference in performance during the pre-crisis and the crisis period is sufficient to buttress our findings. Nonetheless, our findings are consistent with the views of the Bank Supervision Annual Report (2009, p.4) that broadly described the financial sector as "remaining vigilant" despite the difficult circumstances that came with the crisis. This vigilance was attributed inter alia to effective bank supervision and regulation of the banking sector (SARB, 2009). Mabwe & Webb (2010) also states that for an economy that is integrated into the global financial system, South Africa weathered the global financial crisis well compared to other countries in the region.

Second-Stage Analysis:

One of the primary motivations driving this study was exploring the relationship between gains in bank efficiency and access to bank services in South Africa. The importance of making such an enquiry arises from the fact that various studies (World Bank, 2006; 2009; 2010) and many others have identified financial inclusion or access to financial services as an important vehicle for lifting the poor out of poverty. Moreover, there has been a concern that banks appetite for good scores on efficiency has the potential of reducing access to services for consumers particularly the low-income clients. This particular study focused on access to bank services as opposed to access to financial services. It was argued in this study that access to bank services is a first necessary step to attaining a host of other financial services such as obtaining an overdraft facility, mortgages, savings, investment, and insurance. The study attempted to answer two central research questions: Firstly, does the quest for banks to improve efficiency preclude access to banking services for some group of consumers? Secondly, do bank efficiency gains necessarily translate to improved accessibility to banking services? We used a fixed effects panel model to investigate the empirical relationship between bank efficiency and access to bank services. The analysis was based on a translog function relating access to bank services to bank efficiency and other determinant factors. Bank TFP efficiency was found to be statistically and economically significant in influencing

access to bank services in South Africa. We found that a 10 percent increase in bank efficiency improves access to bank services by 0.013, an equivalent of 12 deposit accounts per 1000 adults (using 2009 figures for comparative purposes). We also justified the size of the marginal impact (which seemed to be trivial) to be economically significant based on the argument that in 2009 the World as a whole added 65 deposit accounts per 1000 adults. Using 2012 IMF (2013) data reported for South Africa, a banking sector efficiency gain of 10 percent would contribute 18 deposit accounts per 1000 adults. Therefore, bank efficiency is one of the preconditions for access to bank services to expand in the country.

GDP per capita a proxy for socio-economic conditions and banking sector development were found to exert a positive and significant impact on access. We also found bank branch penetration to have a positive influence on access possibly via an expanded range of choice to clients. However, we highlight the need to locate these branches at actual centres of demand as many bank branches are historically clustered in urban areas. Lastly, the study reported that a highly concentrated bank market and high bank operating costs have a detrimental effect on access to bank services.

In an attempt to shed light to the second objective, we used a pooled GLS model to investigate the empirical relationship between bank efficiency and unemployment in South Africa. A translog function relating unemployment and bank efficiency among other determinant factors was used. The results confirmed a negative and significant relationship between bank efficiency and unemployment. Hence there is indeed evidence that improvement in the banking sector efficiency transmit through to reduced unemployment in the economy. We hypothesise that the likely channel through which gains in bank efficiency feeds through to reduced unemployment is via minimised bank operating cost or reduced interest spread. All things being equal, this is expected to stimulate increased loan demands for household consumption and business investment and generating job opportunities in the process. Moreover, the volume of domestic credit from the banking sector was found to exert the strongest influence on unemployment while economic growth had the smallest impact on unemployment.

In summary, this thesis has made three important contributions. Firstly, this study is the first in South Africa (to the writers knowledge) to use DPIN 3.0 developed by O'Donnell (2011) to estimate and decompose TFP changes in the South African banking sector. Many previous

studies have analysed total factor productivity in the banking industry using the Malmquist DEA-based technique which recent developments have shown is not robust under VRS technology. The second and third notable contributions of this study are entrenched in the importance and novelty of research questions addressed. Thirdly, an explored but important issue that we have discussed is whether changes in banking sector efficiency are reflected in unemployment. To the best knowledge of the investigator there has not been any study that has focused on these issues in South Africa.

8.3 POLICY IMPLICATIONS

In view of the fact that large banks outperformed smaller banks in terms of total factor productivity efficiency, we recommend that smaller banks relook their business models in light of the changing economic and financial landscape. We postulated that this disparity in efficiency could be due to the fact that large banks and small banks operate different business models and hence emphasise different focus areas. Hence we suggest that an optimal blend of retail and wholesale banking may help diversify small business operations and so allow them to deliver better efficiency.

Our empirical findings have also cast light on the link between banking sector efficiency and access to bank services. This study suggests that banking sector efficiency plays a crucial role in promoting access to bank services in South Africa. We therefore underscore the need for all banks to maintain high efficiency in order to augment government efforts towards improving accessibility for the unbanked South African people. We also found that the rural population variable exerted the greatest but negative marginal impact on access. In light of this outcome, we propose that the government initiate and support investment in rural infrastructure as one of the ways to broaden access on a larger scale. Nonetheless, we argued that attempts by banks to remain sustainable (profitable) would not necessarily affect outreach. In recent years the success stories of Capitec and African bank the largest providers of unsecured lending have challenged the existence of a trade-off between sustainability and outreach. Hence banks have the potential to extend their services to low-income and to the poor and still remain sustainable.

Based on our results of a negative impact of market concentration we posit that this could be indicative of a relatively low competitive banking environment. It would therefore be useful if the South African Reserve Bank would relax the barriers to entry particularly the licensing

of banks in order to encourage broader participation thereby promoting competition among bank service providers. We also support Simpasa`s (2013) suggestion of stimulating competitive behaviour by encouraging more foreign bank participation in the industry through easing regulatory barriers.

One of the objectives of this study was to investigate the link between banking sector efficiency and unemployment in South Africa. Evidence in this study confirmed a negative and significant relationship between banking sector efficiency and unemployment implying that banking sector efficiency transmits through to reduced unemployment in the economy. However, there is still need for future research to explore the possible transmission channels by which changes in banking sector efficiency transmit through to decrease unemployment in the economy. From a policy standpoint, this result suggests that initiatives designed to reduce unemployment in the country which are founded on improving bank efficiency are of paramount value. In this respect the government should encourage and perhaps put incentive structures designed to promote efficiency within the banking system.

We also found the volume of domestic credit from the banking sector to have a negative influence on unemployment. In our study, the bank credit variable had the largest marginal contribution towards reducing unemployment. Hence, measures targeted at reducing high unemployment via enhanced access to affordable credit would be of great significance. We therefore recommend the South African Reserve Bank (SARB) to consider the following: Firstly, financial deregulation and liberalisation, presuming that they lead to higher credit to GDP ratio, could have a significantly large impact on efforts designed to reduce unemployment in South Africa. Secondly, since real GDP growth and low inflation are important determinants of credit expansion as argued by Guo and Stepanyan (2011), sound policies that foster macroeconomic fundamentals would be conducive to credit growth expansion. However, the SARB would need to fine tune its monetary policy rate to balance the need to provide affordable domestic credit and the need to contain an inflation rate that is consistent with growth. Thirdly, the Reserve Bank needs to develop policies that promote intermediation efficiency by dealing with unjustified wide interest spreads among banks. The SARB may also encourage the entrance of more new players in the banking sector, in particular, those banks that provide unsecured credit to disadvantaged households and small-to-medium enterprises (SMEs). Finally, promoting a health and stable banking sector is an important factor which the SARB needs to embolden.

8.4 LIMITATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

This study has focused on only one dimension of efficiency, that is, TFP efficiency. Further studies may need to focus on profit efficiency. Hence, there is a need to collect data on the average prices per unit of each input and output of sampled banks. This would then facilitate measurement of specific types of efficiencies such as profit efficiency. In addition, while we have utilised the recent methodology to estimate TFP efficiency, however, we have not attempted to investigate the various determinant factors of TFP efficiency. It would be interesting for future researches to explore both bank-specific and macroeconomic factors that influence bank efficiency. A useful extension of this study might be to determine the quality variations and their effect on the efficiency scores. The quality dimension of inputs and outputs must be held constant because provision of greater quality services requires additional inputs per unit of output. However, the banks included in our study were fairly homogeneous in size. This is important because banks offering higher quality of services may require more staff complement and other resources inputs than those offering low quality of services. Consequently higher quality banks may have lower efficiency scores not because they are less efficient but because they provide better quality services to their clients. A crucial extension of this study might be to increase the time period covered and the number of banks in order to obtain more accurate results. It would be necessary to duplicate this study for a longer time period and a larger sample of banks particularly small banks. This follows from the evidence found in this study that small banks were associated with providing a relatively high level of access in general. On that note, fruitful research may then attempt to investigate why this is the case: why are large banks apparently not as proactive as small banks?

In conclusion, this chapter has summarised the empirical literature, findings, recommendations and conclusions to the study. The researcher found out that access to banking services is influenced inter alia by the efficiency with which banks operate. The study revealed that there is a negative relationship between banking sector efficiency and unemployment. This means that, when banks improve their efficiency, unemployment tend to decrease. The study uncovered that the efficiency of banks under study was not significantly affected by the global financial crisis. This study therefore submits that, the nature of the relationship of access to banking services is intricately intertwined with the efficiency with which banks operate. Simply put, if banks enhance their efficiency this leads to improved access.

9.1 REFERENCES

ABDMOULAH, W., and LAABAS, B., 2012. Kuwaiti Banks Efficiency: An Examination of Technical and Allocative Efficiency over the period 1994 – 2009. *International Review of Business Research Papers*, 8(5), pp. 123 – 136.

AGENCE FRANCAISE & DEVELOPPEMENT (AFD), 2011. *Contributing to banking services for disadvantaged populations*. Available [Online]: http://www.afd.fr/lang/en/home/pays/afrique/geo-afr/afrique-du-sud/projets-afrique-du-sud/proparco-south-africa/contribuer-bancarisation-populations-defavorisees. [Accessed 1 March 2014].

AHMAD, E., and MALIK, A., 2009. Financial Sector Development and Economic Growth: An Empirical Analysis of Developing Countries. *Journal of Economic Cooperation and Development*, 30(1), pp. 17 – 40.

AIGNER, D., LOVELL C.A.K., and SCHIMDT P., 1977. Formulation and Estimation of Stochastic Frontier Production Models, *Journal of Econometrics*, 6, pp. 21– 37.

AKHIGBE, A., and MCNULTY, J. E., 2005. Profit efficiency sources and differences among small and large U.S. commercial banks, *Journal of Economics and Finance*, 29(3), pp. 289–299.

ALI, A.I., and SEIFORD, L.M., 1993. The Mathematical Programming Approach to Efficiency Analysis. In FRIED H.O., LOVELL, C.A.K., and SCHMIDT, S.S., (Eds), *The Measurement of Productive Efficiency*, Oxford University Press, New York, pp. 120 – 159.

ALZUBAIDI, H., and BOUGHEAS, S., 2012. The Impact of Global Financial Crisis on European Banking Efficiency, Centre for Finance and Credit Markets (CFCM), *Working Paper 12/05*, Nottingham.

ANAYIOTOS, G.C., and TOROYAN, H., 2009. Institutional Factors and Financial Sector Development: Evidence from Sub-Saharan Africa, *IMF Working Paper*, Monetary and Capital Markets Department, WP/09/258.

ARDINGTON, C., and LEIBBRANDT, M., 2004. *Financial Services and the Informal Economy*, African Development and Poverty Reduction: The Macro-Micro Linkage. Forum Paper 2004.

ARJOMANDI, A., and VALADKHANI, A., 2010. Banks Efficiency and Productivity Analysis Using the Hicks-Moorsteen Approach: A Case Study of Iran, *Economics Working Papers*, wp10/11, School of Economics, University of Wollongong, NSW, Australia.

AVKIRAN, N., 1999. An Application Reference for Data Envelopment Analysis in Branch Banking: helping the novice researcher, *International Journal of Bank Marketing,* 17(5), pp. 206 – 220.

BALMER, N., 2006. Worried Sick: The Experience of Debt Problems and their Relationship with Health, Illness and Disability. *Social Policy and Society*. 1(1) pp.39 – 51.

BALTAGI, B., 2008. *Panel Data Econometrics,* 4th Edition. Chichester: John Wiley and Sons.

BANKER, R.D., CHANG, H., and NATARAJAN, R., 2007. Estimating Technical and Allocative Efficiency using DEA: An Application to the U.S. Public Accounting Industry. *Journal of Productivity Analysis*, 27(2), pp. 115 –121.

BANKER R.D., and MAINDIRATTA, A., 1986. Piecewise Log-linear Estimation of Efficient Production Surfaces. *Management Sciences*. 32, pp.126 – 135.

BANKER R.D., and NATARAJAN, R., 2011. Assessing Bank and Bank Branch Performance. In: Cooper, W.W., SEIFORD, L.M., and ZHU, J., eds. 2011. *Handbook on Data Envelopment Analysis*, International Series in Operations Research & Management Science 164, New York: Springer Science + Business Media. Ch 11.

BANKER R.D., and THRALL, R.M., 1992. Estimation of Returns to Scale Using Data Envelopment Analysis. *European Journal Operations Research*, 62, pp. 74 – 84.

BANKER, R.D., CHARNES, A., and COOPER, W.W., 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management* Science, 30 (9), pp.1078 – 1092.

BANKER, R.D., COOPER, W.W., SEIFORD, L.M., and ZHU, J., 2004. Returns to Scale in DEA. In: COOPER, W.W., SEIFORD, L.M., and ZHU, J. Eds. 2011. *Handbook on Data Envelopment Analysis.* 2nd Edition. New York: Springer. Ch.2.

BANKER, R.D., COOPER, W.W., SEIFORD, L.M., THRALL, R.M., and ZHU, J., 2004 Returns to Scale In Different DEA Models, *European Journal of Operational Research*, 154, pp. 345 – 362.

BANKING ASSOCIATION SOUTH AFRICA (BASA), 2012. *2012 South African Banking Sector Overview*. Available [Online]: http://www.banking.org.za/index.php/our-industry/2012-south-african-banking-sector-overview/ [Accessed 2013 February 24].

BANKING ASSOCIATION SOUTH AFRICA (BASA), 2009. *Annual Review 2009*. Available [Online]: http://www.banking.org.za/ [Accessed 1 September 2010].

BANKING ASSOCIATION SOUTH AFRICA (BASA), 2010. *Annual Review 2010*. Available [Online]: http://www.banking.org.za/ [Accessed 5 November 2012].

BANKING ASSOCIATION SOUTH AFRICA (BASA), 2011. *Mzansi Account Summary*, Available [Online]: www.banking.org.za [Accessed 2012 November 21].

BANKING ASSOCIATION SOUTH AFRICA (BASA), 2010. *Mzansi Account Summary*, Available [Online]: www.banking.org.za [Accessed 2011 March 24].

BANKING COUNCIL OF SOUTH AFRICA, 2003. *Financial Sector Charter*. Available [Online]: www.banking.org.za/Downloads/Charter_Final.pdf. [Accessed 7 February 2012].

BAUER, P.W., BERGER, A.N., and HUMPHREY, D.B., 1991. Inefficiency and Productivity Growth in Banking: A Comparison of Stochastic Econometric and Thick Frontier Methods, *Federal Reserve Bank of Cleveland, Working Paper No. 9117*. Available [Online]: http://www.clevelandfed.org/research/workpaper/index.cfm. [Accessed 3 October 2012].

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 Methods. *Journal of Economics and Business,* 50 (2), pp. 85 – 114.

BECK, T., and DE LA TORRE., A., 2006. *The Basic Analytics of Access to Financial Services*, World Bank. Washington D.C.

BECK, T., DEMIRGUC-KUNT, A., and LEVINE, R., 2004. *Finance, Inequality and Poverty*. Policy Research Working Paper. 3338. World Bank, Washington, D.C.

BECK, T., DEMIRGUC-KUNT, A., and LEVINE, R., 2006, Bank Concentration, Competition, and Crises: First Results, *Journal of Banking and Finance*, 30, pp. 1581– 603.

BECK, T., DEMIRGUC-KUNT, A., and MAKSIMOVIC, V., 2003. *Bank Competition and Access to Finance: International Evidence*: Available [Online]: http://siteresources.worldbank.org/DEC/Resources/84797111443727430 4/concentration_JMCB-rev072803.pdf [Accessed 2013 February 24].

BECK, T., DEMIRGUC-KUNT, A., and MAKSIMOVIC, V., 2008. Financing Patterns around the World: Are Small Firms Different? *Journal of Financial Economics* 89, pp. 467–487.

BECK, T., DEMIRGUC-KUNT, A., and MARTINEZ P, M., 2008. Bank Financing for SMEs around the World: Drivers, Obstacles, Business Models, and Lending Practices. *World Bank Policy Research Working Paper 4785.*

BECK, T., LEVINE, R., and LOAYZA, N., 2000. Finance and the Sources of Growth, *Journal of Financial Economics*, 58 (1–2), pp. 261 – 300.

BEKIARIS, E., and NAKANISHI, Y.J., 2004. Economic Impacts of Intelligent Transportation Systems, Innovations and Case Studies. *Research in Transportation Economics*, 8, pp. 1 – 640.

BERGER, A.N., 1993. "Distribution-Free" Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distributional Assumptions, *Journal of Productivity Analysis* 4 (3), pp. 261 – 292.

BERGER, A.N., 1995. The Profit-Structure Relationship in Banking-Tests of Market-Power and Efficient-Structure Hypotheses, *Journal of Money, Credit and Banking*, 27, pp. 404 – 431.

BERGER, A.N., and DeYOUNG, R., 1997. Problem Loans and Cost Efficiency in Commercial Banks, *Journal of Banking and Finance*, 21(6), pp. 849 – 870.

BERGER, A.N., and HUMPHREY, D.B., 1991. The Dominance of Inefficiencies over Scale and Product Mix Economies in Banking, *Journal of Monetary Economics* 28 (1), pp. 117 – 148.

BERGER, A.N., and HUMPHREY, D.B., 1997. The efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research* 98(2), pp. 175-212.

BERGER, A.N., HUNTER, W. C., and TIMME, S. G., 1993. The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future, *Journal of Banking and Finance*, 17, pp. 221 – 249.

BERNANKE, B.S., 2007. Bank Structure and Competition, *Federal Reserve Bank of Chicago's 43rd Annual Conference*, May 17, 2007, Chicago, Illinois.

BHANDARI, A.K., 2009. *Access to Banking Services and Poverty Reduction: A State-wise Assessment in India*. Indian Institute of Social Welfare and Business Management and IZA. Discussion Paper No. 4132.

BJUREK, H., 1996. The Malmquist Total Factor Productivity Index. *Scandinavian Journal of Economics,* 98(2), pp. 303 – 313.

BONIN, J.P., HASAN, I., and WACHTEL, P., 2005. Privatization Matters: Bank efficiency in Transition Countries. *Journal of Banking and Finance*, 29(8-9), pp. 2155-2178.

BOWLIN, W.F., 1986. Evaluating Performance in Governmental Organisations. *The Government Accountants Journal, 35*(2), pp. 50 – 57.

BRIEC, W., and KERSTENS, K., 2011. The Hicks–Moorsteen Productivity Index Satisfies The Determinateness Axiom, *School of Management*, The Manchester. 79(4), pp. 765 -775.

BROOKS, C., 2008. *Introductory Econometrics for Finance*, 2nd Edition. Cambridge, New York: Cambridge University Press.

CASU, B., and GIRARDONE. C., 2006. Bank Competition, Concentration and Efficiency in the Single European Market' *The Manchester School*, 74 (4), pp. 441- 468.

CASU, B., GIRARDONE, C., and MOLYNEUX, P., 2006. *Introduction to banking*. Harlow: FT Prentice Hall.

CHARNES, A., COOPER, W. W., and RHODES, E., 1978. Measuring the Inefficiency of Decision Making Units. *European Journal of Operational Research*, 2, pp. 429 – 444.

CHARNES, A., COOPER, W.W., LEWIN, A.Y., MOREY, R.C., and ROUSSEAU, J.J., 1985. Sensitivity and Stability Analysis in DEA, *Annals of Operations Research*, 2, pp. 139 – 150.

CHARNES, A., COOPER, W.W., SEIFORD, L.M., and STUTZ, J., 1982. A Multiplicative Model for Efficiency Analysis. *Socio-economic Planning Sciences*, 16, pp. 213 – 224.

CHARNES, A., COOPER, W.W., SEIFORD, L.M., and STUTZ, J., 1983. Invariant Multiplicative Efficiency and Piecewise Cobb-Douglas Envelopments. *Operations Research Letters*, 2, pp. 101–103.

CHERCHYE, L., KUOSMANEN, T., and POST, T., 2001. FDH Directional Distance Functions, with an Application to European Commercial Banks. *Journal of Productivity Analysis,* 15(3), pp. 201 – 215.

CLAASSEN, L., and BROOKS, E., 2003. In: KPMG *Banking Survey 2002*. KPMG: Johannesburg.

CLAESSENS, S., and LAEVEN, L., 2003. *What Drives Bank Competition*? Some International Evidence. World Bank. Washington D.C.

CLAESSENS, S., 2006. *Access to Financial Services: A Review of the Issues and Public Policy Objectives*, Oxford University Press.

COELHO, C.C.S., 2007. *Modes of Governance and Public Service Efficiency*. Ph.D. Birmingham: University of Birmingham.

COELLI, T.J., 1996. A Guide to DEAP version 2.1: A Data Envelopment Analysis (Computer) Program, *CEPA Working Paper* 96/07. Armidale: Department of Econometrics, University of New England.

COELLI, T.J., and RAO, D.S.P., 2005. Total Factor Productivity Growth in Agriculture: A Malmquist Index Analysis of 93 Countries, 1980-2000, *Agricultural Economics*, 32, pp.115 –134.

COLLARD, S., and KEMPSON, E., 2005. *Affordable Credit: The Way Forward*. York: Joseph Rowntree Foundation.

COMPETITION COMMISSION SOUTH AFRICA, 2008. *Banking Enquiry: Report to the Competition Commissioner by the Enquiry Panel.* Available [Online]: http://www.compcom.co.za/assets/Banking/Executive-overview.pdf [Accessed 2011 March 7].

CONSULTATIVE GROUP TO ASSIST THE POOR (CGAP), 2009. *Financial Access: Measuring Access to Financial Services around the World*, Washington DC. USA

CONSULTATIVE GROUP TO ASSIST THE POOR (CGAP), 2010. *Update on Regulation of Branchless Banking in South Africa*, Washington D.C, USA.

CONSULTATIVE GROUP TO ASSIST THE POOR (CGAP), 2003. *Saving as Important as Credit: Deposit Services for the Poor*. CGAP Donor Brief No. 4, Washington D.C. USA.

CONSULTATIVE GROUP AGAINST POVERTY (CGAP), 2011. *Technology Program Country Note, South Africa*. (Accessed 2013 May 7) Available at http://www.cgap.org/publications/technology-program-country-note-south-africa[Accessed 7 April 2012].

CONSULTATIVE GROUP TO ASSIST THE POOR (CGAP), 2003. *Saving as Important as Credit: Deposit Services for the Poor*. CGAP Donor Brief No. 4, Washington D.C. USA.

COOPER, W.W., RUIZ, J.L., and SIRVENT, I., 2009. Selecting Non-Zero Weights to Evaluate Effectiveness of Basketball Players with DEA. *European Journal of Operational Research*, 195(2), pp. 563 – 574.

COOPER, W.W., SEIFORD, L.M., and TONE, K., 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd Edition, New York, Springer.

COOPER, W.W., SEIFORD, L.M., and ZHU, J., (eds.), 2011. *Handbook on Data Envelopment Analysis*, International Series in Operations Research & Management Science 164, New York: Springer Science + Business Media.

COSSET, K. FISCHER, BOUBAKRI, N., J-C., and GUEDHAMI, O., 2005, Privatization and Bank Performance in Developing Countries, *Journal of Banking and Finance,* 29, pp. 2015–41.

COSSET, J.C., FISCHER, K., BOUBAKRI, N., and GUEDHAMI, O., 2005, Privatization and Bank Performance in Developing Countries, *Journal of Banking and Finance,* 29, pp. 2015 – 41.

DAS, A., and KUMBHAKAR, S.C., 2012. Productivity and Efficiency Dynamics in Indian Banking. An Input Distance Function Approach Incorporating Quality of Inputs and Outputs, *Journal of Applied Econometrics*, 27(2), pp. 205 – 234.

DE SOUSA, M., and SCHWENGBER, S., 2005. *Efficiency Estimates for Judicial Services in Brazil: Nonparametric FDH (Free Disposal Hull) and the Expected Order-M Efficiency Scores for Rio Grande Do Sul Courts*. Available [Online]: http://www.anpec.org.br/encontro2005/artigos/A05A053.pdf [Accessed 2011 April 7].

DE WET., A.H., and VAN EYDEN., R., 2005. Capital Mobility in Sub-Saharan Africa: A Panel Data Approach, *South African Journal of Economics*, 73 (1), pp. 22 – 35.

DELOITTE., 2004. *Banking the Unbanked in South Africa*. Global Regulation, Issue II.

DEMAESTRI, E., WALKER, W., and MARTÍN, F., 2006. *Deposit Insurance and Poverty Reduction*, Inter-American Development Bank. Washington, DC.

DEMIRGUÇ-KUNT, A., and DETRAGIACHE, E., 1998. *Financial Liberalisation and Financial Fragility.* IMF Working paper no. 98/83, Washington, DC.

DEMIRGUC-KUNT, A., and KLAPPER, L., 2011. *Measuring Financial Inclusion.* The World Bank Development Research Group, Bill and Melinda Gates foundation.

DEMIRGUC-KUNT, A., and KLAPPER, l., 2012. *Measuring Financial Inclusion: The Global Findex Database*, Policy Research Working Paper 6025, World Bank, Washington DC.

DEPRINS, D., SIMAR, L., and TULKENS, H., 1984. Measuring Labour Efficiency in Post Offices. In MARCHAND, M., PETIEAU, P., and TULKENS, H. (eds), *The Performance of Public Enterprises: Concepts and Measurement*. Amsterdam: North Holland.

DEVLIN, J.F., 2006. *A Detailed Study of Financial Exclusion in the UK*, Nottingham Financial Services Research Forum: University Business School.

DIAS. D., 2011. *Advancing Financial Access for the World`s Poor*, Consultative Group to Assist the Poor (CGAP). Available [Online]: http://www.cgap.org/48211411.pdf [Accessed 11 May 2013].

DIEWERT, W.E., 1992. Fisher Ideal Output, Input, and Productivity Indexes Revisited. *Journal of Productivity Analysis,* 3, pp. 211– 248.

DIEWERT, W.E., and PARKAN, C., 1983, Linear Programming Tests of Regularity Conditions for Production Frontiers, In: EICHORN, W., HENN, R., NEUMANN, K., and SHEPHARD, R. W., eds. 1983. *Quantitative Studies in Production and Prices*, Physica-Verlag, Würzburg, Germany.

DLOUHY, M., 2009. Efficiency and Productivity Analysis in Health Service, *Prague Economic Papers*, 2, pp. 176 – 184.

DORSEY, D., and JACOBS, K., 2005. *Financial Services Trends and Recent Innovations in South Africa*: Lessons for the United States, The Centre for Financial Services Innovation.

DUYGUN, M., SENA, V., and SHABAN, M., 2013. Schumpeterian Competition and Efficiency among Commercial Banks, *Journal of Banking & Finance,* 37 (2013), pp. 5176 – 5185.

EDWARDS, F.R., and MISHKIN, F.S., 1995. The Decline of Traditional Banking: Implications for Financial Stability and Regulatory Policy. *NBER Working Paper Series*, Working Paper No. 4993, New York.

EGWUATU, B. S. C., 2008. Reducing Poverty through Better Credit Delivery: The Asian Experience. *Central Bank of Nigeria Bullion,* 32(1), pp. 8 – 16.

ELYASIANI, E., and MEHDIAN, S., 1990, Efficiency in the commercial banking industry: a production frontier approach, *Applied Economics*, (22) 4, pp.539-551.

EUROPEAN COMMISSION, 2008. *Financial services provision and prevention of financial exclusion*, Brussels: Directorate-General for Employment, Social Affairs and Equal Opportunities.

EVANOFF, D.D.,1988. Branch Banking and Service Accessibility, *Journal of Money, Credit and Banking*, 20(2), pp. 192 – 202.

FALKENA, H., DAVEL, G., HAWKINS, P., LLEWELLYN, D., LUUS, C., MASILELA, E., PARR, G., PIENAAR, J., and SHAW, H., 2004. *Competition in South African Banking*. Task Group Report for the National Treasury and the South African Reserve Bank.

FARRELL, M. J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A (General)* 120(3), pp. 253 – 290.

FARRELL, M.J., and FIELDHOUSE, M., 1962. Estimating Efficient Production under Increasing Returns to Scale. *Journal of the Royal Statistical Society,* Series A 125, pp.252 –267.

FERNANDO, N.A., 2007. *Low-Income Household's Access to Financial Services: International Experience, Measures for Improvement and the Future*. Asian Development Bank.

FIGUERIA, C., NELLIS, J., and PARKER, D., 2006. Does Ownership Affect the Efficiency of African Banks? *The Journal of Developing Areas*, 40 (1), pp. 38 – 63. Available [Online]: http://www.jstor.org/stable/4193015. [Accessed 13 April 2012].

FINMARK TRUST, 2003a. *Finscope: South Africa's Financial Sophistication Measure.* Johannesburg: FinMark Trust.

FINMARK TRUST, 2003b. *Vision 2010: Annual report*. Johannesburg: FinMark Trust.

FINMARK TRUST, 2005. *Themes Arising from the Southern African Workshop on Tiered Banking Regulation.* Available [Online]: http://www.finmarktrust.org.za/documents/2005/JANUARY/Themes_Tiered.pdf [Accessed 2012 February 19].

FINSCOPE, 2011. FinScope South Africa 2011. Available [Online]: www.finmarktrust.org.za. [Accessed 7 October 2012].

FINSCOPE, 2003. *Summary Report: Findings of the FinScope Study into financial access and behaviour of the South Africa population*, Johannesburg. FinMark Trust.

FINSCOPE, 2010. *FinScope South Africa 2010: Making Financial Markets Work for the Poor*. Available [Online] http://www.finmark.org.za/wp-content/uploads/pubs/FinScope-2010_Brochure.pdf [Accessed 2012 February 19].

FRIMPONG, J.M., 2010. Investigating Efficiency of Ghana Banks: A Non-Parametric Approach*, American Journal of Scientific Research*, 7, pp. 64 – 76.

GANLEY, J. and CUBBIN, J., 1992. *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, New York: Elsevier Science Publishers.

GARCIA, G., 1996. *Deposit Insurance: Obtaining the Benefits and Avoiding the Pitfalls*. IMF Working Paper no. 96/83, International Monetary Fund, Washington, D.C.

GITAU, C.M.W., and GOR, S.O., 2011. Measuring Factor Productivity of the Banking Sector in Kenya, *International Journal of Sustainable Development*, 2 (12), pp. 11 – 18.

GLOBAL FINANCIAL INCLUSION (GLOBAL FINDEX) DATABASE, 2012. World Bank, Washington, DC. Available [Online]: http://www.worldbank.org/globalfindex [Accessed 27 May 2014].

GORDHAN, P., 2011. *Keynote Address by Minister of Finance*, The Banking Association's 2011 Banking Summit, National Treasury: South Africa.

GORDHAN, P., 2011. *Keynote Address at the 10th anniversary of the South African Savings Institute and launch of Savings Month 2011*, July 13 2011, Available [Online]: http://govza.gcis.gov.za/node/469326 [Accessed 26 May 2014].

GORDHAN, P., 2013. *2013 Budget Speech*, Department of National Treasury, Pretoria: South Africa. 27 February 2013.

GREENBERG, J.B., and SIMBANEGAVI, W., 2009. *Testing for Competition in the South African Banking Sector*, Faculty of Commerce University of Cape Town.

GRIFELL-TATJÉ, E., and LOVELL, C. A. K., 1995, A Note on the Malmquist Productivity Index, *Economics Letters,* 47, pp.169 –175.

GRIFFITHS, E.W., HILL, R.C., and LIM, C.C., 2008. *Using EViews for Principles of Econometrics*. 3rd Edition, John Wiley & Sons: Hoboken NJ.

GUJARATI, D., 2004. *Basic Econometrics*. 4th Edition. New York: McGraw-Hill.

GUJARATI, D.N., and PORTER, D.C., 2009. *Basic Econometrics*. 5th Edition. New York: McGraw-Hill.

GUJARATI, D.N., and PORTER, D.C., 2010. *Essentials of Econometrics*. 4th Edition International Edition. New York: McGraw-Hill.

GUO, K., and STEPHANYAN, V., 2011. Determinants of Bank Credit in Emerging Market Economies, *IMF Working Paper*, European Department, WP/11/51.

HAO, Q.M., 2005. *Access to Finance and Poverty Reduction an Application to Rural Vietnam*. PhD. Birmingham Business School, University of Birmingham.

HARVEY, L., 1978. *General X-Efficiency Theory and Economic Development*, New York: Harvard University Press.

HAWKINS, P., 2004. *South Africa's Financial Sector Ten Years On: Performance since Democracy*, Development Southern Africa, 21 (1).

HAWKINS, P., 2004. *The Appropriate Framework for South African Financial Regulation‖ Prepared for the Financial Regulatory Issues Standing Committee (FRISC)*, Task Team of National Treasury and the South African Reserve Bank.

HAWKINS, P.A., 2010. *Measuring Consumer Access to Financial Services in South Africa*. IPMs sessions.

HENDERSON, D.J., 2003. The Measurement of Technical Efficiency Using Panel Data, *Working Paper No. 308*, State University of New York, Binghamton. Available [Online]: http://www2.binghamton.edu/economics/research/working-papers/pdfs/wp03/WP0308.pdf [Accessed 16 January 2013].

HERRERA, S., and PANG, G., 2005. Efficiency of Public Spending in Developing Countries: An Efficiency Frontier Approach, *Policy Research Working Paper No. 3645*, World Bank, Washington DC, USA.

HERSI, W.H., 2009. *The Role of Islamic Finance in Tackling Financial Exclusion in the UK.* PhD. Durham University. Available [Online] http://etheses.dur.ac.uk/23/ [Accessed 29 March 2012]

HESS, K., and FRANCIS, G., 2004. Cost Income Ratio Benchmarking in Banking: A Case Study. *Benchmarking: An International Journal*, 11(3), pp. 303 – 319.

HICKS, J. R., 1961. Measurement of Capital in Relation to the Measurement of Other Economic Aggregates. In: F. A. Lutz & D. C. Hague (eds.). *The Theory of Capital*. London, Macmillan.

HOLLINGSWORTH, B., and WILDMAN, J., 2003. The Efficiency of Health Production: Re-estimating the WHO Panel Data Using Parametric and Non-Parametric Approaches to Provide Additional Information. *Health Economics*, 12(6), pp. 493 – 504.

HONOHAN, P., 2004. *Data on Microfinance and Access: Thinking about What is Available and what is Needed.* Mimeo, World Bank. Washington D.C.

HUSSEIN, K.A., 2000. *Efficiency in Islamic Banking*, Unpublished presentation, Available: http://www.sbp.org.pk/dept/ibd/Banking Efficiency.ppt [Accessed 7 April 2012].

IKHIDE, S.I., 1996. Commercial Bank Offices and the Mobilisation of Private Savings in Selected Sub-Saharan African Countries, *The Journal of Development studies*, 33(1), pp. 117 – 132.

IKHIDE, S.I., 2000. Efficiency of Commercial Banks in Namibia, Bank of Namibia, *BON Occasional Paper No. 4*. Available [Online]: https://www.bon.com.na/CMSTemplates/Bon/Files/bon.com.na/b8/b8b60638-88ce-457b-a0c2-554765376e70.pdf [Accessed 20 October 2013].

IKHIDE, S.I., 2008. Measuring the Operational Efficiency of Commercial Banks in Namibia, *South Africa Journal of Economics*, 76(4), pp. 586 – 595.

IM, K.S., PESARAN, M.H., and SHIN, Y., 2003. Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, 115, pp. 53 – 74.

INTERNATIONAL MONETARY FUND (IMF), 2008. South Africa: *Financial System Stability Assessment, Including Report on the Observance of Standards and Codes on the following topic: Securities Regulation*, Country Report No. 08/349.

INTERNATIONAL MONETARY FUND (IMF), 2012. *Financial Access Survey (FAS).* Washington, D.C. Available [Online]: http://www.fas.imf.org/misc/FAS_Dataset.xlsx [Accessed 14 June 2013].

INTERNATIONAL MONETARY FUND (IMF), 2013. *Financial Access Survey (FAS).* Washington, D.C. Available [Online]: http://www.fas.imf.org/misc/FAS_Dataset.xlsx [Accessed 11 February 2014].

INTERNATIONAL MONETARY FUND (IMF), 2010. *Global Financial Stability* Available [Online]: https://www.imf.org/external/pubs/ft/gfsr/2010/02/ [Accessed 14 February 2014].

JACOBY, H.G., and SKOUFIAS, E., 1997. Risk, Financial Markets, and Human Capital in a Developing Country, *Review of Economic Studies*, 64(3) pp.311 – 335.

JAMES, B.R., CAPRIO, G., and LEVINE, R., 2005. *Rethinking Bank Regulation: Till Angels Govern.* New York: Cambridge University Press.

KABLAN, S., 2010. Banking Efficiency and Financial Development in Sub-Saharan Africa, *IMF Working Paper*, African Department, WP/10/136.

KAGISO ASSET MANAGEMENT, 2013, *Unsecured Lending: Gathering Clouds*. Available [Online]: www.kagisoam.com [Accessed 2012 February 19].

KAMAU, A.W., 2011. Intermediation Efficiency and Productivity of the Research in Business, 1(9), pp. 12 – 26.

KEMPSON, E., 2006. Policy Level Response to Financial Exclusion in Developed Economies: Lessons for Developing Countries. *Paper delivered at World Bank/Brookings Institute Conference*, 30-31 May 2006.

KEMPSON, E., and COLLARD, S., 2006. *Do the Poor Pay More? The Costs and Affordability of Goods and Services to People on Low Income*. Report for the Joseph Rowntree Foundation. Bristol: Personal Finance Research Centre, University of Bristol.

KEMPSON, E., WHYLEY C., CASKEY J., AND COLLARD, S., 2000. In or Out? Financial Exclusion: A Literature and Research Review, Financial Services Authority, London.

KERSTENS, K., and VAN DE WOESTYNE, I., 2014. Comparing Malmquist and Hicks-Moorsteen Productivity Indices: Exploring the Impact of Unbalanced vs. Balanced Panel Data, *European Journal of Operational Research*, 233 (2014), pp. 749 – 758.

KHAN, M.S., and SENHADJI, A., 2000. Financial Development and Economic Growth: An Overview. *IMF Working Paper*, WP/00/209.

KIRKPATRICK, C., V. MURINDE., and M. TEFULA., 2008. The Measurement and Determinants of X-efficiency in Commercial banks in Sub-Saharan Africa, *The European Journal of Finance*, 14(7), pp. 625 – 39.

KIYOTA, H., 2009. Efficiency of Commercial Banks in Sub –Saharan Africa: A Comparative Analysis of Domestic and Foreign Banks, *CSAE Conference on Economic Development in Africa*, University of Oxford, 22nd – 24th March 2009.

KOHLER, M., 2013. Does Non-interest income Make Banks More Risky? Retail- versus Investment-oriented Banks, Deutsche Bundesbank Eurosystem, Discussion Paper No. 17/2013. Available [Online]: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2261120 [Accessed 20 November 2013].

KOUTSOMANOLI-FILIPPAKI, A., 2007. *Competition, Efficiency and Productivity Growth In The European Banking Industry.* PhD. Athens University of Economics and Business.

KPMG, 1998. Bank Margins and Related Profitability in South Africa. Available [Online]: http://www.compcom.co.za/banking/Documents/Company%20Submissions/Banking%20Ass-%20KPMG%20Study%20-%20Bank%20Margins.pdf. [Accessed 22 October 2013].

LANG, G., and WELZEL, P., 1998, Technology and Cost Efficiency in Universal Banking: A 'Thick Frontier' Analysis of German Banking Industry, *Journal of Productivity Analysis*, 10, pp. 63 – 84.

LAWSON, B., and SAMSON, D., 2001. Developing Innovation Capability in Organizations: A Dynamic Capabilities Approach, *International Journal of Innovation Management,* 5(3), pp. 377 – 400.

LAZAROV, M., 2009. *The Global Financial Crisis 2007 – 8. Crisis of Human Knowledge and Government Intervention*, 1 March 2009, Cambridge.

LEVIN, A., and LIN, C., 1992. Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, Department of Economics, *University Of California at San Diego, Discussion Paper no. 92-93* (revised 1993).

LEVINE, R., 1997. Financial Development and Economic Growth: Views and Agenda. *Journals of Economic Literature*, 35(2), pp. 688 – 726.

LEVINE, R., LOAYZA, N., and BECK, T., 2000. Financial Intermediation and Growth Causality and Causes. *Journal of Monetary Economics* 46, pp.31 – 77.

LEWIS, W. A., 1955. *The Theory of Economic Growth*, George Allen and Unwin.

LOZANO-VIVAS, A., and PASIOURAS, F., 2008. The Impact of Non-traditional Activities on Estimation of Bank Efficiency: International Evidence. *University of Bath Working Paper Series 2008:01*. http://www.bath.ac.uk/management/research/pdf/2008-01.pdf [Accessed 20 October 2013].

MABWE, K., and WEBB, R., 2010. A Financial Ratio Analysis of Commercial Bank Performance in South Africa, *African Review of Economics and Finance*, 2(1), pp. 31 – 53.

MADDALA, G., and WU, S., 1999. A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

MAHESHWARI, S., 2009. South African Banking Sector Analysis. *Global Press Release Distribution*.

MAREDZA, A., and IKHIDE, S., 2013. The Impact of the Global Financial Crisis on Efficiency and Productivity of the Banking System in South Africa, Economic Research Southern Africa, *ERSA Working Paper Series No 328*.

MARTIC, M.M., NOVAKOVI, M.S., and BAGGIA, A., 2009. Data Envelopment Analysis – Basic Models and their Utilization, *Research Papers*, 42(2), pp. 37 – 43.

MBOWENI, T.T., 2004. *The South African banking sector - an overview of the past 10 years.* Johannesburg.

McKIBBIN, W.J., and STOECKEL, A., 2009. The Global Financial Crisis: Causes and Consequences. *Lowy Institute for International Policy*, Working Papers in International Economics No. 2.09.

MERCAN, M., ARNOLD R., REHA Y., and AHMET B.E., 2003. The Effect of Scale and Mode of Ownership on the Financial Performance of the Turkish Banking Sector: Results of a DEA-Based Analysis. *Socio-Economic Planning Sciences*, 37, pp.185 – 202.

MESTER, L.J., 1997. Measuring Efficiency of U.S Banks, Accounting for heterogeneity is important, *European Journal of Operational Research*, 98, pp. 230–242.

MISHKIN, F.S., 2010. *The Economics of Money, Banking and Financial Markets***.** 8th ed. Pearson Education International, USA.

MISHKIN, F.S., 2013. *The Economics of Money, Banking and Financial Markets***.** Global Edition, 10th Edition. Pearson Education Limited, Edinburgh Gate, Harlow Essex, England.

MLAMBO, K., and NCUBE, M., 2011. Competition and Efficiency in the Banking Sector in South Africa, *African Development Review*, 23 (1), pp. 4 – 15.

MNYANDE, M., 2012. South Africa's macroeconomic outlook in the wake of the sovereign debt crisis and reforms to financial sector and regulation, *South African Reserve Bank, Keynote address at the 6th Annual Risk and Return South Africa Conference*, Cape Town, 1 –2 March 2012, South Africa.

MOODY`S INVESTORS SERVICES, 2013. *Moody's: South Africa's Unsecured Lenders Face Asset Quality Pressures in Challenging Economic Environment*, Global Credit Research, Available [Online]: http://www.moodys.com/research/South-Africasunsecured-lending-market-faces-rising-asset-quality-risks--PBC_145370, [Accessed 2014 February 17].

MOORSTEEN, R. H., 1961. On Measuring Productive Potential and Relative Efficiency. *The Quarterly Journal of Economics,* 75(3), pp. 151 – 167.

MULLINEUX, A.W., and MURINDE, V., (eds)., 2003. *Handbook of International Banking*, Edward Elgar Publishing Limited, UK.

MUSARA, M., and FATOKI, O., 2010. Has Technological Innovations resulted in increased Efficiency and Cost Savings for Banks` Customers? *African Journal of Business Management*, 4(9), pp. 1813 – 1821.

MWEGA, F.M., 2010. Global Financial Crisis Discussion Series, Paper 17: Kenya Phase 2, *Overseas Development Institute* (ODI), London.

NAPIER, M., 2005a. *Provision of Financial Services of South Africa*: Country Case Study, Services Expert Meeting.

NAPIER, M., 2005b. *Engaging the Private Sector: The Case for Financial Access Charters in Sub-Saharan Africa.* In: UNITED NATIONS CAPITAL DEVELOPMENT FUND (UNCDF), 2005. *Microfinance Matters: Building Inclusive Financial Sectors*, Issue 17, Available [Online]: http://www.uncdf.org/mfmatters [Accessed 14 April 2011].

NATIONAL TREASURY, 2011. *Economic Policy and Outlook*. Available [Online]: http://www.treasury.gov.za/documents/national%20budget/2011/review/chapter%202.pdf[Accessed 10 June 2012].

NATIONAL TREASURY, 2011. *A Safer Financial Sector to Serve South Africa Better*, Policy Document, National Treasury: Pretoria.

NCUBE, M., 2009. Efficiency of the Banking Sector in South Africa, *Paper presented at the Fourth African Economic Conference 2009 on Fostering Development in an Era of Financial and Economic Crises*, 11–13 November 2009, Addis Ababa, Ethiopia.

NIERI, L., 2007. *Access to Credit: The Difficulties of Households*, In: ANDERLONI, L., BRAGA, M.D., and CARLUCCIO, E.M., New Frontiers in Banking Services: Emerging Needs and Tailored Products for Untapped Markets. Berlin: Springer Verlag.

O'DONNELL, C.J., and VAN DER WESTHUIZEN, G., 2002. Regional Comparisons of Banking Performance in South Africa, *South African Journal of Economics*, 70(3): pp. 485–518.

OBERHOLZER, M., VAN DER WESTHUIZEN G., and S. VAN ROOYEN., 2010. The influence of banks' internal performance on market performance: a non-parametric approach, *Southern African Business Review*. 14 (2), pp. 65 – 88.

O'DONNELL, C.J., 2010a. *DPIN Version 1.0. A Program for Decomposing Productivity Index Numbers*. Centre for Efficiency and Productivity Analysis Working Papers WP01/2010, University of Queensland.

O'DONNELL, C.J., 2010b. Measuring and Decomposing Agricultural Productivity and Profitability Change. *Australian Journal of Agricultural and Resource Economics*, 54(4), pp. 527 – 560.

O'DONNELL, C.J., 2010c. Nonparametric Estimates of the Components of Productivity and Profitability Change in U.S. Agriculture. *Centre for Efficiency and Productivity Analysis Working Papers WP02/2010*, University of Queensland.

O'DONNELL, C.J., 2011. *DPIN Version 3.0: A Program for Decomposing Productivity Index Numbers*. Centre for Efficiency and Productivity Analysis Working Papers, University of Queensland.

OKEAHALAM, C. C., and ADAMS, D. W., 2000. Banks and Micro-Finance in Namibia, *The South African Journal of Economics*, 68, pp.118–150.

OKEAHALAM, C., 2008. Client Profiles and Access to Retail Bank Services in South Africa. *Applied Financial Economics*, 18, pp. 1131 – 1146.

OKEAHALAM, C.C., 1998. Concentration in the Banking Sector of the Common Monetary Area of Southern Africa. *Paper was presented at the Development Issues in the New Economy Conference*. University of Cape Town, South Africa. March 24 – 26 2002.

OKEAHALAM, C.C., 2001. *Structure and Conduct in the Commercial Banking Sector of South Africa*. Presented at TIPS 2001 Annual Forum.

OKEAHALAM, C.C., 2006. Production Efficiency in the South African Banking Sector: A Stochastic Analysis, *International Review of Applied Economics*, 20, pp. 103 – 23.

ONDIEGE, P., 2010. Mobile Banking in Africa: Taking the Bank to the People, *Africa Economic Brief,* African Development Bank, 1(8), pp. 1 – 16.

OZCAN, Y.A., 2007. *Health Care Benchmarking and Performance Evaluation: An Assessment using Data Envelopment Analysis (DEA),* International Series in Operations Research & Management Science. New York: Springer.

PANZAR, J., and ROSSE, J., 1987. Testing for Monopoly Equilibrium, *The Journal of Industrial Economics*, 35, pp. 443–56.

PARADI, J.C., YANG Z., and ZHU, H., 2004. Assessing Bank and Bank Branch Performance. In: Cooper, W.W., SEIFORD, L.M., and ZHU, J., eds. 2011. *Handbook on Data Envelopment Analysis*, International Series in Operations Research & Management Science 164, New York: Springer Science + Business Media. Ch 13.

PAULSON, J.A., and McANDREWS, J., 1998. *Financial services for the urban poor: South Africa's E Plan.* World Bank Policy Research Working Paper No 2016. Washington, DC: World Bank.

PEACHEY, S., and ROE, A., 2004. *Access to Finance: A study for the World Savings Banks Institute*, Oxford Policy Management.

PETERSEN, M.A., and RAJAN, R.G., 1995. The Effect of Credit Market Competition on Lending Relationships. *The Quarterly Journal of Economics*. 110 (2), pp. 407 – 443.

PORTEOUS, D., and HAZELHURST, E., 2004. *Banking on Change: Democratizing Finance in South Africa 1994-2004 and Beyond.* Cape Town: Double Storey Books.

PORTER, R.C., 1966. The Promotion of the Banking Habit and Economic Development, *The Journal of Development studies*, 2(4), pp. 346 – 366.

PRINSLOO, J.W., 2002. *Household Debt, Wealth and Saving*, Quarterly Bulletin December 2002, South African Reserve Bank.

RAJAN, R., and ZINGALES, L., 2004. *Saving Capitalism from the Capitalists,* New Jersey: Princeton University Press.

RODRÍGUEZ-MEZA, J., 2006. *Enhancing Access to Credit for Low-income Borrowers*, Inter-American Development Bank. Washington, DC.

ROSENZWEIG, M.R., 2001. Savings Behaviour in Low Income Countries, *Oxford Review of Economic Policy*, 17**:** 40-54.

RUTHERFORD, S., 2000a. *Raising the Curtain on the 'Microfinancial Services Era*. CGAP Focus Note No. 15. Washington, D.C.

RUTHERFORD, S., 2000b. *The Poor and Their Money.* New Delhi: Oxford University Press

SADOULET, L., 2006. *Savings and Deposit Services for the Poor*, Inter-American Development Bank. Washington, DC.

SATHYE, M., 2003, Efficiency of Banks in a Developing Economy: The Case of India, *European Journal of Operational Research*, 148(3), pp. 662 – 671.

SCHMIDT, P., and SICKLES, R.C., 1984. Production Frontiers and Panel Data, *Journal of Business and Economic Statistics*, 2, pp. 367 – 374.

SCHOOMBEE, A., 2004. South African Banks and the Unbanked: Progress and Prospects. *South African Journal of Economics*, 72 (3), pp. 581 – 603.

SHARMA, A.K, SHARMA, D., and BARUA, M.K., 2012. Efficiency and Productivity of Indian Banks: An Application of Data Envelopment Analysis and Tobit Regression, *National Conference on Emerging Challenges for Sustainable Business 2012*, pp. 81 – 90.

SHAW, C., WILLIAMS, K., and ASSASSA, R. P., 2000. Information Point: Wilcoxon Signed Rank Test, Journal of Clinical Nursing, 9, pp. 574 – 584.

SHEPHARD, R.W., 1953. *Cost and Production Functions*. Princeton University Press, Princeton.

SHERMAN, H.D., 1982. *Efficiency Evaluation Techniques: A Critical Comparison and Assessment.* Cambridge Massachusetts, Massachusetts Institute of Technology.

SIMPASA, A.M., 2013. Competition and Market Structure in the Zambian Banking Sector, *African Development Bank Group*, Working Paper no. 168.

SIZWEKAZI, J., 2007. *Banking on the Unbanked*, Finweek, Available [Online]:http://eds.b.ebscohost.com.nwulib.nwu.ac.za/eds/pdfviewer/pdfviewer?sid=20254566-6200-46a2-ae50-f72e76630152%40sessionmgr110&vid=1&hid=109. [Accessed 25 May 2014].

SOUTH AFRICAN RESERVE BANK, 2002. Annual Report 2002, *Bank Supervision Department*, Available [Online]: http://www.reservebank.co.za/ [Accessed 10 April 2014].

SOUTH AFRICAN RESERVE BANK, 2009. *Annual Report 2009*, Bank Supervision Department, Statistical Society, Series A (General), 120, pp. 253 –78. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2010. *Annual Economic Report 2010*, 90th Anniversary. Available [Online]: http://www.reservebank.co.za/PublicationsandNotices/Reports [Accessed 12 July 2012].

SOUTH AFRICAN RESERVE BANK, 2010. *Annual Report 2010*, Bank Supervision Department, 90th Anniversary, pp. 1–173. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2011. *Annual Economic Report 2011*, 90th Anniversary. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2009. Annual Report 2009, *Bank Supervision Department*, Statistical Society, Series A (General), 120, pp. 253–78.

SOUTH AFRICAN RESERVE BANK, 2010. Annual Report 2010, *Bank Supervision Department*, 90th Anniversary, pp. 1–173.

SOUTH AFRICAN RESERVE BANK, 2011. Annual Report 2011, *Bank Supervision Department*, Available [Online]: http://www.reservebank.co.za/ [Accessed 11 April 2012].

SOUTH AFRICAN RESERVE BANK, 2013, Quarterly Bulleting June 2013 No. 268. Available[Online]:http://www.durban.gov.za/City_Services/Economic_Development/PSIR/SARB%20Quarterly%20Bulletin/SARB%20Quarterly%20Bulletin%20%20June%202013.pdf [Accessed 26 May 2014].

STANDARD CHARTERED, 2013. *Banking on Innovation*, Available [Online]: http://www.standardchartered.com/en/about-us/our-brand-and-values/here-for good/innovation.html [Accessed on 10/05/2013].

SOUTH AFRICAN RESERVE BANK, 2009. *Selected South African Banking Sector Trends*, Bank Supervision Department. December 2009. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2010. *Selected South African Banking Sector Trends*, Bank Supervision Department. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2011. *Selected South African Banking Sector Trends*, Bank Supervision Department. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

SOUTH AFRICAN RESERVE BANK, 2012. *Selected South African Banking Sector Trends*, Bank Supervision Department. March 2012. Available [Online]: http://www.reservebank.co.za/ [Accessed 15 July 2012].

STATISTICS SOUTH AFRICA, 2013. Quarterly Labour Force Survey, Quarter 3, Available [Online]: http://www.statssa.gov.za/publications/P0211/P02113rdQuarter2013.pdf.[Accessed 5 December 2013].

STATISTICS SOUTH AFRICA, 2013. *Quarterly Labour Force Survey, Quarter 2*, Available [Online]:http://www.statssa.gov.za/publications/P0211/P02112ndQuarter2013.pdf. [Accessed 19 October 2013].

STATISTICS SOUTH AFRICA, 2013. *Quarterly Labour Force Survey, Quarter 1*, Available [Online]: http://www.statssa.gov.za/publications/P0211/P02111stQuarter2013.pdf. [Accessed 12 May 2013].

STATISTICS SOUTH AFRICA, 2010. *Quarterly Labour Force Survey*, *Quarter 2*, Available[Online] http://www.statssa.gov.za/publications/P0211/P02112ndQuarter2010.pdf. [Accessed 7 May 2012].

STATISTICS SOUTH AFRICA, 2012. *Quarterly Labour Force Survey*. Available at: Available at http://www.statssa.gov.za/keyindicators/qlf.asp [Accessed 7 April 2012].

STIROH, K.J., 2002. *Diversification in Banking Is Non-interest income the Answer?* Available [Online]: http://www.newyorkfed.org/research/staff_reports/sr154.pdf [Accessed 20 October 2013].

STUDENMUND, A.H., 2011. *Using Econometrics: A Practical Guide*, 6th Edition. Pearson Education Inc, USA.

SUFIAN, F., 2009. The Impact of the Asian Financial Crisis on Bank Efficiency: The 1997 Experience of Malaysia and Thailand, *Journal of International Development,* 22, pp. 866 –889.

SUFIAN, F., and HABIBULLAH, M,S., 2010. Assessing the Impact of Financial Crisis on Bank Performance Empirical Evidence from Indonesia, *ASEAN Economic Bulletin*, 27 (3), pp. 245–262.

SUFIAN, F., and MAJID, M.Z.A., 2007. *Bank Ownership, Characteristics, and Performance: A Comparative Analysis of Domestic and Foreign Islamic Banks in Malaysia*, MPRA Paper No. 12131, Available [Online]: http://mpra.ub.uni-muenchen.de/12131/ [Accessed 8 June 2013].

SUTTON, C.N., and JENKINS, B., 2007. *The Role of the Financial Service Sector in Expanding Economic Opportunity*, Corporate Social Responsibility Initiative Report No. 19. Cambridge, MA: Kennedy School of Government, Harvard University.

THE ECONOMISTS, 2007. *CSI: Credit Crunch, Central Banks have played a Starring Role*, 18th October 2007. Available [Online]: http://www.economist.com/node/9972489?storyid=9972489 [Accessed 15 June 2011].

TIENHAARA, K., 2010. A Tale of Two Crises: What the Global Crisis Means for the Global Environmental Crisis, *Environmental Policy and Governance*, 20(3), pp. 197 – 208.

TROUTT, M.D., HU, M.Y., and SHANKER, M.S., 2005. A Distribution-free Approach to Estimating Best Response Values with Application to Mutual Fund Performance Modelling, *European Journal of Operational Research*, 166 (2005), pp. 520 – 527.

VAN DER WESTHUIZEN, G., 2008. Estimating technical and scale efficiency and sources of efficiency change in banks using balance sheet data: a South African study, *Studies in Economics and Econometrics,* 32(1), pp. 23 – 45.

VAN DER WESTHUIZEN, G., 2010. The Effect of the Shift in the Composition of Bank Income on Bank Efficiency, *Journal for Studies in Economics and Econometrics*, 34(2), pp. 69 – 81.

VAN DER WESTHUIZEN, G., 2013. Bank Productivity and Sources of Efficiency Change: A Case of the Four Largest Banks in South Africa, *International Business & Economics Research Journal* 12(2), pp. 127 – 138.

VITTAS, D., 1991, Measuring Commercial Bank Efficiency, *Policy Research Working Papers Series, Country Economic Development,* The World Bank, Washington D. C.,

VU, H., and TURNELL, C., 2011. Cost and Profit Efficiencies of Australian Banks and the Impact of the Global Financial Crisis. *The Economic Record*, 87(279), pp. 525 – 536.

WALDEN, J.B., and TOMBERLIN, D., 2010. Estimating Fishing Vessel Capacity: A Comparison of Non-parametric Frontier Approaches, *Marine Resource Economics*, 25, pp. 23 – 26.

WALKER, I., 2010. *Research Methods and Statistics*, Palgrave Insights in Psychology series Basingstoke: Palgrave Macmillan.

WHEELOCK, D.C., and WILSON, P.W., 1995. Evaluating the Efficiency of Commercial Banks: Does our View of what Banks do Matter? *Federal Reserve Bank of St. Louis Review* 77, pp. 39 – 52.

WORLD BANK, 1993. *Poverty Reduction Handbook*, Washington, D.C.

WORLD BANK, 1995. *Poverty*. Washington, D.C., Oxford University Press, London.

WORLD BANK, 2006. *Making Finance Work for Africa*, Washington DC.

WORLD BANK, 2009. *Banking the Poor: Measuring Banking Access in 54 Economies*, The World Bank, Washington D.C.

WORLD BANK, 2010. *Financial Access 2010: The State of Financial Inclusion Through the Access*, The World Bank Group, Washington DC: USA.

WORLD BANK, 2014. World Development Indicators (WDI) database. The World Bank Group, Available [Online]: http://data.worldbank.org/indicator/SI.POV.2DAY. [Accessed 2014 March 7].

ZELLNER, A., 1962. An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias. *Journal of the American Statistical Association,* 57: pp. 348–368.

ZHU, J., 2003. *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spread Sheets and DEA Excel Solver*. Boston: Kluwer.

10.1 APPENDICES

APPENDIX 1: FIRST STAGE BANK LEVEL DATA

YEAR	BANK	LOANS (Rm)	INT (Rm)	NINT (Rm)	LAB	FA (Rm)	DEP (Rm)	OE (Rm)
2003	**ABSA**	222395	28901	10753	31658	2597	234380	30862
2003	**FRB**	209665	22412	10348	32732	4456	146226	23746
2003	**NEDBANK**	227872	28609	6451	24205	2596	239041	32359
2003	**STANDARD**	186529	31517	7672	25911	2149	187435	31879
2003	**CAPITEC**	163.6	398.7	4.3	1402	146.8	48.9	307.3
2003	**SASFIN**	1001.2	192.1	174.1	451	40.5	268.5	280.6
2003	**TEBA**	207	152.8	142.5	692	52	1242	280.5
2003	**AFRICAN**	6314	2439	570	2911	193	884	1407
2004	**ABSA**	268240	27132	11914	32515	2683	278582	29329
2004	**FRB**	228003	23417	14018	36156	4233	152405	26409
2004	**NEDBANK**	237372	22807	7679	21103	2622	258798	26679
2004	**STANDARD**	233435	30677	9345	26437	2069	229729	31781
2004	**CAPITEC**	238.6	544	9	1708	176.4	222.4	410.9
2004	**SASFIN**	1115.3	196.1	210.1	435	43.2	402.1	282.2
2004	**TEBA**	206	134.7	154.1	739	49.3	1464	302.9
2004	**AFRICAN**	6129	2608	585	2672	140	544	1510
2005	**ABSA**	302189	22267	7159	33029	3247	289113	24154
2005	**FRB**	317016	31229	37665	39738	5011	321235	35785
2005	**NEDBANK**	236442	23234	8469	22188	4415	229993	25930
2005	**STANDARD**	257925	31432	11978	26869	2605	316102	36051
2005	**CAPITEC**	547.3	784	16.9	1901	134	537.9	538.6
2005	**SASFIN**	1222.7	205.2	363.6	487	49.4	626.9	338
2005	**TEBA**	200.2	180.2	127.1	760	38.7	1656.2	295.7
2005	**AFRICAN**	6399	2908	631	2845	112	644	1493
2006	**ABSA**	374823	36518	11205	34348	3509	344452	37758
2006	**FRB**	391570	45324	20568	39726	6411	372310	48012
2006	**NEDBANK**	296282	28521	9295	24034	4643	292292	29789
2006	**STANDARD**	331044	38372	14520	29358	2829	405970	40535
2006	**CAPITEC**	914.2	967.5	111.8	2129	155.6	842.1	674.8
2006	**SASFIN**	1552.4	261.8	376.5	496	60.6	745.8	418.7
2006	**TEBA**	338.3	271.4	149.6	854	46.9	1747.4	356.2
2006	**AFRICAN**	7499	3087	870	2727	116	447	1559
2007	**ABSA**	463939	52213	12576	36893	4258	304877	51584
2007	**FRB**	453669	55475	21147	38863	8859	418750	58022
2007	**NEDBANK**	361668	42001	10239	26522	3929	339562	41649
2007	**STANDARD**	424473	52751	14646	30041	3577	455896	53175
2007	**CAPITEC**	2192.1	740.1	663.1	2800	196.2	1475.7	863.2
2007	**SASFIN**	1850.5	336	393.5	542	94.6	1108.1	541.5

2007		506.8	354.3	155.2	899	48.8	1595.2	371.8
2007	**AFRICAN**	10644	3268	1449	3011	155	808	1765
2008	**ABSA**	540144	73475	15289	33074	5431	373176	72053
2008	**FRB**	426079	52005	19789	39177	10220	318202	57474
2008	**NEDBANK**	428189	57986	10288	27570	4327	429426	56244
2008	**STANDARD**	487777	82797	15444	29663	4284	542586	79746
2008	**CAPITEC**	3238.1	1212.9	1038.2	3414	240.1	3298.9	1343.7
2008	**SASFIN**	1867.2	371.1	445.5	573	187.6	881.4	610.8
2008	**TEBA**	672.4	401.7	184.5	938	81.1	2361.4	474.8
2008	**AFRICAN**	20828	4627	5126	3426	496	3779	5405
2009	**ABSA**	518451	62533	15482	30627	6010	349371	61044
2009	**FRB**	443765	39054	26525	34904	10018	351394	47778
2009	**NEDBANK**	448155	50537	11850	27037	4967	427774	49769
2009	**STANDARD**	481678	66443	16640	29477	5144	520027	64821
2009	**CAPITEC**	5607.4	1764	1282.7	4154	281.6	7107.4	1875.6
2009	**SASFIN**	1982.8	352	415.9	563	184.4	911.6	608.8
2009	**TEBA**	467.6	303.3	297	919	92	2656.3	509.8
2009	**AFRICAN**	26147	5804	6123	3476	586	2758	6608
2010	**ABSA**	506483	52264	14667	36770	6987	372644	52091
2010	**FRB**	472615	38437	28202	34612	10542	338709	48333
2010	**NEDBANK**	469021	44377	13174	27525	5612	454135	44814
2010	**STANDARD**	488020	50650	17728	30396	7908	557385	53415
2010	**CAPITEC**	10916.2	2808.6	1683.6	5331	375.2	10449.9	2563.9
2010	**SASFIN**	2429.1	359.3	411.2	583	175.4	1215.4	620.1
2010	**TEBA**	840.1	309.3	217.2	783	73.8	2810.5	493.5
2010	**AFRICAN**	30968	6340	6065	3935	622	1038	6830
2011	**ABSA**	505462	49210	16514	35200	7268	431762	49517
2011	**FRB**	533347	41455	28578	36398	12026	472283	47875
2011	**NEDBANK**	490539	42880	15033	28494	6312	472740	44270
2011	**STANDARD**	561552	48196	18071	28422	8430	623295	50195
2011	**CAPITEC**	18408.2	4346.9	2319.6	7194	543.1	11660.1	3508.7
2011	**SASFIN**	2931	434	425.4	664	57.4	1787.3	706.6
2011	**TEBA**	1055.9	367.5	230	793	47.8	2882.3	479.2
2011	**AFRICAN**	41787	7647	7295	4978	852	1666	7780
	MEAN	**195072**	**22008**	**8078**	**16581**	**2766**	**179537**	**22908**
	MEDIAN	**114158**	**14957**	**6805**	**14149**	**1461**	**78943**	**15763**
	MAX	**561552**	**82797**	**37665**	**39738**	**12026**	**623295**	**79746**
	MIN	**164**	**135**	**4**	**435**	**39**	**49**	**281**
	STD. DEV	**207069**	**23148**	**8563**	**15145**	**3178**	**195640**	**23684**

APPENDIX 2 LARGE BANKS: *LEVELS* COMPUTED USING HICKS-MOORSTEEN AGGREGATOR FUNCTIONS															
Period	**Firm**	**TFP**	**TFP***	**TFPE**	**OTE**	**OSE**	**OME**	**ROSE**	**OSME**	**ITE**	**ISE**	**IME**	**RISE**	**ISME**	**RME**
2003	*ABSA*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2003	*FRB*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2003	*NEDBANK*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2003	*STANDARD*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2004	*ABSA*	1.268	1.997	0.635	1.000	1.000	1.000	0.635	0.635	1.000	1.000	1.000	0.635	0.635	0.635
2004	*FRB*	1.207	1.368	0.882	1.000	1.000	1.000	0.882	0.882	1.000	1.000	1.000	0.882	0.882	0.882
2004	*NEDBANK*	1.356	2.663	0.509	1.000	1.000	1.000	0.509	0.509	1.000	1.000	1.000	0.509	0.509	0.509
2004	*STANDARD*	1.279	1.449	0.883	1.000	1.000	1.000	0.883	0.883	1.000	1.000	1.000	0.883	0.883	0.883
2005	*ABSA*	1.660	6.866	0.242	1.000	1.000	1.000	0.242	0.242	1.000	1.000	1.000	0.242	0.242	0.242
2005	*FRB*	1.966	2.445	0.804	1.000	1.000	1.000	0.804	0.804	1.000	1.000	1.000	0.804	0.804	0.804
2005	*NEDBANK*	1.198	3.411	0.351	1.000	1.000	1.000	0.351	0.351	1.000	1.000	1.000	0.351	0.351	0.351
2005	*STANDARD*	1.301	2.681	0.485	1.000	1.000	1.000	0.485	0.485	1.000	1.000	1.000	0.485	0.485	0.485
2006	*ABSA*	1.410	1.939	0.727	1.000	1.000	1.000	0.727	0.727	1.000	1.000	1.000	0.727	0.727	0.727
2006	*FRB*	1.889	2.457	0.769	1.000	1.000	1.000	0.769	0.769	1.000	1.000	1.000	0.769	0.769	0.769
2006	*NEDBANK*	1.163	2.012	0.578	1.000	1.000	1.000	0.578	0.578	1.000	1.000	1.000	0.578	0.578	0.578
2006	*STANDARD*	1.102	1.263	0.873	1.000	1.000	1.000	0.873	0.873	1.000	1.000	1.000	0.873	0.873	0.873
2007	*ABSA*	1.402	1.820	0.771	1.000	1.000	1.000	0.771	0.771	1.000	1.000	1.000	0.771	0.771	0.771
2007	*FRB*	1.297	1.958	0.662	1.000	1.000	1.000	0.662	0.662	1.000	1.000	1.000	0.662	0.662	0.662
2007	*NEDBANK*	1.364	1.882	0.725	1.000	1.000	1.000	0.725	0.725	1.000	1.000	1.000	0.725	0.725	0.725
2007	*STANDARD*	1.322	1.700	0.777	1.000	1.000	1.000	0.777	0.777	1.000	1.000	1.000	0.777	0.777	0.777
2008	*ABSA*	1.372	1.972	0.696	1.000	1.000	1.000	0.696	0.696	1.000	1.000	1.000	0.696	0.696	0.696
2008	*FRB*	1.234	2.028	0.609	1.000	1.000	1.000	0.609	0.609	1.000	1.000	1.000	0.609	0.609	0.609
2008	*NEDBANK*	1.336	2.530	0.528	1.000	1.000	1.000	0.528	0.528	1.000	1.000	1.000	0.528	0.528	0.528
2008	*STANDARD*	1.504	2.529	0.595	1.000	1.000	1.000	0.595	0.595	1.000	1.000	1.000	0.595	0.595	0.595
2009	*ABSA*	1.247	2.616	0.477	1.000	1.000	1.000	0.477	0.477	1.000	1.000	1.000	0.477	0.477	0.477
2009	*FRB*	1.540	2.368	0.650	1.000	1.000	1.000	0.650	0.650	1.000	1.000	1.000	0.650	0.650	0.650
2009	*NEDBANK*	1.309	3.035	0.431	1.000	1.000	1.000	0.431	0.431	1.000	1.000	1.000	0.431	0.431	0.431
2009	*STANDARD*	1.408	2.867	0.491	1.000	1.000	1.000	0.491	0.491	1.000	1.000	1.000	0.491	0.491	0.491
2010	*ABSA*	1.282	1.657	0.774	1.000	1.000	1.000	0.774	0.774	1.000	1.000	1.000	0.774	0.774	0.774

2010	*FRB*	1.087	2.319	0.469	1.000	1.000	1.000	0.469	0.469	1.000	1.000	1.000	0.469	0.469	0.469
2010	*NEDBANK*	1.260	2.451	0.514	1.000	1.000	1.000	0.514	0.514	1.000	1.000	1.000	0.514	0.514	0.514
2010	*STANDARD*	1.615	2.490	0.649	1.000	1.000	0.959	0.676	0.649	1.000	1.000	1.000	0.649	0.649	0.649
2011	*ABSA*	1.207	2.120	0.570	1.000	1.000	1.000	0.570	0.570	1.000	1.000	1.000	0.570	0.570	0.570
2011	*FRB*	1.252	1.376	0.910	1.000	1.000	1.000	0.910	0.910	1.000	1.000	1.000	0.910	0.910	0.910
2011	*NEDBANK*	1.160	2.031	0.571	1.000	1.000	1.000	0.571	0.571	1.000	1.000	1.000	0.571	0.571	0.571
2011	*STANDARD*	1.203	1.732	0.694	1.000	1.000	1.000	0.694	0.694	1.000	1.000	1.000	0.694	0.694	0.694
MEAN		**1.311**	**2.167**	**0.675**	**1.000**	**1.000**	**0.999**	**0.676**	**0.675**	**1.000**	**1.000**	**1.000**	**0.675**	**0.675**	**0.675**
MAX		**1.966**	**6.866**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
MIN		**1.000**	**1.000**	**0.242**	**1.000**	**1.000**	**0.959**	**0.242**	**0.242**	**1.000**	**1.000**	**1.000**	**0.242**	**0.242**	**0.242**
STDEV		**0.220**	**1.006**	**0.192**	**0.000**	**0.000**	**0.007**	**0.192**	**0.192**	**0.000**	**0.000**	**0.000**	**0.192**	**0.192**	**0.192**

APPENDIX 3: SMALL BANKS: *LEVELS* COMPUTED USING HICKS-MOORSTEEN AGGREGATOR FUNCTIONS															
Period	**Firm**	**TFP**	**TFP***	**TFPE**	**OTE**	**OSE**	**OME**	**ROSE**	**OSME**	**ITE**	**ISE**	**IME**	**RISE**	**ISME**	**RME**
2003	*CAPITEC*	1.000	1.265	0.790	1.000	1.000	1.000	0.790	0.790	1.000	1.000	1.000	0.790	0.790	0.790
2003	*SASFIN*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2003	*TEBA*	1.000	1.222	0.819	1.000	0.819	1.000	0.819	0.819	1.000	0.819	1.000	0.819	0.819	1.000
2003	*AFRICAN*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2004	*CAPITEC*	2.318	67.858	0.034	1.000	0.767	1.000	0.034	0.034	1.000	0.767	0.955	0.036	0.034	0.045
2004	*SASFIN*	1.340	8.307	0.161	1.000	1.000	1.000	0.161	0.161	1.000	1.000	1.000	0.161	0.161	0.161
2004	*TEBA*	0.815	1.466	0.556	0.712	0.960	1.000	0.780	0.780	0.932	0.734	1.000	0.597	0.596	0.813
2004	*AFRICAN*	1.386	1.738	0.798	1.000	1.000	1.000	0.798	0.798	1.000	1.000	1.000	0.798	0.798	0.798
2005	*CAPITEC*	1.770	8.411	0.210	1.000	0.747	1.000	0.210	0.210	1.000	0.747	0.992	0.212	0.210	0.282
2005	*SASFIN*	1.514	11.101	0.136	1.000	1.000	1.000	0.136	0.136	1.000	1.000	1.000	0.136	0.136	0.136
2005	*TEBA*	1.009	4.669	0.216	1.000	0.530	1.000	0.216	0.216	1.000	0.530	1.000	0.216	0.216	0.408
2005	*AFRICAN*	1.257	1.394	0.902	1.000	1.000	1.000	0.902	0.902	1.000	1.000	1.000	0.902	0.902	0.902
2006	*CAPITEC*	2.372	31.605	0.075	1.000	0.724	0.479	0.157	0.075	1.000	0.724	0.965	0.078	0.075	0.104
2006	*SASFIN*	1.224	2.687	0.456	1.000	1.000	1.000	0.456	0.456	1.000	1.000	1.000	0.456	0.456	0.456

2006	*TEBA*	1.223	17.361	0.070	1.000	0.579	1.000	0.070	0.070	1.000	0.579	1.000	0.070	0.070	0.122
2006	*AFRICAN*	1.398	1.986	0.704	1.000	1.000	1.000	0.704	0.704	1.000	1.000	1.000	0.704	0.704	0.704
2007	*CAPITEC*	0.963	1.496	0.644	0.988	0.947	0.736	0.885	0.652	0.989	0.946	0.991	0.657	0.651	0.688
2007	*SASFIN*	1.309	3.024	0.433	1.000	1.000	1.000	0.433	0.433	1.000	1.000	1.000	0.433	0.433	0.433
2007	*TEBA*	1.233	13.462	0.092	1.000	0.515	1.000	0.092	0.092	1.000	0.515	1.000	0.092	0.092	0.178
2007	*AFRICAN*	1.605	2.276	0.705	1.000	1.000	1.000	0.705	0.705	1.000	1.000	1.000	0.705	0.705	0.705
2008	*CAPITEC*	1.014	1.291	0.786	1.000	1.000	0.946	0.831	0.786	1.000	1.000	1.000	0.786	0.786	0.786
2008	*SASFIN*	1.461	4.075	0.358	1.000	0.793	1.000	0.358	0.358	1.000	0.793	1.000	0.358	0.358	0.452
2008	*TEBA*	1.302	4.797	0.271	1.000	0.952	1.000	0.271	0.271	1.000	0.952	0.996	0.273	0.271	0.285
2008	*AFRICAN*	3.205	3.205	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2009	*CAPITEC*	1.089	1.255	0.867	1.000	1.000	1.000	0.867	0.867	1.000	1.000	1.000	0.867	0.867	0.867
2009	*SASFIN*	1.094	3.528	0.310	1.000	0.823	1.000	0.310	0.310	1.000	0.823	1.000	0.310	0.310	0.377
2009	*TEBA*	1.610	10.320	0.156	1.000	0.667	1.000	0.156	0.156	1.000	0.667	0.974	0.160	0.156	0.234
2009	*AFRICAN*	1.327	1.720	0.771	1.000	1.000	1.000	0.771	0.771	1.000	1.000	1.000	0.771	0.771	0.771
2010	*CAPITEC*	1.098	1.358	0.808	1.000	1.000	1.000	0.808	0.808	1.000	1.000	1.000	0.808	0.808	0.808
2010	*SASFIN*	1.299	8.412	0.154	1.000	0.864	1.000	0.154	0.154	1.000	0.864	0.944	0.164	0.154	0.179
2010	*TEBA*	1.745	9.796	0.178	1.000	0.602	1.000	0.178	0.178	1.000	0.602	1.000	0.178	0.178	0.296
2010	*AFRICAN*	1.897	3.147	0.603	1.000	1.000	1.000	0.603	0.603	1.000	1.000	1.000	0.603	0.603	0.603
2011	*CAPITEC*	1.104	1.496	0.738	1.000	1.000	1.000	0.738	0.738	1.000	1.000	1.000	0.738	0.738	0.738
2011	*SASFIN*	1.751	7.901	0.222	1.000	1.000	1.000	0.222	0.222	1.000	1.000	1.000	0.222	0.222	0.222
2011	*TEBA*	1.355	4.486	0.302	1.000	0.857	1.000	0.302	0.302	1.000	0.857	1.000	0.302	0.302	0.353
2011	*AFRICAN*	1.257	1.257	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MEAN		**1.398**	**6.982**	**0.509**	**0.992**	**0.893**	**0.977**	**0.526**	**0.516**	**0.998**	**0.887**	**0.995**	**0.511**	**0.510**	**0.547**
MAX		**3.205**	**67.858**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
MIN		**0.815**	**1.000**	**0.034**	**0.712**	**0.515**	**0.479**	**0.034**	**0.034**	**0.932**	**0.515**	**0.944**	**0.036**	**0.034**	**0.045**
STDEV		**0.472**	**12.066**	**0.326**	**0.048**	**0.155**	**0.096**	**0.333**	**0.329**	**0.011**	**0.156**	**0.013**	**0.326**	**0.327**	**0.315**

APPENDIX 4: LARGE BANKS: HICKS-MOORSTEEN INDEXES COMPARING FIRM *i* IN PERIOD *t* WITH FIRM *i* IN PERIOD *t-1*															
Period	**Firm**	**ΔTFP**	**ΔTECH**	**ΔTFPE**	**ΔOTE**	**ΔOSE**	**ΔOME**	**ΔROSE**	**ΔOSME**	**ΔITE**	**ΔISE**	**ΔIME**	**ΔRISE**	**ΔISME**	**ΔRME**
2003	*ABSA*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*FRB*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*NEDBANK*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*STANDARD*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2004	*ABSA*	1.001	0.957	1.046	1.000	1.000	1.025	1.020	1.046	1.000	1.000	1.000	1.046	1.046	1.046
2004	*FRB*	1.134	1.180	0.961	1.000	1.000	1.000	0.961	0.961	1.000	1.000	1.000	0.961	0.961	0.961
2004	*NEDBANK*	1.065	1.218	0.874	1.000	1.000	1.000	0.874	0.874	1.000	1.000	1.000	0.874	0.874	0.874
2004	*STANDARD*	1.016	0.957	1.062	1.000	1.000	1.000	1.062	1.062	1.000	1.000	1.000	1.062	1.062	1.062
2005	*ABSA*	0.824	1.678	0.491	1.000	1.000	1.000	0.491	0.491	1.000	1.000	1.000	0.491	0.491	0.491
2005	*FRB*	1.244	0.979	1.271	1.000	1.000	1.000	1.271	1.271	1.000	1.000	1.000	1.271	1.271	1.271
2005	*NEDBANK*	0.981	1.744	0.562	1.000	1.000	1.000	0.562	0.562	1.000	1.000	1.000	0.562	0.562	0.562
2005	*STANDARD*	0.969	1.292	0.750	1.000	1.000	1.000	0.750	0.750	1.000	1.000	1.000	0.750	0.750	0.750
2006	*ABSA*	1.119	1.117	1.001	1.000	1.000	1.000	1.001	1.001	1.000	1.000	1.000	1.001	1.001	1.001
2006	*FRB*	0.769	0.407	1.889	1.000	1.000	1.000	1.889	1.889	1.000	1.000	1.000	1.889	1.889	1.889
2006	*NEDBANK*	1.005	0.471	2.132	1.000	1.000	1.000	2.132	2.132	1.000	1.000	1.000	2.132	2.132	2.132
2006	*STANDARD*	1.073	0.456	2.353	1.000	1.000	1.000	2.353	2.353	1.000	1.000	1.000	2.353	2.353	2.353
2007	*ABSA*	1.152	0.851	1.354	1.000	1.000	1.000	1.354	1.354	1.000	1.000	1.000	1.354	1.354	1.354
2007	*FRB*	0.965	0.857	1.126	1.000	1.000	1.000	1.126	1.126	1.000	1.000	1.000	1.126	1.126	1.126
2007	*NEDBANK*	1.074	0.851	1.263	1.000	1.000	1.000	1.263	1.263	1.000	1.000	1.000	1.263	1.263	1.263
2007	*STANDARD*	1.017	0.849	1.198	1.000	1.000	1.000	1.198	1.198	1.000	1.000	1.001	1.197	1.198	1.198
2008	*ABSA*	1.144	1.370	0.835	1.000	1.000	1.000	0.835	0.835	1.000	1.000	1.000	0.835	0.835	0.835
2008	*FRB*	1.001	0.916	1.093	1.000	1.000	1.000	1.093	1.093	1.000	1.000	1.000	1.093	1.093	1.093
2008	*NEDBANK*	0.994	0.945	1.052	1.000	1.000	1.000	1.052	1.052	1.000	1.000	1.000	1.052	1.052	1.052
2008	*STANDARD*	1.057	0.945	1.119	1.000	1.000	1.000	1.119	1.119	1.000	1.000	1.002	1.116	1.119	1.119
2009	*ABSA*	0.959	1.295	0.741	1.000	1.000	1.000	0.741	0.741	1.000	1.000	1.000	0.741	0.741	0.741
2009	*FRB*	1.047	0.909	1.152	1.000	1.000	1.000	1.152	1.152	1.000	1.000	1.000	1.152	1.152	1.152

2009	*NEDBANK*	0.994	1.399	0.711	1.000	1.000	1.000	0.711	0.711	1.000	1.000	1.005	0.707	0.711	0.711
2009	*STANDARD*	0.941	1.437	0.655	1.000	1.000	1.000	0.655	0.655	1.000	1.000	1.000	0.655	0.655	0.655
2010	*ABSA*	0.893	0.764	1.168	1.000	1.000	1.000	1.168	1.168	1.000	1.000	1.000	1.168	1.168	1.168
2010	*FRB*	1.017	0.682	1.490	1.000	1.000	1.000	1.490	1.490	1.000	1.000	1.000	1.490	1.490	1.490
2010	*NEDBANK*	0.980	1.051	0.932	1.000	1.000	1.000	0.932	0.932	1.000	1.000	1.000	0.932	0.932	0.932
2010	*STANDARD*	0.801	0.682	1.174	1.000	1.000	0.959	1.224	1.174	1.000	1.000	1.000	1.174	1.174	1.174
2011	*ABSA*	0.981	1.023	0.959	1.000	1.000	1.000	0.959	0.959	1.000	1.000	1.000	0.959	0.959	0.959
2011	*FRB*	0.910	0.727	1.252	1.000	1.000	1.000	1.252	1.252	1.000	1.000	1.000	1.252	1.252	1.252
2011	*NEDBANK*	0.996	1.023	0.974	1.000	1.000	1.000	0.974	0.974	1.000	1.000	1.000	0.974	0.974	0.974
2011	*STANDARD*	1.023	0.813	1.259	1.000	1.000	1.000	1.259	1.259	1.000	1.000	1.000	1.259	1.259	1.259
MEAN		**1.004**	**0.995**	**1.122**	**1.000**	**1.000**	**1.000**	**1.123**	**1.122**	**1.000**	**1.000**	**1.000**	**1.122**	**1.122**	**1.122**
MAX		**1.244**	**1.744**	**2.353**	**1.000**	**1.000**	**1.025**	**2.353**	**2.353**	**1.000**	**1.000**	**1.005**	**2.353**	**2.353**	**2.353**
MIN		**0.769**	**0.407**	**0.491**	**1.000**	**1.000**	**0.959**	**0.491**	**0.491**	**1.000**	**1.000**	**1.000**	**0.491**	**0.491**	**0.491**
STDEV		**0.0998**	**0.3202**	**0.4043**	**0.0000**	**0.0000**	**0.0086**	**0.4048**	**0.4043**	**0.0000**	**0.0000**	**0.0010**	**0.4044**	**0.4043**	**0.4043**

APPENDIX 5: SMALL BANKS: HICKS-MOORSTEEN INDEXES COMPARING FIRM *i* IN PERIOD *t* WITH FIRM *i* IN PERIOD *t-1*															
Period	**Firm**	**ΔTFP**	**ΔTECH**	**ΔTFPE**	**ΔOTE**	**ΔOSE**	**ΔOME**	**ΔROSE**	**ΔOSME**	**ΔITE**	**ΔISE**	**ΔIME**	**ΔRISE**	**ΔISME**	**ΔRME**
2003	*CAPITEC*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*SASFIN*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*TEBA*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2003	*AFRICAN*	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2004	*CAPITEC*	0.6954	1.2071	0.5761	1	0.7665	1	0.5761	0.5761	1	0.7665	0.9547	0.6034	0.5761	0.7516
2004	*SASFIN*	0.934	0.8757	1.0665	1	1	1	1.0665	1.0665	1	1	1	1.0665	1.0665	1.0665
2004	*TEBA*	1.0015	1.1998	0.8347	0.712	1.1721	1	1.1723	1.1723	0.9317	0.8958	1	0.896	0.8959	1.0002
2004	*AFRICAN*	1.2536	0.588	2.1321	1	1	1	2.1321	2.1321	1	1	1	2.1321	2.1321	2.1321
2005	*CAPITEC*	1.0546	1.0633	0.9918	1	0.975	1	0.9918	0.9918	1	0.975	0.9916	1.0002	0.9918	1.0173
2005	*SASFIN*	1.0209	0.7736	1.3197	1	1	1	1.3197	1.3197	1	1	1.0577	1.2477	1.3197	1.3197
2005	*TEBA*	1.1844	1.3938	0.8498	1.4044	0.5521	1.1269	0.537	0.6051	1.0734	0.7224	1.0972	0.7216	0.7917	1.096
2005	*AFRICAN*	1.1087	1.2292	0.902	1	1	1	0.902	0.902	1	1	1	0.902	0.902	0.902
2006	*CAPITEC*	2.2557	0.8472	2.6626	1	0.9689	0.479	5.5591	2.6626	1	0.9689	0.9649	2.7594	2.6626	2.7481
2006	*SASFIN*	1.0233	1.1075	0.924	1	1	1	0.924	0.924	1	1	1	0.924	0.924	0.924
2006	*TEBA*	1.1876	1.1529	1.0301	1	1.0928	1	1.0301	1.0301	1	1.0928	1	1.0301	1.0301	0.9426
2006	*AFRICAN*	1.4211	1.9864	0.7154	1	1	1	0.7154	0.7154	1	1	1	0.7154	0.7154	0.7154
2007	*CAPITEC*	1.6519	0.8589	1.9233	0.9882	1.3077	1.537	1.2663	1.9463	0.9894	1.3062	1.0073	1.9298	1.944	1.4883
2007	*SASFIN*	0.8973	1.0517	0.8532	1	1	1	0.8532	0.8532	1	1	1	0.8532	0.8532	0.8532
2007	*TEBA*	1.2189	1.0436	1.168	1	0.8888	1	1.168	1.168	1	0.8888	1.1224	1.0406	1.168	1.3141
2007	*AFRICAN*	0.9399	0.864	1.0878	1	1	1	1.0878	1.0878	1	1	1	1.0878	1.0878	1.0878
2008	*CAPITEC*	0.9805	0.842	1.1645	1.012	1.0561	1.1412	1.0084	1.1508	1.0108	1.0573	1.0253	1.1237	1.1521	1.0897
2008	*SASFIN*	0.9567	1.0867	0.8804	1	0.7933	1	0.8804	0.8804	1	0.7933	1	0.8804	0.8804	1.1098
2008	*TEBA*	0.8883	0.8349	1.064	1	1.8494	1	1.064	1.064	1	1.8494	0.9956	1.0687	1.064	0.5753
2008	*AFRICAN*	0.9702	0.7899	1.2283	1	1	1	1.2283	1.2283	1	1	1	1.2283	1.2283	1.2283

2009	*CAPITEC*	1.0091	1.0085	1.0006	1	1	1	1.0006	1.0006	1	1	1	1.0006	1.0006	1.0006
2009	*SASFIN*	0.9877	1.4231	0.6941	1	1.0376	1	0.6941	0.6941	1	1.0376	1	0.6941	0.6941	0.6689
2009	*TEBA*	0.987	1.0906	0.905	1	0.7008	1	0.905	0.905	1	0.7008	0.9735	0.9296	0.905	1.2914
2009	*AFRICAN*	1.2964	1.6806	0.7714	1	1	1	0.7714	0.7714	1	1	1	0.7714	0.7714	0.7714
2010	*CAPITEC*	1.1587	1.0911	1.0619	1	1	1.0215	1.0395	1.0619	1	1	1.0179	1.0432	1.0619	1.0619
2010	*SASFIN*	0.9677	1.6027	0.6038	1	1.0496	1	0.6038	0.6038	1	1.0496	0.9441	0.6396	0.6038	0.5752
2010	*TEBA*	1.2834	1.1158	1.1502	1	0.9029	1	1.1502	1.1502	1	0.9029	1.0286	1.1183	1.1502	1.2738
2010	*AFRICAN*	1.6594	2.7535	0.6026	1	1	1	0.6026	0.6026	1	1	1	0.6026	0.6026	0.6026
2011	*CAPITEC*	1.1452	1.1394	1.0051	1	1	1	1.0051	1.0051	1	1	1.0427	0.964	1.0051	1.0051
2011	*SASFIN*	1.613	0.8246	1.9562	1	1.1575	1	1.9562	1.9562	1	1.1575	1.3463	1.453	1.9562	1.6901
2011	*TEBA*	1.4322	1.0256	1.3965	1	1.4223	1	1.3965	1.3965	1	1.4223	1.0664	1.3096	1.3965	0.9818
2011	*AFRICAN*	0.9422	0.9422	1	1	1	1	1	1	1	1	1	1	1	1
MEAN		**1.160**	**1.140**	**1.110**	**1.004**	**1.022**	**1.010**	**1.175**	**1.113**	**1.000**	**1.018**	**1.020**	**1.086**	**1.110**	**1.103**
MAX		**2.2557**	**2.7535**	**2.6626**	**1.4044**	**1.8494**	**1.537**	**5.5591**	**2.6626**	**1.0734**	**1.8494**	**1.3463**	**2.7594**	**2.6626**	**2.7481**
MIN		**0.6954**	**0.588**	**0.5761**	**0.712**	**0.5521**	**0.479**	**0.537**	**0.5761**	**0.9317**	**0.7008**	**0.9441**	**0.6026**	**0.5761**	**0.5752**
STDEV		**0.3058**	**0.4112**	**0.4638**	**0.0891**	**0.2170**	**0.1384**	**0.8702**	**0.4690**	**0.0182**	**0.2065**	**0.0696**	**0.4505**	**0.4650**	**0.4413**

APPENDIX 6: COST TO INCOME RATIO

Year/Bank	ABSA	FRB	NED	STAN	CAP	SAS	TEBA	AFRI	AVG
2005	66.90	37.23	65.50	57.97	65.44	50.97	95.96	31.33	**58.91**
2006	60.62	52.50	59.87	51.39	59.84	55.07	83.65	27.67	**56.33**
2007	56.54	56.38	55.72	51.30	57.83	60.88	71.10	48.48	**57.28**
2008	53.28	58.98	53.74	46.17	54.18	62.45	79.10	46.28	**56.77**
2009	52.03	57.77	54.98	48.41	54.17	68.84	83.37	44.37	**57.99**
2010	58.61	58.93	56.75	56.65	48.44	72.44	93.14	40.77	**60.72**
2011	57.97	54.98	57.99	54.67	44.04	72.45	78.95	37.62	**57.33**
AVG	**57.99**	**53.82**	**57.79**	**52.37**	**54.85**	**63.30**	**83.61**	**39.50**	**57.90**
MAX	**66.90**	**58.98**	**65.5**	**57.97**	**65.44**	**72.45**	**95.96**	**48.48**	**95.96**
MIN	**52.03**	**37.23**	**53.74**	**46.17**	**44.04**	**50.97**	**71.1**	**27.67**	**27.67**

Source: Bankscope database www.bvdinfo.com

APPENDIX 7: NET-INTEREST MARGIN

NET-INTEREST MARGIN									
Year/Bank	**ABSA**	**FRB**	**NED**	**STAN**	**CAP**	**SAS**	**TEBA**	**AFRI**	**AVG**
2005	3.62	3.33	2.78	2.52	110.59	9.04	9.83	42.63	23.04
2006	3.51	3.25	3.15	2.47	70.71	9.40	15.98	35.29	17.97
2007	3.51	3.37	3.43	2.87	32.40	9.76	17.63	26.04	12.38
2008	3.21	2.37	3.37	2.93	34.44	8.52	14.98	20.23	11.26
2009	2.96	2.51	3.09	2.73	26.37	8.96	10.06	17.18	9.23
2010	3.31	2.84	3.08	2.68	23.40	7.37	10.22	15.86	8.60
2011	3.44	3.28	3.18	2.59	22.84	6.74	11.66	15.92	8.71
AVG	**3.37**	**2.99**	**3.15**	**2.68**	**45.82**	**8.54**	**12.91**	**24.74**	**13.03**
MAX	**3.62**	**3.37**	**3.43**	**2.93**	**110.59**	**9.76**	**17.63**	**42.63**	**110.59**
MIN	**2.96**	**2.37**	**2.78**	**2.47**	**22.84**	**6.74**	**9.83**	**15.86**	**2.37**

Source: Bankscope database www.bvdinfo.com

APPENDIX 8: NON-INTEREST INCOME TO GROSS REVENUE

NON-INTEREST INCOME/GROSS REVENUE									
Year/Bank	**ABSA**	**FRB**	**NED**	**STAN**	**CAP**	**SAS**	**TEBA**	**AFRI**	**AVG**
2005	43.86	70.39	49.42	53.88	2.22	76.16	44.25	24.91	45.64
2006	44.15	51.33	45.49	54.51	11.06	71.71	37.75	35.51	43.94
2007	41.24	47.21	41.76	45.62	50.35	66.99	32.58	60.73	48.31
2008	42.66	53.23	38.54	40.41	52.37	70.58	34.61	61.84	49.28
2009	43.71	61.53	42.01	42.58	50.17	66.46	54.58	60.52	52.70
2010	40.79	61.55	43.86	45.81	45.00	67.31	45.13	60.33	51.22
2011	42.76	56.50	44.88	41.35	41.09	66.50	40.93	56.18	48.77
AVG	**42.74**	**57.39**	**43.71**	**46.31**	**36.04**	**69.39**	**41.40**	**51.43**	**48.55**
MAX	**44.15**	**70.39**	**49.42**	**54.51**	**52.37**	**76.16**	**54.58**	**61.84**	**76.16**
MIN	**40.79**	**47.21**	**38.54**	**40.41**	**2.22**	**66.46**	**32.58**	**24.91**	**2.22**

Source: Bankscope database www.bvdinfo.com

APPENDIX 9: NET FEES & COMMISSIONS TO NON-INTEREST INCOME

NET FEES & COMMISSIONS/NON-INTEREST INCOME									
Year/Bank	**ABSA**	**FRB**	**NED**	**STAN**	**CAP**	**SAS**	**TEBA**	**AFRI**	**AVG**
2005	88.67	31.88	68.13	73.26	88.17	77.06	78.76	47.24	69.15
2006	80.08	65.38	70.34	67.37	99.73	88.76	71.86	46.45	73.75
2007	79.88	78.83	73.52	70.46	98.54	100.00	90.08	32.29	77.95
2008	76.66	78.04	76.89	75.99	99.76	87.92	72.63	31.85	74.97
2009	79.10	64.31	72.43	79.81	99.91	88.87	84.75	36.04	75.65
2010	84.65	65.08	74.07	77.45	100.00	93.12	88.31	35.87	77.32
2011	81.10	62.99	73.38	80.34	99.45	84.96	93.78	39.74	76.97
AVG	**81.45**	**63.79**	**72.68**	**74.95**	**97.94**	**88.67**	**82.88**	**38.50**	**75.11**
MAX	**88.67**	**78.83**	**76.89**	**79.81**	**100.00**	**100.00**	**90.08**	**47.24**	**100.00**
MIN	**76.66**	**31.88**	**68.13**	**67.37**	**88.17**	**77.06**	**71.86**	**31.85**	**31.85**

Source: Bankscope database www.bvdinfo.com

APPENDIX 10: IMPAIRED LOANS TO GROSS LOANS

IMPAIRED LOANS (NPLS)/GROSS LOANS									
Year/Bank	**ABSA**	**FRB**	**NED**	**STAN**	**CAP**	**SAS**	**TEBA**	**AFRI**	**AVG**
2005	1.83	1.33	2.94	1.33	n/a	5.24	n/a	29.51	7.03
2006	1.28	1.66	2.61	1.29	11.58	4.59	n/a	28.22	7.32
2007	1.54	2.89	2.74	1.88	11.24	5.19	21.21	29.95	9.58
2008	3.34	5.69	4.04	4.39	10.06	7.96	24.90	31.55	11.49
2009	6.75	5.41	6.03	7.53	6.24	7.36	24.40	29.39	11.64
2010	7.50	4.31	5.71	7.21	5.73	7.76	44.94	26.61	13.72
2011	6.80	3.50	4.73	4.54	5.06	6.42	35.75	27.65	17.50
AVG	**4.15**	**3.54**	**4.11**	**4.02**	**8.32**	**6.36**	**30.24**	**28.98**	**11.22**
MAX	**7.50**	**5.69**	**6.03**	**7.53**	**11.58**	**7.96**	**44.94**	**31.55**	**14.94**
MIN	**1.28**	**1.33**	**2.61**	**1.29**	**5.06**	**4.59**	**21.21**	**26.61**	**1.28**

Source: Bankscope database www.bvdinfo.com

APPENDIX 11: EQUALITY OF PRE-CRISIS AND CRISIS MEAN TFPE TEST

Test for Equality of Means: TFPE
Sample: 2003 2011
Included observations: 72

Method	df	Value	Probability
t-test	54	0.504386	0.6160
Satterthwaite-Welch t-test*	53.71554	0.530884	0.5977
Anova F-test	(1, 54)	0.254406	0.6160
Welch F-test*	(1, 53.7155)	0.281838	0.5977

*Test allows for unequal cell variances

Category Statistics

Variable	Count	Mean	Std. Dev.	Std. Err. of Mean
PRE	32	0.614534	0.327123	0.057828
CRISIS	24	0.575229	0.226494	0.046233
All	56	0.597689	0.286622	0.038301

APPENDIX 12: EQUALITY OF PRE-CRISIS AND CRISIS MEAN CIR TEST

Test for Equality of Means Between Series
Sample: 2005 2011
Included observations: 56

Method	df	Value	Probability
t-test	38	-0.063307	0.9499
Satterthwaite-Welch t-test*	21.46326	-0.056927	0.9551
Anova F-test	(1, 38)	0.004008	0.9499
Welch F-test*	(1, 21.4633)	0.003241	0.9551

*Test allows for unequal cell variances

Category Statistics

Variable	Count	Mean	Std. Dev.	Std. Err. of Mean
CRISIS	24	57.34792	9.806547	2.001753
PRE	16	57.61938	17.31217	4.328041
All	40	57.45650	13.11512	2.073682

APPENDIX 13: BALANCE SHEET SIZE AS AT 31 DECEMBER 2011

BANK	TOTAL ASSETS AT 31 DECEMBER (R MILLIONS)
The Standard Bank of South Africa Limited	889 250
Absa Bank Limited	725 679
FirstRand Bank Limited	665 525
Nedbank Limited	585 033
Investec Bank Limited	253 514
African Bank Limited	49 236
Capitec Bank Limited	22 230
Mercantile Bank Limited	6 136
Bidvest Bank Limited	4 062
Ubank[16] Limited	3 586
HBZ Bank Limited	3 530
Grindrod Bank Limited	3 481
Albaraka Bank Limited	3 238
Sasfin Bank Limited	2 767
The South African Bank of Athens Limited	1 653
Habib Overseas Bank Limited	1 032

Source: SARB Supervision Department, Annual Report, 2011

[16] Formerly known as TEBA Bank limited

APPENDIX 14: PANEL UNIT ROOT TESTS: EFFICIENCY & ACCESS MODEL

Pool unit root test: Summary
Series: ABSA_TFPE, FNB_TFPE, NED_TFPE, STAN_TFPE, CAP_TFPE, SAS_TFPE, TEBA_TFPE, AFRI_TFPE, ABSA_PY, FNB_PY, NED_PY, STAN_PY, CAP_PY, SAS_PY, TEBA_PY, AFRI_PY, ABSA_BDEV, FNB_BDEV, NED_BDEV, STAN_BDEV, CAP_BDEV, SAS_BDEV, TEBA_BDEV, AFRI_BDEV, ABSA_BRA1, FNB_BRA1, NED_BRA1, STAN_BRA1, CAP_BRA1, SAS_BRA1, TEBA_BRA1, AFRI_BRA1, ABSA_HHI, FNB_HHI, NED_HHI, STAN_HHI, CAP_HHI, SAS_HHI, TEBA_HHI, AFRI_HHI, ABSA_OE, FNB_OE, NED_OE, STAN_OE, CAP_OE, SAS_OE, TEBA_OE, AFRI_OE
Date: 10/16/13 Time: 23:49
Sample: 2004 2011
Exogenous variables: Individual effects
Automatic selection of maximum lags
Automatic lag length selection based on SIC: 0
Newey-West automatic bandwidth selection and Bartlett kernel
Balanced observations for each test

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-10.4511	0.0000	48	288
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-2.70597	0.0034	48	288
ADF - Fisher Chi-square	138.358	0.0030	48	288
PP - Fisher Chi-square	151.575	0.0003	48	288

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

APPENDIX 15: BANK SECTOR EFFICIENCY AND ACCESS TO BANKING SERVICES

Dependent Variable: ((?ACC+0.458918*?ACC(-1)))
Method: Pooled Least Squares
Date: 11/02/13 Time: 14:23
Sample (adjusted): 2005 2011
Included observations: 7 after adjustments
Cross-sections included: 8
Total pool (balanced) observations: 56
White diagonal standard errors & covariance (no d.f. correction)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	185.1434	9.924387	18.65540	0.0000
(?TFPE+0.458918*?TFPE(-1))	0.001246	0.000512	2.435717	0.0193
(?PY+0.458918*?PY(-1))	2.493894	0.058201	42.84979	0.0000
(?BDEV+0.458918*?BDEV(-1))	1.070796	0.021049	50.87114	0.0000
(?BRA1+0.458918*?BRA1(-1))	0.749867	0.012624	59.39775	0.0000
(?HHI+0.458918*?HHI(-1))	-4.086575	0.076178	-53.64479	0.0000
(?OE+0.458918*?OE(-1))	-0.009522	0.000577	-16.49727	0.0000
(?NRUR+0.458918*?NRUR(-1))	-33.17327	1.534176	-21.62285	0.0000
Fixed Effects (Cross)				
ABSA_--C	-0.003170			
FNB_--C	-0.002603			
NED_--C	-0.001396			
STAN_--C	-0.000319			
CAP_--C	0.001720			
SAS_--C	0.001839			
TEBA_--C	0.001865			
AFRI_--C	0.002064			

Effects Specification

Cross-section fixed (dummy variables)

R-squared	0.999949	Mean dependent var	4.193768
Adjusted R-squared	0.999932	S.D. dependent var	0.156146
S.E. of regression	0.001291	Akaike info criterion	-10.24295
Sum squared resid	6.83E-05	Schwarz criterion	-9.700441
Log likelihood	301.8025	Hannan-Quinn criter.	-10.03262
F-statistic	57471.94	Durbin-Watson stat	2.094058
Prob(F-statistic)	0.000000		

APPENDIX 16: PANEL UNIT ROOT TESTS: EFFICIENCY & UNEMPLOYMENT MODEL

Pool unit root test: Summary
Series: ABSA_TFPE, FNB_TFPE, NED_TFPE, STAN_TFPE, CAP_TFPE, SAS_TFPE, TEBA_TFPE, AFRI_TFPE, ABSA_SPREAD, FNB_SPREAD, NED_SPREAD, STAN_SPREAD, CAP_SPREAD, SAS_SPREAD, TEBA_SPREAD, AFRI_SPREAD, ABSA_CI, FNB_CI, NED_CI, STAN_CI, CAP_CI, SAS_CI, TEBA_CI, AFRI_CI, ABSA_BDEV, FNB_BDEV, NED_BDEV, STAN_BDEV, CAP_BDEV, SAS_BDEV, TEBA_BDEV, AFRI_BDEV, ABSA_GDPG, FNB_GDPG, NED_GDPG, STAN_GDPG, CAP_GDPG, SAS_GDPG, TEBA_GDPG, AFRI_GDPG
Date: 05/29/14 Time: 15:24
Sample: 2004 2011
Exogenous variables: Individual effects
Automatic selection of maximum lags
Automatic lag length selection based on SIC: 0 to 1
Newey-West automatic bandwidth selection and Bartlett kernel

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-7.19476	0.0000	40	266
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-2.00141	0.0227	40	266
ADF - Fisher Chi-square	106.656	0.0249	40	266
PP - Fisher Chi-square	123.749	0.0012	40	280

** Probabilities for Fisher tests are computed using an asymptotic Chi -square distribution. All other tests assume asymptotic normality.

APPENDIX 17: BANK SECTOR EFFICIENCY AND UNEMPLOYMENT

Dependent Variable: ?UNEM
Method: Pooled EGLS (Cross-section SUR)
Date: 05/29/14 Time: 15:19
Sample: 2004 2011
Included observations: 8
Cross-sections included: 8
Total pool (balanced) observations: 64
Linear estimation after one-step weighting matrix
White diagonal standard errors & covariance (no d.f. correction)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.468239	0.079347	31.10681	0.0000
?TFPE	-0.022681	0.001721	-13.17896	0.0000
?SPREAD	0.031993	0.015563	2.055626	0.0443
?CI	0.003174	0.000516	6.153166	0.0000
?BDEV	-0.487387	0.032117	-15.17544	0.0000
?GDPG	-0.002662	0.000449	-5.931023	0.0000

Weighted Statistics

R-squared	0.909531	Mean dependent var	18.74901
Adjusted R-squared	0.901732	S.D. dependent var	214.5048
S.E. of regression	0.893163	Sum squared resid	46.26897
F-statistic	116.6207	Durbin-Watson stat	2.075844
Prob(F-statistic)	0.000000		

Unweighted Statistics

R-squared	0.346332	Mean dependent var	1.375261
Sum squared resid	0.013855	Durbin-Watson stat	1.439761