University of Reading

The Impact of banking reforms on competition and efficiency of Ghana's banking sector

PhD

Economics

John Kwame Dadzie

June 2017

THE IMPACT OF BANKING REFORMS ON COMPETITION AND EFFICIENCY OF GHANA'S BANKING SECTOR

University of Reading

The impact of banking reforms on competition and efficiency of Ghana's banking sector

BY

JOHN KWAME DADZIE

SUPERVISED BY

DR. ALESSANDRA FERRARI

DR. CHRISTIAN ANDI NYGAARD

A THESIS SUBMITTED TO THE UNIVERSITY OF READING IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD) IN ECONOMICS

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Name: John Kwame Dadzie

Date: June 2017

Department of Economics
University of Reading

Acknowledgements

To God be the Glory for the grace and strength given me to complete my PhD studies.

I would like to express my deepest appreciation to my supervisor, Dr. Alessandra Ferrari for the immense and selfless support she has given me from the start to the end of my PhD studies. The transfer of knowledge and skills from her has been enormous; and her patience and professionalism are very much appreciated. I also want to acknowledge the tremendous support I received from my co-supervisor, Dr. Christian Andi Nygaard. Without the guidance and assistance of my supervisors, this thesis would not have been possible. The encouragement and support they offered even before I commenced the programme deserve commendation.

I also want to sincerely thank Dr. Mark Guzman of the Economics Department for the constructive feedback and advice he offered me throughout my studies.

A special gratitude also goes to my wife, *Victoria*, and my lovely daughters, *Ivana, Eleora* and *Odelinda* for the love, kindness and support that they have given me during the programme. I also thank my parents and siblings immensely for their encouragement and support throughout my academic journey.

I wish to also sincerely appreciate the Commonwealth Scholarship Commission (UK) for the scholarship and financial support they provided to enable me pursue the programme.

To all others: relatives, friends, tutors and colleagues who provided assistance in one way or the other, I say ‘God bless you’. I really appreciate your support.

Abstract

African countries are pursuing financial reforms to address inhibitions to competition and efficiency of their banking sectors. This thesis focuses on Ghana which recently implemented deregulation reforms including the introduction of universal banking, the adoption of an open licensing policy to enhance contestability and competition, and the abolition of secondary reserves; and examines the impact of these reforms on banking competition and efficiency. The study uses a comprehensive and unique panel dataset of 25 banks for the period 2000-2014 which captures the pre- and post-reform periods.

The study employs the persistence of profit and Boone indicator models of competition to analyse competitive conditions in the loans market. The empirical results suggest that competition initially increased following the reforms but subsequently declined as a result of macroeconomic weaknesses, in particular high interest rates, which was partly impacted by the indirect effects of the global financial crisis. The study also uses stochastic frontier analysis to examine the efficiency impacts of the reforms, as well as the role of bank ownership and size in influencing efficiency levels. Different deregulation reform indices are constructed using survey data and coding rules from two international databases on banking regulations and reforms, and captured as inefficiency covariates together with ownership and bank size in the one-step Battese-Coelli (1995) model. The findings point to an overall increase in cost efficiency following the reforms although there is non-uniformity in efficiency-impacts from the different policies. Foreign and regional banks are found to be marginally more efficient than private domestic and state-owned banks. Bank size was found to positively impact cost efficiency while the global financial crisis had an adverse impact on efficiency.

The policy implications are that for African countries to benefit from financial deregulation reforms, there is the need for the reforms to be anchored on strong macroeconomic fundamentals, institutional initiatives which support these reforms, strong credit environments and appropriate sequencing of reforms.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AfDB	AFRICAN DEVELOPMENT BANK
BI	BOONE INDICATOR
BOG	BANK OF GHANA
CFCFGH	CAUDILL & FORD (1993), CAUDILL, FORD, & GROPPER (1995) AND HADRI (1999)
DEA	DATA ENVELOPMENT ANALYSIS
POP	PERSISTENCE OF PROFIT
SSA	SUB-SAHARA AFRICA
CAR	CAPITAL ADEQUACY RATIO
CEE	CENTRAL AND EASTERN EUROPE
CR_N	*N*-FIRM CONCENTRATION RATIO
CV	CONJECTURAL VARIATIONS
DEPN	CURRENCY DEPRECIATION
DFA	DISTRIBUTION FREE APPROACH
DMU	DECISION-MAKING UNIT
ERP	ECONOMIC RECOVERY PROGRAMME
ESH	EFFICIENT STRUCTURE HYPOTHESIS
FINSAP	FINANCIAL SECTOR ADJUSTMENT PROGRAM
FINSSIP	FINANCIAL SECTOR STRATEGIC PLAN
FISCAL	FISCAL DEFICIT TO GDP RATIO
GDP	GROSS DOMESTIC PRODUCT
GHC	GHANA CEDIS
HHI	HERFINDAHL- HIRSCHMAN INDEX
IMF	INTERNATIONAL MONETARY FUND

IO	INDUSTRIAL ORGANISATION
KGMHLBC	KUMBHAKAR, GHOSH, & MCGUCKIN (1991), HUANG & LIU (1994) AND BATTESE & COELLI (1995) MODELS
LI	LERNER INDEX
LOC	LOAN OVERCHARGE
MC	MARGINAL COST
ML	MAXIMUM LIKELIHOOD
MPC	MONETARY POLICY COMMITTEE
MPR	MONETARY POLICY RATE
MS	MARKET SHARE
NBFI	NON-BANK FINANCIAL INSTITUTIONS
NEIO	NEW EMPIRICAL INDUSTRIAL ORGANISATION
NPL_S	NON-PERFORMING LOANS
OBS	OFF-BALANCE SHEET ACTIVITIES
OEA	OTHER EARNING ASSETS
P-R	PANZAR-ROSSE H-STATISTIC
RGDPG	REAL GDP GROWTH
SCP	STRUCTURE-CONDUCT-PERFORMANCE
SEE	SOUTH EASTERN EUROPE
SFA	STOCHASTIC FRONTIER ANALYSIS
TBR	TREASURY BILL RATE
TFA	THICK FRONTIER APPROACH
WTO	WORLD TRADE ORGANISATION

CHAPTER 1 INTRODUCTION

1.1 Background and motivation of the study

Banking in Africa has been undergoing gradual but notable changes since the turn of the 1990s. These changes have been triggered by the initiation of financial policy reforms in most African banking markets with the view to removing the numerous challenges affecting the banking sector's competition, efficiency, growth and development. The dominant role of the state in the financial sector was emphasized by most African countries in their economic development agenda following the attainment of political independence during the 1960s to the 1980s. Most state banks were set up with some restrictions placed on the entry of foreign banks and private domestic banks. Restrictive policies such as interest rate control; subsidised lending rates to priority sectors; directed credit allocation to sectors deemed crucial to the economy; and huge reserve requirements were also implemented.

A first wave of financial reforms was accordingly ushered in Africa under the auspices of the International Monetary Fund (IMF) and the World Bank during the late 1980s to 1990s as part of broader structural adjustment programs. These reforms, under the Financial Sector Adjustment Program (FINSAP), led to a gradual liberalisation in interest rates and exchange rates; abolition of directed lending; clean-up of non-performing loans of banks; recapitalisation or closure of insolvent banks; and the development of other components of the financial sector, including the capital markets, insurance and other non-bank financial institutions. The reforms however seem to have had a limited impact on enhancing banking efficiency and competition due to many lingering challenges that constrain effective competition and efficient intermediation (Beck and Cull, 2014).

Competition and efficiency in banking have been identified as key drivers of financial and economic development through their expected impact on lowering the cost of financial intermediation, improving access to banking services, facilitating technological progress and innovation, and contributing to the growth of the wider economy (Allen and Gale, 2000); (Claessens, 2009). More specifically, competition stimulates efficiency among banks as the competing banks innovate and improve on their product offerings, distribution channels and service delivery. These efficiency gains translate into lower costs of financial services and therefore have the potential to increase access to financial services by individuals and firms.

This helps deepen financial intermediation, facilitate optimal resource allocation and contribute to economic growth. Competition in banking also fosters industrial growth as more competitive banks are able to offer financing on better terms to firms (Claessens and Laeven, 2004). An efficient banking system facilitates financial intermediation and contributes to optimal financial resource allocation needed to support the real sector for economic growth.

Further, due to the role of banks as the primary conduit of monetary policy, a competitive and efficient banking system makes banks interest rates more responsive to changes in policy rates, implying that the transmission mechanism of monetary policy is more effective in a competitive banking system and in this way contributes to enhancing stability of financial systems (van Leuvensteijn, Sørensen, Bikker, and van Rixtel, 2013; Brissimis, Delis, and Papanikolaou, 2008). Fostering banking competition and efficiency accordingly remain key policy objectives of all bank regulators and supervisors, hence the pursuit of policy reforms built around these aims (Berger and Humphrey, 1997).

In the context of Africa's financial services sector, banking competition and efficiency assume greater significance due not only to the dominance of the banking sector, but also to its relatively low development even by developing economies' standards. Like most developing economies, Africa's banking sector plays a more important role in financial resource allocation due to the fact that other aspects of the financial sector are underdeveloped. The low level of development of the continent's banking sector is attributed to the dominant state ownership of banks, financial repressive policies pursued during 1960s–1980s, weak financial institutional infrastructure, macroeconomic instability, and the slow pace of reforms which commenced in the late 1980s and early 1990s and contributed to the continent's economic development (Allen, Otchere, and Senbet, 2011; Beck and Cull, 2014).

It has been rightly suggested that in order to enhance the competitiveness of the financial services sector and improve financial development, African countries must undertake some further reforms to address inhibitions to competition and efficiency of the banking sector so as to foster the continent's overall economic development (Kasekende, Mlambo, Murinde, and Zhao, 2009). It is against this background that Ghana recently pursued significant reforms under its Financial Sector Strategic Plan (FINSSIP) during the early 2000s, aimed at addressing lingering challenges to competition and efficiency in the banking industry, and the development of the financial sector (Acquah, 2006).[1]

[1] Dr. Paul Acquah was Ghana's Central Bank Governor during 2001–2009, and under whose governorship those reforms were implemented.

The deregulation policies undertaken during 2003–2006 included: (i) the introduction of universal banking in 2003 to remove product and geographical restrictions on banking activities; (ii) the adoption, in 2006, of an open licensing policy to enhance contestability and competition through the licensing of new banks; and (iii) the abolition of huge secondary reserve requirements in 2006 which hitherto compelled banks to hold 35% of deposits in government securities, which constrained financial intermediation. To minimise potential risk-taking behaviour usually associated with such deregulation policies, the Bank of Ghana[2] complemented these policy reforms with prudent stability-enhancing policies such as increasing minimum capital requirements of banks from $2 million to $18 million during 2009-2012 and raising capital adequacy ratio from 6% to 10%. Further, to help strengthen the credit environment, a Credit Reporting Act was passed in 2007 resulting in the licensing of three credit reference bureaus by the end of 2012.

The structural transformation of the banking sector following these reforms was quite substantial, and included the licensing of new banks; a rapid growth in branch network; new product offerings such as electronic banking, assets-backed lending, and mortgage financing. The entry of new banks has resulted in a sharp decline in banking concentration levels. A notable feature worth noting is that most of the new banks are pan-African regional banking groups, while the dominance of state banks has significantly been curtailed. The growth in the number of banks and branches, the consequent reduction in concentration levels, the creation of a unified banking platform, and the diversity of the ownership base has positioned the banking sector in a relatively stronger footing. Empirical research on the impact of these banking reforms on competition and efficiency would therefore serve as a test of the effectiveness of such deregulation policy initiatives and as a guide to policy makers in Africa. It is against this background that this thesis is undertaken, as discussed in the next section.

1.2 Aims of the study

In spite of the structural changes that followed the recent banking reforms in Ghana, there are key issues that regulators, academics and bankers alike are still grappling with and that have not yet been addressed. These relate directly to questions regarding the impact of these reforms on banking competition and efficiency, which were the primary objectives of the reforms.

[2] Official name of Ghana's Central Bank

There are very few empirical studies on banking competition and efficiency in Africa in general, and specifically on the effectiveness of reforms aimed at improving them. Studies on Ghana are even fewer. The underlying reasons for these include the non-availability of quality data on banks to carry out such empirical work, the fact that financial reforms are a recent phenomenon in most African countries including Ghana, coupled with the slow pace of the reforms in African countries. Earlier empirical work on Ghana shows that banking reforms pursued during the late 1980s and early 1990s under FINSAP had a limited impact on competition; although they had a positive impact on banks' operational efficiency, as shown by an improvement in various financial ratios that proxied it (Antwi-Asare and Addison, 2000). More recent studies on banking competition (Buchs and Mathisen, 2005; Biekpe, 2011) also suggest that Ghana's banking sector is not competitive. As detailed in Chapter 2, these two studies, which analyse the periods 1998–2002 and 2000-2007 respectively, conclude that Ghana's banking system exhibited a non-competitive market structure which hampers financial intermediation. Both studies identify the highly concentrated banking market structure, fragmented banking system, huge mandatory reserve requirements (used to accommodate government fiscal deficits) and the low penetration of new banks as the key factors which constrained banking competition. However in each study, the study period does not fully cover the recent post-reform period, thus preventing a meaningful assessment of the effects of these recent reforms. Indeed at a first glance the recent deregulation reforms would seem to have addressed the competition-constraining challenges found by the literature: concentration levels have significantly declined following the entry of new banks; the introduction of universal banking seems to have given banks a level playing field which is expected to facilitate competition; and the removal of the mandatory secondary reserves also provides opportunities for increased competition, as it releases funds for intermediation.

As of efficiency, Isshaq and Bokpin (2012) analyse only the branch expansion outcome of the reforms on banking efficiency, while Adjei-Frimpong, Gan, and Hu (2014) do not explicitly examine the impact of deregulation. Further, in spite of the ownership dynamics, no study exists that investigates how the significant ownership changes are influencing efficiency.

A detailed review of the literature on Ghana's banking reforms is provided in Chapters 3 and 4. From the above though it can be anticipated that this literature is extremely limited and its theoretical predictions are mixed and inconclusive. While earlier studies opined an uncompetitive banking sector in Ghana, the recent reforms might have reduced the un-competitiveness and enhanced the efficiency of the sector. Second, due to the lagged effects of reforms, a relatively longer study period is required to fully assess their impact. Third, as

explained more in detail in Chapters 6 and 7, more reliable and credible measurement tools will also have to be used to guarantee the robustness of results. Fourth, the dynamic ownership changes require an assessment of how ownership status affects efficiency. Fifth, it is unclear whether attempts to induce consolidation through higher capital increases will positively impact efficiency as the relationship between bank size and efficiency is unclear.

These are the identified gaps which this thesis seeks to fill. To date, no comprehensive study has, to the best of our knowledge, been carried out which looks specifically at these recent deregulation reforms and their impact on both competition and efficiency of Ghana's banking sector, and over a period of time long enough to meaningfully assess such impact. We expect this to be one of the few studies on Africa that comprehensively analyses the effects of deregulation, ownership and size on firm-efficiency and market competition, based on a bank-level dataset spanning 2000–2014, covering both the pre- and post-reform periods, and using reliable models that have been applied in the context of developed and other developing countries. The study also assesses the extent to which the global financial crisis impacted on banking competition and efficiency in Ghana.

Ghana makes a useful case study for the evaluation of banking sector reforms in a lower middle-income country to draw lessons for other African countries. The choice of Ghana is also premised on the fact that it is a stable democracy in the African context which facilitates the implementation and therefore the analysis of reforms. Banking reforms have been comprehensive and systematically pursued and appear to have resulted in a unified banking system, well-diversified by ownership class, comprising of state banks, private domestic banks, pure foreign banks, and regional pan-African banks. This makes Ghana a good representative of other African countries of similar layout as well as serving as a model for poorer, less developed African countries that are embarking onto reforms to enhance financial and economic development.

The specific research questions that this thesis seeks to answer are the following:

- What has been the impact of the deregulation reforms on the competitiveness of the banking sector in Ghana; that is, how has competition evolved during the period before and after the implementation of these reforms and to what extent is the observed pattern attributable to them?

- Has banking efficiency been impacted by the deregulation reforms?

- What role does bank ownership play in efficiency? A notable feature of the outcome of the reforms is the high entry of foreign banks, predominantly pan-African (regional) banking groups.[3] The banking ownership-efficiency relationship is an ongoing debate, and the significant changes in the ownership structure make it imperative for a study of their relative efficiencies. Due to the preponderance of these pan-African banks in Ghana, and their current penetration into other African markets, we classify them separately as regional banks and distinguish them from other foreign banks. This could serve as an important policy guide to the class of banks to be attracted to domestic banking sectors in Africa.

- How does bank size influence banking efficiency? The sharp increase in capital requirements of banks announced in 2008 was seen by some analysts as a way of driving consolidation (PricewaterhouseCoopers, 2010).[4] Although this did not occur as all banks were able to meet the new high capital requirements, it nevertheless raises the issue of how bank size affects efficiency.

These are pertinent policy-related questions of relevance to Ghana and other African countries pursuing or contemplating the pursuit of similar deregulation reforms.

1.3 Research methodology

We employ different methodologies to address our research questions. In investigating the impact of the reforms on competition in the banking sector, we use both the persistence of profit (POP) and the Boone indicator (BI) models to measure competition in the loans market. We focus on the loans market because it is the most important segment of the banking system in line with banks' primary role of financial intermediation, and also because the reforms were specifically targeted to address the challenges in this market. The two models are selected because they overcome the various limitations of the widely used Panzar-Rosse model (in African literature), as discussed in detail in Chapter 3. Our chosen models can

[3] Pan-African banking groups are defined as banks whose parent company's headquarters are in African countries. Notable among the new banks are United Bank of Africa (UBA), Access Bank, Guaranty Trust Bank and Zenith Bank, all based in Nigeria; and. Bank of Africa which is based in Mali. The main existing regional banks prior to the reforms are Ecobank and Stanbic Bank whose headquarters are based in Togo and South Africa respectively.

[4] This is against the backdrop that Nigeria pursued a similar policy during the mid-2000s which resulted in massive consolidations, scaling down the number of banks from about 89 to 25.

reliably measure competition levels over time and therefore facilitate the analysis of the impact of policy reforms on competition. The use of the two models also enhances the robustness of the empirical results.

The POP model analyses the persistence of the loan overcharge over time.[5] A higher persistence of the loan overcharge indicates a lower level of competition, while a lower persistence reflects increasing competition. Where competition in the loans market intensifies in a particular year, banks will reduce the loan overcharge in order to win new customers or at least retain existing ones. Hence the persistence of profits will reduce. Following Zhao, Casu, and Ferrari (2010), we formulate the POP model as a partial adjustment model of loan price overcharge, interacted with a policy reform dummy variable to measure both the pre- and post- reform persistence parameters in the loans market. We also account for macroeconomic and industry-specific variables as well as the global financial crisis in the model.

To check the robustness of the results, we also use the relatively new BI to establish the relationship between bank performance and efficiency, to analyse the evolution of competition in the loans market. The BI is receiving growing attention in banking sector studies due to its desirable properties, one of which is the ability to measure competition on an annual basis so as to track the evolution of competition over time (Delis, 2012; Schaeck and Cihák, 2014), as discussed in Chapter 6. It measures the competitive environment by examining the extent to which relatively more efficient firms are able to gain market share or increase profits at the expense of less efficient firms – the so-called 'reallocation process' (Boone, 2008b). The higher the intensity of competition, the greater the reallocation of market shares from inefficient firms to more efficient ones. The model accordingly examines the relationship between performance and efficiency to infer the competitive environment.

On efficiency, we use the stochastic cost frontier model to measure cost efficiency levels of banks. As explained in detail in Chapter 5, cost efficiency is the most extensively applied concept of economic efficiency in empirical banking studies. This is due to the fact that cost functions are a preferred means of representing production technologies in a multiple-input and multiple-output production context with an economic optimisation assumption embedded. Also, cost minimisation is a necessary and sufficient condition not only for short term profitability, but also for long term survival of banks and represents an important measure of efficiency (Koetter and Meesters, 2013). Cost efficiency measures how close a bank's actual

[5] The loan overcharge, as explained in detail in Chapters 3 and 6, is defined as the ratio of the loan price to the marginal cost of loans, and represents the margin of profitability in the loans market after accounting for credit risk.

cost is to what a best-practice bank's cost would be to produce the same output bundle (Berger and Mester, 1997). The rationale for the choice of stochastic frontier analysis (SFA) as an estimation technique is that it has a stronger appeal than its popular non-parametric alternative of Data Envelopment Analysis (DEA) as it allows for statistical noise and enables the conduct of statistical tests of hypotheses on both the production structure and the degree of inefficiency. SFA is also a powerful tool for quantifying effects of policies, such as deregulation on efficiency. The SFA model used is the Battese and Coelli (1995) model of time-varying inefficiency, which combines the estimation of the stochastic cost frontier with possible covariates of inefficiency in one-step using maximum likelihood within a panel data context. This facilitates the effective examination of deregulation, ownership and bank size on banking efficiency, and overcomes the econometric challenges associated with the two-step approach.

Deregulation is captured using deregulation indices or liberalisation scores for Ghana derived from the World Bank's internationally recognised cross-country surveys on banking regulation by Barth, Caprio and Levine (2001, 2003, 2007 and 2012), and the IMF's database on financial reforms by Abiad, Detragiache, and Tressel (2010). This approach facilitates the examination of the impact of individual reform policies on efficiency. Ownership dummies, total assets as a proxy for bank size, and a financial crisis dummy variable are also included. Banking industry average efficiency trends over the sample period are analysed. Similarly, efficiency levels per ownership class of banks and by bank sizes are also discussed.

The thesis uses an unbalanced panel dataset of 25 banks for the period 2000-2014. The bank-level data was compiled from audited financial statements of all the mainstream banks in operation in each year during the sample period. The data sources for the bank-data were the banks' published annual reports and compilations by Ecobank Research and the Ghana Bankers Association. Upon cleaning the data by accounting for missing values, mergers and acquisitions, we have a total of 321 observations. Macroeconomic data was compiled from the Bank of Ghana annual reports and the World Bank's world development indicators database, while deregulation indices were compiled based on work by Barth et al. (2001, 2003, 2007, 2012) and Abiad et al. (2010).

1.4 Contributions to the literature

The thesis makes several contributions to the existing literature and enhances the understanding of the role of banking reforms in facilitating a competitive and efficient banking system. We summarise them as follows.

- Most existing studies on banking reforms, ownership, size, competition and efficiency have considered only aspects of these inter-relationships. For instance, studies have examined competition and efficiency without explicit analysis of the role of reforms, or the relationship between reforms and competition without analysing efficiency. This study comprehensively and explicitly analyses the impact of reforms on both competition and efficiency, and the role of ownership and size on efficiency based on a consistent theoretical framework.

- Previous studies have also used aggregate banking industry data and/or data for relatively short periods due to challenges in obtaining credible bank-level data over long periods on African countries. Such data limitations largely account for the low empirical study on African banking markets, and also sometimes affect the reliability of empirical results and policy recommendations from the few studies undertaken. We make a useful contribution by using a rich and unique panel dataset of 25 banks for a sufficiently long period (2000-2014) which captures both the pre- and post-reform periods. This comprehensive bank-level data enriches the empirical analysis of bank-level efficiency which facilitates efficiency analysis of banks with different ownership types and bank sizes. Secondly, the relatively long data period enriches the estimation and analysis of both efficiency and competition over time and provides a more effective way of assessing policy impacts. Thirdly, the dataset enhances the reliability of the results and the associated quality of the policy recommendations arising therefrom.
- Another important contribution we make is our use of deregulation indices or liberalisation scores to capture the potential different efficiency impacts of each deregulation policy introduced in Ghana using the internationally recognised survey data and database on banking regulation and reforms by Barth, Caprio, and Levine (2001, 2003, 2007, 2012) and Abiad et al. (2010). This very informative and detailed approach has, to our knowledge, not been applied in the African context. By treating deregulation as a general category, previous studies have assumed a uniform impact of

different deregulation policies which might be erroneous since different policies might have different impacts (Barth et al., 2013)

- We focus on the loans market, which is the most important yet uncompetitive segment of Ghana's banking sector, and the main target of the deregulation reforms. We also adopt two dynamic competition models - the POP and BI measures. The estimation of competition dynamics using the BI, to our knowledge has not been used in any country-specific African study.[6] Both measures are chosen for reasons of completeness (the models allow for a different analysis of the dynamics of competition) and robustness (competition models can often lead to conflicting results).

- The estimation of banking efficiency using stochastic frontier analysis, and based on the Battese and Coelli (1995) time-varying inefficiency model, which has only been used in a very few studies in its panel set up, contributes to methodological issues on efficiency estimation in the African banking literature. Most African banking studies have applied DEA to measure efficiency, which has serious limitations, or the two-stage SFA which suffers from serious econometric drawbacks.

- The distinct classification of regional banks in the analysis of ownership status on banking efficiency is innovative and makes an important contribution as the study of this class of banks has not been explored in Africa, to the best of our knowledge. The study therefore provides policy insights as to the type of banks that could foster higher efficiency in African's emerging banking sector.

In summary, the main contribution of the thesis will be to emphasize that deregulation reforms could be important in enhancing banking efficiency, but different policies might have different impacts on efficiency levels. In addition, while deregulation is important in stimulating a competitive banking environment, such reforms must be anchored on sound macroeconomic fundamentals and institutional initiatives which are supportive of the reforms if the full benefits of the reforms are to be exploited. Since most African countries are fragile and open-economy based and exposed to external shocks by being commodity-dependent, banking reforms should be consistently evaluated and appropriate changes be made to accommodate weaknesses in macroeconomic fundamentals and/or external shocks. The thesis

[6] Amidu and Wilson (2014) applied the Boone indicator in a cross-country African study.

also recognises the important contribution that foreign and regional banks play in enhancing efficiency through technological spill-over effects they bring to domestic banking markets.

1.5 Structure of the thesis

The rest of the thesis, besides this introductory chapter, is organised into seven chapters. To put the study into perspective, Chapter 2 provides a country profile of Ghana and describes the evolution of its financial sector. It details politico-economic developments which have shaped financial sector policy, as well as the evolution of the banking sector since the country's independence in 1957. It then discusses the recent banking reforms implemented and the structural changes arising from them. Chapter 3 provides a literature review on banking competition as well as a description of the models proposed for measuring banking competition. In particular, it discusses the pros and cons of competition as it relates to its impact on stability, efficiency and financial access. It also reviews empirical work on the impact of deregulation reforms on banking competition. The models of banking competition measurement are also discussed. Chapter 4 reviews relevant literature on the relationship between deregulation and banking efficiency, and the role of bank ownership and bank size on banking efficiency. Chapter 5 presents the analytical foundations and measurement of frontier efficiency. The concepts of efficiency and production theory and cost theory are briefly discussed. The chapter also reviews the main parametric and non-parametric models of measuring efficiency, DEA and SFA. Chapter 6 presents the empirical analysis of the impact of banking reforms on banking competition using both the POP model and the Boone model of competition. Chapter 7 examines the empirical work on the impact of deregulation on efficiency, and the role of ownership and bank size in influencing efficiency in Ghana's banking sector. The methodology adopted, empirical results and analysis are discussed therein. Chapter 8 summarises the main findings of the thesis and draws on general conclusions and policy implications. It also points out some limitations of the study and offers suggestions for future research.

CHAPTER 2 COUNTRY PROFILE, FINANCIAL SECTOR DEVELOPMENTS AND BANKING REFORMS

2.1 Introduction

This chapter discusses the evolution of Ghana's financial sector and its banking reforms in the context of the politico-economic developments of the country from independence in 1957 to 2014. It examines the various banking reforms that have been implemented in the country. It finally analyses the existing empirical studies on banking reforms, competition and efficiency. The chapter is organized as follows. Section 2.2 provides a brief profile of the country and an overview of its recent economic performance over the study period of 2000–2014. Section 2.3 examines in detail major politico-economic developments during 1960–2014 which shaped the evolution of the financial system. Section 2.4 elaborates on the development of the country's financial sector and banking industry during the period 1960-2014, including earlier financial reforms undertaken and the recent deregulation reforms in detail. The effects of the recent reforms on the structure of the banking sector are discussed in Section 2.5. Section 2.6 concludes.

2.2 Country profile and overview of recent economic performance

Ghana is a relatively stable and peaceful country within the West African region. The first country in Sub-Sahara Africa (SSA) to gain political independence (from Great Britain) in March 1957, Ghana has made considerable progress in its political and economic development over the last three decades compared to the challenging years of political instability and economic policy inconsistencies that characterized its initial years after independence. With Gross Domestic Product (GDP) of US$113 billion in 2015[7], Ghana ranks as the second largest economy in West Africa after Nigeria, and Africa's twelve largest economy (African Economic Outlook, 2016).[8] Ghana has achieved consistent year-on-year economic growth, with an average yearly real GDP growth of 7.9% during 2009-2014 surpassing SSA's average growth rate of 5% over the same period. On account of its per

[7] GDP based on Purchasing Power Parity (PPP) as at 2015.Population of about 27.4 million; and a land area of 239,000 km^2

[8] Based on GDP at PPP as at 2015, the 12 largest economies are Nigeria, Egypt, South Africa, Algeria, Morocco, Angola, Sudan, Ethiopia, Kenya, Tanzania, Tunisia, and Ghana.

capita GDP of US$1,480 in 2015, Ghana is classified as a lower middle income economy as per the World Bank's country classifications (World Bank, 2016).[9]

The country's economy has traditionally depended on two key sectors – the agriculture sector (Ghana is the world's second largest producer of cocoa) and the mining industry (principally gold). The services sector has, in recent times, overtaken the agriculture sector thanks to the growth in the financial services, information and communication sub-sectors. The country's economic fortunes have also been boosted by the discovery of oil deposits since 2007 and subsequent production of oil in commercial quantities since 2011. Ghana is therefore in an enviable league of a few commodity-dependent countries that are endowed with all the three principal commodities of interest to African trade – agriculture, oil and minerals, and thus expectations for accelerated growth and structural transformation of the economy are high.

Table 2.1 shows a snapshot of key economic indicators for Ghana for 2000-2014, which are commented thereafter.

Table 2.1 Ghana: Key Economic Indicators, 2000 - 2014

	2000	**2002**	**2004**	**2006**	**2008**	**2009**	**2010**	**2012**	**2014**
Real GDP Growth (%)	3.7	4.5	5.6	6.2	7.3	4.0	8.0	9.3	4.0
Inflation (Year-End)	40.5	15.2	11.8	10.5	18.1	15.9	8.6	8.8	17
BoG Prime Rate (%)	n/a	24.5	18.5	12.5	17	18	13.5	15	21
91-day Treasury Bill (%)	42	26.6	17.1	10.2	24.7	23.7	12.3	22.9	25.8
Exchange Rate (GHC/US$)	0.70	0.84	0.91	0.92	1.21	1.43	1.47	1.88	3.2
Fiscal Balance (% of GDP)	-9.7	-6.7	-3.7	-7.55	-11.48	-5.6	-6.8	-11.5	-10.2
Balance of Payments (US$ Million)	-117	40	-11	415	-941	1,159	1,463	-669	-85
Months of Imports Cover	0.8	2.2	3.8	3	1.8	2.9	3.7	3.1	3.8

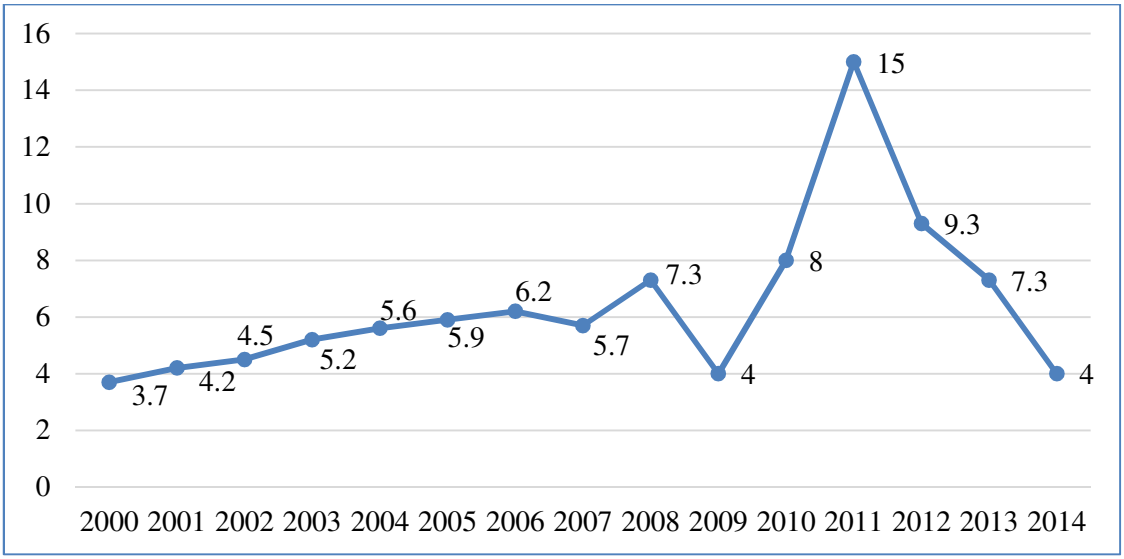
***Source*:** Bank of Ghana Annual Report (various issues)

As Table 2.1 and Figure 2.1 show, Ghana experienced consistent economic growth during 2000-2008, with annual real GDP growth rate increasing steadily from 3.7% in 2000 to 7.3% in 2008. This was anchored on a relatively strong macro-economy, sound fiscal and monetary

[9] GDP per capita based on World Bank's Atlas method. Details at https://blogs.worldbank.org/opendata/new-country-classifications-2016

policies, and strong growth in agriculture and mining production which were supported by stable prices of its major export commodities - cocoa and gold - during that period. Economic growth declined to 4% in 2009 due to the adverse impact of the global economic crisis on export commodity prices, the decline in remittances, tourism receipts and other capital flows into the country, coupled with high crude oil import prices. Domestic economic challenges such as a severe energy crisis and huge fiscal deficits also compounded its external balance of payments deficits to constrain economic growth in 2009.

Figure 2.1 Real GDP growth rate (%), 2000–2014

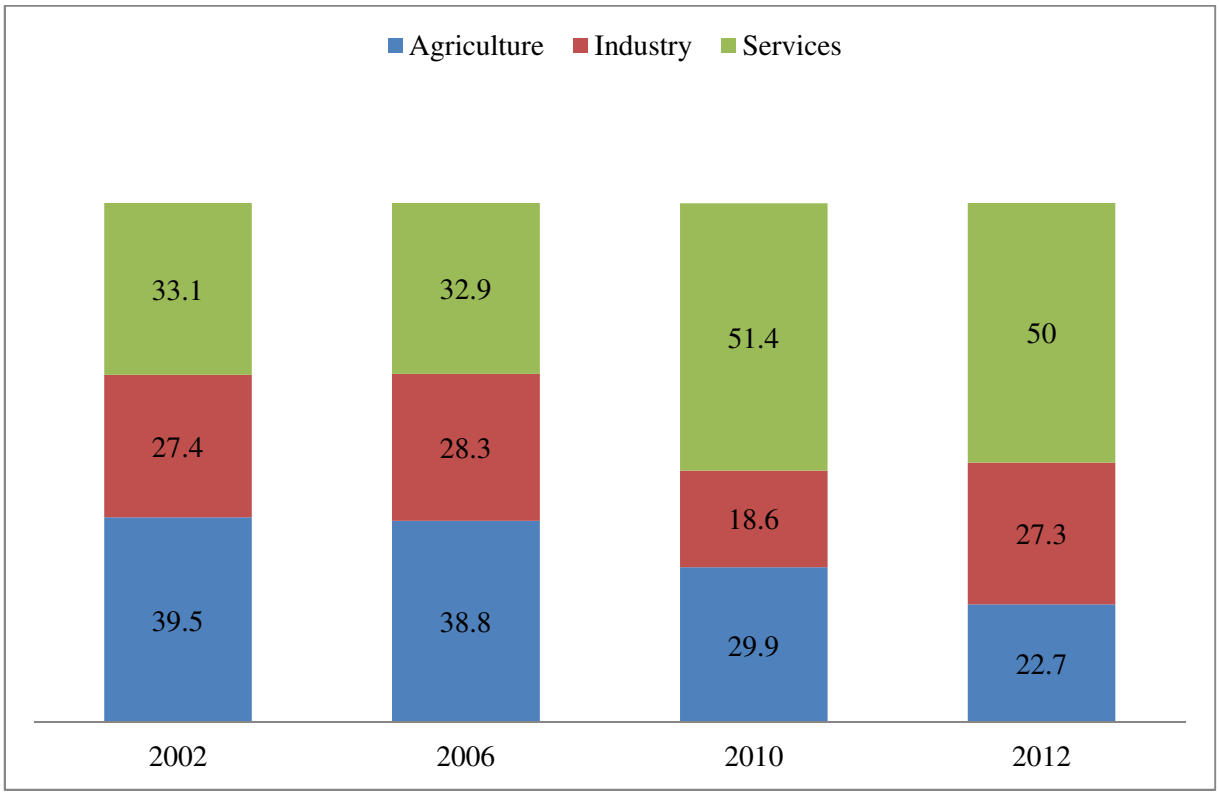


Source: Bank of Ghana, *Annual Reports* (various issues)

Economic growth rebounded strongly thereafter and the sharp rise seen in 2011 was driven by the oil sector when the country commenced full production of crude oil in commercial quantities in 2011, following discovery of oil deposits in 2007.

In terms of sector contribution to GDP, the agricultural sector had traditionally been the major contributor to the country's GDP. In the recent past, though, it has been overtaken by the services sector as the major contributor (Figure 2.2). The services sector accounted for 50% of total GDP on average during 2010-2012, with agriculture and industry contributing respective levels of 26% and 24%. The strong growth in the services sector in recent years has been spurred by strong growth in the information and communication and the financial services sub-sectors (Ghana Statistical Services, 2014).

Figure 2.2 % Shares of Economic Sectors in GDP, Selected Years

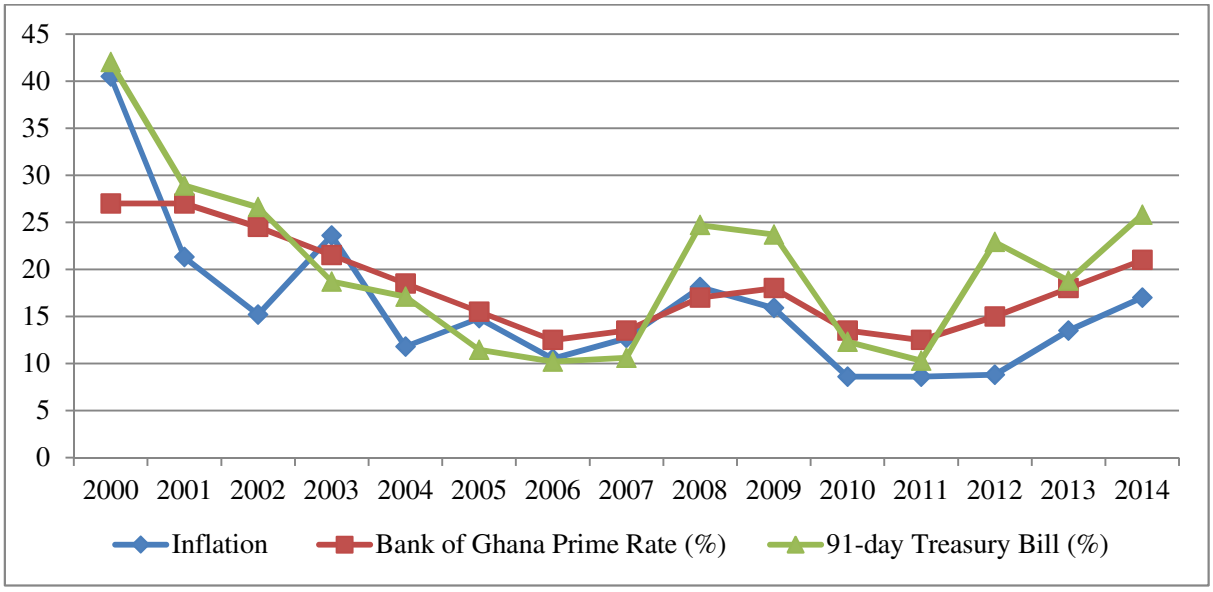


Source: Ghana Statistical Service, *National Accounts Statistics* (various issues)

On the macroeconomic front, Ghana's economic policy objectives during 2000-2014 have been to achieve and sustain macroeconomic stability and fiscal discipline so as to support the development of the private sector in the economic growth agenda of the country. These objectives have however not been achieved on a sustainable basis as the economy is usually plagued by high inflation and interest rates, and steep depreciation of the local currency against major international currencies. These unstable macroeconomic episodes have usually occurred during national elections years, such as in 2008 and 2012, principally due to huge fiscal deficits from expenditure overruns by governments to retain power. As shown in Table 2.1, the budget deficit worsened in both 2008 and 2012, exceeding 10% of GDP. Such huge fiscal deficits are financed by domestic borrowing in the banking sector and this leads to high interest rates, crowding out private sector credit. Further, as a commodities-dependent country, occasional worsening terms of trade and reductions in capital inflows and remittances especially during election years also create balance of payments challenges for the country. Thus, the country also recorded severe balance of payments deficits in 2008 and 2012. The twin deficits in the fiscal and external positions accordingly impact adversely on

the macroeconomic front via high inflation and interest rates and steep local currency depreciation against major international currencies.

Figure 2.3 Inflation, BoG Prime rate and Treasury bill rates, 2000 – 2014



As shown in Figure 2.3, inflation broadly declined during 2000–2007 (except for a temporary increase in 2003 resulting from domestic fuel price hikes early that year) due to a combination of prudent fiscal and monetary policies, the introduction of an automatic domestic fuel price adjustment mechanism, as well as significant foreign exchange savings made as a result of external debt cancellation enjoyed under the HIPC initiative.[10] Inflation was thereafter relatively stable within the 10%-15% range until 2008 when a combination of high crude oil prices, food crisis and election-related spending caused it to shoot up. Inflation eased and remained within single digit since 2010 but increased in 2013–2014.

Interest rates developments have moved in tandem with the trend in inflation. The Bank of Ghana prime rate witnessed a consistent reduction during 2000-2007 after which it was raised in 2008-2009 to curb the rising inflation. Following a temporary decline in 2010-2011, the prime rate has increased since 2012. Other market rates have followed that trend. Interest rates on government securities as well followed a similar pattern prior to 2008. Due to the huge fiscal deficits in 2008 and 2012, and the resulting increased domestic borrowing by government to finance the deficits, interest rates on government securities increased significantly in 2008 and 2012.

[10] Ghana signed up to the Highly Indebted Poor Country (HIPC) initiative of the IMF/World Bank which resulted in significant debt cancellation once key macroeconomic and other policy reform targets were achieved.

The review shows that economic growth performance has been strong during the past few years but the surge in fiscal deficits and the resulting huge government borrowing have created inflationary pressures and raised interest rates which crowd out private sector credit. It is worth noting that the current economic growth and macroeconomic performance are a significant improvement over the previous decades as will be seen shortly. Political stability, democratic dividends and policy consistency have also played a key role in sustaining the current economic performance.

We now review politico-economic developments since independence to provide a basis of understanding how financial sector development in the country has evolved.

2.3 Politico-economic developments in Ghana: 1960 – 2014

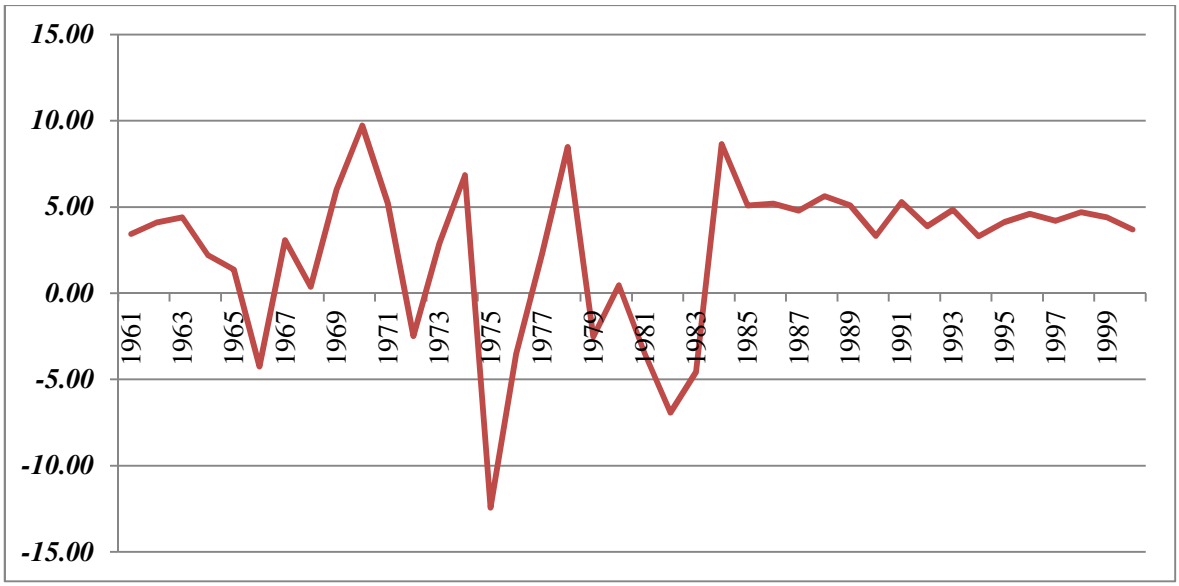
Ghana's politico-economic history provides a good basis to understand how the financial system in general, and the banking sector in particular, have evolved. The review of recent economic developments shows a relatively strong performance. History, however, shows that the country experienced swings in its economic growth performance for almost two decades since the post-independence political process was interrupted by a coup d'état in 1966. It has nonetheless managed to achieve a consistent economic growth pattern over the last three decades, helped by political stability, pro-market economic reforms, broad base policy consistency, a growing services sector, and in particular a financial sector which consistently contributes to economic growth. Major economic and financial reforms have contributed to this stabilization in economic growth.

The review of the historical politico-economic developments is intended to show how these developments shaped economic policies pursued and the evolution and development of the financial system and banking sector. The review does not attempt to assess or critique the economic philosophies and policies pursued by different governments over the years, nor delve into details of the underlying causes of economic performance over those years. Several papers and research work have been done on these fronts (see, for example, Aryeetey and Kanbur (2006) for a detailed discussion), and any attempt to focus on these issues could derail the essence of the review, which is fundamentally to help understand how the banking and financial system have evolved over the years. Further, the discussion does not examine in detail policies of each political regime except for major economic and financial policies which had a profound impact on monetary policy, financial system and the banking sector.

For ease of analysis, the review is carried out under three time periods: 1960–1981, 1982–2000, and 2001–2014. The first phase (1960-1981) marks the immediate post-independence era which was unfortunately characterised by political instability and inconsistent economic policies. It was a predominantly state-led economy with strong government involvement in the financial system. The second phase (1982-2000) witnessed political stability albeit under a military regime for most part of the period until multi-party democracy was restored in 1992. It also marks the beginning of economic reforms and liberalisation policies in the financial system. The final phase of 2001–2014 could be described as a period of sustained democracy and political stability, strengthening the macroeconomic environment and further reforms in the financial sector to support private sector participation in the economy.

Figure 2.4 gives a snapshot of real GDP growth over the 1960-2000 period, with sharp volatilities experienced during 1960–1983, reflecting the period of political instability. Thereafter stability in the political environment and the pursuit of economic reforms yielded sustained economic growth of around 5% on average during the second phase. The acceleration in growth during the last phase was aided by deepening democracy and improved macroeconomic environment as discussed earlier.

Figure 2.4 Real GDP growth rate (%), 1960 - 2000



Source: World Bank, *World Development Indicators* Database

2.3.1 Politico-economic developments during 1960-1981

It makes interesting reading that the sharp volatility in real GDP growth during 1960–1981 mirrors the rather chequered political history during that period. This is not to suggest that the political environment was the main driver of economic performance. External factors such as

commodity price shocks, especially oil price hikes and declines in cocoa and/or gold prices severely affected the country's terms of trade and balance of payments position and had a dampening effect on growth prospects. Internally, economic mismanagement and structural weaknesses also contributed to the sluggish performance of the economy. Nonetheless, political instability and economic policy reversals seem to have played a major role in the economic challenges the country faced during that period. As noted by Aryeetey and Kanbur (2006), the periods of negative economic growth occurred mostly in years that saw changes in government and culminated in explosive policy reversals. In this regard, it is noticeable that the negative real GDP growth recorded in 1966, 1972, 1979 and 1981 coincided with years of military takeovers (Aryeetey and Kanbur, 2006).

Following independence in March 1957, Ghana's first president, Dr. Kwame Nkrumah, pursued a socialist economic policy agenda in which the role of the State was paramount. The government made significant investments in developing social and economic infrastructure in the areas of energy, education, health, transport, agriculture and industry. For instance, state farms were set up to boost agricultural production. Industries in garments and textiles, food processing, pharmaceuticals etc. were also set up by the government. In addition, several major infrastructure and construction works were undertaken during that period, including a major hydro-electric power generating plant (Akosombo Dam), a major port (Tema Harbour), an oil refinery (Tema Oil Refinery), Africa's largest aluminium smelter (Volta Aluminium Company), major hospitals and universities were established. The financial system was also controlled by the State as it was used to facilitate financing of the development projects of the government. State banks were set up to support sectors considered as priority to the developmental needs of the country and interest rate controls and credit ceilings were set by the government for the banks.

A military overthrow of the government – the first – in 1966, ushered in a military government, the National Liberation Council (NLC), which was opposed to Nkrumah's socialist economic ideology and sought to pursue a liberalized market agenda. It accordingly abandoned most state projects and disposed of some state-owned enterprises to private investors. There was however little room for the government to make policy changes in the financial system as it was forced to hand over power to a civilian government after three years due to internal and external pressures to restore democracy.

The new civilian government, under Dr. K. A. Busia installed in 1969 also pursued a private sector driven economic strategy, by trying to encourage domestic entrepreneurship and attract

foreign investment. Free market policies were introduced but plans to liberalize the financial sector were yet to be implemented when the civilian government was also overthrown by the military in 1972.

The economic policy paradigm was to be reversed when the military took over in 1972, as the new military government, the National Redemption Council (NRC)[11] was opposed to the Busia administration's free market philosophy, and resorted to reverse the free market economic policies. It re-emphasized the role of the State in economic planning and resource allocation and so sought to indigenize businesses that were privately owned or foreign owned through the passage of various decrees.[12] State intervention in the banking industry continued as new state banks were set up while interest rate controls and credit ceilings continued to be implemented.

The military government was however overthrown by junior officers of the army in 1979, led by Flight Lieutenant Rawlings and the Armed Forces Revolution Council (AFRC). The AFRC handed over to a civilian government, in September 1979 but this civilian government was short–lived as it was also overthrown in December 1981 by Rawlings and set up the Provisional National Defence Council (PNDC) government.

The unstable political environment that characterised the post-independence era to 1981 partly accounted for the sluggish economic performance during the period. Although economic growth performance was initially strong, averaging 4% during 1961-1963, the disruption in economic activities following the military take-over in 1966 led to the –4.2% growth rate in 1966. Economic growth rebounded thereafter with an average growth of 5% during 1967-71. With a negative growth in 1972 following the military takeover, economic growth was restored during 1973-4 but the country moved into a recession in 1975-1976 due to severe terms of trade and economic mismanagement, before picking up in 1977-1978. The military interruptions in 1979 and 1981 saw the economy enter its lowest ebb as it registered negative growth in all the years during 1979-1983 except for a 0.5% growth in 1980 and with sky-rocketing inflation during that period.

As indicated earlier, although the economic growth declines coincided with military takeovers, the political situation was not the sole cause of the poor economic performance. Indeed, in almost all cases, the military takeovers were precipitated by harsh economic

[11] The NRC was in power from January 1972 until October 1975 when it was transformed into the Supreme Military Council (SMC), but under the same leadership.

[12] For instance, an Investment Policy Decree was passed in 1975 that required 51% Ghanaian equity participation in most foreign firms, while the government took over 40% equity stake in specific industries.

conditions prevailing at that time. Other contributing factors of the highly volatile and dismal economic performance during 1960-1983 included external shocks such as high crude oil prices and declining export prices of Ghana's major exports (cocoa and gold) and internal challenges such as structural weaknesses and economic mismanagement (Sowa, 2002).

Table 2.2 Ghana, Selected Economic Indicators, 1960 – 1981

	1960	**1965**	**1966**	**1970**	**1972**	**1975**	**1980**	**1981**
GDP (US$ million)	1,218	2,054	2,127	2,215	2,113	2,800	4,445	4,222
Real GDP Growth (%)	3.4	1.4	-4.2	9.7	-2.5	-12.4	0.5	-3.5
Inflation (annual, %)	n/a	26.4	13.2	3.0	10	29.8	50	116.5
Reserves (US$ million)	278	116	111	42	104	147	330	268

Source: World Bank, World Economic Indicators

The review of the politico-economic developments above shows that the clash of ideologies and policy inconsistencies contributed not only to the sluggish economic growth performance, but also to the non-development of the financial sector. The dominance of the State's involvement in the banking sector via the establishment of state banks, interest rate controls and credit ceilings and other regulatory restrictions to entry were still in place.

2.3.2 Politico-economic developments during 1982-2000

A stable political environment was ushered in during this period albeit under military rule until 1992 when multi-party democracy was restored. The stable political climate enabled the government to focus its attention on undertaking reforms to reverse the declining economic fortunes of the country. The PNDC government launched the Economic Recovery Programme (ERP) in April 1983 under the auspices of the Bretton Woods institutions, the IMF and the Word Bank. The ERP aimed at reversing the protracted period of economic decline by adopting a market-oriented approach and reducing extensive government involvement in the economy. The reforms involved macroeconomic stabilization measures comprising fiscal, monetary and exchange rate policies; liberalization of interest rates; restructuring of the public sector including divestiture of state-owned enterprises and reforms in the financial sector.

The financial sector reform programme was dubbed the Financial Sector Adjustment Program (FINSAP), which is discussed in detail in the next section. The 1990s accordingly saw the emergence of greater private sector involvement in the establishment of banks in the country, while state involvement was gradually scaled back. Thanks to the stable political environment during the PNDC regime, the economic ideology of a market-based private sector-led economy agenda was carried through until multi-party democracy was restored following elections in December 1992. The ERP yielded some gains at least in terms of delivering positive real GDP growth during the period. Ghana's economy has not experienced any negative growth since 1984, with an average rate of 5.2% during 1984-1992 on account of the economic reforms implemented.

Multi-party democracy was restored with general elections held in December 1992. Although the elections were won by Flight Lieutenant Rawlings and his National Democratic Congress (NDC), and was re-elected in December 1996 for a second term, it is worth noting that it marked the first time that an elected government had completed its term of office under a multi-party constitutional regime. The government continued with the liberal market-based economic policies. The return to multi-party democracy however initiated a cycle of huge government spending in elections years which led to significant budget deficits. Macroeconomic performance during 1992-2000 went through this cycle of high fiscal and external sector deficits, accompanied by high inflation and currency depreciation in those election years (IMF, 2011).

Table 2.3 Ghana, Selected Economic Indicators, 1982 – 2000

	1982	**1986**	**1990**	**1994**	**1998**	**2000**
GDP (US$ million)	4,036	5,728	5,889	5,446	7,482	4,983
Real GDP Growth (%)	-6.9	5.2	3.3	3.3	4.7	3.7
Inflation (annual, %)	22.3	24.6	37.2	24.8	14.6	25.2
Foreign Reserves (US$ million)	314	624	309	689	457	309

Source: World Bank, *World Economic Indicators*

2.3.3 Politico-economic developments during 2001-2014

Political-economic developments significantly improved during this period. A smooth and successful transfer of power from one civilian government to another from a different political party was first witnessed in the history of the country following the 2000 elections.

This was a landmark event not only in Ghana but also in the otherwise politically turbulent West Africa region. After serving two terms of four years each, the NDC government handed over power to the New Patriotic Party (NPP) after the 2000 elections. The new government also continued the agenda of a liberal economic system and carried out significant policies and programmes to enhance private sector participation in the economy. Private sector response was strong and the external financing support from the international donor community as a result of 'democracy dividends' was massive (CEPA, 2003). Ghana opted for the Highly Indebted Poor Country (HIPC) initiative of the World Bank /IMF and made substantial gains from external debt forgiveness and macroeconomic restructuring. The financial sector also witnessed significant reforms which impacted on the monetary and financial system. Further reforms were also undertaken in the banking sector. After also serving two terms, the NPP government was voted out of power in 2008, and the successful and peaceful transfer of power back to the NDC enhanced further the country's democratic credentials. The NDC government also continued with market liberalisation policies. It is worth noting that the political stability since 1983 and the economic ideology of a free market economy which had been embraced by successive governments contributed positively in ensuring that necessary pro-market reforms were undertaken.

So how did the country's financial system evolve in the light of the different shades of political environments experienced over the years? What policies were undertaken to develop the financial sector and what was the response of the banking sector? The review of the politico-economic environment above provides a reasonable basis for understanding the financial policies that shaped the evolution of the country's financial sector and banking industry, which is elaborated in the next section.

2.4 Evolution of Ghana's financial sector and banking industry, 1960-2014

The evolution of Ghana's financial system and banking sector reflects deliberate policies pursued in line with the economic ideologies of the different governments discussed above. Different monetary policy regimes and financial sector policies were implemented under each of the three phases to meet the financial development needs of the country, and are discussed hereafter within the context of the three phases identified above.

2.4.1 Financial sector and banking industry developments, 1960-1981

At independence in 1957, Ghana's financial sector consisted mainly of the Central Bank (Bank of Ghana), a wholly-owned government bank (Ghana Commercial Bank), and two foreign banks (Barclays Bank and Standard Chartered Bank). The foreign banks were operating in the country for many years prior to independence with their focus on financing trade between Ghana and the UK and other European countries. Following independence, and in line with the social-oriented developmental agenda of the new government, the government pursued a policy of using the financial sector to support its development policies and programmes. Accordingly, a new Bank of Ghana Act was passed in 1963 that required the Central Bank to operate in consultation with the government. Further a monetary policy system of direct controls was adopted which included interest rate controls, credit ceilings and directed credit (Bawumia, 2010). The Bank of Ghana determined the structure of bank interest rates and specified floors and ceilings for deposits and lending rates, with priority sectors receiving preferential rates. In addition, credit ceilings were imposed to determine the level of maximum financing each sector received while directed credit was implemented to ensure that credit was available to sectors such as agriculture, agro-processing, manufacturing, and housing which were considered priority sectors. There were also restrictions on entry of banks as the government set up state banks to support sectors deemed critical to the economy. For instance, the National Investment Bank (NIB) was established in 1963 to support manufacturing and agro-processing; the Agricultural Development Bank (ADB) to promote the agriculture sector; and the Bank for Housing and Construction (BHC) to support the real estate sector. Other state banks set up during the 1960s and 1970s included the National Savings and Credit Bank (NSCB), the Merchant Bank Ghana and the Social Security Bank (SSB). In fact all the banks established between the 1950s and the late 1980s were wholly or majority-owned owned by the government, either directly with equity or indirectly through public institutions such as the Bank of Ghana, State Insurance Corporation or the Social Security and National Insurance Trust (SSNIT) (Antwi-Asare and Addison, 2000; Bawumia, 2010). The only privately sponsored bank set up during that period was the Bank for Credit and Commerce, which began operations in 1978, while the government also acquired 40% minority shares in the two already established foreign banks following the indigenisation decree enacted in 1975 (Brownbridge and Gockel, 1996). Table 2.4 shows the ownership status of banks in Ghana and the focus of banking activities they were engaged in at the end of the period.

Table 2.4 Structure of Ghana's banking sector, 1981

Bank	Year of establishment	Ownership	Focus of banking activity
1. Standard Chartered Bank (SCB)	1896	Foreign	Commercial
2. Barclays Bank of Ghana (BBG)	1917	Foreign	Commercial
3. Ghana Cooperative Bank (Co-op)	1948	Private domestic	Commercial: Financing of Cocoa Cooperatives
4. Ghana Commercial Bank (GCB)	1957	State	Commercial
5. National Investment Bank (NIB)	1963	State	Development: Industrial sector
6. Agricultural Development Bank (ADB)	1965	State	Development: Agriculture
7. Merchant Bank Ghana (MBG)	1972	State	Merchant Banking
8.Bank for Housing and Construction (BHC)	1973	State	Development: Estate development and Construction
9. National Savings and Credit Bank (NSCB)	1975	State	Commercial: Small-scale businesses
10. Social Security Bank (SSB)	1977	State	Commercial: Consumer Credit
11. Bank for Credit and Commerce (BCC)	1978	Foreign	Merchant Banking

Ownership classification based on entity with majority shareholding (51% or more)

Source: Antwi-Asare and Addison (2000)

In terms of structure, the banking industry at the end of the period consisted of only 11 banks, with 7 state banks, 3 foreign banks and 1 private domestic bank, with no regional bank.[13] Reflecting the fragmented nature of the banking industry, 3 banks (SCB, BBG and GCB) operated as pure commercial (retail) banks while 3 other banks (Co-op, NSCB and SSB) also operated as commercial banks but with specific focus on cocoa cooperatives, small scale businesses, and consumer credit/hire purchase operations. 2 banks (MBG and BCC) operated as merchant banks, while 3 state banks (NIB, ADB and BHC) operated as development banks. Due to the long presence of the two foreign banks that had been operating in the country prior to Independence (SCB and BBG) and the operation of GCB as a pure

[13] Indeed, there was no single bank without state shareholding as the state had shares even in the foreign and private domestic banks.

commercial bank, these banks dominated the banking sector. Available data suggests that the two foreign banks (SCB and BBG) and two state banks (GCB and SSB) had a combined market share of 88% of the industry's deposit base and 83% of total assets in 1981.

So what were the implications of the above for the competitiveness and efficiency of the banking sector during that period? While there are no specific studies on banking competition and efficiency in Ghana during the pre-reform period, it could be the case that, as noted by Kasekende et al. (2009), in the case of Africa's banking sector the post-independence era was characterised by the absence of competition. As noted above this was a period characterised by the predominance of state-owned banks with restriction on the entry of new banks, whether private domestic or foreign. Thus there was limited competitive pressure on the existing, mostly state-owned banks and the banking system was heavily concentrated. As mentioned above, two foreign banks (SCB and BBG) and 2 state banks (GCB and SSB) accounted for a total 88% of deposits market share and 83% market share in total assets in 1981 (Antwi-Asare and Addison, 2000). In addition, with the direct controls monetary policy regime, lending rates and deposit rates were determined and regulated by the government and thus banks could not independently set rates to attract more deposits or grow their lending portfolios. The directed lending and credit ceilings also meant there was little commercial consideration in the selection and financing of projects. There was therefore limited competition in the banking industry, which meant that efficiency implications were also limited. As a matter of fact, due to the lack of commercial consideration in the selection of projects to finance, the long-term nature of some of the government projects financed, and managerial incompetency, most of the public sector banks were saddled with huge non-performing loans (NPLs) and operational losses which led to financial distress in the 1980s. The economic challenges of the late 1970s also compounded the deterioration in the banks' asset portfolios. The banking sector was on the verge of collapse and required some reforms, hence the adoption of FINSAP.

2.4.2 Financial sector reforms and developments, 1982-2000

Financial sector development during this period was shaped by the first financial reforms, dubbed FINSAP, which was carried out as part of the broad Economic Recovery Programme (ERP) and Structural Adjustment Programme (SAP) undertaken by the Government in collaboration with the IMF and the World Bank.

The reforms, most of which were implemented during the late 1980s and the early 1990s, saw the abolition of the direct controls monetary policy regime. A new monetary policy regime of monetary targeting, where the Bank of Ghana uses indirect instruments as the primary instruments of monetary policy, was introduced (Bawumia, 2010). Interest rates were liberalised and floors and ceilings on lending and borrowing rates were abolished. The fixed exchange rate system was also replaced with a flexible exchange rate system with an interbank foreign exchange market established.

Reforms aimed at enhancing the soundness of the banking system through an improved regulatory and supervisory framework were undertaken, including the enactment of a new Banking Law 1989 (Act 227), a non-bank financial institutions (NBFI) Law 1993 (Act 328), the introduction of standardized reporting and accounting procedures, and the strengthening of the supervisory capacities of the Bank of Ghana (Brownbridge and Gockel, 1996). The new Banking law established minimum capital requirements for various classes of banks: commercial, development and merchant banks. Commercial banks focused on retail banking services provision for small and medium sized businesses; while merchant banks were to focus on corporate clients and providing corporate finance, syndications, and trade finance, with restrictions on branch banking. The development banks were the state-owned banks engaged in supporting specific sectors of the economy.

The state-owned banks were also restructured, with the removal of the huge non-performing loans (NPLs) from the banks' balance sheet and their replacement with Bank of Ghana bonds. A specialized government agency, the Non-performing Assets Recovery Trust (NPART) was set up to take over the NPLs and recover as many of them as possible. This was to ensure sanity into the banking system. To improve on prudent lending practices, lending policies were overhauled and internal controls and risk management policies were either introduced or strengthened in the public sector banks.

As part of the reforms, under FINSAP, the banking sector also witnessed some changes in the composition and diversity of banks. Following relaxation of bank entry restrictions as part of the reforms, the industry witnessed the entry of four new private domestic banks, three foreign banks and two regional banks.[14] No new state bank was established as two existing state banks merged.[15] Furthermore, the regulatory environment and supervisory powers of the

[14] The two regional banks were Ecobank Ghana whose parent company was based in Togo and Stanbic Bank, a subsidiary of Standard Bank of South Africa.
[15] These were the Social Security Bank (SSB) and the National Savings and Credit Bank (NSCB).

Central bank were strengthened, and two insolvent state banks which could not meet capital adequacy ratios had their licenses withdrawn by the regulator. The two largest state banks were partially privatised with minority stakes floated on the Ghana Stock Exchange in 1995 and 1996. Following these developments during the 1990s, the structure of the banking sector at the end of FINSAP is captured in Table 2.5.

Table 2.5 Structure of Ghana's banking sector, 2000

Bank	Year of Establishment	Ownership type	Banking business
1. Standard Chartered Bank	1896	Foreign	Commercial
2. Barclays Bank of Ghana	1917	Foreign	Commercial
3. Ghana Commercial Bank	1957	State	Commercial
4. National Investment Bank	1963	State	Development
5.Agricultural Development Bank	1965	State	Development
6. Merchant Bank Ghana	1972	State	Merchant
7. SSB Bank	1977	Foreign	Commercial
8. Ecobank Ghana	1990	Regional	Merchant
9. CAL Merchant Bank	1990	Foreign	Merchant
10. The Trust Bank	1994	Foreign	Commercial
11. Metropolitan and Allied Bank	1995	Private domestic	Commercial
12. First Atlantic Bank	1995	Private domestic	Merchant
13.International Commercial Bank	1996	Foreign	Commercial
14. Prudential Bank	1996	Private domestic	Development
15. Unibank Ghana	1997	Private domestic	Commercial
16. Stanbic bank	1999	Regional	Commercial

Source: *Bank of Ghana Reports (various issues)*

In addition to the gradual reduction in the dominance of state banks, another notable feature of the above developments was the increase in the number of merchant banks from one to four. The merchant banks were to focus on investment banking services provision for large corporates, while the commercial banks had their line of business primarily in the retail banking segment. Merchant banks accordingly had branch restrictions as they could not have more than a branch in a major city. The development (mainly state) banks continued to operate to support the specific sectors for which they were established. In terms of banking business though, banks were primarily engaged in the provision of short-term loans and overdraft facilities, as well as investments in government securities and placements with other local banks. These served as the major components of the banking sector's assets. These assets were funded largely by deposits which were mobilised in the local market.

In terms of assets and liabilities composition, Table 2.6 shows that investment in government securities and placement with other banks accounted for 42% of total assets; while loans and advances represented 38% of the assets size in 2000. The major liabilities component is deposits which accounted for 62% of total assets. The banking sector operates at the short-term end with the maturity profile of government securities and loans being, on average, one year, while deposits maturities are, on average, three months. The major sources of deposits are individuals, firms and companies operating in Ghana, while the loan portfolio is similarly to retail and corporate clients operating in Ghana.

Table 2.6 Composition of Assets and Liabilities of Ghana's banking industry, 2000

	GHC '000	
Cash and Short-term Funds	82,934	7%
Investments in Treasury Bills and due from banks	512,560	42%
Loans and Advances	465,391	38%
Fixed and Other Assets	168,272	14%
Customer Deposits	760,407	62%
Borrowings and due to banks	60,700	5%
Shareholders' Funds	131,237	11%
Other Liabilities	276,814	23%

Source: Computed by Author from Banks Annual Reports, 2000

Beyond the banking sector, the financial sector reforms also saw the development of a capital market, with the establishment of the Ghana Stock Exchange in 1990, stock brokerage firms, and the capital market regulator, the Securities and Exchange Commission. Other financial reforms in the non-banking sector were carried out in the insurance industry, savings and loans companies, and rural banks. Appendix 2.1 summarises the key reforms undertaken during FINSAP.

The gradual liberalisation of the banking sector resulted in a reduction in the level of concentration. However, the four largest banks remained unchanged with the two foreign banks, Standard Chartered Bank and Barclays Bank and the two state banks Ghana Commercial bank and SSB accounting for about 68% market share of total assets. In terms of

the impact of these reforms on banking competition and efficiency, as discussed in detail in Chapters 3 and 4, Antwi-Asare and Addison (2000) observe that the reforms helped in expanding the size and diversity of the banking sector, and enhanced operational efficiency using various accounting ratio measurements. The authors note that although concentration levels declined, they still remained high and signalled that competition was low. A study by Buchs and Mathisen (2005) based on data for the period 1998–2002 finds that the banking industry still exhibited an uncompetitive market structure and that the high domestic financing requirements of the government fostered inefficiency in the banking industry. Banking concentration remained high in spite of the fall in concentration levels, and this was seen as fostering anti-competitive behaviour by the big banks. The huge mandatory reserve requirements to accommodate government fiscal deficits and low penetration of new banks were also seen as factors which constrained banking competition and efficiency. Thus, although FINSAP created an improved environment for the banking sector, it was seen as making only a limited impact on enhancing the competitiveness and efficiency of the industry due to those lingering challenges.

2.4.3 Banking sector reforms and developments, 2001-2014

It was to address the challenges highlighted above that the Central Bank embarked on a comprehensive banking sector deregulation reform programme under its Financial Sector Strategic Plan (FINSSIP) in 2001. Unlike FINSAP which was undertaken under the auspices of the IMF and World Bank, the recent reforms under FINSSIP during the 2000s were initiated by the Central Bank. The aims of these reforms were to deepen the financial sector and also to increase the competitiveness and efficiency of the banking sector (Acquah, 2006). Details of the key financial liberalisation or deregulation reforms are discussed in chronological order below, while a summary of all the reforms are highlighted in Appendix 2.2.

2002: New Bank of Ghana Act, conferring Operational Independence on BoG

Prior to 2002, the Banking Law (PNDC Law 225) enacted during FINSAP and the operational modalities of the Bank of Ghana suggested that the Central Bank was an appendage of the Ministry of Finance, which made it a conduit for the financing of fiscal deficits. Accordingly, monetary policy was not independently undertaken. In January 2002, however, a new Bank of Ghana Act (Act 612, 2002) was passed. The new Act conferred

operational independence to the Bank of Ghana and also assigned monetary policy formulation to an independent Monetary Policy Committee (MPC) of the Bank. Central Bank financing of fiscal deficits was also limited to 10% of fiscal revenues in the preceding year. The MPC was inaugurated in September 2002, and meets bi-monthly or quarterly to review economic developments and set the Prime rate. The outcomes of its meetings and monetary policy decisions are communicated to the public through press releases.

2002: Introduction of BoG Policy Rate, the Prime Rate

In line with its new monetary policy setting framework, the Bank of Ghana introduced a new policy rate, the *BoG Prime Rate,* in September 2002 as an instrument to signal the Bank's assessment of inflationary pressures and its monetary policy stance. The Prime Rate also serves as the benchmark rate for the setting of interest rates (base rates) by banks in the country. Following the MPC's periodic review meetings on the assessment of the economy and decisions on the policy rate, adjustments are made by banks to their lending and deposit rates which are also published.

2003: Introduction of Universal Banking to remove restrictions on banking activities

Banking activities were restricted in terms of scope (what business lines banks could engage in) and geography (where banks could operate). The existing law had classified banks into three categories: commercial (retail), merchant (corporate) and development banks. The Bank of Ghana introduced universal banking in 2003 with a view to removing such restrictions on banking activity and integrating the fragmented banking sector. Universal banking was introduced to allow banks to choose the type of banking services they would like to offer in line with their capital, risk appetite and business orientation. This deregulation policy therefore abolished the segmented commercial, merchant and development banking categorisation that existed following the banking reforms in the 1990s to create a uniform playing field for all banks. Universal banking was also intended to embrace mortgage financing, insurance business, among others as enshrined in a new banking law. The Bank of Ghana announced new capital requirements to be complied by banks to receive the universal banking license. All existing banks met this new capital requirement by the deadline date of end of 2006.

2004: Passage of new Banking Law, expanding the definition of banking activities to embrace universal banking

A new Banking Law, the Banking Act 2004 (Act 673) was passed in 2004 to replace the existing Banking law 1989 (PNDCL 225). Some of the significant changes introduced by the new Banking Act included the expansion of the definition of banking activities to include insurance business, mortgage financing, securities, finance leasing, portfolio management, advisory services such as capital restructuring, mergers and acquisitions, credit reference services and the keeping and administration of securities. This regulatory change was to give credence to the universal banking concept of relaxing bank activity restrictions. The deregulation policy was also intended to enhance scope economies in the financial services industry.

2005: Relaxation of bank entry restrictions with an open licensing policy

As part of the deregulation reforms, the central bank relaxed bank entry restrictions. It accordingly adopted an open but gradual licensing policy which allows the entry of new banks into the industry to enhance market contestability. The entry of such new banks is expected to enhance competition, encourage faster modernization of banking operations and facilitate efficiency of the banking system, to better support the growth and diversification of the financial services industry.

2006: Abolishment of secondary reserve requirements

The Bank of Ghana reduced the secondary reserve requirements of banks from 35% to 15% in July 2005, and further abolished the remaining 15% in August 2006, leaving only primary reserve requirements of 9% held in cash. The high secondary reserve requirement, which was held in government securities, was the legacy of high fiscal deficits and served as a captive market to finance these deficits. The consequence was that it crowded out private sector finance. This deregulation policy of scrapping the secondary reserves is to increase the supply of loanable funds to the private sector, encourage competition in the loans market and help deepen financial intermediation.

2007: Passage of Credit Reference Act

As a means of managing potential risk-taking behaviour of banks associated with such deregulation reforms in developing countries, the Bank of Ghana sought to strengthen the credit environment through the passage of this law. This Act paved the way for the licensing of credit reference bureaus in the country. Credit referencing is critical due to the problems of adverse selection and moral hazard caused by information asymmetry between banks and borrowers. In Ghana, it was also felt that the absence of this service resulted in the banks over-pricing risks in the economy on account of the high real interest rate spreads. Since the passage of the law, three credit reference bureaus have been licensed by the Bank of Ghana and are operational to provide credit information sharing to ameliorate the problem of information asymmetry and facilitate credit risk assessment.

2007: Currency Redenomination

Another policy implemented by the Bank of Ghana in 2007 was a re-denomination of the national currency (the Cedi). Although not directly a deregulation policy, this policy was to enhance efficiency of the payments system. The legacy of past episodes of high inflation and macroeconomic instability had been the rapid increases in the numerical values of prices of goods and services, which had imposed significant deadweight burden on the economy. This was in the form of high transaction costs at bank tellers/cashiers, high risks involved in carrying loads of currency for transaction purposes, and a strain on the payments system, particularly at the ATMs. The currency re-denomination was implemented in July 2007, with the new currency, the Ghana Cedi (*GH¢)* equivalent to 10,000 Old Cedis (*¢).* Banks re-calibrated their ATMs and other banking software to accommodate the change and actively engaged in a comprehensive customer and public education programme. Some services by banks, such as cash collection services, were dispensed off and banks re-engineered some of their operational services.

2008: Increase in minimum capital requirements for banks

To enhance stability of the banking system and support the rapid growth in credits anticipated by the deregulation policies, the Central Bank announced increases in the minimum capital requirements of banks from GHC7 million to GHC 60 million in 2008. For existing banks, there was a deadline of December 2009 for foreign banks to comply and December 2012 for local banks to comply, while all new banks were to meet the new capital requirements before being granted a banking license. Although it was envisaged that the sharp increase in the

minimum capital would force some form of consolidations, this was not the case as all existing banks met the new capital requirements by their respective deadlines. The capital adequacy ratio (CAR) requirement of banks had also been raised from 6% to 10% in line with international standards and practices.

Besides these key policy reforms, other initiatives and banking laws were passed to support other segments of the financial sector, including venture capital, pensions, home finance, etc. as captured in Appendix 2.

So what were the effects of these reforms on the structure of the banking sector? To what extent did they enhance the growth and development of the sector? These issues are explored in the next section.

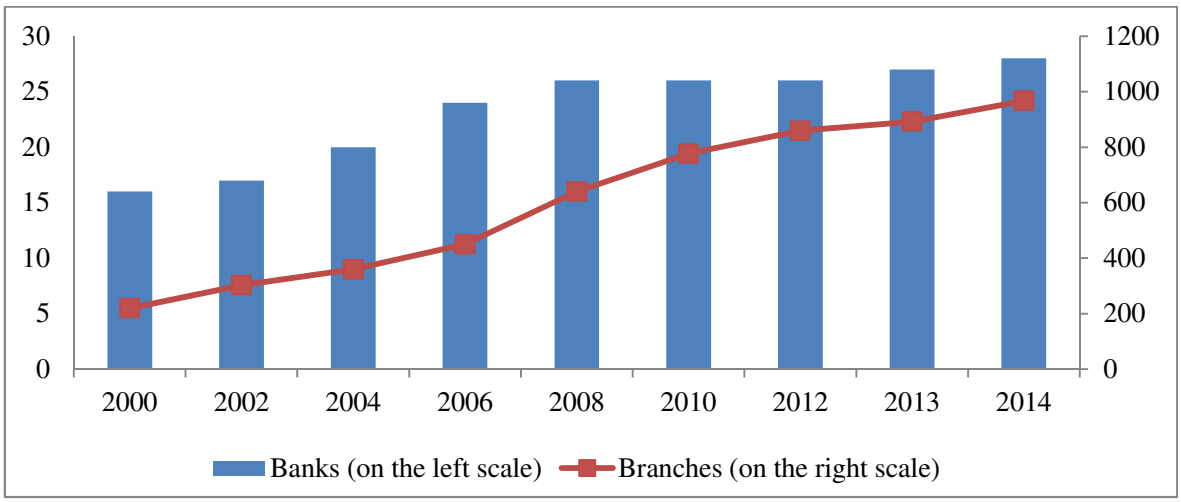
2.5 Structural changes in the banking sector following the reforms

In this section, we analyse the structural changes that have taken place in the banking sector following the implementation of the deregulation reforms. We examine the evolution of some key banking sector metrics during the period 2000–2014 to analyse any transformation of the sector as a result of the reforms. These metrics are not necessarily measures of competition or efficiency but provide some initial evidence as to how the banking sector has responded to the deregulation reform initiatives pursued by the Central Bank. These are discussed under the separate headings of industry growth; ownership changes; concentration levels; financial intermediation; and financial soundness.

2.5.1 Industry growth

The relaxation of bank entry restrictions and subsequent licensing of new banks resulted in the number of banks increasing from 16 in 2000 to 28 in 2014. Most of the new entrants are foreign banks, predominantly pan-African banks, and a few private domestic banks. Aggressive branch expansion by both new and existing banks following the removal of branch restrictions via the introduction of universal banking led to a significant growth in the branch network, from less than 250 in 2000 to almost 970 at the end of 2014 as shown in Figure 2.5.

Figure 2.5 Number of banks and branches, 2000-2014



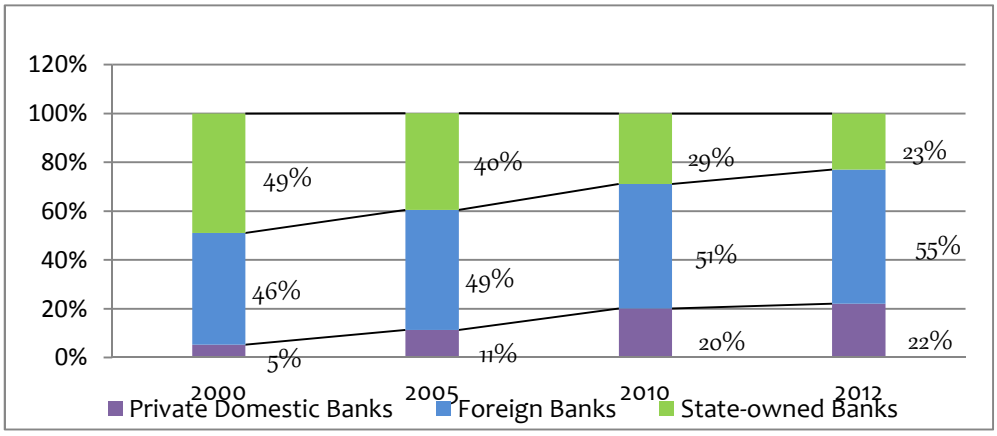
Source: Bank of Ghana, *Annual Report* (various issues)

The increase in the number of banks and branch network impacted significantly on the growth in banking business in terms of deposits, loans and total assets. The banking industry's assets size increased consistently from less than US$2 billion in 2000 to almost US$16 billion in 2014, with consistent growth in deposits and loans.

2.5.2 Ownership changes

The reforms led to a gradual waning of the dominance of state-owned banks in the share of assets of the banking sector. As shown in Figure 2.6, the share of total banking assets controlled by state banks declined from almost half (49%) in 2000 to less than a quarter (23%) in 2012.

Figure 2.6 Share of assets by ownership class



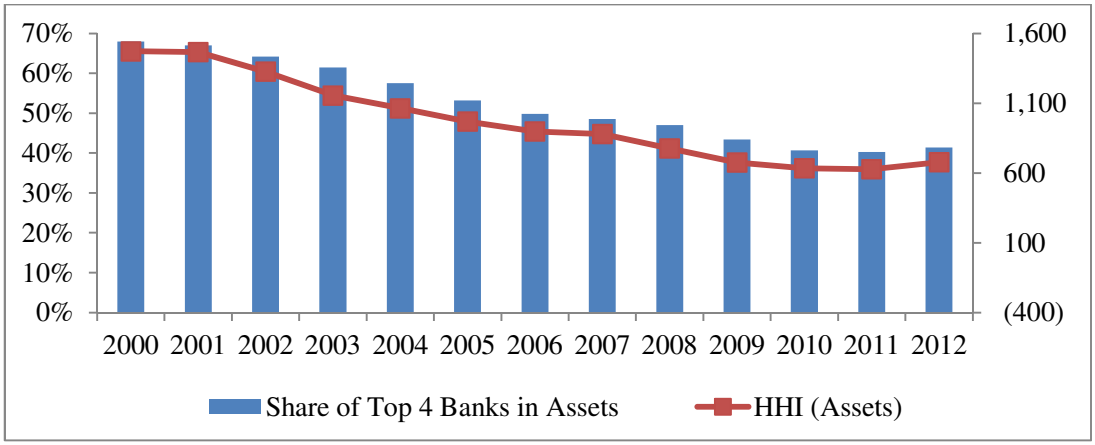
Source: IMF (2013)

The share of assets by private domestic banks increased significantly from 5% to 22% over the same period. The share of foreign banks inched up from 46% to 55%, but reflects a more diversified ownership base than previously as the new foreign banks have been predominantly regional banking groups from South Africa and Nigeria, with others from India and Libya.[16]

2.5.3 Concentration levels

Figure 2.7 shows the trend in concentration of the banking sector, with two measures of concentration: the CR_4 ratio and the Herfindahl- Hirschman Index (HHI). The CR_4 ratio of total assets shows a gradual but consistent decline in the market share of the Top 4 banks from 68% in 2000 to 41% in 2012 while the HHI more than halved from 1,473 to 678 over the same period.

Figure 2.7 Concentration Ratios, CR_4 and HHI, 2000 – 2012



Source: Computed from Banks' Annual Reports and Central Bank Reports

This is attributable to gradual taking of market shares from the new banks and gains made by existing banks. While declining concentration gives a sense of a growing incidence of competition, it cannot be assumed to represent intensification of competition as concentration measures have been found not to be necessarily appropriate proxies for competition in banking (Delis, 2012). It is worth noting however that the decline in concentration is a consequence of the deregulation reforms.

[16] The existing foreign banks were subsidiaries of banks based in Europe.

2.5.4 Financial intermediation

The sharp increases in the two measures of financial intermediation especially between 2004 and 2008 (see Table 2.7) show the strong immediate impact of scrapping the secondary reserves in deepening intermediation with significant growth in loans.

Table 2.7 Financial intermediation indicators, 2002 - 2014

Intermediation Measures	2002	2004	2006	2008	2010	2012	2014
Loans-to-Deposits Ratio	49%	53%	67%	79%	57%	60%	68%
Loans-to-Total Assets Ratio	31%	36%	45%	52%	40%	44%	44%

Source: Computed from Banks' Annual Reports and Central Bank Reports

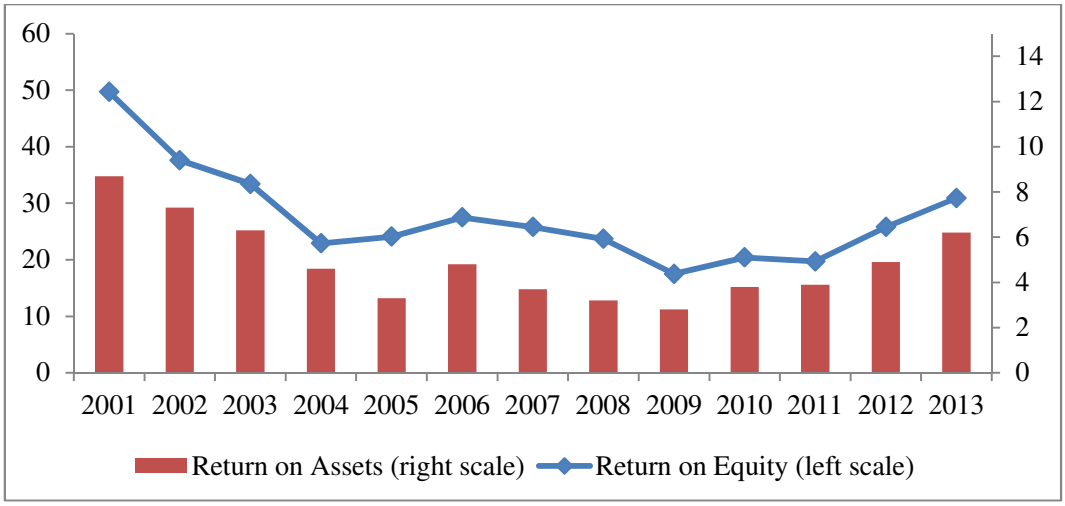
The subsequent fall in the ratios in 2010 and marginal rise in 2012 seems to suggest that credit growth is still somehow being constrained by huge investment in government securities. It could also be that banks did not manage the credit expansion well in the aftermath of the scrapping of the secondary reserves and had to curtail the rapid growth subsequently.

2.5.5 Financial soundness

We review the trend in key financial soundness indicators in relation to profitability, capital adequacy and asset quality. As shown in Figure 2.8, the return on assets and return on equity are extremely high in Ghana compared to international benchmarks in developed countries. Even among developing regions, the Sub-Sahara African region is reckoned to have one of the highest levels of bank profitability indicators (Beck and Cull, 2014), and Ghana's banking sector ranks as one of the most profitable in the region. The high profitability ratios reflect the high interest rate regime in Ghana. Borrowing by government from the banking sector to finance budget deficits have led to the historically high interest rates on risk-free government securities. Macroeconomic weaknesses with high inflation rates and the need to maintain positive real interest rates also account for the prevalent high interest rates regime. The 91-day Treasury bill rate was 42% p.a. at end of 2000, and although declined by the end of 2014, stood at 24% p.a. which is high by internal standards. Based on the high (benchmark) Treasury bill rates, lending rates have also remained high resulting in wide interest rate

spreads between loans and deposits. Average interest rates spreads stood at high levels of 29% in 2000 and 22.35% in 2014.

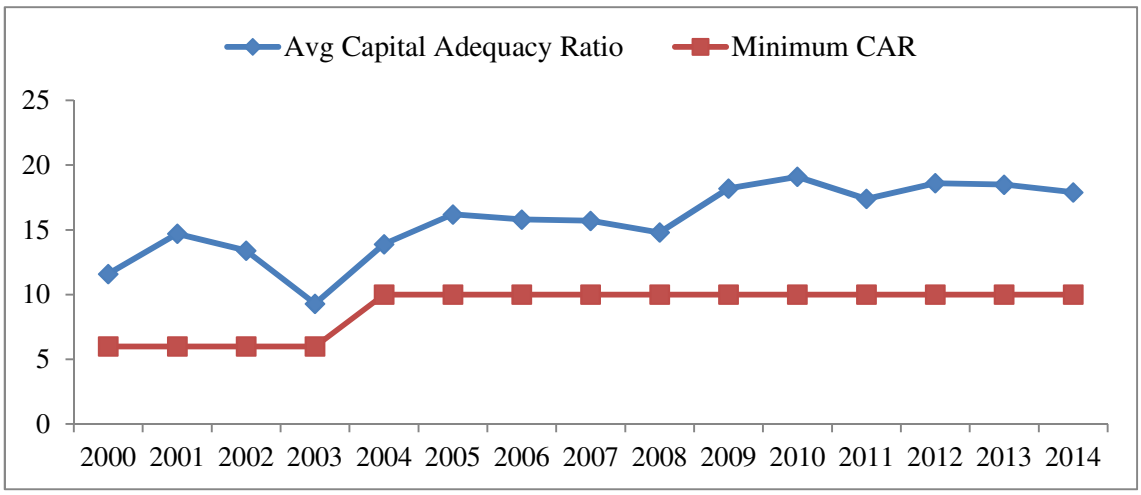
Figure 2.8 Return on Assets and Return on Equity, 2001 – 2013



Source: Computed from Banks' Annual Reports and Central Bank Reports

From Figure 2.8, the trend in profitability indicators shows a sharp decline during 2001-2004; relative stability thereafter to 2010 followed by a rising trend during the last three years, and reflects the trend in interest rates due to macroeconomic conditions during those periods. This notwithstanding, the profit indicators are deemed to be strong and higher than other African banking sectors (IMF, 2013).

Figure 2.9 Capital adequacy ratios, 2000–2014

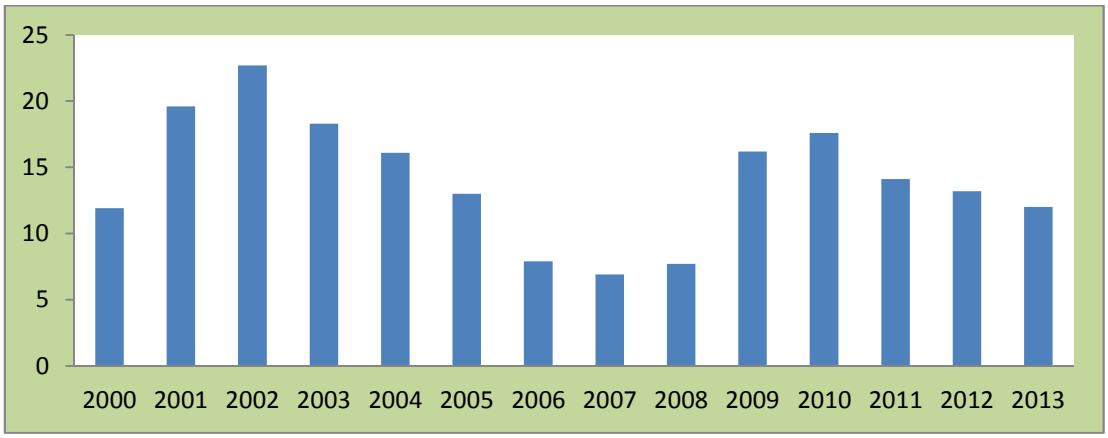


Source: Bank of Ghana and IMF Country Report on Ghana

Compliance with capital adequacy does not seem to be a problem in Ghana, as the banking sector's average capital adequacy ratio (CAR) has been consistently above the statutory minimum as shown in Figure 2.9. Thus, notwithstanding the raising of the CAR from 6% to 10% by the Central Bank in 2004 in line with Basle requirements, the industry average has remained over 17% in each of the last six years. The capital raise during 2009-2012 provided additional buffer for banks' capital adequacy.

On asset quality, although there is no international defined benchmark on NPL to gross loans ratio, the trend in the ratio depicted in Figure 2.10 seems to suggest that the NPL ratio is high. The high NPL has been attributed to the generally high interest rates in Ghana, coupled with the lack of credit reporting bureaus (until recently). Except for the period 2006-2008 during which the NPL ratio was below 10%, the ratio has averaged about 15% for most part of the study period. The rapid credit growth experienced during 2006–2008 following the abolishment of the secondary reserve requirements could be the underlying reason for the cosmetic sharp drop in the ratio during that period, as the NPL rose up sharply in 2009 and 2010 before easing gradually during 2011–13. The sharp growth in NPL during 2009-2010 reflects the adverse effects of excessive risk taking through the strong growth in credit expansion during 2006-2008 immediately following the scrapping of the secondary reserve requirements, coupled with sharp interest rate rises during 2007-2009.

Figure 2.10 Non-performing loans to gross loans, 2000 – 2013



Source: Bank of Ghana and IMF Country Report on Ghana

Lack of credit information was adduced as one of the major reasons for high NPLs in Ghana and thus the passage of the credit reference Act in 2007 and the subsequent licensing of three credit reference bureaus during 2008-2010 have been expected to enhance the credit environment and improve asset quality. It may be that the decline in the NPL ratio since 2011 could be attributed to the impact of these credit reference bureaus but this requires further investigation.

2.5.6 Composition of Assets and Liabilities

Table 2.8 shows the trend in the composition of the banking industry's assets and liabilities, which reveals no significant changes in the structure of the industry's balance sheet, in terms of the asset and liability groupings. Following the deregulation reforms however, we observe some changes in the assets composition, with loans and advances accounting for relatively higher percentage of assets during the post-reform period. The composition of Shareholders' funds has been relatively small over the years, although the increase in capital requirements in 2009 have resulted in higher shareholders' funds to assets ratios observed in 20012 and 2014 in Table 2.8.

Table 2.8 Trends in Assets and Liabilities Composition, 2000-2014

Assets Composition	**2000**	**2004**	**2008**	**2012**	**2014**
Cash and Short-term Funds	7%	10%	10%	13%	16%
Investments in Treasury Bills and due from banks	42%	45%	27%	36%	30%
Loans and Advances	38%	36%	52%	44%	44%
Fixed and Other Assets	14%	8%	10%	7%	10%
Liabilities Composition	**2000**	**2004**	**2008**	**2012**	**2014**
Total Deposits	62%	68%	66%	73%	64%
Borrowings and due to other banks	5%	8%	13%	8%	16%
Shareholders' Funds	11%	12%	10%	14%	14%
Other Liabilities	23%	12%	10%	4%	5%

Source: Computed by Author from Banks Annual Reports

The above review shows the structural transformation of the banking sector following the implementation of the reforms. The industry's growth has been strong, concentration levels have declined and the dominance of state banks has been significantly curtailed. Although financial deepening has overall increased, loan quality seems to remain a challenge to financial stability. High profitability and capital adequacy indicators seem to suggest that the banking sector's financial soundness is adequate.

These were the major structural changes that occurred in the banking industry. In terms of the composition of the Top 4, there was only a marginal change as Ecobank Ghana, one of the regional banks, became the largest bank at the end of 2014. The Top 4 banks still account for about 41% of total market share although this has declined from about 68% in 20100. In terms of size, the banking industry grew from less than US$2 billion in 2000 to about US$16 billion in 2014. There were no banking failures during the period and the global financial

crisis did not have any direct impact on the survival of the banks. In addition, there were no mergers and acquisitions except for the acquisition of The Trust Bank by Ecobank Ghana.

These structural changes notwithstanding, key questions that remain unanswered have been whether the intended positive impact of these reforms on industry competition and bank-level efficiency have been achieved.

2.6 Conclusions

This chapter has provided the country-context of the thesis and reviewed the politico-economic environments which shaped the pursuit of different financial and banking sector policies. The review shows that economic and financial sector reforms are relatively new in Ghana and commenced only during the late 1980s under the auspices of the IMF and World Bank. The reforms undertaken during that period under FINSAP seem to have had only a limited impact on competition and efficiency. This is due to the fact that although interest rate controls and credit ceilings were abolished, the banking industry was still fragmented and highly concentrated due to restrictions on bank entry, banking activity and high reserve requirements. It also emerged that notwithstanding reforms in the non-bank financial sector, including the capital markets and the insurance industry, the banking industry remains the dominant component of the overall financial sector, accounting for about 75% of the total financial system (IMF, 2013).

It is against this background that the Central Bank implemented further deregulation reforms during the early part of the 2000s under FINSSIP. These reforms included the enactment of a new Bank of Ghana Act (2002) which gave operational independence to the Central Bank, the introduction of universal banking, the abolition of the secondary reserve requirements, the adoption of an open licensing requirement with the resultant increase in the number of banks, which have all contributed to significant structural changes in the banking sector. Besides these major deregulation reforms, other policy measures aimed at improving the regulatory framework and strengthening bank supervision were implemented to facilitate increased competition and efficiency of banks. In spite of these far-reaching reforms, which have structurally transformed the banking industry, an assessment of the impact of these comprehensive reforms on banking competition and efficiency is yet to be carried out. This is the policy research gap that this study seeks to fill.

In the next chapter, we explore the issue of banking competition by reviewing relevant literature on the concept of banking competition, financial reforms, and the models used in measuring banking competition. Summaries of the main reforms implemented during FINSAP and FINSSIP are provided in Appendices 2.1 and 2.2.

Appendix 2.1: Reforms: Financial Sector Adjustment Program (FINSAP), 1988 – 1998

- Liberalisation of interest rates and abolition of credit ceilings and directed credit
- Restructuring of financially distressed banks
- Strengthening of the regulatory and supervisory framework
- Promotion of non-bank financial institutions through establishment of
    - Discount houses
    - Finance houses
    - Leasing companies
    - Mortgage finance companies
- Liberalisation of the foreign exchange market and licensing of forex bureaux
- Establishment of the Ghana Stock Exchange
- Banking Act 1989, PNDCL 225
- Bank of Ghana Law 1992, PNDCL 291
- Securities Industry Law 1993, PNDCL 333
- Non-Bank Financial Institutions (NBFI) Law 1999, PNDCL 328
- Insurance Act 1989, PNDCL 227
- Social Security Act 1991 PNDCL 247

Appendix 2.2: Reforms: Financial Sector Strategic Plan (FINSSIP), 2001–2008

- Bank of Ghana Act 2002 – Autonomy to Central Bank
- Monetary Policy Committee (MPC) process – Transparency
- Introduction of Universal Banking in 2003
- Abolishing secondary reserve requirements in 2006
- Banking Act 2004, and Banking Amendment Act 2007 – Offshore Banking
- Long Term Savings Act 2004 and Venture Capital Trust Act 2004
- Payment System Act 2003
- Foreign Exchange Act 2008
- Anti-Money Laundering Act 2008
- Credit Reporting Act 2008, and licencing of Credit Reference Bureaux
- Borrowers and Lenders Act 2008
- Insolvency Act 2003
- Home Finance Act 2008
- Non-Bank Financial Institutions Act 2008
- Central Securities Depository Act 2007
- Insurance Act 2006
- National Pensions Act 2008
- Strengthening Regulatory and Supervisory Framework
    - Risk based Supervision
    - Electronic Financial Analysis and Surveillance System
    - Stress Testing
    - Publication of Annual Percentage Rates (APRs) of Banks
- Redenomination of the Currency

CHAPTER 3 BANKING COMPETITION: A REVIEW OF THE LITERATURE AND MEASUREMENT TECHNIQUES

3.1 Introduction

This chapter reviews the literature on banking sector competition and examines the models for measuring competition. Considering the broad nature of the literature on banking competition, we will focus the discussion on the literature regarding the desirability of banking competition in terms of its expected impact on other banking policy objectives. We also review the expected impact of banking reforms on competition in general, as well as the proposed models for measuring competition in banking. The rest of the chapter is structured as follows. In Section 3.2, we review the literature on the potential impact of competition on financial stability, banking efficiency and financial access. This is followed by a review of the literature on the impact of banking reforms on banking competition in Section 3.3. Section 3.4 examines the various methods proposed in the literature for measuring banking competition. Section 3.5 concludes the chapter.

3.2 Desirability (pros and cons) of banking competition

One key issue that often comes up in the banking competition literature is the potential impact of a competitive banking sector on a country's financial system. Is competition desirable and intense competition in banking always good for the economy? While an obviously positive response seems inevitable from conventional wisdom, or at least, from a standard microeconomic perspective, the literature surveyed does not provide such a conclusive view. What is the potential impact of banking competition on the financial system of a developing country, such as Ghana? What level of competition is good and beyond what level does competition seem to be excessive or detrimental? Should regulators and policymakers be concerned with deepening competition through policy or restraining excessive competition in Ghana? An attempt at addressing these issues is made in the review of existing literature since the potential impact of competition on the financial system and policy recommendations deriving therefrom seem to have been a subject of controversy especially following the global financial crisis of 2007-2009.

A major debate cited in the literature regarding the desirability of competition in banking relates to its impact on financial stability. One school of thought argues that competition in banking can destabilise the financial system; another school of thought posits that competition in banking rather enhances stability by minimising the fragility that is instead associated with an uncompetitive banking sector. Stability matters as one of the central objectives of banking policy (Allen and Gale, 2000). Competition in banking also matters for other reasons beyond stability, especially if stability can be guaranteed via other, usually regulatory means. Banking efficiency and access to financial services are also fundamental banking policy objectives especially for developing countries and have been identified as important outcomes that competition impacts on but with different predictions (Demsetz, 1973; Beck, Demirgüç-Kunt, and Maksimovic, 2004; Claessens and Laeven, 2004; World Bank, 2012).

3.2.1 Banking competition and financial stability

As mentioned above there are two opposing theories on the impact of banking competition on the stability of the financial system, namely the **competition–fragility hypothesis** and the **competition-stability hypothesis**. The competition–fragility theory sees competition in banking as detrimental to the financial system as it destabilises the sector. The competition-stability view however argues that an uncompetitive banking sector breeds fragility in the financial system and hence competition policy is required to minimise such fragility and stabilise the sector.

The key argument of the competition-fragility school is that competition among banks adversely affects their net interest margins and profitability due to increases in deposit rates and cuts in lending rates. As profits are eroded, banks are pushed into taking excessive risks to help maintain their profitability levels. Such excessive risk-taking behaviour by banks include imprudent lending behaviour, poor screening of potential clients and projects, loosening credit standards, extending lower quality loans, and financing riskier projects (Boot and Thakor, 1993; Cetorelli, 2001) . The risky behaviour of banks makes them vulnerable to high loan default rates, thereby potentially creating instability in the financial system. Risk is considered as endogenously determined by the bank and competition is accordingly seen as inducing fragility in the financial system through higher loan default, bank losses and equity erosion. Conversely, in less competitive banking environments, it is argued that banks' net interest margins and profitability are usually not under threat, and so banks behave more

prudently and have no incentives to engage in any risky ventures. Hence the threat to stability is muted in such uncompetitive banking environments (Keeley, 1990; Allen and Gale, 2000).

Competition–stability theorists on the other hand point to the fact that lack of banking competition tends to create a concentrated market structure in which large banks do exploit customers by charging higher lending rates.[17] The high lending rates constrain the capability of borrowers to service their loans, exacerbate moral hazard incentives of borrowers to shift to riskier projects which increases credit default risk of borrowers, and thereby make the banking system susceptible to instability. Risk as a consequence is seen as exogenous to the bank since it is determined by the behaviour of the borrowers. The higher interest rates could also lead to problems of adverse selection of riskier borrowers (Boyd and Nicoló, 2005). Increased competition in this context by lowering lending rates makes loan repayment more affordable to borrowers and accordingly minimises bank default risks and any potential risky behaviour by borrowers, thereby reducing any threat of instability. Thus, the competition-stability view predicts that bank actions will result in more risk-taking and greater fragility in less competitive banking systems than in competitive banking environments, and thus competition policy is required to foster stability.

The two theories thus both see credit risk as the key channel through which instability enters the financial system. In the competition-fragility hypothesis this risk is endogenous while in the competition-stability hypothesis, it is exogenous.

Beside the above key arguments, there are other transmission mechanisms through which competition leads to fragility or stability. According to the competition-fragility theorists for example, a competitive banking sector, characterised by a large number of relatively small banks, is prone to instability due to the fact that small banks have a greater incentive to undertake risky behaviour (Allen and Gale, 2000). Such risky behaviour can have contagion effects in the banking sector. A concentrated banking sector with larger banks is therefore seen as inherently more stable because of the ability of larger banks to spread risks. The counter argument by competition-stability is that the contagion effect of a relatively small bank in a competitive banking environment is less pronounced than the contagion effect of a large bank which goes burst in a highly uncompetitive banking system.

[17] Although these arguments seem to link banking competition to a non-concentrated banking market, it has been shown that concentration measures are not necessarily good proxies for competition from both theoretical and empirical viewpoints (Liu et al., 2013). In particular, there could high competition in a highly concentrated banking industry, while competition could be low even in an industry with many banks and low concentration. This distinction between concentration and competition is important and explains why the structure-conduct performance model which relates competition to concentration has been widely discarded in recent empirical work as discussed later in this chapter.

Another channel proposed by competition-fragility theory is that since a more concentrated and uncompetitive banking system tends to have a few large banks, it is easier to monitor and supervise such banks. The reduced supervisory burden could enhance the quality of supervision, and help foster stability of the banking system. On the contrary, a competitive market with many small banks makes supervision burdensome and ineffective, and could adversely affect financial stability (World Bank, 2012). Proponents of competition-stability counter this argument by indicating that this is not the case as larger banks in a concentrated market can be more complex and diversified, and hence more difficult to monitor and supervise than small banks.

Earlier work in support of the competition-fragility view includes Keeley (1990); Hellmann, Murdock, and Stiglitz (2000); and Allen and Gale (2000). In a study of the US banking sector, Keeley (1990) attributes the surge in bank failures in the US during the 1980s to intense competition which reduced monopoly rents and profit margins, and caused banks to engage in excessive risky behaviour with the view to maintaining profitability levels. The paper observes that prior to the 1980s, regulatory restrictions on bank entry and branching as well as other anti-competitive measures made banks profitable and bank charters valuable. Thus banks had an incentive not to engage in risky behaviour. The relaxation of regulatory controls and pursuit of deregulation policies in the 1980s, led to declining profitability and bank charter values due to increased competition. Thus, increased competition reduced incentives for prudent bank behaviour and led to excessive risk-taking which contributed to bank failures during the 1980s.

Hellmann et al. (2000) also attribute the crisis in the US Savings and Loans market as well as the Japanese crisis to excessive risky behaviour by banks following the deregulation policies of relaxation of restrictions to bank branching and bank entry as well as deregulating interest rates. Increased competition in the deposits market led to higher deposit rates which caused banks' profitability to be under pressure and led to a reduction of their charter or franchise values, fuelling moral hazard in their behaviour.

At the other end of the spectrum, and using a panel data set of 69 countries over the period 1980–1997, Beck, Demirgüç-Kunt, and Levine (2006) establish that greater bank concentration is associated with a lower likelihood of suffering a systemic banking crisis. Their result is consistent with the concentration-stability, with the negative relationship between concentration and crises found to be robust. An interesting observation by the paper is that they find no evidence that banking system concentration is a proxy for a less

competitive banking environment. This is crucial as the theories presume that a concentrated market is a sign of uncompetitiveness or that the presence of a large number of banks in a banking sector is synonymous with intense competition which may not be the case. The authors however find support that concentrated banking systems have larger, better diversified banks with a correspondingly lower probability of failure, but no evidence that they are easier to monitor and hence more stable than less concentrated banking systems (Beck et. al., 2006).

Schaeck, Cihak, and Wolfe (2009), in a cross-country study of the relationship between competition and banking system fragility find support for the competition–stability view. The paper establishes that more competitive banking systems are less prone to systemic crises than less competitive banking systems. The results suggest that competitive behaviour of financial institutions not only significantly decreases the probability of systemic banking risk but also increases the survival time of banking systems. Based on various specification models they find no empirical support for the competition–fragility theory.

Ariss (2010) investigates how different degrees of market power affect efficiency and bank stability in a cross-country study of 821 banks in 60 developing countries over the period 1999-2005.[18] The paper establishes that an increase in the degree of market power leads to greater bank stability and reduces risk potential, in support of the competition-fragility theory. The paper notes that the findings could provide a rationale for the growth in mergers and acquisitions of banks in developing countries. The paper opines that increased market power may be a welcome development as it will facilitate financial stability in the relative stressed banking sectors in most developing countries in general.

For African countries, Moyo, Nandwa, Council, Oduor, and Simpasa (2014) explore the competition-stability-fragility nexus in a cross-country study of 16 African countries during 1995–2010, and the role macroeconomic and institutional factors play in the relationship. Specifically, the study examines the proposition that increased competition in the banking sector resulting from financial liberalization enhances financial stability. The results show that financial liberalization enhanced competition in Africa's banking sector, and that increased competition in the post liberalisation period corresponded to higher lead times to bank distress episodes. Their result is therefore in support of the competition-stability theory. The authors note, however, that stability of Africa's banking system in a liberalized and competitive

[18] The developing countries are from Africa, East/South Asia & Pacific, Eastern Europe and Central Asia, Latin America & Caribbean and the Middle East.

environment is contingent on the pursuit of sound macroeconomic policies and strong institutional support to enable the banking sector thrive.

The review of the empirical literature across different regions clearly shows that the relationship between competition and stability is neither clear-cut nor conclusive. This observation is also noted by Carletti and Hartmann (2003), who opine that the theoretical literature on competition-stability is not conclusive, as theories of bank runs and systemic risk largely disregard the implications of different bank market structures for the safety of the financial sector. The authors also contend that while some empirical work support the influential 'charter value' hypothesis of a negative relationship between competition and bank stability, others do not, and therefore conclude that the stability effects of changes in market structures and bank competition are case-dependent, and therefore require different institutional approaches to address them in different countries.

Another important consideration in this debate is that both competition and risk are measured in very different ways across the empirical work, which might partly explain the contradictory results. Different measures of bank competition (concentration ratios, Panzar-Rosse H-statistic, Lerner Index, etc.) can lead to different outcomes as we discuss in Section 3.4. Similarly, different measures of risk and fragility (non-performing loans, z-score, systemic risk, etc.) have different implications for stability. In addition, while some studies define fragility in terms of individual bank risk others define it in terms of the co-dependence of those risks, that is, systemic risk (Anginer, Demirgüç-Kunt, and Zhu, 2013).

The global financial crisis during 2007/8 reignited the debate on the competition-stability/fragility nexus and the design of pro-competition policies, that is, regulatory and deregulatory policies that influence the way and extent to which banks compete. Some believe that increases in banking competition and financial innovation led to distortions such as subprime lending, thus contributing to financial instability and crisis. However, others contend that the assertion that competition increased before the crisis does not necessarily suggest that greater competition in itself spurred the crisis. Beck (2008) observes that notwithstanding the conflicting empirical results, banking system fragility arising from increased competition is mostly the consequence of regulatory and supervisory failures, rather than competition per se. Thus, the benefits of increased competition for an efficient and inclusive financial system are strong, and regulatory and supervisory policies should focus on an incentive compatible environment for banking rather than try to fine-tune market structure or the degree of competition. This view is shared by Anginer et al. (2013) who suggest that

supervisory lapses and inadequate risk management supervision contributed to the crisis, rather than increased competition, as the financial crisis was preceded by an increase in market power.

In a more recent study, Beck, De Jonghe, and Schepens (2013) provide further empirical evidence which suggests that the relationship between competition and stability varies across different countries with different regulatory frameworks, market structures and levels of institutional development. While the authors show, on average, a positive relationship between banks' market power and banks' stability, an increase in competition is associated with a larger rise in banks' fragility in countries with stricter activity restrictions, lower systemic fragility, better developed stock exchanges, more generous deposit insurance and more effective systems of credit information sharing.

In conclusion, the likely impact of increased competition in Ghana's banking sector on financial stability is an empirical question that ought to be examined. The paper by Moyo et al. (2014) however argues that competition following deregulation or financial liberalisation in African banking sectors fosters stability if underpinned by macroeconomic environment and strong regulatory and institutional capacity. Most African banking sectors have been generally stable. Financial soundness indicators in most African countries point to relatively stable banking sectors with high liquidity, profitability and capital adequacy ratios. High capital to risk-weighted asset ratio averaging 19% in Africa is higher compared to other developing countries and significantly higher than international benchmark rates (Beck & Cull, 2014), partly due to the huge investment by banks in government securities, with moderate investment in risky assets. In Ghana, banks have consistently exceeded the minimum capital to risk-weighted asset ratio of 10% set by the Central Bank, with the industry average of over 17% for the past six years (Bank of Ghana, 2014). These large values are the combinations of large minimum capital requirements introduced in 2009, and the fact that government treasury bills continue to be an attractive investment for banks notwithstanding the scrapping of the secondary reserves. The stability of the banking sector in recent times however is no guarantee that it will continue in the light of these financial liberalisation policy reforms especially in the face of macroeconomic shocks and weak institutions. We do not examine the competition-stability-fragility theories in this thesis, but have discussed the potential impact that competition can have on financial stability based on these theories. The expected impact of competition in enhancing banking stability or creating

fragility for African countries, such as Ghana, with relatively less developed banking systems, is therefore an open empirical question, and one that needs to be explored in future research.

3.2.2 Banking competition and banking efficiency

Banking efficiency is another important policy objective that competition is expected to achieve. Accordingly, the impact of competition and pro-competitive policies on banking efficiency is also of paramount importance to regulators, policy-makers and academics.

Two separate theories are discussed in the theoretical literature regarding the impact of competition on efficiency: the **competition-efficiency theory** which emphasizes the efficiency-enhancing role of competition, and the **competition-inefficiency theory**, which postulates a negative relationship between competition and efficiency (World Bank, 2012). The literature seems to be much less ambiguous than the one on the competition-stability-fragility debate as most of the empirical literature lends credence to the competition-efficiency hypothesis.

The competition-efficiency theory, also referred to as the quiet life hypothesis, has its origins in Hicks' assertion that 'the best of all monopoly profits is a quiet life'. In the case of banking markets the theory would predict that an uncompetitive banking system will enjoy high interest margins and supernormal profits thus providing no real incentives for managers to strive to be efficient. Managers enjoy a 'quiet life' on account of their market power, and thus market power and uncompetitive banking systems breed managerial incompetence (Berger & Hannan, 1998). Increased competition in banking however drives down interest margins and profitability and this induces managers to improve operational efficiency so as to reduce cost in order to improve profitability. In this way, competition enhances managerial and operational efficiency (Schaeck & Čihák, 2008). Another channel through which competition enhances efficiency is that competition serves as a threat to managers of inefficient banks of losing market shares (Shirley & Walsh, 2000), as there is a re-allocation of market share from inefficient banks to more efficient banks (Boone, 2008b). This could be a direct threat from efficient incumbents or new entrants. Managers of inefficient banks are therefore incentivised to improve on their efficiency. In regards to deregulation-induced competition, it is argued that deregulation reforms open up the market place to new entrants; the superior management practices and know-how brought in by the new entrants will facilitate efficiency through technology and skill transfer.

The rather less-popular competition-inefficiency hypothesis, on the other hand, suggests that due to perception of ease of customer–switching in a more competitive environment, increased competition might result in less stable and short-term customer-bank relationships. This will amplify information asymmetry which requires additional resources for screening and monitoring of borrowers. Further, there could be limited reusability and value of information on account of the anticipated short-term nature of bank-customer relationships in a competitive environment (Schaeck & Čihák, 2008). These arguments suggest a reduction in the value of proprietary information held by banks, and thus banks incur greater costs in customer-retention efforts through huge investments in ATMs, new information systems, and aggressive marketing, which could constrain cost-efficiency.

This theory has however received little empirical support as an overwhelming majority of studies point to increased competition enhancing banking efficiency. The empirical literature on the efficiency-enhancing role of banking competition is extremely vast and it would be impossible to review it in its entirety, so we will focus only on some of the main contributions.

Berger and Hannan (1998) in a study of over 5,000 banks in the US find that banks in more concentrated and uncompetitive markets exhibited lower cost efficiency, lending support to the 'quiet-life' effect. The lower cost efficiency is attributable to non-minimisation of costs due to shirking of managerial responsibility, the pursuit of objectives other than profit maximization or managerial incompetence, which is obscured by the high profits resulting from the exercise of market power. (Schaeck and Čihák, 2008) also find evidence of the quiet life hypothesis in which competition has a positive effect on profit and cost efficiency in a study of 12,500 banks in the US and ten European countries from 1995–2005. They use Granger causality tests to establish a positive effect of competition on different measures of profit efficiency. The study further establishes that efficiency served as the transmission mechanism through which competition positively impacted financial soundness. In a more recent study of some European banks, and using different measures of competition, the authors find that bank capital and profitability increase as a result of accelerating competition, confirming that competition enhances efficiency (Schaeck and Cihák, 2014).

Ariss (2010) examines the impact of banking competition (proxied by market power) on banking efficiency in developing countries for the period 1999-2005. The results show a significant negative relationship between bank market power and cost efficiency, but the opposite effect on profit efficiency. In other words, when banks enjoy a greater degree of

market power, they do not manage their costs effectively, but are able to achieve higher profit efficiency levels. This confirms the notion that efficiency suffers in an uncompetitive banking environment due to the high price mark-up over marginal costs enjoyed by banks with significant market power.

We provide a detailed literature review of the impact of deregulation reforms on banking efficiency in Chapter 4. As detailed in that review, most improvements in banking efficiency arising from banking reforms pass through the channel of increased competition, pointing to the efficiency-enhancing role of banking competition.

3.2.3 Banking competition and financial access

Does competition in banking increase access of firms and individuals to finance, and what are the channels through which this can be achieved? Similar to the competition-stability and competition-efficiency debates discussed earlier, the impact of competition on access to finance is also a much debated subject in the literature. Here too we find two distinct viewpoints: the **market power hypothesis** and the **information hypothesis**. The market power hypothesis predicts a positive relationship between competition and access, and states that increased competition in banking leads to a reduction in the cost of finance and thereby increases credit availability and access by firms and individuals. The information hypothesis sees a negative relationship between competition and access as it contends that due to the presence of information asymmetries in banking, competition can reduce access to finance as increased competition makes it more difficult for banks to internalize the returns from investing in relationship banking (Love and Peria, 2012).

The market power hypothesis therefore sees the impact of competition on accessibility as a demand-driven phenomenon where cost of credit by firms is the main constraint to access. Thus, since an uncompetitive banking system is characterized by banks with significant market power who charge higher prices on loans, increasing competition will not only lead to lower prices through enhanced efficiency but also result in enhancing access to credit (Claessens, 2009). However, it can be argued that this would be the case if the cost of finance were the only or major constraint to access in the demand and supply of loanable funds. This seems to be the case in most African countries where cost of credit and interest rates spreads are very high, although there are other challenges such as lack of credit history, improper financial record keeping by firms, and low financial literacy on the part of individuals.

The information hypothesis theory, on the other hand, sees access to credit as a supply-side phenomenon which hinges on banks' long-term relationship-building incentives. In particular, banks with greater market power are deemed to have more of an incentive to establish long-term relationships with new firms and extend financing to such firms due to the benefits of long-term relationship banking. In that case, financial access is seen to increase with market power. Similar to the underlying rationale for the competition-inefficiency hypothesis discussed earlier, the perceived short-term relationships in a highly competitive environment is seen to constrain credit extension by banks. Petersen and Rajan (1995) argue that lenders are more likely to increase financing to firms in a more concentrated banking environment because it is easier to deepen relationship banking in such markets.

The empirical work related to the competition-access issue has also yielded mixed results. Fischer (2000) and Petersen and Rajan (1995) find that market concentration leads to more information acquisition and greater credit availability in the study of manufacturing firms in Germany and SMEs in the US respectively. Beck et al. (2004) in a cross-country study instead find that market power is associated with less access, especially for developing countries. As noted by Love and Peria (2012), the mixed results in empirical studies is partly attributed to the different measures of competition used, while differences in the nature of the countries and their levels of financial and institutional development seem to impact on the outcomes of these studies (Beck, Demirgüç-Kunt, and Pería, 2011).

A closely related issue in the competition-access debate has to do with credit access and quality of loan portfolio. It is argued that increased competition leads to more access not only on account of lower cost but due to weaker credit lending standards, as was observed in the US sub-prime mortgage market which triggered the global financial crisis. On the other hand, while concentration may reduce the total amount of loanable funds, it may also increase incentives to effectively screen borrowers, thereby enhancing the quality of the loan portfolio.

In Ghana, like other African countries, access to credit is constrained by high interest rates, a symptom of lack of competition in the banking sector, but also due to high interest rates on government securities from increased borrowing from the banking sector by government, and lack of credit reference infrastructure. Although the use of credit reference services is in its infancy stage, it is expected to yield significant benefits in the medium term. It is therefore anticipated that with improvement in the credit environment, fiscal discipline and strong macroeconomic, competition in the banking industry should impact positively in increasing access.

To sum up, the literature suggests that banking competition especially in the context of African countries is expected to have a positive impact on banking efficiency and increase financial access. This could arise either directly from competition pressures which induces or encourages efficiency in the production and allocation of financial services or from technology enhancements and innovations that are usually associated with the entry of foreign banks with such superior technology. While its impact on stability remains debatable, strong regulatory and supervisory framework can be adopted to minimize any potential instability that can be triggered by increased competition. We now turn our attention to examine policies that enhance competition in banking.

3.3 Pro-competition policies and impact on competition

Having reviewed the main theoretical and empirical literature on the desirability of bank competition for stability, efficiency and access to financial services, this section reviews the existing literature on pro-competitive banking policies, that is, policies that could be pursued to enhance competition in banking. Enhancing competition in banking has been one of the main goals of the numerous financial sector reforms worldwide (Delis, 2012).

However, a major question is whether these reforms are able to engender the expected level of competition. This section reviews the literature on policies that can be adopted to remove barriers to competition in banking and also empirical work on the response of such reforms on banking competition. This is very useful to understand the extent to which the recent financial reforms undertaken in Ghana are pro-competitive and also what the expected impact on banking competition could be. At the basis of competition policies are the same principles of liberalization policies: (i) interest rate deregulation; (ii) relaxing entry and exit restrictions; (iii) foreign bank entry; (iv) access to credit information and institutional environment.

3.3.1. Interest rate deregulation

Interest rate deregulation refers to the removal of government control of interest rates in the banking industry. Interest rate ceilings and floors on lending and borrowing activities as well as credit allocation have been used by most governments at some point in time with the view to maintain financial stability and support economic development. This usually results in

financial repression and leads to inefficient financial intermediation and constrained growth. The uncompetitive nature of interest rate controls is that it does not allow market fundamentals to determine interest rates and distorts efficient allocation of resources.

In a review of the impact of interest rates deregulation in China, Porter, Takáts, and Feyzioglu (2009) noted that interest rate deregulation in the early 1980s enhanced competition in the deposits markets, leading to a surge in interest rates. Soon after though the country experienced a banking system crisis, triggered by weak supervisory and regulatory control, non-existent capital adequacy requirements, inadequate classification and provisioning rules, underdeveloped interbank markets, and inadequate monetary and exchange rate policies. Interest rates controls were therefore reinstated until the weak financial infrastructure was addressed, to be liberalised again later on (Porter et al., 2009). The authors therefore conclude that effective regulation and supervision and strengthening the financial infrastructure are critical factors for the successful implementation of interest rate deregulation.

3.3.2 Bank privatisation

State ownership of banks is one of the direct forms of control that a government can have, along with credit rationing and interest rate control. While in some cases state ownership is the result of nationalisation following a banking crisis (e.g. Indonesia in 1998), it is most often the result of a conscious policy decision by the government. Such state control of banks facilitates the government's directed credit schemes and interest rate control indirectly as these policies are implemented through these state banks especially where state banks dominate the banking system. In addition, management incompetence and ineffective supervision of such banks breeds lack of competition which is transmitted to the entire banking sector. Privatisation of such state banks accordingly reduces governments' dominance of the banking sector and promotes competition in the banking sector. We discuss in detail the theoretical framework and empirical work on the relative performance of different classes of banks by ownership status in Section 4.3 of Chapter 4.

3.3.3 Relaxing entry restrictions

This goes usually hand in hand with privatization. Market concentration in the banking system is often cited as a feature of an uncompetitive banking system, and arises due to regulatory restrictions on bank entry policies. Accordingly, relaxing entry restrictions, it is

argued, will lead to an increase in the number of market participants and promote competitive behaviour by reducing market power (Pasadilla and Milo, 2005). It is therefore inferred that lifting of entry restrictions for the banking industry are critical for competition to thrive in the sector.

However, some empirical studies do not support this assertion of a negative relationship between concentration and competition, and that competition outcomes might be observed in concentrated markets, while monopoly power might be sustained in un-concentrated markets (Claessens and Laeven, 2004; Casu and Girardone, 2006). Concentration is seen as a measure of market structure, while competition is a measure of market conduct. There can be competition in concentrated markets, if there is a credible threat of entry and exit (i.e., if markets are contestable). It is argued that the threat of entry and exit into the industry is the real driver of competition as it compels banks to operate competitively, rather than the actual number of market participants. Accordingly, restrictions on bank entry policies should be curtailed by relaxing regulations on licensing of new banks and appropriate exit mechanisms should be put in place to revoke the licenses of insolvent banks (World Bank, 2012). For most developing African countries with relatively less competitive banking systems, there is evidence to suggest that the opening up of markets to new entrants could be a major driver of competition (Beck and Cull, 2014).

In analysing bank entry conditions and competitive conduct in a cross-section of highly concentrated U.S. banking markets, Cetorelli (2002) found that entry of new banks, or the threat of it, improves competition. Estimating entry thresholds for a cross-section of U.S. local banking markets, the author finds no evidence of collusive behaviour of banks leading to maximization of joint monopoly profits, even in those markets with only two or three banks. Instead, the evidence shows substantial increases in the intensity of competition as markets see the entry of a third or fourth bank and gradual convergence toward more competitive behaviour as more banks enter. He concluded that by eliminating important barriers to entry, the process of deregulation in banking enhanced the conditions for market competition.

3.3.4 Foreign bank entry and opening up to foreign competition

Theoretically, foreign banks have the potential to enhance competition in the host country's banking sector through higher efficiency and improved quality of service delivery. Claessens and Laeven (2004) found that financial systems with greater foreign bank entry and fewer entry and activity restrictions tended to be more competitive. There is also some empirical

evidence that foreign bank penetration can result in the deployment of modern banking technology, superior risk management skills, high product quality, human resource development, and also stimulate the regulatory and legal frameworks of the host country (Hawkins and Mihaljek, 2001). Such attributes of foreign banks enhance the general competitive climate and compel local banks to also adopt policies and measures that will enhance their competitiveness. However, foreign banks can reduce the franchise value of domestic banks and lead to the creation of unstable local banking conditions. The recent financial crisis highlighted that there can be risks associated with cross-border banking and foreign banks penetration. Regulatory reforms are accordingly required to address challenges associated with foreign bank entry (Domanski, 2005); (Claessens and Van Horen, 2012).[19]

3.3.5 Access to credit information and institutional environment

The problems of adverse selection and moral hazard caused by information asymmetry between lenders and borrowers can be ameliorated through the creation of credit reporting institutions to provide credit information sharing services (Bruhn, Farazi, and Kanz, 2013). Such access to credit history information about potential borrowers also facilitates competition in the banking sector as the unavailability of such credit history information might enable incumbent banks to exercise market power and limit competition (World Bank, 2012). Greater disclosure of information regarding the terms of banking products will generate greater awareness by bank clients and promote bank competition. Thus promoting the establishment and operation of credit bureaus as well as consumer protection regulations and practices will enhance the information environment and influence the extent of bank competition.

Regarding the impact of pro-competition or deregulation reforms on competition in empirical work, a detailed review is provided in Chapter 4 on the impact on banking efficiency. As indicated earlier, the conduit through which deregulation impacts efficiency is mostly through competition or through technological spill-overs from new banks. Accordingly, we discuss in detail the impact of deregulation on efficiency (via these channels) in detail in the next chapter. We only briefly review empirical studies on the impact of banking reforms on competition in African countries as well as studies on Ghana. We also discuss the models of competition used in these studies, to set the stage for the detailed discussion of competition

[19] We discuss in detail theories underlying foreign and domestic banks as they relate to efficiency in Chapter 4.

models in the next section. Kasekende et al. (2009) analyse banking competition in Africa's four largest economies: South Africa, Algeria, Nigeria and Egypt, following different financial reforms pursued in those countries. They use the Panzar-Rosse H-statistic and the Conjectural Variation measures of bank competition, and found that reforms in general had positive impact on competition notwithstanding the contrasting approaches to financial reforms implemented in each of these countries. Simpasa (2013) examines the degree of banking competition in Zambia following reforms such as entry of foreign banks and privatisation of the largest state-owned bank. The Panzar-Rosse H-statistic and Lerner index measures of competition are used on a panel dataset. The results show that the banking sector was characterised by monopolistic competition although increased foreign bank entry and privatisation of state ownership seemed to have heightened competition. Mwenga (2011) in a study of Kenya's banking sector during 1998-2007 notes that the implementation of universal banking and removal of restrictions on banking activity led to a reduction in industry concentration and presumably more competition with the concentration ratio (CR_4) reducing from 57% to 45% over the study period. The estimated Panzar-Rosse H-statistic of 0.38 suggests conditions of monopolistic competition. Like the above studies show, most deregulatory financial reforms have impacted positively on banking sectors although in most cases the impact is marginal.

In the case of studies on the impact of reforms on banking competition in Ghana, an early empirical work carried out by Antwi-Asare and Addison (2000) that uses data for the pre- and post- reform periods of 1980-1986 and 1990-1996 shows that the banking reforms under FINSAP had a positive impact on bank operational performance. The analysis is based on the computation of various financial ratios as indicators of operational efficiency. Profitability ratios, intermediation ratios, net interest spreads and net interest margins were found to have improved during the post-reform period compared to the pre-reform period. The study notes however that although the market share of total assets by the Top 4 banks had declined from 84% in 1980 to 65% in 1996, it was still high, and coupled with the high interest rate spreads concluded that competition in the banking industry was not strong enough. Recent empirical studies also suggest that Ghana's banking sector was not competitive in spite of the earlier reforms implemented under FINSAP.

Buchs and Mathisen (2005) analysed competition in Ghana's banking sector using the Panzar-Rosse model on data for the period 1998–2002. The study found that Ghana's banking sector exhibited a non-competitive market structure which hindered effective financial intermediation. They concluded that the industry was monopolistically competitive and

argued that the highly concentrated nature of the industry facilitated collusive behaviour by the large banks. In addition, the study observed that the economic cost of the non-competitive behaviour of banks was also exacerbated by the high domestic financing requirements of the government, which made it captive to the banks' behaviour and fostered inefficiency in the banking industry.

Using data covering 2000-2007, Biekpe (2011) empirically investigated the degree of bank competition using the Panzar-Rosse, POP and Conjectural Variation methods. The paper also found evidence that banks in Ghana were monopolistically competitive, and direct and indirect barriers to entry existed. The study alluded to the highly concentrated market structure of the banking system, the huge mandatory reserve requirements to accommodate government fiscal deficits and low penetration of new banks as factors which constrained banking competition.

It is worth noting that the reference periods of these studies do not fully cover the recent post-deregulation reform period to enable a meaningful assessment of the impacts of these reforms. Indeed the recent deregulation reforms seem to have addressed the competition-constraining challenges observed in these studies. As we noted earlier, concentration levels have significantly declined following the entry of new banks; while the relaxation of banking activity restrictions via the introduction of universal banking should have given banks a level playing field which is expected to facilitate competition. The removal of the huge mandatory secondary reserves also provides opportunities for increased competition in the loans market. In addition to the identified gaps in the literature above, there are several limitations to the empirical analyses described above. The time span is not long enough to properly cover for the post reform period especially if the effect of policies is lagged over time. The impact of changes in ownership is not analysed at all. There are also methodological limitations. The Panzar-Rosse measure of competition has been criticised as it always finds monopolistic competition and performs poorly as a model when it comes to measuring competition between time periods to assess policy impacts. Like other African banking studies, concentration ratios and the Panzar-Rosse model have been extensively used. We discuss in detail the various proposed models of competition in the next section, and explain the choice of the models we use with detailed justification given in Chapter 6.

3.4 Measurement of competition in banking

The literature shows different methods proposed for measuring banking competition. The somewhat conflicting results and ambiguous inference relating to the impact of competition on stability, and to some extent on efficiency, can also be attributed to the different measures of competition used. Banking competition is a complex phenomenon which cannot be easily and directly observed, and many models have been developed in the literature to attempt to measure it. Credible measures of banking sector competition are however crucial to accurately measure competition levels for an effective analysis of the impact of deregulatory policies. The accuracy and predictive power of the models will depend strongly on the precision of the measure used. Put differently, policy measures following from such predictions could be misleading if the measure of competition is inaccurate (Liu, Molyneux, and Wilson, 2013).

The literature on measuring competition in banking has evolved around two broadly distinct approaches. These are the Industrial Organisation (IO) and the New Empirical Industrial Organisation (NEIO) approaches. Under the IO approach, competition is measured indirectly by establishing relationships between market structure and bank performance and drawing inferences on the competitive conduct of banks from such relationships. The NEIO approach instead makes direct observations of conduct and then draws inference about what they might mean for structure (Dick and Hannan, 2010). We examine the main measures of banking competition under these two broad approaches as discussed in detail in Girardone, Molyneux, and Casu (2015).

3.4.1 Industrial Organisation (IO) approach

Earlier empirical research on banking competition was based on the IO approach which focussed on market structure and performance linkages to infer competitive behaviour. Sometimes referred to as the structural approach to competition, it relates the existence of competitive behaviour to the structure of the banking industry and the performance of banks. Two models developed under this approach are the Structure-Conduct-Performance (SCP) model and the Efficient Structure Hypothesis (ESH). These models link competition to market concentration and investigate whether a highly concentrated market causes collusive behaviour among larger banks which results in superior performance (SCP), or whether it is the efficiency of larger banks that enhances their performance (ESH) to grow market share.

(a) The Structure– Conduct– Performance (SCP) Model

The SCP model seeks to explain competitive conduct by relating market structure to performance of banks. It hypothesises that in a highly concentrated market, large banks are able to enjoy superior performance (such as higher profitability) through collusive and anti-competitive behaviour in pricing and other market practices. In this model, competition is indirectly inferred by examining the relationship between an exogenous market structure and the performance of banks. It argues that concentration weakens competition through collusive behaviour among large banks and that enables them to reap above normal profits (Liu et al., 2013).

A positive relationship between market structure (measured by the *n*-firm concentration ratio (CR_n) or the Herfindahl-Hirschman Index (HHI) and performance (measured by profitability) is interpreted as evidence of collusive behaviour of banks to achieve high profitability.[20] In assessing market structure, there is the need to define the market from which the concentration ratio is being measured as banks operate in different markets – retail banking vs. corporate banking, loan market vs. deposit market etc. However, due to lack of disaggregated data on business lines, most empirical work use deposit or loan market, or total banking market (total assets).

The standard SCP model is usually specified as follows:

$$\pi_{ij} = a_0 + a_1 CR_j + a_2 \sum BS_{ij} + a_3 \sum \mathrm{MKT_j} + \mathrm{u_{ij}} \qquad (3.1)$$

where: π_{ij} is the chosen profitability measure for bank *i* in market *j*[21];

CR_j is a measure of concentration or market structure in market *j*;

BS_{ij} is a vector of bank-specific variables for bank *i* in market *j;*

$\mathrm{MKT_j}$ is a vector of market-specific variables (interest rates, inflation, etc.) that can influence bank performance in market *j;* and

u_{ij} is the error term.

[20] **CR_n** is the sum of the market shares of the n largest firms in the industry and **HHI** is the sum of squares of the market shares of all banks in the industry, where market share is measured in terms of loans, deposits or total assets.

[21] Profit can be measured in terms of Return on Assets (RoA) or Return on Equity (RoE)

From the above model specification, a positive and statistically significant value of (a_1) is taken as an indication that banks operating in a concentrated market with greater market power adopt collusion and uncompetitive pricing behaviour to reap high profits. Another implicit assumption underlying the SCP model is the exogeneity of market structure and the one-way causality running from concentration to performance.

Against the background of the interpretation of a positive relationship between concentration and performance as evidence of collusive behaviour among large banks, the policy recommendation of proponents of the SCP model is for government intervention and regulation to check the abuse of market power as they tend to view most markets as imperfect in terms of their competitive structure.

This type of model was very popular in banking studies throughout the 1990s except that the model produced mixed and conflicting results (see Berger, Demirguc-Kunt, Levine, and Haubrich (2004) for a good review of the SCP approach). It has however been criticised from different points of view. First, the model is static in nature and therefore assumes that all observed characteristics are derived from long-term equilibrium conditions. Second, the exogeneity of market structure assumed by the SCP paradigm and the unidirectional causality from market structure to performance inherent in the model has been criticised. It is argued that market structure may not be exogenous but could be endogenously shaped by banks' conduct. For instance, banks providing better quality services can gain market shares, and this could change the market structure endogenously. In addition, a bank operating with a higher efficiency level will be able to achieve profitability and market share. Thus the positive relationship between market share and performance might not necessarily be as a result of anti-competitive collusion but be the result of increased efficiency, leading to the efficiency structure hypothesis, which we discuss below.

Another criticism of the SCP is that the model ignores banks' reaction to competitive pressures, that is, it does not take explicit account of the conduct of banks. It is also contended that it is not concentration levels or market structure which determines competitiveness, but market contestability, that is the threat of entry of new banks. Another criticism is the fact that the SCP model is a reduced form approach and assumes banks' pricing conduct could be modelled as a function of market structure without analysing any information on the underlying mechanism of such banks' pricing conduct, namely, the industry-specific demand conditions, firm specific demand and cost characteristics, and the interdependence among market participants (the degree of collusion).

(b) The Efficient Structure Hypothesis (ESH) Model

The Efficient Structure Hypothesis (ESH), sometimes referred to as the Chicago School model, also examines market structure-performance relationships but offers a competing reasoning behind the observed relationship.[22] According to ESH, a positive relationship between concentration and profitability may not necessarily be due to collusive anti-competitive behaviour but to greater efficiency of some banks, which translates into increased profitability and higher market share. As a consequence market structure is not exogenous but shaped endogenously by banks' performance, and concentration is the result of superior efficiency of some banks. Not surprisingly the policy recommendations arising from the ESH approach are at variance to those of the SCP. Government interference in the banking market is seen to constrain competition rather than enhance it and accordingly the approach does not recommend government intervention.

The ESH can be modelled by including in the original SCP equation a separate market share variable (MS_{ij}) alongside with the concentration variable, as follows:

$$\pi_{ij} = a_0 + a_1 CR_j + a_2 MS_{ij} + a_3 \sum BS_{ij} + a_4 \sum \mathrm{MKT_j} + \mathrm{u_{ij}} \qquad (3.2)$$

If the coefficient a_2 is positive and its introduction lowers the significance of a_1 this can be taken as an indication of the prevalence of the ESH over the SCP hypothesis. Berger (1995) however criticizes this interpretation as it does not explicitly account for direct measures of efficiency. The author proposes the inclusion of direct efficiency measures such as x-efficiency and scale economies in equation (3.2) and examines the sign and significance of the two coefficients in addition to a_1 and a_2 [see (Berger, 1995) for details].

Despite the popularity of these approaches in the early literature the issue of whether high profitability is attributable to collusion or to superior efficiency has not been satisfactorily resolved [(Dick and Hannan, 2010)]. A general criticism of both the SCP and ESH is that performance measures such as profitability, interest margins, etc. could be influenced by macroeconomic conditions (inflation, interest rates), regulatory environment (reserve requirements), and the legal and institutional framework (deposit insurance, quality of judicial system, credit reference bureaus), rather than market structure (Demirgüç-Kunt, Huizinga,

[22] The model was developed by Chicago School Economists, including Stigler (1968) and Demsetz (1973).

and Claessens, 1998; Claessens and Laeven, 2004). In view of the limitations of the SCP model, its popularity in empirical work has declined in more recent empirical studies.

3.4.2 New Empirical Industrial Organisation (NEIO) approach

The New Empirical Industrial Organisation (NEIO) approach rejects the notion that competitive behaviour should be inferred indirectly through the relationship between market structure and banks' performance, as this does not take explicit account of the conduct of banks. The NEIO attributes competitive behaviour to factors other than market structure and thus competition should be directly measured by the actions of firms in response to competitive pressures. One of the major strengths of the NEIO approach is that it is grounded in microeconomic theory as it seeks to estimate competition using more flexible models of profit-maximising behaviour (Dick and Hannan, 2010). Different direct measures of competition have been developed under the NEIO approach and are discussed below.

(a) Panzar-Rosse H Statistic (P-R) Model

One of the most widely applied direct competition measurement models in earlier empirical banking work is the Panzar-Rosse H-statistic (P-R) model, which was pioneered by Panzar and Rosse (1987). It is a reduced form model which measures competition directly by examining the effect of changes in input prices on revenues of a profit-maximising firm in long run equilibrium. The theory underlying the P-R model is that changes in input prices of a profit-maximising firm affect output prices and quantities sold (and hence revenues) differently depending on the competitiveness of the industry. The model therefore examines the elasticities of total revenue to changes in input prices to measure the degree of competition.

The H-statistic is accordingly modelled by estimating the reduced-form revenue equation as follows:

$$\log R_{i,t} = \alpha + \sum_{j=1}^{J} \beta_i w_{ji,t} + \Theta' X_{i,t} + U_{i,t} \qquad (3.3)$$

where:

$\log R_{i,t}$ is the log of gross revenue of bank *i* in year *t*;

$w_{j\,i,t}$ is the price of input factor *j* for bank *i* in year *t*;

$X_{i,t}$ is a vector of exogenous control variables; and

$U_{i,t}$ is a random disturbance term.

The H-statistic is computed as the sum of partial elasticities of revenue to input price changes:

$$H = \sum_{j=1}^{J} \beta_j \tag{3.4}$$

where j = 1,…, J is the number of inputs included in the model

In its application to banking studies, a bank's output is usually defined as loans, total earning assets, or total assets (using the intermediation approach), and hence output prices are proxied by interest income on loans, total interest income or gross revenue.[23] Banks' inputs are usually defined as deposits, labour and fixed assets, and hence input prices are proxied by interest expense on deposits, personnel costs and other non-interest cost.

Based on the explicit assumption of profit maximizing behaviour of firms, and assuming the market is in long run equilibrium, it can be shown that under pure monopoly, total revenue decreases when input prices increase. Since the monopolist operates at the price elastic portion of the demand curve, an increase in output price, in response to an increase in input prices, leads to a more than proportionate fall in quantities sold, and hence total revenue falls. Thus the H-statistic is negative (H < 0) for pure monopoly. In the case of a monopolistic competitive market, an increase in input prices increases average and marginal costs and leads to the exit of loss-making firms and an increase (but a less than proportionate increase) in total revenue, hence 0<H<1. Under conditions of perfect competition, due to the free entry and exit assumption, an increase in input prices causes total revenue to increase proportionately, with H = 1.

The H-statistic definition for the competitive environment is summarised in Table 3.1.

[23] The literature distinguishes between the production approach and intermediation approach in identifying outputs and inputs in the banking model. However many models use the intermediation approach in which deposits are considered as inputs, together with labour and physical capital. Outputs are loans, earning assets or total assets.

Table 3.1 Panzar-Rosse H-Statistic and competitive conditions

H-Statistic	Competitive Condition
$H < 0$	Pure Monopoly
$0 < H < 1$	Monopolistic Competition
$H = 1$	Perfect Competition

The P-R model has been widely used in empirical research including studies on African banking systems. The popularity of the P-R model is due to its simplicity and less stringent data requirements as a reduced form equation which requires only input prices and revenues. For banks in developing countries with relatively smaller banks and few observations, the P-R model is a very convenient approach to use, and probably explains its overwhelming use in the African banking competition literature. Further, since only data from banks included in the sample are required to estimate revenue equations, it facilitates cross-country studies (Claessens and Laeven, 2004).

Despite its simplicity, and probably because of it, the P-R model suffers from quite a few shortcomings that have eventually led to a sharp decrease in its popularity in academic work. First, as a reduced form model, it suffers the same shortcoming of a non-structural model as the SCP and does not provide information on the underlying mechanism of competitive behaviour and interdependence among market participants. Another criticism relates to the assumption of long-run equilibrium, which means that (risk-adjusted) returns are not statistically significantly correlated with input prices for the results to be plausible (Shaffer, 1982). According to Bikker, Shaffer, and Spierdijk (2012), the coexistence of firms of different sizes within the same market is strong evidence of disequilibrium which undermines the reliability of the P-R test, making the P-R unsuitable for firms of widely differing sizes within a single industry.

This has been traditionally addressed by an explicit testing for long-run equilibrium condition in the market prior to estimation of the H-statistic but has been criticised since banks are not usually observed in long-run equilibrium due to the static nature of the model.[24] As pointed out by Goddard and Wilson (2009), the static nature of the model leads to mis-specification bias, but the use of recent advances in dynamic panel data modelling seems to overcome this

[24] The test consists of replacing total revenue by profit rates in the equation, since profits are not correlated with the input prices in the long-run equilibrium.

shortcoming of establishing long run equilibrium and could improve the predictive power of the model (Goddard and Wilson, 2009).

The econometric interpretation of the magnitude of the H-statistic has also been questioned as theoretical studies show that the H-statistic can be negative even under highly competitive conditions and positive for a monopoly in contrast to the set standard criteria above, while challenges in interpretation of changes in the value of the H-statistic makes it difficult to analyse the impact of policy changes on competition as the H-statistic does not assume a continuous nature Leon (2015). Thus an increase in H from 0.4 to 0.6 is no indication of increasing or decreasing competition except to say that it exhibits monopolistic competition.

(b) Conjectural Variations Model and the Lerner Index

The Conjectural Variations (CV) model is attributed to the pioneering work of Iwata (1974), Bresnahan (1982) and Lau (1982). The model is founded on oligopoly theory, on the basis that profit-maximising firms can influence prices and output and thus can influence the behaviour of competitors due to their interdependence. Such interdependence among firms in the market means that each firm formulates its pricing and output strategies in anticipation of the potential or likely reactions of its competitors. Thus, each firm anticipates or conjectures competitors' reactions to its price-output decisions. Such anticipated reactions of other firms in the industry to a firm's pricing and output decision are termed "Conjectural Variations".

The estimation of the model and the computation of the CV parameters however require the estimation, in addition to the profit-maximising condition, of a market demand function and a cost function to derive the marginal cost function embedded in the profit maximisation condition. The Conjectural Variation model is thus a structural model based on estimating a system of equations involving a market demand function, a cost (supply) function and a specification of the interdependence of market participants (the degree of collusion). The degree of competition in the market is evaluated on the basis of the estimated interdependence of market participants.

The modelling of the CV is as follows:

Assuming firm i produces output q_i, with total industry output Q; then the market demand is given by

$$p = p(Q, z) \tag{3.5}$$

where z is a set of exogenous variables affecting demand, and the cost function is given by

$$c_i = c_i(q_i, w_i) \tag{3.6}$$

where w_i is a vector of prices of inputs employed by firm *i*.

The profit maximisation condition of firm *i* firm is given by:

$$\text{Maximise } \pi_i = p(Q, z)q_i - c_i(q_i, w_i) \tag{3.7}$$

On the assumption of a homogenous commodity market where the firm seeks to choose output level q_i to maximise profits, solving the above profit maximisation condition for firm *i* yields[25]:

$$\frac{d\pi_i}{dq_i} = p + f'(Q)\,(\frac{dQ}{dq_i})q_i - c_i'(q_i) = 0 \tag{3.8}$$

which can be re-written as

$$p + f'(Q)\,(1 + \lambda_i)q_i - C'(q_i) = 0 \tag{3.9}$$

where $\lambda_i = d\sum_{j\neq i}^{n} x_j / dx_i$ is the conjectural variation of firm *i* with respect to all other firms in the industry, that is the change in output of all other firms anticipated by firm *i* in response to an initial change in its own output dx_i.

For instance, in the case of *perfect competition,* an increase in output by firm *i* has no effect on market output, so that $^{dQ}/_{dq_i} = 0$, and thus $1 + \lambda_i = 0$ or $\lambda_i = -1$.

[25] A similar derivation can be made where the firm chooses price as the decision variable on the assumption of a heterogeneous commodity market. However, most empirical work on the CV model tends to assume output as the decision-making variable.

Under *Cournot oligopoly*, a firm does not expect retaliation from its competitors, when it decides to increase its own output, so that an increase in output by firm *i* leads to an increase in total industry output by the same amount: thus, $^{dQ}/_{dq_i} = 1 = (1 + \lambda_i)$ and therefore $\lambda_i = 0$.

Where there is *perfect collusion* among firms in the industry (acting like a cartel or pure monopoly), firm *i* will expect full retaliation from its competitors, so as to protect their market share in response to an increase in output by Bank *i*: hence $^{dQ}/_{dq_i} = {}^{Q}/_{q_i} = (1 + \lambda_i)$, thus an increase in output by Bank i by one unit leads to an increase in market output by $^{Q}/_{q_i}$ units so that $\lambda_i = (Q - q_i)/q_i = \sum_{j \neq i}^{n} q_j / q_i$.

If we define $\eta \equiv -\left(\frac{p}{Q}\right) * (\frac{\partial Q}{\partial p})$ as the market price elasticity of demand for output; and $\theta_i = (1 + \lambda_i)(\frac{q_i}{Q})$ as the degree of collusion in the industry, then equation (3.9) can be expressed in the form:

$$\frac{p - c_i'(q_i, w_i)}{p} = \frac{1}{\eta}\theta_i \qquad (3.10)$$

In equation (3.10),

$\frac{p - c_i'(q_i, w_i)}{p}$ is defined as the Lerner Index (LI) which is a measure of a bank's market power by measuring deviation of price from marginal cost. We shall return to the LI shortly.

From the relation $\theta_i = (1 + \lambda_i)(\frac{q_i}{Q})$, the CV parameter, λ_i can therefore be interpreted by the degree of collusion in the industry parameter, θ_i, in defining the competitive condition in the industry. For the specific of values of λ_i identified above we can find corresponding values of θ_i. Where $\lambda_i = -1$ (the case of perfect competition), $\theta_i = 0$; where $\lambda_i = 0$ *(Cournot oligopoly),* θ_i will be equal to q_i/Q; and in the case of *perfect collusion*, with $\lambda_i = \sum_{j \neq i}^{n} q_j / q_i$ means θ_i will take on the corresponding value of 1. This can be summarised in the Table 3.2.

Table 3.2 Conjectural Variation parameters and competitive conditions

Equilibrium Situation	CV parameter, λ_i	Degree of Collusion, θ_i
Perfect Competition	-1	0
Cournot Oligopoly	0	q_i/Q
Perfectly Collusive oligopoly (Pure Monopoly)	$\sum_{j\neq i}^{n} q_j/\, q_i$	1

As shown in Table 3.2, the range of values that θ_i can take lies between zero and one which corresponds to the opposite ends of the perfect competition – pure monopoly spectrum. Accordingly, where firms have θ_i statistically not significant from 0, then the hypothesis of competitive behaviour cannot be rejected.

A major advantage of the CV model is that by being a structural model, it does not only offer a measure of competition but also provides insights into the sources of the estimated competitive conduct as it models the explicit demand, cost and profit-maximising conditions faced by banks, and estimates the degree of competition from the nature of the equilibrium of the banks. Another advantage attributed to the model is that θ_i is a continuous variable and so we can be used to measure the trend in competitive conditions over a period of time. Accordingly, it is useful in analysing the impact of policy changes on competition. Notwithstanding these advantages, the main challenges with the CV model is the data intensive requirements of the model and the functional form the structural demand and supply equations should take. For instance, should marginal costs be proxied by average costs, or be derived from the estimation of a translog cost function.

The CV has not been applied widely in empirical banking studies. A few notable studies using this approach include Angelini and Cetorelli (2003), for measuring competition in Italian banks; Uchida and Tsutsui (2005) for analysing competition in Japanese banks; Brissimis et al. (2008) for assessing competition of newly acceded EU countries; and Kasekende et al. (2009) for banking competitiveness of Africa's 4 largest economies. These studies assume a homogenous-product quantity setting game and estimates the CV parameters in quantities as illustrated above.

As noted in equation (3.10) above, $\frac{p-c_i'(q_i,w_i)}{p} = \frac{1}{\eta}\theta_i$, the Lerner Index (L.I.), $\frac{p-c_i'(q_i,w_i)}{p}$ can be derived as part of the CV model. Some empirical studies explicitly compute the L.I. as a measure of competition. It is based on the profit-maximising equilibrium condition that price equals marginal cost in a perfectly competitive market so that any divergence between price and marginal is seen as a measure of the deviation from a perfectly competitive position. It is an indication of market power as it shows the extent by which a bank could charge price above marginal cost. The Lerner index has become a frequently used measure of competition in recent research work for reasons of simplicity and less stringent data requirements. Its major limitation is that it is a measure of pricing market power and not necessarily a proxy for competition while it does not offer insights into the sources of market power.

(c) Persistence of Profits (POP) Model

The persistence of profits (POP) model sees competition as a dynamic process in which the entry and exit of firms drives short run excess profits to their long run average levels. It therefore measures the degree of competition from changes in the level of the persistence of profits over time. The speed of adjustment or convergence of firm- or industry-level profits to the long-run values reflects the degree of competition in the industry: the faster the speed of adjustment, the higher the intensity of competition. Conversely, the slower the speed of adjustment, the stronger the persistence of profits and hence the lower the level of competition. The POP model acknowledges that there is no guarantee that the profit and market structure observed at any point in time represents equilibrium, and therefore does not draw the competitiveness in the market place from the observable market structure nor relies on the long-run equilibrium concept of competition. It therefore rejects the static nature inherent in the SCP and the P-R models discussed above, and sees competition as a dynamic process.

The modelling of the POP in empirical work is usually carried out by estimating a first-order autoregressive AR (1) model:

$$\pi_{it} = \alpha_i + \lambda_i \pi_{it-1} + u_{it} \tag{3.11}$$

Where π_{it} is the profit of bank i at time t (for example, the return on assets); and π_{it-1} is the profit of bank i at time t-1. λ_i represents the strength of profit persistence and can be interpreted as follows: if $\lambda_i = 0$, there is no association between $\pi_{i,t}$ and $\pi_{i,t-1}$, and represents

perfect competition. If $0 < \lambda_i < 1$, there is a positive association between $\pi_{i,t}$ and $\pi_{i,t-1}$, or evidence of persistence of profit.

Empirical studies using the POP model include Genay, Udell, DeYoung, and Berger (2000) who used the model to measure banking competition in the US; Goddard, Molyneux, and Wilson (2004) who assess profitability in European banks; Zhao et al. (2010) in measuring competition in India's banking sector. In the case of studies on Africa, Biekpe (2011) and Poshakwale and Qian (2011) used this model to measure competitive conditions in Ghana and Egypt respectively. The need for and sometimes the challenge of obtaining sufficient time series banking data is attributed to the rather limited use of this model in empirical banking research studies. Various specifications for profitability are used in the POP model depending on whether competition is being measured in the loan market (in which case loan overcharge, proxied by the ratio of implicit price of loans to marginal cost of loans is used); or the entire banking industry (overall return on assets or return on equity is used).

(d) The Boone Indicator

A more recently developed measure of competition is the Boone indicator (BI), which is based on a slightly modified version of the efficiency structure hypothesis discussed earlier. The basis of this model is that efficient firms are highly rewarded in more competitive markets (Boone, 2008a). The BI measure therefore asserts that competition drives performance but benefits efficient firms more than inefficient firms. Efficient firms are able to grow their businesses and thereby increase their market share at the expense of less efficient banks, and that these growth and reallocation effects are greater, the stronger the intensity of competition in the market. As competition intensifies, the profit and market share levels of the more efficient banks increase relative to those of less efficient banks.

The BI is estimated from the following profit equation:

$$ln\, \pi_i = \alpha_0 + \beta ln\, MC_i + \varepsilon_i \qquad (3.12)$$

where

π_i is the profit of Bank i

MC_i is marginal cost of Bank i, which is a proxy for efficiency

β is the BI, profit elasticity w.r.t. cost and a measure of competition

The estimated β should be negative as higher efficiency levels are associated with lower MC and higher profit levels. The magnitude of β measures the level of competition such that the higher the absolute value ofβ, the greater the intensity of competition.

Although profits and marginal costs can be estimated in levels (to allow for negative profit levels), the benefit of the log specification (which ignores negative profit values though) is the ease of interpreting the coefficient as elasticity. Empirically, the computation of marginal cost can be derived from the estimation of a cost function.

One of the advantages of the Boone indicator is the fact that the relationship between costs and profits is continuous, so that higher competition implies that the value of β is larger in absolute terms. Second, it is easier to estimate as it can be obtained by a simple linear econometric specification. Third, it requires data on only profits and costs, and makes it suitable for developing countries, where data can be a challenge. One of its unique attributes is that it is possible to use it to evaluate competition in different markets such as the loans market. Another appealing property is that it can measure competition on an annual basis so as to track the evolution of competition over time. It is not without criticism though. The assumption that higher efficiency leads to increase in market share without reducing prices means efficiency gains are not even passed onto customers (van Leuvensteijn, Bikker, van Rixtel, and Sørensen, 2011). Secondly, it is argued that the empirical relationship between a firm's efficiency and its profitability does not always correspond to Boone's theoretical framework, where the most efficient firm is always, by design, also the biggest.

The Boone measure is gaining attention in empirical work. For instance, van Leuvensteijn et al. (2011) and Schaeck and Čihák (2008) used the Boone indicator in measuring competition in the Euro area; while Xu, Rixtel, and Leuvensteijn (2014) used it as one of the models in assessing bank competition in China. The Boone indicator has also been used in cross-country studies, notably by Delis (2012) and Clerides, Delis, and Kokas (2015). The review of the above models of competition measurement shows that there is no consensus on the best model for empirical estimation of banking competition and that the selection of a particular model depends largely on the purpose of the research study, the underlying research questions being investigated, the availability of data and the nature of such data – bank level or industry data. Given the differences in the theoretical underpinnings of the models, it is not surprising that these models sometimes provide conflicting results in empirical work regarding the level and pattern of competition observed. As a consequence, in order to improve robustness of results, researchers are increasingly focussing on the adoption of two or more measures of

competition, as evidenced in recent empirical banking competition studies (see for instance, Demirgüç-Kunt and Martínez Pería, 2010; Delis, 2012; Clerides et al., 2015; Xu et al., 2014; Brissimis, Iosifidi, and Delis, 2014). This approach has also been applied in recent African banking competition studies as in the case of Kasekende et al. (2009); Biekpe (2011); Poshakwale and Qian (2011); Mwega (2011); Simpasa (2013); and Leon (2015). It is worth noting that while in some cases the empirical results based on different competition measures could have different outcomes, in other cases they do tend to show similar results. For instance, Liu et al. (2013), in an empirical estimation of competition in nine EU banking markets observe that the Lerner index and the P-R models showed a strong correlation compared to other models while the Boone indicator did not correlate with these competition models. Xu et al. (2014) however point out in an empirical assessment of competition in China that the Boone model provided a more accurate measure of competition compared to the conventional Lerner Index and P-R measures. On the other hand, Clerides et al. (2015) in estimating competition in the banking sectors of 148 countries over the period 1997–2010 found that the three methods used: the Lerner index, the adjusted Lerner index, and the Boone indicator produced similar patterns of competition over time. While cross-country and country-specific empirical applications of different models yield conflicting and inconclusive results, the authors emphasize that it is important to realise that different measures can yield different outcomes, and thus the need for not only the use of multiple models but also for the same interpretation of the type of competition being measured.

Thus, this multiple-model approach does not only contribute to corroborate results of different models but also helps to unearth the underlying reasons for conflicting empirical competition outcomes. In line with this trend, we will employ two measures, the POP and the BI models for our empirical study of the impact of the reforms on competition in the banking sector. We provide detailed justification for the choice of these models in favour of the extensively used P-R model used in most African banking studies in Chapter 6.

3.5 Conclusions

This chapter has provided a review of the literature on different aspects of banking competition. The pros and cons of banking competition on stability, efficiency and financial access were examined. As noted in the review, many theoretical papers and empirical research sought to analyze the (sometimes rather ambiguous) impact of competition on the stability of the financial system, efficiency of the banking system and access to banking services and

credit. In spite of the different theoretical views on the effect of banking competition, most of the empirical literature supports the efficiency-enhancing role of banking competition as well as its positive impact on enhancing financing access. While the competition-stability vs. competition-fragility theories is an ongoing debate, there seems to be some consensus that the outcome depends very much on the regulation, supervision and other institutional factors in a particular country (Delis, 2012; Beck et al., 2013). The review also shows that deregulatory or financial liberalization policies in general are pro-competitive in nature, and are usually expected to enhance efficiency by creating a more competitive banking environment. The deregulatory reform policies pursued in Ghana are therefore pro-competition and were intended to stimulate competition and enhance banking efficiency. In Chapter 4, we discuss in detail the underlying theories and empirical work regarding the impact of deregulation policies on banking efficiency. The review of the measurement of banking competition also shows that most recent empirical studies employ the direct measures of competition in the NEIO framework. In our empirical study on competition in Chapter 6, we use the POP and BI models and provide justifications for their use.

CHAPTER 4 LITERATURE REVIEW ON THE IMPACT OF DEREGULATION REFORMS, OWNERSHIP AND SIZE ON BANKING EFFICIENCY

4.1 Introduction

This chapter reviews relevant literature on the relationship between deregulation and banking efficiency, and the role of bank ownership and bank size on banking efficiency. Broadly, we look at three main streams of literature relevant to this study. These are: (i) those relating to the impact of deregulation policies on bank efficiency; (ii) those that attribute efficiency differences to differences in ownership types: state-owned banks, private domestic banks and foreign banks; (iii) those relating differences in efficiency to size: large vs. small banks. We discuss both theoretical frameworks and empirical work across developed, transition economies, developing countries including African countries, as well as on Ghana. Section 4.2 discusses the theoretical framework and empirical work on deregulation reforms and banking efficiency. In section 4.3, we discuss theoretical perspectives and review of empirical work on ownership type and banking efficiency. This is followed by a review of the theory and empirical literature on bank size and efficiency in Section 4.4. We summarise and conclude this chapter in section 4.5.

4.2 Deregulation and banking efficiency

4.2.1Theoretical perspectives

The financial sector is the most highly regulated sector in any economy, with the banking industry being the most heavily regulated industry in most countries (Casu, Girardone, and Molyneux, 2015). Accordingly, the role of banking regulation (and deregulation) in facilitating a sound, competitive and efficient financial system is of paramount importance. Banking sector regulation stems from the need to safeguard depositors' funds, ensure stable banking systems, facilitate security and stability of the payments system, and to prevent or minimise bank failures. The positive effects of banking stability as well as the pervasive negative effects of banking crises on national and regional economies are well known. These underscore the important role of regulation and deregulation reforms.

There are two broad underlying theories regarding banking regulation: the public interest theory and the private interest theory. The public interest theory to regulation argues that governments regulate or intervene in the banking industry to facilitate the efficient functioning of the industry in the interest of the public and to mitigate possible bank failure (Kroszner and Strahan, 1999). The existence of market failures in the banking industry is the rationale for government regulation, according to this theory, but with the assumption that government has the incentives and capabilities to eliminate or minimise such market failures and maximise social welfare. Regulation is therefore expected to exert a positive effect on bank behaviour and performance by influencing competition, enhancing efficiency and facilitating a better functioning of the banking sector for the good of the public and to the benefit of the economy as a whole. Deregulation on the other hand will have an adverse impact on the banking sector, according to this theory.

The private interest theory to regulation, while acknowledging the existence of market failures, sees regulation as a product of the private interests of various stakeholder groups. These stakeholder groups lobby to influence the setting of banking regulations (Kroszner and Strahan, 1999; Barth et al., 2006). The crucial role of banks in financial resource allocation makes the sector attractive to different interest groups, and so the well-organised lobby groups try to use the coercive power of the state to influence banking regulations in their favour.[26] This theory therefore suggests that regulation is used to foster the interest of only the few in power or well-organised groups to the detriment of the wider society. This creates distortions, which hinder the competitiveness of the banking sector and negatively affects efficiency of banks as they constrain socially optimal resource allocation in favour of special-interest groups. From the private interest view therefore, one would expect regulation to impede efficiency, rather than enhance it. Conversely, deregulation is expected to impact positively on the efficiency of the banking sector, according to this theory.

Following from these broad theoretical perspectives on regulation (and deregulation), we relate them to the three specific areas affected by the deregulation policy reforms undertaken in Ghana. These relate to removal of banking activity restrictions, relaxation of bank entry, and easing of credit controls through abolishing of secondary reserve requirements.

[26] These interest groups could be bankers themselves through their Association, non-financial sectors such as Industrial Associations, Trade Association, Small and Medium sized enterprise Associations, and government itself.

(a) Banking activity restrictions and banking efficiency

Following the work of Barth et al. (2006) and Abiad et al. (2010), we summarise below the expected impact of deregulation in these specific areas on banking efficiency from both the public interest and private interest viewpoints.

The public interest view favours banking activity restrictions (restricting banks from engaging in non-core banking services such as securities, insurance and real estate activities). It argues that engaging in such broad financial service activities could create conflict of interest among different business segments, give banks more opportunities for risk-taking activities, and can lead to the creation of financial conglomerates which are difficult to monitor, discipline or "too big to fail". Furthermore, these large financial conglomerates may engage in anti-competitive behaviour, which will reduce competition and adversely affect efficiency of the banking sector. Activity restriction is therefore expected to positively impact banking efficiency; while deregulation of activity restrictions will have an adverse influence on efficiency.

The private interest view on the other hand argues that banking activity restriction is detrimental to banking efficiency, as it constrains possible exploitation of economies of scale and scope opportunities from broad financial services provision. The restrictions could limit risk and income diversification opportunities, reduce the franchise value of banks, lower incentives for prudent behaviour by banks, and thereby create greater instability. Activity restriction is therefore seen as constraining efficiency while removal of such restrictions deregulation is seen as enhancing efficiency, according to this theory.

(b) Entry restrictions and banking efficiency

The public interest view sees bank entry restrictions as enhancing the efficiency and stability of the banking system by preventing entry of dubious banks, who can exploit bank customers through fraudulent activities and expose the banking sector to contagion risk and fragility. Excessive entry of even genuine banks could drive down loan interest rates through excessive competition and encourage more risk-taking behaviour, which could adversely affect the efficiency and stability of the banking system. Furthermore, bank entry restrictions can potentially increase the franchise power of banks, and thereby motivate banks to behave more prudently (Hellmann et al., 2000; Keeley, 1990).

The private interest view argues that entry restrictions create an uncompetitive banking system, as incumbent banks seek to exploit customers to make excessive profits in the absence of threat to entry. Existing banks will therefore favour entry restrictions in order to limit potential competition from new entrants. Deregulation of bank entry is therefore expected to enhance market contestability, encourage competition and enhance efficiency of banks in general. Competition has been found to be very important in facilitating efficiency improvements in the banking sector (Berger and Humphrey, 1997).

(c) Restrictions on lending through high reserve requirements and banking efficiency

In most developing economies, banks are usually mandated to hold high reserves in the form of government securities or bonds, as an indirect means of financing government debt. The public interest view favours such domestic financing of government debt as an indirect means of financing important development projects for the benefit of society. It also sees the holding of such reserves as a means of controlling potential risk-taking behaviour of banks through excessive credit expansion. The private interest view however argues that bank financing of government debt crowds out private sector financing. Although some level of reserves might be necessary for liquidity and prudential purposes, excessively high reserves over and above reasonable levels for prudential purposes constrain lending to the private sector, limits intermediation role of banks and could restrain efficiency (Abiad et al., 2010).

Similar to these theoretical distinctions, there is little consensus on which theory dominates empirically in promoting banking efficiency, as regulatory (or deregulatory) policies that enhance competition and efficiency of banks in one country may have opposite outcomes in another country (Barth et al., 2013). As will be seen from the empirical review below, both theories seem to hold in particular cases and shows that the empirical evidence of deregulation reforms on banking efficiency is mixed, although there seems to be strong empirical support of a positive impact of deregulation on banking efficiency of banks in most cases. There are however situations where there is no evidence of any impact of such reforms on efficiency, and cases where there are conflicting impacts between the short-term vs. medium term as banks react and adjust differently to these policies over time.

4.2.2 Empirical literature on deregulation and banking efficiency

There is an enormous empirical literature on the study of deregulation reforms and bank efficiency especially on developed economies as they embraced deregulation and other forms of banking reforms earlier in their financial development. A comprehensive review of this

literature would be impossible, given its size, and we refer the reader to see Ferrier and Lovell (1990), Berger and Humphrey (1997), Maudos, Pastor, and Serrano (2000), Kumbhakar, Lozano-Vivas, Lovell and Hasan (2001); Berger and Mester (2003); just to mention a few. We review selected country-specific empirical work on developed countries, which are relevant to this study. We also review some empirical work on transition economies of Central and Eastern Europe (CEE) and developing countries (notably in Asia and Latin America).[27] We conclude the section with a review of empirical literature on African countries.

The review of the empirical work (summarised in Table 4.1) shows that deregulation had varied impacts on banking efficiency – in some cases it influenced positively by enhancing efficiency, but in other cases, the impact was negative. We also observe that in a few cases deregulation did not have any impact on efficiency, or had a non-uniform impact over time or across all the policies.

(a) Empirical studies in the United States

In the case of the United States (US), deregulation policies implemented since the mid-1970s include interest rate deregulation on deposits in the early 1980s, relaxation of branch restrictions to allow inter-state banking and branching[28], and removal of barriers to facilitate integration of banking and insurance businesses[29], among others.[30]

Elyasiani and Mehdian (1995) analyse the impact of deregulation on banking efficiency and the relative efficiencies of large and small banks between 1979 and 1986 as proxies for the pre- and post-deregulation periods respectively. Using nonparametric linear programming techniques to measure overall technical efficiency, the paper finds that banking efficiency was relatively unchanged between the two periods by the deregulation of interest rates in the early 1980s.

[27] We note that the state of banking sector in the CEE countries is markedly different from other developing countries as those countries moved from centrally planned economies to market-based economies. We review these studies as some of the deregulation policies are of relevance and similar to those of developing countries, including African countries.
[28] under the Interstate Banking Efficiency Act of 1994
[29] under the Financial Service Modernization Act, also known as the Gramm-Leach-Bliley (GLB) Act of 1999
[30] See Kroszner and Strahan (2014) for a detailed discussion on the origins of the regulatory and deregulatory changes in the US.

Table 4.1 Summary of empirical literature on impact of deregulation on banking efficiency

Author	Country/region study period	Deregulation policies implemented and examined	Impact on efficiency
Elyasiani and Mehdian (1995)	US 1979–86	Interest rate deregulation	No change
Hughes, Lang, Mester and Moon (1996)	US 1994	Relaxation of branch restriction	Positive
Humphrey and Pulley (1997)	US 1977–88	Interest rate deregulation	Negative
Yuan and Phillips (2008)	USA 2003-05	Activity restrictions removal	Negative
Berg, Forsund, and Jansen (1992)	Norway 1980-89	Relaxation of interest rate and lending restrictions	Positive
Kumbhakar, Lozano-Vivas, Lovell and Hasan (2001)	Spain 1986-1995	interest rates liberalisation; removal of credit ceilings; lowering of reserve requirements; relaxation of branch restrictions	Negative
Girardone, Molyneux and Gardener (2004)	Italy 1993-1996	Privatisation of public-owned banks; regulatory changes to facilitate bank consolidations through mergers and acquisitions	Positive
Sturm and Williams (2004)	Australia 1988–2001	Removal of restrictions on foreign bank entry	Positive
Fries and Taci (2005)	Eastern Europe 1994–2001	interest rate liberalisation; opening up of banking systems to private domestic and foreign banks; privatisation of state banks.	Initially positive; but negative later
Yildirim and Philippatos (2007)	Central and Eastern Europe 1993–2000	liberalisation, privatisation, foreign participation	Positive

Fang et al. (2011)	South Eastern Europe (SEE) 1998–2008	privatisation of state-owned banks opening up of banking sector to foreign investors foreign banks acquisition of state-owned banks	Positive
Chen et al. (2005)	China 1993–2000	Transfer of banking activities to state-owned policy banks. Joint-equity (public-private) banks established; marginal relaxation on foreign banks' participation.	Positive impact initially; but later declined
Dong (2009)	China 1994–2007	After China's accession to the WTO in 2001: gradual relaxation of restrictions on foreign banks' operations, including acquisition of domestic banks and product/geographical expansion	Negative initially; but rebounded after
Mahesh and Bhide (2008)	India 1985–2004	Reduction in reserve requirements; removal of activity restrictions; gradual reduction in government ownership of banks; deregulation of interest rates	Positive
Barros and Williams (2013)	Mexico 1998–2006	repeal of restrictions on foreign entry; acquisition by foreign banks allowed; domestic (A) bank consolidations; removal of activity restrictions with universal banking (B)	No impact by (A) Positive impact by (B)
Barros and Wanke (2014)	Brazil 1998–2010	Banking sector opened to foreign participation; privatisation of state banks to foreign entities	Positive

Hughes, Lang, Mester, and Moon (1996) examine evidence of large and increasing scale economies that apply to geographic expansion following the relaxation of branch restrictions in the US. Obtaining bank-specific risk-return estimates of banking efficiency relative to a stochastic risk-return frontier, the paper finds that depositor diversification geographically enhances expected return; and that an increase in branches improves efficiency by moving inefficient institutions closer to the efficient frontier in both the return and risk dimensions. The paper also established that the enhanced efficiency is due to both the diversification of risk due to spreading assets over a larger branch network as well as due to large and increasing scale economies.

Humphrey and Pulley (1997) analyse the impact of interest rate deregulation in the early 1980s on banking profitability and efficiency in the US. The policy led to higher funding costs and lower profits, and in response, banks raised service fees, cut branch-operating costs, and shifted to higher earning assets to offset reduced profitability. In analysing profit efficiency using a panel data covering the period 1977–88, the authors found that deregulation had an immediately negative impact on efficiency as measured bank productivity fell. The interest rate deregulation induced a scramble by banks to offer higher and competitive interest rates on consumer deposits. Accordingly, productivity benefits, which would have been captured by the banks, were passed onto customers at least in the short term. It was also observed that large banks were able to adjust output prices and input usage to mitigate the adverse impact of deregulation. However, the adjustment to deregulation was essentially completed after four years, with subsequent changes in bank profitability during the late 1980s driven primarily by changes in the business environment. Thus, deregulation had a negative short-run effect on banks' profit efficiency.

Yuan and Phillips (2008) investigate efficiency effects from possible economies of scope arising from the integration of the banking and insurance businesses following the passage of the Financial Service Modernization Act in the US in 1999, which largely removed the barriers on the separation of banking, and insurance businesses. Using multi-product revenue, cost, and profit functions, the authors find empirical evidence of significant revenue scope economies, but rather cost scope diseconomies and resultant weak profit scope economies in the post–deregulation period. The results suggest that the cost savings from sharing inputs generally do not offset the extra costs possibly incurred in joint production of banking and insurance businesses. The revenue scope economies suggest consumption complementarities on the demand side, which implies efficiency gains arise through cross-selling of banking and insurance products especially at the retail end. The weak profit scope economies are due to the offsetting of revenue scope efficiency gains by the cost scope efficiency losses. However, the authors allude to the fact that in spite of the statistically significant efficiency gains due to the integrated banking and insurance businesses, the economic significance seems small. The study therefore opines that only the most well run financial institutions would be able to achieve the benefits of integration and conglomeration. Thus, as they put it "the allure of universal bank still exists – it is just not easy to achieve (Yuan & Phillips, 2008)".

(b) Empirical studies in Europe

In the case of Europe, Berg, Førsund, and Jansen (1992) investigate the impact of deregulation of the Norwegian banking industry during 1980-89. With the volume of credit and interest rates on bank lending strictly regulated prior to 1984, this regulation was gradually phased out from 1984 to 1988. The authors measure productivity growth in the banking sector using DEA techniques to construct a frontier production function and Malmquist indices to decompose productivity growth into both frontier growth and the spread of productivity levels. The authors find little productivity growth at the frontier during the period but substantial improvement in the relative efficiencies of most banks. They observe convergence of efficiency as the reforms-induced efficiency gains via productivity growth throughout the period occurred mainly among the least efficient banks, which indicates that the deregulation may indeed have created a more competitive banking industry. The authors' observation that productivity growth was particularly rapid in the largest banks also suggests that large banks were initially less efficient. Their improved efficiency was seen as preparation to meet anticipated competitive pressures from the impending common European market. The authors find productivity regress at the average bank prior to the deregulation (that is, 1980-83), suggesting that banks created idle capacity in anticipation of deregulation which they utilised immediately after deregulation while led to the rapid growth post-deregulation.

Kumbhakar et al. (2001) examine the impact of deregulation on the performance of Spanish savings banks during the period 1986-1995. The deregulation was carried out during the mid-1980s and was intended to keep the competitive edge of the Spanish banking and financial markets at par with other European countries following the removal of barriers to inter-country competition in financial services within the European Economic Community. The deregulation policies included interest rates liberalisation in 1987; removal of credit ceilings by 1990; lowering of reserve requirements in 1992; relaxation of branch restrictions on savings banks in 1989. Using panel data and a flexible variable profit function that incorporates time-varying technical efficiency, the authors establish declining levels of output technical efficiency along with a significantly high rate of technical progress and productivity growth.

Girardone, Molyneux, and Gardener (2004) analyse the determinants of bank efficiency of Italian banks during 1993-1996 following deregulation reforms during the early 1990s, in line with the European Commission Single Market '1992' programme. The deregulation reforms

included privatisation of public-owned banks and regulatory changes that facilitated bank consolidations through mergers and acquisitions, with the aim of improving competition and facilitating efficiency. Using stochastic cost frontier to measure cost efficiency, and employing the Fourier-flexible functional form on panel data for the period 1993–1996, the paper finds evidence of cost efficiency gains as the mean inefficiency levels of banks declined during the period for all sizes of banks, suggesting a positive impact of the deregulation reforms on banking efficiency. Scale economies are also found to be present and significant, an interesting observation as the reform programme sought to increase bank sizes through consolidation. The authors find bank efficiency to be positively related to capital strength but negatively related to non-performing loans, although they find no clear relationship between asset size and bank efficiency.

(c) Empirical studies in Australia

On Australia, Sturm and Williams (2004) analyse the impact of foreign bank entry on banking efficiency in Australia during the post-deregulation period 1988–2001. This deregulation policy markedly transformed Australia's financial sector from a highly regulated banking system with limited foreign participation to a deregulated banking system with 15 foreign-owned banks commencing operations in 1986.[31] Employing DEA, Malmquist Indices and SFA based on parametric distance functions, the authors establish that bank efficiency increased post-deregulation and find evidence to suggest that increased competition from the entry of foreign banks was a key contributor to the efficiency improvements. The authors establish that the main source of the productivity gains post-deregulation was technological change rather than technical efficiency.

The review of the above literature on the experiences of developed countries shows the diversity of the impact of deregulation policies such as removing branch restrictions, relaxing activity restrictions, removal of restrictions on lending volumes and interest rate controls, and relaxation of entry restrictions. The review also shows the conflicting outcomes of deregulation on banking efficiency as in some cases efficiency increased, in others it declined, and in others it remained unchanged. In particular, we observe the positive impact on

[31] The foreign bank entry policy was part of a broader deregulation package which also included (i) the removal of quantitative controls on bank balance sheets, (ii) the floating of the Australian dollar in 1983, and (iii) the use of market based operations for monetary policy.

efficiency of geographical expansion; integration of banking and insurance businesses; removal of lending restrictions and interest rate controls; and foreign banks' presence in domestic banking systems from the surveyed papers. In contrast, we also note that efficiency was unchanged and even declined following deregulation in other cases. Thus, while one of the goals of deregulation is to improve banking efficiency, the results are mixed. As noted by Berger and Humphrey (1997), the effects of deregulation on measured efficiency depend on the intended objectives, industry conditions prior to deregulation, and other incentives which may intervene.

(d) Empirical studies in transition economies

Given the developing country context of this study, we provide an extensive review of empirical studies in both transition and developing economies. The transition economies in CEE had a different banking set-up in the centrally-planned economic system. Nonetheless, their deregulation policies are similar to those of other developing countries, including African countries. Most developing economies (in Asia and Latin America) share similar financial development paradigms with African countries, in which the dominant role of governments was initially emphasized, but have in recent years embarked on various deregulatory policies of their banking systems. We now turn our attention to empirical studies on transition economies.[32]

Banking markets in European transition economies have undergone massive deregulation as part of their transformation from socialist economies to market-based economies. Following the first phase of banking reforms during the late 1980s and early 1990s, involving the decentralisation of banking activities,[33] the second phase saw significant deregulation of the banking sector during the mid-1990s to mid-2000s. The deregulation policies included: (i) interest rate liberalisation; (ii) restructuring and privatisation of state banks; and (iii) opening up of banking systems to private domestic and foreign banks.[34]

Against the background of these deregulation reforms, Fries and Taci (2005) analyse the efficiency performance of a sample of 289 banks in 15 Eastern European countries during

[32] An important aspect of these studies is the cross-country nature of most of them due to the reforms implemented simultaneously across countries within the same region.
[33] This involved transfer of commercial banking activities hitherto undertaken by the central bank to state banks.
[34] The banking deregulation reform was only a component of broader structural economic, financial markets development, investment, business and enterprising restructuring reforms that were pursued. See Fries and Taci (2005), Fang, Hasan and Marton (2011) and Bonin, Hasan and Wachtel (2015) for details.

1994–2001. The study uses cost efficiency as the efficiency indicator as relative cost efficiency is associated with changes in incentives and constraints in banking arising from such structural and institutional deregulation reforms. The authors model cost efficiency using the stochastic frontier approach and based on the one–step Battese and Coelli (1995) model, with the deregulation variable captured using the index of banking reform computed by EBRD.[35] The empirical results show average cost efficiencies of between 0.72 and 0.85 for most banks, with the exception of four banks whose average efficiencies were between 0.47 and 0.62. Regarding the impact of deregulation on cost efficiency, the authors observe a non-uniform relationship. Cost efficiency gains made during the early stages of deregulation were reversed during the later stages of the reform process. The authors attribute this pattern to two different strategies that banks pursued in response to the policy changes. Banks initially adopted a defensible cost-reduction strategy as they took time to adjust to these deregulation reforms, but subsequently pursued a more aggressive restructuring to increase quality, innovation and value addition of banking services, in line with the higher competitive conditions in the industry once they fully adjusted to the reforms.

Yildirim and Philippatos (2007) also examine cost and profit efficiency of banking sectors in twelve transition economies of Central and Eastern Europe (CEE) during 1993–2000.[36] Using the stochastic frontier and distribution-free measures of efficiency, the paper establishes average cost efficiencies of 72% and 77% respectively. Deregulation reforms (liberalisation, privatisation, foreign participation) was measured indirectly using a 'competition' proxy (the Panzar-Rosse H-statistic), and the results show that this had a positive impact on cost efficiency, suggesting higher efficiency from the deregulation reforms.

In a similar cross-country study, Fang, Hasan, and Marton (2011) examine the role of deregulation reforms, enterprise-restructuring and privatisation, bank ownership and competition on banking efficiency during 1998–2008 in six South Eastern European (SEE) countries.[37] Following the financial crisis in these countries during the late-1990s,[38] deregulation policies including opening up the banking sector to foreign investors and privatisation of state-owned banks were pursued. These measures improved significantly the

[35] The EBRD index of banking reform is an average measure of various reform indices and ranges between 1 and 4, with higher values reflecting convergence to performance norms and regulatory standards of the more advanced European economies.

[36] The countries were the Czech Republic, Estonia, Croatia, Hungary, Latvia, Lithuania, FYR of Macedonia, Poland, Romania, Slovenia, the Slovak Republic, and the Russian Federation.

[37] The countries are Albania, Bulgaria, Croatia, Macedonia, Romania and Serbia

[38] resulting from inadequate bank regulation and supervision as well as large build-up of non-performing loans

investment climate and encouraged foreign banks to acquire state-owned banks, and to a lesser extent, set up de novo operations. With average cost efficiency of 69% and average profit efficiency of 54%, the authors find a positively strong evidence of deregulation reforms leading to substantial gains in cost efficiency.

The empirical studies on transition economies show an overwhelmingly positive impact of deregulation on banking efficiency. It is worth noting in these cases however that other structural changes and reforms contributed to the overwhelmingly positive deregulation outcomes.

(e) Empirical studies in Asia

Developing countries in Asia and Latin America have also in recent times pursued deregulation reforms of their banking sectors focusing on interest rate liberalisation; privatisation of state–owned banks and opening up banking sectors to foreign participation; removing restrictions on banking activities; restoring and enhancing bank stability; and strengthening regulation, supervision and financial infrastructure.

Chen, Skully, and Brown (2005) examine changes in banking efficiency in major Chinese banks during 1993–2000 on the back of major reforms. These reforms included decentralisation of banking activities to state-owned policy banks following cessation of the Central Bank's role as a mono-bank in 1995;[39] establishment of joint-equity banks;[40] and gradual easing of restrictions on foreign banks' operations. Using DEA methodology in estimating cost efficiency, the authors note that deregulation had an initial positive impact on enhancing cost efficiency, but the efficiency gains were not sustained during the later stages (1997–2000). Estimated mean efficiency of banks increased from 0.46 in 1993 to 0.58 in 1995 but declined to 0.43 in 2000. The authors attribute the decline in efficiency partly to the adverse effects of the Asian financial crisis with huge NPLs on Chinese banks. An important implication of this study is that although deregulation can enhance cost efficiency, external factors can interfere with the sustainability of such efficiency gains.

[39] Prior to the reforms, China operated a mono-bank system where the People's Bank of China (PBOC) operated as both the Central Bank as were as performing all commercial banking activities. The reforms, ushered in by new Central bank and commercial banking laws, saw the transfer of the banking functions to newly-established state-owned policy banks, with the PBOC focusing on its central bank's mandate.

[40] Joint-stock commercial banks are banks in which the majority of shares are owned by non-government entities.

Dong (2010) also analyses the impact of further deregulation reforms in China following its accession to the WTO in 2001 on banking sector efficiency during 1994–2007. Using the one-step BC95 stochastic frontier model, the paper finds much higher cost efficiency levels averaging 91% during the period.[41] Deregulation had an initial negative effect with the pre-reform average efficiency of 95% declining to 86% during 2002-4 before increasing to 92% in 2007. The paper attributes the immediate decline in efficiency to regulatory challenges that foreign banks were still facing, and which created little incentive for domestic banks to enhance their efficiencies. The anticipated competition from foreign banks might have caused domestic banks to invest heavily in preparing to meet these competitive challenges and this could have resulted in the fall in cost efficiency. The rebound in efficiency during 2005–7 suggests improved banking efficiency in the medium term in line with the easing of more restrictions on foreign banks.

Mahesh and Bhide (2008) examine the impact of deregulatory reforms on efficiency levels of Indian banks during 1985–2004. The reforms included gradual reduction in reserve requirements from 53.5 per cent in 1991 to 29.5 per cent in 2005; removal of activity restrictions; gradual relaxation of bank entry restrictions, which allowed entry of private domestic banks and a gradual entry of foreign banks; and deregulation of interest rates (Mahesh and Bhide, 2008; Zhao et al., 2010). Using the BC95 one-step stochastic frontier model to compute cost and profit efficiencies, the authors find deregulation to be statistically significant in enhancing cost efficiency in the banking sector during the post-reform period.[42] In particular, the reforms induced cost-saving restructuring policies as competition intensified following the reforms. Thus, deregulation enhanced both banking competition and efficiency during the period of study.

In a similar study, Zhao et al. (2010) analysed the impact of both deregulation reforms and concomitant prudential re-regulation reforms on the Indian banking sector during 1992–2004.[43] The study period is quite similar (1992–2004) to that of Mahesh and Bhide (2008), but with the deregulation phase of 1992–97, and the concomitant prudential re-regulation

[41] Results using alternative measures of DEA show cost efficiency estimates of 88% on average.

[42] Authors also computed loan efficiency; but we focus the review on cost and profit efficiency in line with the two main objectives of cost minimisation and profit maximisation in the established literature.

[43] The re-regulation policies, which commenced in 1998 and aimed at ensuring banking stability, included increases in capital adequacy ratio, prudential norms on assets classification, income recognition, provisioning on non-performing loans and risk-based capital requirements, which became progressively more important, particularly against the backdrop of the Asian crisis (Zhao et al, 2010).

phase of 1998-2004. Employing the one-step BC95 stochastic frontier model to measure cost efficiency, the authors find technological regress with increasing costs during the initial deregulation period until 1996 after which costs decreased. The reduction in costs was temporarily interrupted by the re-regulation policy in 1998 with an upward shift in the cost frontier. This was nonetheless a one-off rather than persistent increase in costs as banks subsequently adjusted to the new regulatory environment.

(f) Empirical studies in Latin America

Barros and Williams (2013) evaluate the impact of banking deregulation on banking efficiency during 1998–2006, using the stochastic frontier model to estimate cost efficiency. Deregulation policies implemented included the repeal of restrictions on foreign entry, which led to substantial acquisitions by foreign banks; domestic bank consolidations; removal of activity restrictions through universal banking; and risk management measures to enhance asset quality. Using different proxies for each of the deregulation reforms, the authors find non-uniformity in the impact of the deregulation policies on cost efficiency. In particular, acquisition by foreign banks did not significantly affect their cost efficiency. Although foreign ownership curtailed inefficient management practices at those acquired banks, the authors alluded to "forces in Mexico which prevent foreign banks from generating efficiency gains" as the reason for the non-impact on cost efficiency. However, consolidation of domestic banks impacted significantly on efficiency, which suggests domestic banks, reacted positively to the threat of foreign bank entry. Diversification through universal banking is also found to be an important contributor to cost efficiency by facilitating competition.

Barros and Wanke (2014) analyse efficiency of Brazilian banks during 1998–2010 using a Bayesian dynamic frontier model to analyse the evolution of efficiency on the back of deregulation policies pursued, mainly opening up the banking system to foreign participation with privatisation of state banks to foreign entities being a major component. The authors find that on average, banks improved cost efficiency levels over the study period with the average efficiency of 0.75. Deregulation was however found to negatively influence cost efficiency.

(g) Empirical studies in Africa

Like other developing countries, many African countries emphasised the role of the state in the banking sector during the 1960s to 1980s. As a result of the restrictive policies on interest

rate, bank entry, and other regulations, coupled with macroeconomic difficulties, the banking system became financially distressed with huge non-performing loans, loss-making state-banks, and generally inefficient and uncompetitive banking systems. The weak economic performances compelled these countries to seek balance of payments support from the IMF and other external financing support from the World Bank during the late 1980s and 1990s. As part of the conditions for such financing support, the countries had to agree to structural economic reforms with deregulation of the financial sectors as a major component. The financial reforms included liberalisation of interest rates; gradual privatisation of state-owned banks; relaxation of controls on foreign exchange to facilitate foreign direct investment; and the development of other components of the financial sector, including capital markets and insurance sectors. These IMF/World Bank-sponsored financial reforms during the 1990s were the first wave of financial reforms in Africa and somehow led to a gradual liberalisation in interest rates and exchange rates; abolition of directed lending; cleaning up non-performing loans of banks; recapitalisation or closure of insolvent banks; and the development of insurance and capital markets. The reforms however seem to have had a limited impact on enhancing banking efficiency and competition due to lingering challenges that constrain efficient intermediation (Beck and Cull, 2014). Since the early 2000s, a few African countries have embarked on further deregulation reforms aimed at driving competition and enhancing efficiency of banks so they can play their intermediation role more effectively. Empirical studies on the impact of these deregulation reforms in Africa are few. We review published papers on deregulation reforms, ownership and size on banking efficiency on African countries, including Ghana.

Hauner and Peiris (2008) investigate the impact of deregulation reforms on banking efficiency in Uganda during the 5-year period, 1999–2004. The major deregulation policy pursued was the privatisation of the state-owned bank, Ugandan Commercial Bank (UCB), to a foreign bank, Stanbic Bank in 2002.[44] Other reform policies were the recapitalisation or closure of distressed banks and the adoption of a risk-based banking supervision approach. Using DEA techniques and the two-stage methodology, the study finds only a marginal enhancement in efficiency from 91.6% (pre-reform) to 92.9% after the reforms. The marginal change in efficiency could also reflect the limited scope of deregulation and the rather short study period of only five years, so short for a meaningful impact of deregulation as efficiency levels are usually persistent during short periods.

[44] UCB was the largest bank with more than 20% of industry deposits but became insolvent in 1998. Its banking activities were frozen and only traded in government securities until its eventual privatisation in 2002.

Zhao and Murinde (2011) examine the effect of deregulation on competition, risk taking and productive efficiency of Nigeria's banking industry during 1993–2008.[45] Two major deregulation policies: (i) interest rate deregulation in 1997 and (ii) the removal of regulatory barriers between merchant and commercial banking via the introduction of universal banking in 2001 were undertaken during the period. A major regulatory policy, a more than ten-fold increase in bank capital in 2004-5, led to massive consolidations through mergers and acquisitions with the number of banks significantly declining from 89 in 2004 to 25 at the end of 2005. Productive efficiency scores were estimated using DEA techniques and then regressed on three separate policy reform dummy variables representing each of the policy reforms undertaken, together with other co-variates such as competition and risk-taking measures. Estimated average efficiency was 0.54 but a significant improvement for post-2004 compared to pre-2004 was observed arising from all three major reforms. Interest rate deregulation facilitated competitive pricing of loans and deposits, which enhanced efficient intermediation. Universal banking offered cross-selling opportunities in both retail and corporate banking and provided both scale and scope opportunities to increase productivity. The capital-induced consolidations were found to have also produced scale economies, which reduced inefficiency. It was established that the reforms had a direct impact on productive efficiency, and also an indirect impact through increased competition, which resulted in a decline in risk taking behaviour of banks and further increase in efficiency.

Mwega (2011) analyses banking efficiency in Kenya over the 10-year period, 1998–2007, against the background of new deregulation reforms implemented in the early 2000s, including introduction of universal banking which led to a rapid branch roll-out and banking business expansion with insurance and stock brokerage services. In addition, significant capital increases led to policy-induced bank consolidations. Using both DEA and SFA techniques (with deregulation captured using a single policy reform dummy variable), the authors note some improvement in cost efficiency, arising from increased competition following the reforms.

In the case of Egypt, Poshakwale and Qian (2011) analyse efficiency developments during 1992–2007, a period marked by two phases of deregulation reforms. Liberalisation of interest and exchange rates, and the removal of lending ceilings were implemented in the first phase during the early 1990s, while privatisation of state banks and restructuring and consolidation

[45] This study, together with the subsequent ones reviewed in this subsection, were commissioned by the African Development Bank (AfDB) and published in a special edition of the African Development Review (2011) journal of the AfDB.

of small banks were carried out in the second phase during the early 2000s. The authors used both DEA and SFA models in a two-stage estimation technique and captured deregulation using two separate dummy variables for each phase. It was observed that cost efficiency improved marginally after the first phase of reforms, but did not significantly change after the second phase.

Following recent reforms including opening up the banking sector to more foreign participation; establishing credit information bureaus; strengthening regulation and supervision; and developing financial infrastructure in Zambia, Mwenda and Mutoti (2011) investigate the impact of these reforms on Zambia's banking sector during 1999–2008. Using the two-step SFA approach to estimate cost efficiency and its determinants, the authors show that the reforms had significantly positive effects on cost efficiency. Although cost efficiency scores fluctuated over the sample period, it broadly showed a general upward trend from 64% in 1999 to 73% in 2008.

In the case of Ghana, Isshaq and Bokpin (2012) examine the impact on cost and profit efficiency of the banking sector of rapid branch expansion during 2006-2007 following the removal of geographical restrictions and the introduction of universal banking. The authors use the Battese and Coelli (1995) time-varying inefficiency model to derive cost and profit efficiency estimates, which are then regressed on a variable measure of expansion (measured by the distance of outlying branches from a bank's head office). The authors note a decline in profit efficiency during 1999–2007. Cost efficiency however improved during the same period and was found to be positively related to distance, but not related to bank size. The authors conclude that improvement in cost efficiency is attributable to the branch expansion and especially to the technological gains arising from the recent electronic networking of bank branches.

Adjei-Frimpong et al. (2014) analyse cost efficiency during 2001–2010, using the two-step approach of computing efficiency scores in a first stage using DEA techniques and examining the underlying drivers of efficiency in the second stage regression. The study finds an average cost efficiency score of 51% but with a gradual improvement during the period. Bank size seems to have no influence on cost efficiency suggesting that large banks have no cost advantages over their smaller counterparts although capitalisation was found to be significant, with well-capitalised banks being less cost-efficient.

To summarise our review on the African literature, we observe that the main findings point to improvement in efficiency in most cases following implementation of the deregulation policies, although in some cases the efficiency gains were minimal. This suggests that the private interest view which recognises deregulation as efficiency-enhancing seems to be the dominant theory in the case of Africa's banking sector.

4.3 Bank ownership and banking efficiency

4.3.1 Theoretical perspectives

Most deregulation reforms include the opening up of banking systems to both private domestic and/or foreign investors' participation to reduce state ownership of the banking sector. This can be done through privatisation of existing state-owned banks to domestic or foreign banks or the licensing and establishment of *de novo* private or foreign banks. The issue of whether ownership matters in influencing performance and efficiency has also received considerable attention both in theoretical and empirical research. There are three broad classes of bank ownership types: state–owned banks, private domestic banks, and foreign banks. For ease of discussing the theoretical arguments, we distinguish between: (i) state vs. private banks, and (ii) domestic vs. foreign banks.[46]

(i) Theoretical framework of state-owned banks vs. private banks[47]

Before examining the theoretical issues regarding efficiency differences between state and private ownership of banks, we first explore the theories underlying state ownership of banks. This is due to the dominance of state-owned banks especially in developing countries during the 1960s to 1980s, which although has declined in more recent years following deregulation and privatisation policies, still remains a major component of the banking sector (Porta et al., 2002; Barth et al., 2012).

Two theories underlying state ownership of banks are the 'development' view and the 'political' view. The earlier 'development' view advocated for government ownership of

[46] This classification is adopted just for convenience. In the state vs private ownership discussion, most of the issues relating to private ownership is applicable to foreign ownership, as in almost all cases, foreign banks are private in nature. Similarly, the discussion on domestic vs foreign ownership assumes the context of private banks. In the subsequent discussion on the empirical literature however, we examine the efficiency impacts of these 3 distinct classes of banks separately.

[47] The distinction is based on who controls majority shareholding. State-owned banks are banks with at least 50% majority shareholding by government directly or through quasi-government or public institutions. Private domestic banks have at least 50% ownership by private sector investors.

banks as a vehicle to directly finance government projects and develop specific sectors, which are crucial for economic development (Gerschenkron, 1962). It was felt that the private sector was not capable or willing to finance medium- to long-term projects in sectors such as agriculture, industrialisation, construction and real estate development, among others. The role of the state in the pursuit of socio-economic development meant that state ownership of banks was a good strategy to mobilise domestic financial resources for development. This view was popularised during the 1960s–80s and widely adopted by most developing countries especially on attainment of political independence. The more recent 'political' view argues that government control of the banking system is an effective tool not necessarily to finance projects with national or economy-wide benefits but rather specific projects that would enhance the politicians' chances of winning votes to stay in power. This therefore results in politicization of resource allocation (Porta et al., 2002).

Although both theories underlie the prevalence of state-owned banks in developing countries with less developed financial systems and in Europe until not that long ago, their point of departure is the expected efficiency impacts. The development theory asserts that, other things equal, government ownership of banks should benefit subsequent financial and economic development, factor accumulation, and especially productivity growth and so enhance efficiency of the banking sector. The political theory, in contrast, suggests that, other things equal, government ownership of banks should displace (crowd out) the financing of private firms, and suggests that state-owned banks tended to be less efficient than private banks. Cross-country studies covering both developed and developing countries over different time periods suggest that government ownership of banks seems more to be in line with the political view rather than the development view, with state ownership of banks associated with lower efficiency than private banks (Porta et al., 2002; Barth et al., 2013).

One of the underlying theoretical reasons for the differences in the performance and efficiency of state banks and private banks is the corporate governance structure. The relative inefficiency performance of state banks is attributed to their weaker corporate governance structure compared to private banks. Under agency theory, the separation of ownership and control can create problems as agents (management) can manage banks in their own interest and not necessarily in the interest of the principal (owners).

First is the issue of monitoring of managers by owners. For private banks, ownership is usually concentrated, with clear incentives and an ability to effectively monitor performance of their managers. Thus, they are better able or more likely to align the managers' interests

with their own interests. However, ownership is in most cases highly diffused in state-owned banks, with no clear responsibility for effective monitoring of managers' performance. This disparity in monitoring leads to the state owned banks likely to have lower efficiency levels than private banks (Alchian, 1965).

Secondly, private banks are expected to be more efficient because they have well-defined incentives compared to state banks. For private banks, bonuses, promotions and other incentives which are linked to the objective of profit-maximisation makes managers to strive for greater efficiency, unlike state banks managers who might not have such performance-based incentives. Managers of state banks might be motivated by socio-political considerations, which are unrelated to efficiency-performance; hence they have less incentive to strive for efficiency.

Closely related to the incentives argument is the threat of job loss by managers of private banks for non-performance. Managers might lose their jobs if key performance targets are not met. The lack or minimal threat of loss of jobs by managers of state banks for non-performance, due to their usually politically-connected appointments rather than on merit contributes to their lack of motivation to perform efficiently.

State-owned banks are expected to be less efficient as they are used by governments to undertake socio-political projects, which might not be profitable nor economically viable. The direct interference of governments in state-owned banks to finance politically-induced projects means that due diligence, project screening and credit appraisals are relaxed and not subject to the prudent requirements that those of their private counterparts are subjected to. Risk management policies are weak and this creates huge loan losses and higher costs, undermining any efforts at cost minimisation or profit efficiency.

The above theoretical arguments and empirical evidence of low efficiency levels of state owned banks formed the basis for the rapid privatisation of banks embarked on by many developing countries. In a review of studies on bank privatisation in developing countries, Clarke, Cull, and Shirley (2005) observe that there is strong evidence in support of private banks operating more efficiently than state banks in most cases. As will be seen in the empirical reviews below, the impact of state and private ownership of banks on banking efficiency seem however not to be conclusive although in most cases private banks are found to be more efficient than state banks.

(ii) Theoretical framework of foreign vs. domestic banks' relative efficiencies

As noted earlier, the issues related to private bank ownership in the private vs. public ownership debate is applicable to foreign banks in most cases. However, we distinguish domestic private ownership from foreign (private) ownership and consider some of the theoretical underpinnings of their impacts on banking efficiency. Genay et al. (2000) discuss two alternative hypotheses underlying the relative efficiency of foreign banks vs. domestic banks: the *home field advantage* versus the *global advantage* hypotheses. According to the *home field advantage* hypothesis, domestic banks are relatively more efficient than foreign banks; while the *global advantage* hypothesis postulates that foreign banks are relatively more efficient than domestic banks.

In the *home field advantage* theory, challenges faced by foreign banks in domestic banking sectors adversely affect their operations and performance, making them less efficient than their domestic counterparts. These challenges include issues with personnel transfer to work abroad; ineffective cross-border monitoring and supervision; difficulties in establishing and maintaining banking relationships with retail clients; language and cultural differences; different regulatory regimes; and bias against foreign banks, among others. These difficulties, if they are not effectively addressed give domestic banks an advantage over foreign banks. These challenges could lead to higher operating costs or lower revenues by foreign banks from the provision of financial services than their domestic counterparts.

With the *global advantage* theory, however, efficiently managed foreign banks are able to overcome these cross-border challenges, and leverage on their superior managerial skills, best-practice policies and procedures, effective risk management, etcetera to provide superior quality of banking services, thereby operating more efficiently than domestic banks in domestic banking markets.

In an empirical analysis of domestic and foreign bank ownership and efficiency in the banking sectors of France, Germany, Spain, the U.K., and the US, Genay et al. (2000) find evidence of the prevalence of the *home field advantage*, that is, on average, domestic banks have higher profit efficiency than foreign banks. The only exception was the U.S. where U.S. banks appear to operate with relatively high efficiency both at home and abroad.

In the case of developing countries, however, the *global advantage hypothesis* seems to be the dominant paradigm due to the more developed and efficient nature of banks in the advanced countries than those in developing countries. Accordingly, foreign banks from developed

countries are expected to be more efficient than domestic banks, all things being equal, due to their superior technology, product innovation, superior managerial capabilities and stronger corporate governance. It is argued that the efficiencies of these foreign banks have positive spill-over effects on domestic banks who in turn increase their efficiency levels to remain competitive; and therefore enhances the overall efficiency of the banking system (Claessens and Laeven, 2004).

Others however argue that foreign banks derive efficiency gains in developing countries as they cherry-pick the best customers, usually multinational firms and large credit-worthy local firms, operating in the most profitable sectors of the economy. This leaves small businesses and less credit-worthy customers operating in less profitable sectors for the domestic banks, which makes them less competitive and less efficient. Thus, an increase in the presence of these foreign banks, it is argued, has an adverse impact on the efficiency of domestic banks and the entire banking sector (Claessens, Demirgüç-Kunt, and Huizinga, 2001).

4.3.2 Empirical Studies on ownership and banking efficiency

The empirical literature surveyed on the role of ownership on banking efficiency following deregulation seems to suggest that in most cases, foreign banks tend to be more efficient than private domestic and state-owned banks.

The study by Sturm and Williams (2004) on the impact of foreign bank entry on banking efficiency in Australia during 1988–2001, showed that the foreign banks were more efficient than the existing domestic banks. The authors note that the major domestic banks had engaged in mergers and branch expansions to serve as (indirect) barriers to entry of the foreign banks following the policy change, the foreign banks were able to effectively compete based on their technological superiority and found to be more efficient than the domestic banks.

Fries and Taci (2005), on the study of the efficiency performance of banks in the Eastern European countries during 1994–2001, found that foreign banks were the most efficient class of banks, followed by private domestic banks. State-owned banks were found to be the least efficient. The study noted the efficiency benefits of the deregulation reforms, in particular, the privatisation of state banks to foreign and private domestic investors, whose efficiency levels in terms of cost minimisation were superior.

Similarly, foreign banks were found to be more cost efficient relative to both private domestic banks and state-owned banks in the study by Yildirim and Philippatos (2007) on cost and profit efficiency of banking sectors in the transition economies of Central and Eastern Europe (CEE) during 1993–2000.

On the other hand, Fang et al. (2011) in their study on banking efficiency during 1998–2008 in six South Eastern European (SEE) countries found domestic private banks to be marginally more cost efficient than foreign banks. This was attributed to the huge investments in branch modernisation and training by the foreign banks on acquiring state-owned banks; challenges in breaking into existing relational networks; and costs incurred in gathering and processing locally based information as foreign banks sought to familiarise themselves with the local environment (*the home field advantage hypothesis*). These challenges notwithstanding, cost efficiency of foreign banks were found to have improved over time suggesting that once the initial investments are made and they gain more experience in the local market, foreign banks are able to improve their efficiency. The authors found no significant difference in cost efficiency between state-owned banks and private domestic banks, but noticed some convergence in efficiency of banks across the different ownership types with the efficiency gaps among foreign, state-owned and private domestic banks narrowing over time, arising from the increased competition.

It however seems that foreign banks thrive and are able to compete effectively to raise their efficiency levels where there is an even playing field for all banks. This seems to be the case in the above cases. In other cases, there are different regulations in place for different ownership classes of banks, especially with gradual easing of restrictions on foreign banks, which constrains their operational efficiency.

For instance, Mahesh and Bhide (2008) in their study on Indian banks during 1985–2004, found that state–owned banks were the most cost efficient, followed by private domestic banks, with foreign banks being the least cost efficient. The low efficiency of foreign banks was attributed to the relatively small scale of operations of these foreign banks during the study period as only a handful of foreign banks had been able to enter the market.

The study on India by Zhao et al. (2010), during the deregulation phase of 1992–97, and the concomitant prudential re-regulation phase of 1998- 2004, show that foreign banks initially enjoyed efficiency and technological advantage during the deregulated phase, but with domestic banks catching up eventually in the post re-regulation phase. Foreign banks were

found to have higher efficiency than public banks initially, but this eventually declines over time. In particular, by the time the policy focus changes in 1998, public banks have increased their efficiency whiles foreign banks have worsened their position.

Chen et al. (2005) in their study on banking efficiency of major Chinese banks during 1993–2000 on the back of deregulation reforms during the mid-1990s observed that the incumbent state policy banks ('Big Four'), in terms of cost efficiency, outperformed the newly established joint equity banks as well as foreign banks who still faced some restrictions.[48]

The study by Dong (2010) following further reforms embarked since 2001 on banking efficiency for the period 1994–2007 found that both state-owned banks and foreign banks were more efficient than private domestic banks. Although they found that the state-owned policy banks ('Big Four') were the most efficient, efficiency of foreign banks gradually increased with the phased removal of restrictions on them in addition to their superior technology and managerial expertise.

Interestingly, in the study of efficiency of Brazilian banks during 1998–2010 by Barros and Wanke (2014), the authors established that ownership did not play any significant role in influencing efficiency. The authors noted that while state banks were inefficient in Brazil, foreign banks also faced challenges against the cultural background of a closed market in Brazil, which constrained their optimum realisation of efficiency. Accordingly, there were no differences in the efficiency performance of the different classes of banks.

In the case of the studies on African countries, Hauner and Peiris (2008) and Mwega (2011) Kenya in their respective studies on the East African countries of Uganda and Kenya established that foreign banks are the most efficient banks while state banks are the least efficient. Poshakwale and Qian (2011) however do not find any efficiency differential between state-owned banks and private domestic banks in their study on Egypt, but surprisingly find domestic banks to be marginally more efficient than foreign banks. They attribute this to high business development costs incurred by the foreign banks for whom Egypt was a new market for them, and this adversely affected the efficiency levels (*home field advantage hypothesis*). Mwenda and Mutoti (2011) in their study on Zambia's banking sector establish that private banks were also found to be less cost efficient than state-owned banks.

[48] The "Big Four" banks are: Agricultural Bank of China, Bank of China, China Construction Bank, and Industrial and Commercial Bank of China

4.4. Bank size and efficiency

4.4.1 Theoretical perspectives

Theoretically, the relationship between bank size and banking efficiency is also not clear-cut as there are arguments for large banks to be more efficient than small banks and vice versa. Intuitively, one will expect a positive relationship between bank size and efficiency as relatively bigger banks can enjoy scale economies, which reduces their average cost of production making them more cost efficient. Such large banks are able to attract and leverage on highly productive workforce, strong managerial competencies and advanced technology to enhance operational efficiency. There is some evidence on the existence of scale economies in banking with bank size enhancing banking efficiency (Hughes *et al.,* 2001). Policy initiatives such as capital raises intended to promote bank consolidations are usually premised on the assumption of economies of scale benefits for larger banks including larger consumer bases and wider distributional channels, which will enhance productive efficiency. This is expected to be achieved through unit cost reduction, innovative products delivery, risk diversification and effective internal controls, policies and procedures.

There are however arguments to suggest that large banks increase banking concentration and reduce competition in the banking industry. Such large banks can leverage on their size to engage in collusive behaviour, which undermines competition and makes them less efficient.

Another argument against efficiency of large banks is that large banks might tend to overly focus on lengthy processes, cumbersome procedures and bureaucracy, which can adversely affect product and delivery turnaround and make them less productive. Lack of effective supervision and coordination can also create dysfunctional problems for large banks and can further constrain efficiency. In such situations, there are diseconomies of scale in such large banks and these tend to make them less efficient than relatively smaller banks.

Small banks could be associated with lower agency costs with the direct participation of owners in such small banks. Such small banks do not have the multi-layered delegation sometimes found in large banks, which create potential adverse selection and moral hazard problems. Managers of small banks therefore have a greater incentive to maximize efficiency, and this can be enhanced through fostering unity, cohesion and loyalty. In addition, small banks especially small domestic banks can exploit local monopolies and other local advantages to enhance their efficiency levels.

In the context of deregulation, it is argued that the resulting competitive pressures could threaten the survival of small banks, implicitly assuming that small banks are less efficient. However, it is argued that even in deregulated environments, small banks can achieve efficiency by focusing on a niche of customers about whom they have informational advantages and offer differentiated products from those of large banks (Elyasiani and Mehdian, 1995).

An important question that arises is what the optimal size of banks is as the distinction between small and large banks is relative. There is no rule regarding the optimal size of banks as this varies from one country to another, and should be largely dependent on the relative size of the economy. However, concentration ratios such as the largest three, four or five banks are usually used to refer to large-sized banks.

To summarise this section, we note that there is really no conclusive prediction of theory on efficiency on the basis of bank size, and thus the relationship between bank size and banking efficiency is an empirically testable hypothesis.

4.4.2 Empirical studies on bank size and banking efficiency

In spite of the non-conclusive prediction of theory on efficiency on the impact of bank size on efficiency, most of the empirical evidence seems to suggest that bank size is positively related to efficiency as discussed below.

In the study by Elyasiani and Mehdian (1995) on the US banking system to analyse the impact of deregulation on banking efficiency and the relative efficiencies of large and small banks between 1979 and 1986, the authors observe that small banks were more efficient than large banks in the pre-deregulation period. Following deregulation however, there was no evidence of efficiency differentials between small and large banks, as large banks had increased their efficiency levels. Small banks, though maintained their efficiency levels by focusing on niche markets to compete effectively with the large banks.

Yildirim and Philippatos (2007) found higher efficiency levels associated with large banks (by assets size) and well-capitalized banks in their study of CEE countries during 1993–2000. The relaxation of restrictions on banking activities facilitated the derivation of economies of scale and scope from larger banks and enhanced their cost efficiencies. Fries and Taci (2005) also establish that larger banks (in terms of deposit size) are more efficient than other banks, positing a positive relationship between bank size and cost efficiency.

Mahesh and Bhide (2008) also find bigger banks to be more cost efficient than smaller banks in India during 1985–2004. The study by Chen et al (2005) on banks in China during 1993–2000 suggests that large banks and small banks are the most efficient with medium-sized banks being the least efficient. Dong (2010) on the other hand establishes that large banks are more cost efficient than small banks in the study on China's banking sector during 1994–2007. Barros and Wanke (2014) establish that bank size positively affects cost efficiency with big banks being more cost efficient than small banks in their study of Brazilian banks during 1998–2010. The size effect reflects the positive impact of consolidations and economies of scale in banking efficiency in Brazil.

In the case of the studies on African countries, the studies on Uganda, (Hauner and Peiris, 2008); Nigeria (Zhao and Murinde, 2011) and Zambia (Mwenda and Mutoti, 2011) established bank size has a positive impact on efficiency, with large banks being more efficient than small banks. Only the study on Kenya (Mwega, 2011) observes that large banks are the least efficient compared to small and medium sized banks.

4.5 Conclusions

This chapter has reviewed relevant empirical studies and theoretical perspectives on the expected impact of deregulation reforms, ownership and size on banking efficiency across developed, transition, developing and African countries. The theoretical expectations on the outcome of deregulation policies are mixed. The empirical studies reviewed show that in most cases deregulation tends to have a positive impact on banking efficiency although others do not find any evidence of efficiency enhancements. In the case of developing countries, the lingering effects of financial repression, inefficient and uncompetitive banking systems arising from the restrictive regulations and excessive government role characterise initial conditions in the banking sector prior to deregulation. Accordingly, deregulation reforms positively impact banking efficiency in most cases. Some empirical results showed that while an initial worsening of efficiency might sometimes be experienced as banks take time to adjust to reforms, improvements in efficiency are subsequently realised in the medium term. However, stable macroeconomic conditions, other institutional capacities and a proper sequencing of the reforms are pre-requisites for an effective impact of reforms on efficiency. Beside the initial conditions, we also note that the nature of the specific deregulation policies implemented could have opposing effects on efficiency, while both internal and external

factors may intervene (Berger and Humphrey, 1997; Sturm and Williams, 2004). The review also confirms that apart from deregulation and the associated competitive conditions, which impact on efficiency, bank characteristics such as ownership and size also do have a role to play in influencing efficiency. In spite of the different predictions by theory regarding ownership impacts on efficiency, we observe for most transition and developing countries that foreign banks (from developed countries) seem not to only have efficiency advantages but are able to transmit efficiency gains to the entire banking sector. In a few cases where there are different regulations for different bank ownership types especially unfavourable to foreign banks or, where culturally there is a bias against foreign banks, these constrain the efficiency performance of foreign banks. While state banks were in some cases the most inefficient banks in some countries, in other countries well established state banks remain very efficient in the midst of deregulation reforms. In the case of size, we also noted different theoretical arguments regarding the impact of bank size on bank efficiency. The empirical studies however show that in most cases, bank size has a positive effect on efficiency.

From the review of the literature on African banks, we observe the growing importance of deregulation reforms in African countries, although it seems to be rolled out on a piecemeal basis. This could be due to uncertainties regarding the long-term impacts of such deregulation reforms especially in the context of African countries where there are macroeconomic challenges and other financial institutional capacity constraints. Furthermore, the limited research on the impact of such reforms on banking efficiency leaves policy-makers with limited options on policy guidance in this regard. Ghana's banking deregulation is quite comprehensive and a very good test case for other African countries contemplating similar reforms.

The review brings up pertinent issues for consideration in our study. The first one is the concept of efficiency used. Most of the empirical papers used cost efficiency and profit efficiency as the efficiency concept. The second issue is the technique for measuring efficiency, in which the predominant use of SFA and in some cases DEA was observed. The third issue relates to how to incorporate determinants or covariates of efficiency in the efficiency estimation model. These issues are examined in detail in the next chapter on analytical foundations and measurement of efficiency. Another issue is how to effectively capture or measure the deregulation variable(s). While studies on the transition economies use a banking reform index, other studies (on developing, including African countries) capture deregulation using a (policy reform) dummy variable to distinguish the pre- and post-reform

periods. The implications of these additional issues are discussed in detail in our empirical modelling of efficiency in Chapter 7. We now turn our attention to the concepts of efficiency and the techniques for its measurement in the next chapter.

CHAPTER 5 ANALYTICAL FOUNDATIONS AND MEASUREMENT OF FRONTIER EFFICIENCY

5.1 Introduction

In this chapter we are going to review the main efficiency measurement models used by the empirical literature. The review will cover them in chronological order ending with the ones that we selected for our empirical analysis. We precede this however with the analytical framework on frontier efficiency measurement.

In the assessment of banking efficiency, traditional efficiency analysis involves the use of financial ratios such as returns on assets (ROA), returns on equity (ROE), cost-income ratio (C-I-R), capital adequacy ratio (CAR) and non-performing loans ratio (NPL) to assess a bank's efficiency level. Regulators, bank managers and industry analysts often use this approach to measure bank performance. Although simple to compute and use, the usefulness of this approach for policy-related purposes is limited as described shortly below.

Frontier efficiency is an alternative performance measurement technique that measures a bank's efficiency in terms of deviations of a bank's input usage, output, cost or profit from an optimally estimated level, known as the 'frontier'. It assumes that banks are not able to fully optimise their production or behavioural objectives and therefore inherently have some level of inefficiency. Frontier efficiency is regarded as superior to the traditional approach as it is based on a sound theory of optimization behaviour, and deviations from the frontier reflect a true interpretation of the inefficiency associated with achieving behavioural objectives (Bauer, 1990). Second, it is an objectively determined quantitative measure, which controls for the effect of market prices and other exogenous factors, and therefore facilitates the separate analysis of quantitative effects on performance from changes in reform policies such as deregulation (Bauer, Berger, Ferrier, & Humphrey, 1998). Third, the structure of the frontier and computation of bank-level efficiencies have policy implications for regulators and bank management. Regulators can use changes in the frontier to design new policies or assess the impact of existing policies, while bank management can use bank level efficiencies to benchmark performance with other banks as well as trends over time within one bank against specific initiatives undertaken. It is for these reasons that frontier efficiency is being extensively used for policy-based research on efficiency performance assessment.

There are various concepts of frontier efficiency and different approaches in measuring it. The main objective of this chapter is to provide the analytical foundation that underpins frontier efficiency, discuss the various concepts of efficiency, and review the two commonly used non-parametric data envelopment analysis (DEA) and the parametric stochastic frontier approach (SFA) to measuring efficiency.[49] Based on our choice of SFA for the empirical work, we examine how to incorporate exogenous and bank-specific variables as determinants of inefficiency in the SFA model in line with the objectives of our study.[50]

We organise the chapter as follows. In Section 5.2, we discuss the conceptual framework of frontier efficiency, including the concept of production technology, technical efficiency, and economic efficiency. These entail a set-theoretic representation of production technology to define the production frontier, efficiency measurement using the notion of distance functions, input and output-oriented measures of technical efficiency, and the concept of economic efficiency. Section 5.3 discusses in detail the parametric and non-parametric efficiency measurement techniques, focusing on the DEA and SFA. We then review the proposed approaches to incorporating exogenous determinants of inefficiency in the SFA model in Section 5.4. We summarise the chapter and draw some conclusions in Section 5.5.

5.2 Conceptual framework of frontier efficiency

5.2.1 Production technology

The literature on frontier efficiency usually distinguishes frontier efficiency from productivity; two related but fundamentally different concepts used in evaluating production performance (Coelli et al. 2005). Productivity is defined as a descriptive performance measure of the ratio of output(s) produced to input(s) usage in the production process. It is more of a resource utilisation measure, without reference to the production technology. The production technology defines the maximum output(s) than can be produced from an input bundle or alternatively minimum input(s) required to produce an output bundle.

[49] Other less frequently used techniques such as the non-parametric Free Disposal Hull (FDH) and the parametric techniques of the Distribution Free Approach (DFA) and Thick Frontier Approach (TFA) are not discussed. These are well discussed in Coelli et al. (2005), Kumbhakar et al. (2015), among others.

[50] The basis for the selection of SFA over DEA is discussed later in the chapter.

Efficiency, on the other hand, measures the ratio of output(s) produced to the maximum possible output for a given input set and production technology. Efficiency is therefore a normative measure of the observed ratio of outputs produced or inputs used against an optimal measure, usually referred to as the 'frontier'.

Changes in productivity growth are therefore measured to capture changes in total output relative to inputs over a period of time, which can cause shifts in efficiency frontier. This is also sometimes referred to as total factor productivity and usually measured by using Malmquist productivity indices. Efficiency changes relate actual production of outputs to the most optimum level reflected in the efficiency frontier. In the context of this study, our focus is on efficiency analysis and therefore on the construction of the efficiency frontier and the derivation of inefficiency levels of observed units from the frontier.

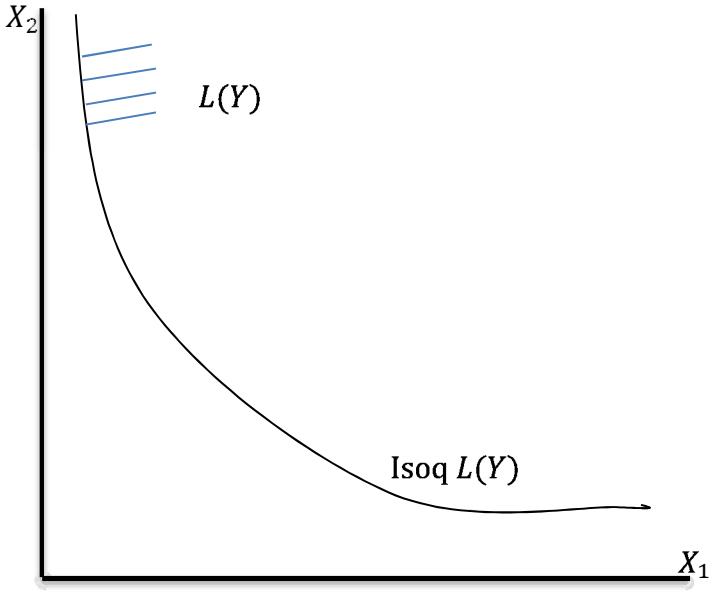
The production technology defines the maximum output(s) than can be produced from an input bundle or alternatively minimum input(s) required to produce an output bundle. A conventional way to present a multiple-input, multiple-output production technology is to use the technology set. The production technology describes the set of feasible input-output combinations.

Given that there are *n* producers using an input vector $\boldsymbol{X} \in R_+^J$ to produce a vector $\boldsymbol{Y} \in R_+^M$ of outputs, where $\boldsymbol{X}$ is a J-dimensional and non-negative input vector and $\boldsymbol{Y}$ is an M-dimensional non-negative output vector, the set of feasible production activity can be characterised by either an *input set* $L(Y)$ or an output set $P(\boldsymbol{X})$. The input set of production technology describes the feasibility set of inputs, that is, all input levels that are feasible for producing output vector Y.

The *input set* $L(Y)$ is formally defined as:

$$L(Y) = \{X : (Y, X) \text{is feasible}\} \tag{5.1}$$

Figure 5.1 Input set of production technology



Similarly, the *output set* $P(\boldsymbol{X})$ of the production technology represents a feasibility set of outputs that can be produced from a given level of inputs; formally defined as:

$$P(X) = \{Y: (Y, X) \text{is feasible}\} \tag{5.2}$$

The boundary of the input set is defined by the input isoquant and in a more general representation by the input efficient subsets.

The boundary of the input set defines the input isoquant, Isoq $L(Y)$ [51]

$$\text{Isoq } L(Y) = \big\{X: X \in L(Y)\big), \lambda X \notin L(Y), if\ 0 \leq \lambda < 1\} \tag{5.3}$$

Formally, "the input isoquant, Isoq $L(Y)$ describes the set of input vectors capable of producing the output vector Y, but which, when radially contracted become incapable of -producing the output vector Y" (Kumbhakar and Lovell, 2003).

The isoquant is the lower boundary of the input set $L(Y)$ as illustrated in Figure 5.1 and is a standard against which efficiency can be measured.

We also define the input efficient subset, $ES(Y)$ as

$$ES(Y) = \{X: X \in L(Y),\ X' \notin L(Y)\ for\ X'\ when\ X'_j \leq X_j \forall j\ and\ X'_j < X_k\ for\ some\ k\} \tag{5.4}$$

[51] A similar analysis can be carried out for the output set. See Kumbhakar and Lovell (2003).

The efficient subset describes the sets of input vector capable of producing each output vector but which when contracted in any dimension becomes incapable of producing output vector Y.

If the input set is strictly convex, the input isoquant and input efficient set coincide and represent the minimum level of inputs required to produce a given level of output.

The input isoquant therefore constitutes what is called the 'frontier' of technology, and reflects the structure of the production technology of best practices, and represents the standard against which efficiency is measured. Efficiency in this context is the comparison of actual (observed) inputs level to an optimal one (frontier).

5.2.2 Production frontiers and measures of technical efficiency

An alternative characterisation of the production technology and feasible technology set can be illustrated using a production function. A production frontier is a functional or mathematical representation of the production technology. This can be expressed as:

$$F(\boldsymbol{X}, \boldsymbol{Y}) = 0 \tag{5.5}$$

Where, as defined above, $\boldsymbol{X}$ is a J-dimensional and non-negative input vector and $\boldsymbol{Y}$ is an M-dimensional non-negative output vector.

In the case of a single output (or an aggregate of multiple outputs), we can represent the production technology by the following production function:

$$Y = f(X_1, X_2, \ldots . X_n) \equiv f(\boldsymbol{X}) \tag{5.6}$$

where $f(.)$ specifies the production technology that characterises the input-output relationship. Thus $f(\boldsymbol{X})$ is the production function which shows the maximum possible output, for a given $\boldsymbol{X}$, and is represented in Figure 5.2.

The measurement of technical efficiency requires knowledge first of the feasible input-output combinations, and then of the production technology. The economic feasible region is the feasible input-output combinations, and is defined if the following conditions of the production function are met (Coelli et al., 2005; Kumbhakar, Wang, and Horncastle, 2015):

1. Non-negativity: Output (Y) and input (X) are non-negative.

2. Weak essentiality: At least one unit of input needs to be used to produce one unit of output.

3. Monotonically increasing in input (x): Output cannot decrease if an additional input is used. This implies positive marginal product. Formally, strong monotonicity implies $f(X_1) \geq f(X_0)$ *for* $X_1 \geq X_0$

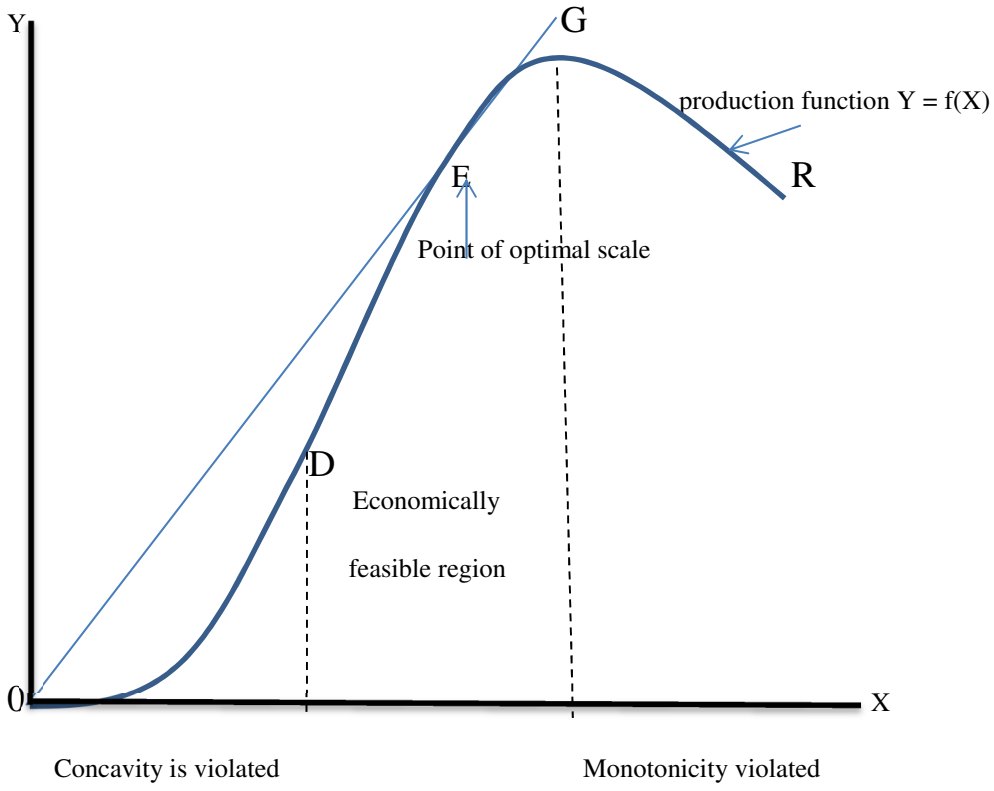
4. Concavity in input (X): any linear combination of vectors X_0 *and* X_1 will produce an output that is no less than the *same linear combination of* $f(X_0)$ *and* $f(X_1)$*. Formally, concavity is defined as*

$$f(\theta X_0 + (1-\theta)X_1) \geq \theta f(X_0) + (1-\theta)f(X_1), \quad for\ all\ 0 \leq \theta \leq 1$$

Concavity implies non-increasing marginal products, as expressed by the law of diminishing marginal productivity.

From the graph of the production function $Y = f(\boldsymbol{X})$ in Figure 5.2, segment 0D shows increasing marginal productivity of **x** and therefore violates the law of diminishing marginal productivity or the concavity condition.

Figure 5.2 Production function and the economic feasible region



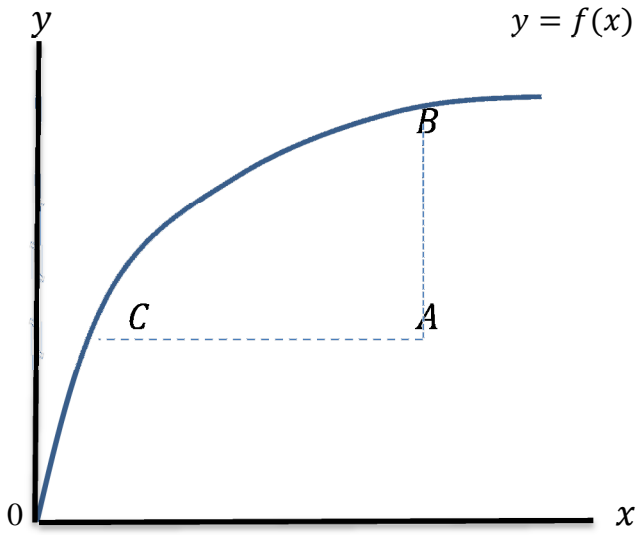
Source: Coelli et al. (2005, p.14)

Similarly, the segment GR depicts negative marginal product and therefore violates the monotonicity condition. Accordingly, the only segment that meets all the above conditions is

the segment DG, and thus defines the *economically feasible region* of production. This region defines the production frontier against which efficiency measures are derived.[52]

We represent the economically feasible production frontier in Figure 5.3 to illustrate the concept of technical efficiency.

Figure 5.3 Technical inefficiency (input-oriented and output-oriented)



From the production frontier in Figure 5.3 which defines the current state of technology, and shows the maximum output that can be produced at each input level based on the current level of production technology, firms B and C are said to be technically efficient as they lie on the frontier. **Technical efficiency** is therefore defined as the minimum level of inputs that can be used to derive a given level output; or *vice versa* the maximum level of output that can be obtained from a given level of inputs and represents all point on the production frontier. Firm A operating below the production frontier is said to be technically inefficient and inefficiency can be measured by the distance from A to B or from A to C.

Technical inefficiency can therefore be classified as either *input-oriented* or *output-oriented*. Input-oriented measure of technical inefficiency is where the observed output can be produced with lower levels of inputs. In Figure 5.3, firm A's inefficiency relative to that of C (distance AC) is known as input-oriented measure of technical efficiency ($\overline{AC}$) and shows the input wastage due to technical inefficiency at A.

[52] The implication of the above conditions is that an empirically estimated production function must pass the monotonicity and concavity tests.

Output-oriented measure of technical inefficiency is where a higher level of output is technically attainable for the given inputs. Firm A's measure of inefficiency relative to Firm B is a case of output-oriented measure of technical efficiency as it can feasibly increase its output to that of B without using any additional input. The distance $\overline{AB}$ shows the output loss due to the technical inefficiency.

The input- and output- oriented measures of technical efficiency of a firm are in general different, except in the case of constant returns to scale. In practice, whether the input- or output- oriented measure is more appropriate would depend on whether input-conservation is more important than output-augmentation and also on what the firm can control best, whether input usage or output produced.

The above illustration of technical inefficiency and input-oriented and output-oriented technical inefficiency from the production frontier (case of one output) can be generalised for the multiple input, multiple output case using distance functions. Distance functions can also be used to characterise the production technology, and we discuss it below.

5.2.3 Distance functions and measures of technical efficiency[53]

Efficiency can be measured as the distance from the efficient frontier by means of a distance function. Both input distance function (from input minimisation perspective) and output distance function (from output maximisation perspectives are discussed in detail below.

(a) Input distance function and measures of input-oriented technical efficiency

Formally, given the input set $L(Y)$ defined in Section 5.2, we define an input distance function as the function:

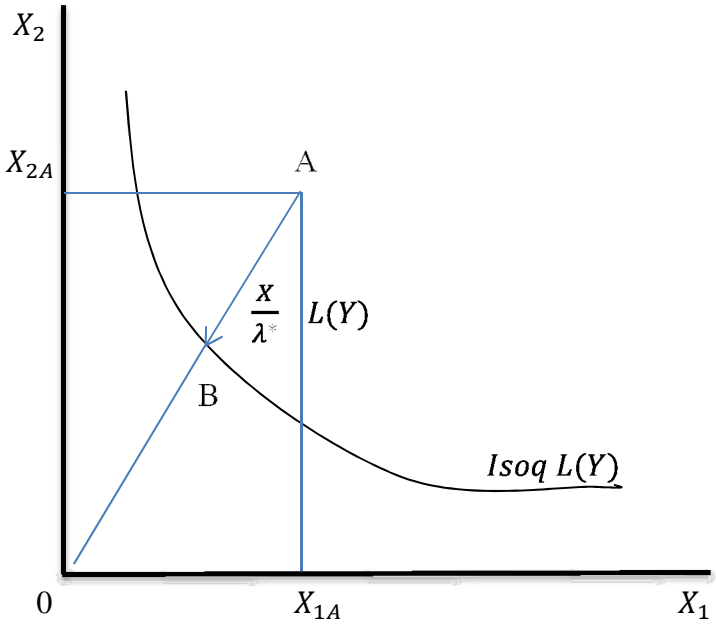
$$D_I(Y, X) = \max\{\lambda: \frac{X}{\lambda} \in L(Y)\} \text{ where } \lambda \geq 1. \quad (5.7)$$

An input distance function (uses an input-conserving approach) examines the production technology by looking at minimal proportional contraction of the input vector, given an output vector. It therefore measures the maximum amount by which input usage can be radially reduced but remains feasible to produce a given vector of outputs (Kumbhakar and Lovell, 2003).

[53] Distance functions were introduced by Shephard (1953), who used it as a tool to establish duality between production technology and cost functions. In terms of empirical estimation of efficiency, distance functions are less frequently used. Similar to production frontiers, distant functions do not require price data and imposition of economic behavioural objectives, and thus are used to measure only technical efficiency.

We illustrate the input distance function in the 2-input space X_1 and X_2 – used in producing output vector Y in Figure 5.4a, using the input set, $L(Y)$ and input isoquant $Isoq\ L(Y)$. We observe that the input vector A is feasible for output vector Y, but Y can be produced with the radially contracted input vector $\frac{X}{\lambda^*}$. The distance function for the production point A where the firm uses X_{1A} of input X_1 and X_{2A} of input X_2 to produce the vector output Y is equal to the ratio $0A/0B$.

Figure 5.4a Input distance function



Since distance functions are defined in terms of input sets, which have to satisfy certain properties, discussed in the context of the economic feasible region of the production frontier in Section 5.3, the input distance function (D_I) must satisfy the following properties summarised below (Kumbhakar et al., 2015):

1. D_I(Y, X) is decreasing in each output level;
2. D_I(Y, X) is increasing in each input level;
3. D_I(Y, X) is homogenous of degree 1 in the feasible input vector **X**;
4. D_I(Y, X) is concave in **X**.

It is clear from the above and earlier discussions under Section 5.2 that $L(Y) = \{X: D_I(Y,X) \geq 1\}$ and that Isoq $L(Y) = \{X: D_I(Y,X) = 1\}$. Thus, the isoquant which serves as a standard against which to measure technical efficiency corresponds to the set of input vectors having an input distance function value of unity.

Input distance functions therefore also serve as a standard against which we measure input-oriented measure of technical efficiency, and it is given as the reciprocal of the input distance function, that is,

$$TE = 1/D_I(Y, X) \tag{5.8}$$

A firm is technically efficient if it is on the frontier, in which case TE=1 and $D_I(Y, X)$ is also equal to 1. From Figure 5.4a, we define the input-oriented technical efficiency by the function

$$TE_I(Y, X) = \min \{\theta: D_I(Y, \theta X), \geq 1\}^{-1} \tag{5.9}$$

(b) Output distance function and measures of output -oriented technical efficiency

Similarly, an output distance function adopts an output-expanding approach and considers a maximal proportional expansion of the output vector, given an input vector. It gives the minimum amount by which an output vector can be deflated and still remain producible with a given input vector.

Based on the output sets of $P(\boldsymbol{X})$, we define an output distance function as:

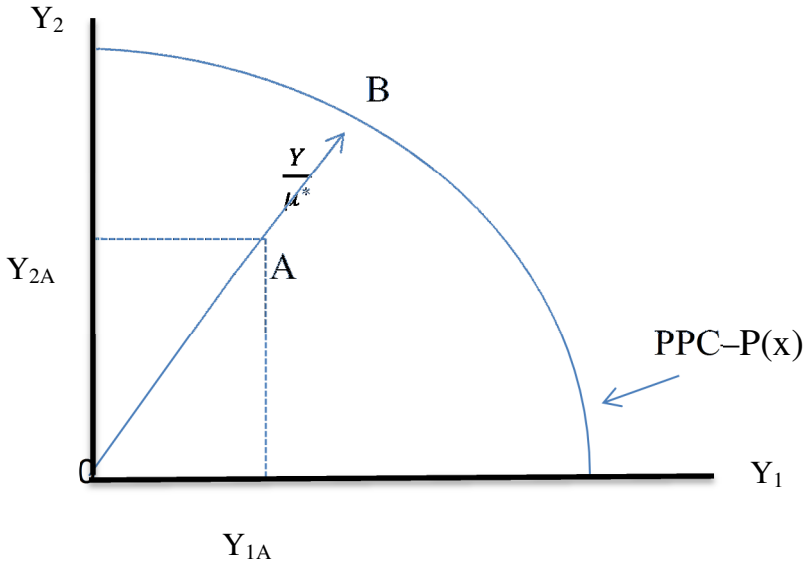
$$D_O(X, Y) = \min\{\mu: \frac{Y}{\mu} \in P(\boldsymbol{X})\}, \tag{5.10}$$

where, as defined earlier, $P(\boldsymbol{X})$ is the set of output vectors that are feasible for each input vector **X**.

We illustrate the output distance function in a 2-commodity output case, Y_1 and Y_2 produced by a vector of inputs X, in figure 5.4b. The production possibility set, P(X), is the area bounded by the production possibility frontier, PPC – P(X), and Y_1 and Y_2 axes.

We note that output Y_A can be produced with input X but so can a large output $\frac{Y}{\mu^*}$ so that $D_O(X, Y) < 1$. The value of the distance function for the firm using input level X to produce the outputs, defined by the point A, is equal to the ratio 0A/0B.

Figure 5.4b Output distance function



Similar to the properties of the input distance function, we summarise the properties of the output distance function (D_O) below (Kumbhakar et al., 2015):

1. D_O (**X**, **Y**) is decreasing in each input level;
2. D_O (**X**, **Y**) is increasing in each output level;
3. D_O (**X**, **Y**) is homogenous of degree 1 in the feasible input vector **Y**;
4. D_O (**X**, **Y**) is concave in **Y**.

The output-oriented measure of technical efficiency coincides with the output distance function, that is,

$$TE_O = D_O(x, y) \tag{5.11}$$

In other words, the output-oriented measure of technical efficiency is given by the function

$$TE_O\,(x, y) = [\max\,\{\varphi: D_0\,(x, \varphi y) \leq 1\}]^{-1} \tag{5.12}$$

The measure of technical inefficiency using production frontiers or distance functions is however limited as it does not address whether the observed combination of inputs are allocatively optimal. **Allocative efficiency** relates to the combination of inputs and outputs that meets some behavioural objective such as cost minimization, revenue or profit maximization and therefore takes into account input and/or output prices. Thus a firm may achieve technical efficiency but not allocative efficiency. Where a behavioural criterion is assumed, efficiency measures can be altered to accommodate *allocative efficiency*. Allocative

and technical efficiencies together provide a measure of *economic efficiency*, which we discuss in the next section.

5.2.4 Economic efficiency

Economic efficiency is the measure of efficiency where a behavioural assumption such as cost minimisation, revenue maximisation or profit maximisation is imposed. Thus, depending on which behavioural objective is assumed, we can measure cost efficiency, revenue efficiency or profit efficiency. Cost minimisation is implied by profit maximisation.

(a) Cost minimisation and cost efficiency

Cost minimisation assumes that a firm seeks to incur the least cost of input combination to produce a given output bundle. In such situations, the production technology is adjusted to include input prices in addition to the input(s) and output(s). Inputs are the choice variables in the context of cost minimisation, hence an input-oriented measure of technical inefficiency, which focuses on input overuse, is assumed in the computation of cost efficiency. We note however that the attainment of technical efficiency is necessary, but not sufficient for the attainment of cost efficiency. This is because a technical efficient firm could use an input mix which, based on the input prices it faces, does not represent the least cost.

Accordingly, assessment of efficiency moves from the production frontier to a cost frontier. The cost frontier defines the minimum cost required to produce any level of output(s) based on input prices. The cost frontier serves as the standard against which to measure cost efficiency by comparing actual cost incurred by a firm to the minimum cost.

Given the objective of minimising cost subject to the production technology constraint, we can write the cost minimisation problem as:

$$C^{*}(\mathbf{W},\mathbf{Y}) = \text{Minimise W'X subject to } \mathrm{F}\,(\mathbf{Y},\mathbf{X}) = 0 \qquad (5.13)$$

where $\mathrm{C}(\mathbf{W},\mathbf{Y})$ is the minimum cost; $\mathbf{X}$ = vector of inputs; $\mathbf{W}$ = vector of input prices; $\mathbf{Y}$ = output vector; and $F(\mathbf{Y},\mathbf{X})$ is the production function.

The cost minimisation problem in equation (5.13) is solved to derive the input demand equations which can be substituted in the cost function to obtain the minimum cost, $C^*(\boldsymbol{W}, \mathbf{Y})$. Hence $C^*(\mathbf{W}, \mathbf{Y})$ is the cost frontier which gives the minimum cost based on input prices and observed output(s), and is the benchmark against which cost efficiency is measured.

Similar to the underlying properties of the production frontier and distance functions discussed in Sections 5.2.2 and 5.2.3, the cost frontier must satisfy the following properties (or regularity conditions) (Coelli et al., 2015):

l. Non-negativity: Costs can never be negative, that is $\mathrm{C}(\mathrm{W}, \mathrm{Y}) > 0$.

2. *Monotonicity in input prices: An increase in input price will not decrease costs. If* $W_1 \geq W_0$ *then* $C(W_1, Y) \geq C(W_0, Y)$.

3. *Monotonicity in output: An increase in output will not decrease costs. If* $Y_1 \geq Y_0$ *then* $C(W, Y_1) \geq C(W, Y_0)$.

4. *Homogeneity: A k-fold increase in input prices will cause a k-fold increase in costs. Formally,* $C(kW, Y) = kC(W, Y), for\ k > 0$.

5. *Concavity: the cost function* $C(W, Y)$ *is concave in input prices, that is, the input demand functions cannot slope upwards.*

These regularity conditions must be satisfied in the empirical cost function estimation. Linear homogeneity is usually imposed prior to estimation of the cost function while monotonicity and concavity conditions can be tested post estimation.

Cost efficiency (CE) is therefore expressed as the ratio of minimum cost defined by the cost frontier to the actual cost of the firm as follows:

$$\mathrm{CE} = C^*(W, Y)/\mathrm{W'X} \tag{5.14}$$

Cost efficiency therefore occurs at the point where a firm is both technically and allocatively efficient. If a firm is not on the cost frontier, then it is either technically inefficient or allocatively inefficient, or both.

The analysis of revenue efficiency and profit efficiency is similar to cost efficiency, except that the objective of cost minimisation is replaced by revenue and profit maximisation. We accordingly provide just a brief description of the main elements of revenue efficiency and profit efficiency below.

(b) Revenue maximisation and revenue efficiency

Where revenue maximisation is the assumed behavioural objective, a revenue frontier is constructed as the benchmark to measure revenue efficiency. In revenue maximisation, outputs become the choice variables and so an output-oriented technical inefficiency is implicitly assumed, although we note that such technical efficiency is only necessary but not sufficient for achieving revenue efficiency.

Assume that a firm faces a positive vector of exogenous output prices **P** and a given input vector **X**. If the firm seeks to maximise its total revenues, R, then it will choose the output levels which will maximise revenue.

We express the revenue maximisation problem as follows:

$$R^*(X,P) = \text{Maximise} \quad \mathbf{P'Y} \text{ such that } F(\boldsymbol{X},\boldsymbol{Y}) = 0 \tag{5.15}$$

Revenue-maximising output supply equations can be derived from the optimisation problem in equation (5.15) and substituted into the revenue function to obtain the revenue frontier in the form $R^*(X,P)$. The revenue frontier provides a standard against which to measure revenue efficiency.

Revenue efficiency (RE) is therefore expressed as the ratio of a firm's observed revenue to the maximum revenue:

$$RE(X,P,Y) = \frac{\mathbf{P'Y}}{\boldsymbol{R(X,P)}} \tag{5.16}$$

Revenue efficiency can similarly be decomposed into (output-oriented) technical efficiency and allocative efficiency. Revenue inefficiencies can arise from either technical inefficiency, or allocative inefficiency (misallocation of outputs in the face of prevailing output prices), or both.

(c) Profit maximisation and profit efficiency

In cases where profit maximisation is assumed by producers, we can measure profit efficiency using the profit frontier. For profit maximisation however, both inputs and outputs become

choice variables as producers must choose an appropriate input mix and also produce an appropriate output mix to maximise profit.

Firms face positive input prices (W) and output prices(P) and choose inputs (X) and outputs (Y) in order to maximise profits ($\mathrm{P'Y} - \mathrm{W'X}$). The profit maximisation problem of the firm is specified as follows:

$$\pi(P, W) = \text{Maximise}\, P'Y - W'X \; such \; that \; F(X, Y) = 0 \tag{5.17}$$

The optimisation problem can also be solved to obtain the profit frontier which becomes the standard against which profit efficiency is measured. The measure of profit efficiency (PE) is given by the ratio of observed profit to maximum profit:[54]

$$PE = \frac{P'Y - W'X}{\pi(P,W)} \tag{5.18}$$

Profit efficiency requires (either input-oriented or output-oriented) technical efficiency and both input allocative efficiency and output allocative efficiency.

To sum up, the first part of this section has reviewed the conceptual framework of frontier efficiency using the structure of the production technology and its characterisation by production frontiers and distance functions to serve as standards for analysing technical efficiency. We then reviewed the concept of economic efficiency including cost, revenue and profit efficiency measures where behavioural objectives are imposed.

Most empirical applications have however focused on cost efficiency (see for instance Ferrier and Lovell, 1990; Berger and De Young, 1995; Fries and Taci, 2005; Zhao et al., 2010; Das and Drine, 2011; Mwega, 2011; and Molyneux and Williams, 2013, just to mention a few). Revenue efficiency is the least considered in empirical work based on the criticism that, by ignoring costs, it fails to effectively measure managerial capacity to optimally manage economic resources. Further, a revenue-efficient firm may derive such efficiency solely on the basis of high market concentration or market power which can enable it to exploit output prices through collusive behaviour. Profit efficiency, on the other hand, has received some attention in the empirical banking literature (see for instance, Maudos et al., 2002; Ariff and Can, 2008; Bonin et al., 2005; Kasman and Yildirim, 2006; Isshaq and Bokpin, 2011). It is

[54] The above specification of profit $\pi(P, W)$ is referred to as the standard profit measure as it assumes exogeneity of output prices in a perfectly competitive banking market. An alternative profit measure ($\pi_A(Y, W)$) is used which assumes that due to market power of banks, output prices are not exogenous and therefore uses output levels instead of output prices (Berger and Mester, 1997; Maudos et al., 2002)

however also argued that profit efficiency could be attributable to the effects of market power or collusive behaviour by dominant banks or by the macroeconomic environment which result in high interest rate spreads, and so does not reflect true efficiency, which is particularly the case in most African countries, including Ghana. Further, it can also be argued that for some banks, especially state-owned banks, the objective may not be one of profit maximisation but to support the development of particular sectors or to foster financial inclusion in less developed areas. In all classes of banks, cost efficiency should remain a relevant objective. As noted by Koetter and Meesters (2013), cost minimisation is a necessary and sufficient condition not only for short term profitability, but also for long term survival of banks. It is in this context, that this study uses cost efficiency as the relevant economic efficiency measure in line most of the empirical literature.

The above analytical review forms the basis upon which models on the construction of efficiency frontiers and measurement of efficiency are developed, and to which we turn our attention to in the next section.

5.3 Measurement techniques of efficiency in banking

The literature on the construction of efficiency frontiers and the measurement of efficiency distinguishes between two broad techniques: parametric and non-parametric techniques. Both techniques measure the efficiency of a firm as the radial distance from a frontier. The main differences however are in how the frontier is constructed and how inefficiency is derived from the constructed frontier.

Parametric or econometric techniques involve specification of a functional form for the frontier, while non-parametric techniques do not impose any functional form for the production process. Within the parametric techniques, we can distinguish between deterministic and stochastic econometric approaches. The deterministic econometric approach does not account for statistical noise and includes models by Aigner and Chu (1968) and Afriat (1972) (see (Førsund, Lovell, & Schmidt, 1980) for a detailed review). A major criticism of these models and indeed all deterministic models is that no account is taken of the possible influence of measurement errors and other noise on the shape and positioning of the estimated frontier, and so all observed deviations from the estimated frontier are assumed to

be technical inefficiency. We do not discuss the deterministic econometric frontier in detail as their usefulness is very limited.

The stochastic parametric approaches mainly coincide with the stochastic frontier analysis (SFA), as well as other techniques such as the Distribution Free Approach (DFA) and the Thick Frontier Approach (TFA), with the SFA being the most widely used approach. The stochastic frontier model accordingly requires an explicit functional form for the frontier, as well as distributional assumptions about random error and inefficiency since SFA allows for the existence of random, stochastic factors. It accordingly attributes the distance between a firm and the frontier to both random noise and actual inefficiency. In the econometric estimation approach also, the parametric models normally include a time trend as a proxy for disembodied technical change in the estimation of the frontier and the derivation of efficiency levels. This is intended to capture technological change over time if there are observations over time, as technological advances alter production or cost functions to change over time. As indicated, the SFA requires an explicit functional form, and the choice of the functional form also pre-supposes the assumption underlying technological change. The Cobb-Douglas function assumes technological changes to be constant while the translog function allows technological change effects to increase or decrease with time with the inclusion of the time trend in quadratic form.

The non-parametric techniques on the other hand rely on mathematical linear programming tools and therefore do not need either of these assumptions of functional form specification or distributional assumptions on inefficiency. Furthermore, deterministic non-parametric methods such as Data Envelopment Analysis (DEA) and the Free Disposal Hull (FDH) attribute the entire deviation from the frontier as inefficiency, without accounting for random errors.

The mostly widely used approaches in the efficiency literature are the non-parametric DEA and the parametric SFA. These two methods have been extensively used in the empirical literature, the DEA mostly in the management sciences and operations research discipline and the SFA in the economics and econometric discipline. We examine the DEA and SFA approaches in more detail in sections 5.6.1 and 5.6.2 respectively.

5.3.1 Data Envelopment Analysis (DEA)

DEA was introduced by Charnes, Cooper, and Rhodes (1978), hereinafter referred to as CCR, for measuring efficiency. The original model proposed by CCR in measuring the efficiency of a decision-making unit (DMU) was in the context of an input-oriented measure of technical efficiency with a production frontier exhibiting constant returns to scale (CRS).[55]

The objective of DEA is to construct a non-parametric envelopment frontier over data points representing output/input ratios (adjusted by output and input weights) such that all observed points lie on or below the production frontier. Although it does not require any functional form to be specified, it must satisfy the general assumptions about the production technology discussed earlier, including feasible input-output combinations, convexity of the production possibility set, and free disposability of both inputs and outputs.

The DEA frontier is thus constructed as the piecewise linear combinations that connect the set of the best-practice observations (of the output/input ratios), yielding a convex production possibilities set. Efficiency scores are then derived by measuring how far an observation is positioned from the 'envelope' or frontier.

We briefly illustrate the formulation of the original DEA model (CCR, 1978) as discussed in detail in Coelli (1995) below.

Suppose we have N firms who produce M outputs from J inputs. For the i-th firm, we represent the input and output vectors by $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$, respectively. The JX x N input matrix, $\boldsymbol{X}$ and the MX x N output matrix, $\boldsymbol{Y}$, represent the data of all N firms. For each firm, we obtain a measure of the ratio of all outputs over all inputs, such as $u'y_i/v'x_i$ where u is an $M\ X\ 1$ vector of output weights and v is a $J\ X\ 1$ vector of input weights.

A linear programming (LP) problem which seeks to find optimal weights for u and v such that the efficiency measure of the i-th firm is maximised, subject to the constraint that all efficiency measures must be less than or equal to one is specified as follows:

$$\text{Maximise } (u'y_i/v'x_i), \tag{5.19}$$

$$\text{subject to } \quad u'y_i/v'x_i \leq 1$$

[55] The term DMU was used to include firms as well as non-market agencies like schools, hospitals, courts, etcetera with lack of market prices for their outputs.

$$u, v \geq 0$$

This formulation has an infinite number of solutions and so the constraint $vx_i = 1$ is imposed, so we have: [56]

$$\text{Maximise } (\mu' y_i) \text{ [57]} \tag{5.20}$$

$$\text{subject to } v'x_i = 1$$

$$\mu' y_j - v'x_j \leq 0, j = 1,2, \dots, N,$$

$$\mu, v \geq 0$$

Using duality in linear programming, an equivalent *envelopment* form of this problem can be derived as follows:

$$Minimise_{\theta,\lambda}\, \theta, \tag{5.21}$$

$$subject\ to \quad -y_i + Y\lambda \geq 0,$$

$$\theta x_i - X\lambda \geq 0,$$

$$\lambda \geq 0$$

where θ is a scalar to be estimated and λ is an $N\ X\ 1$ vector of constants. The linear programming problem is solved N times, once for each firm in the sample, and a value of θ is then obtained for each firm.

The parameter θ is the estimation of technical efficiency *(TE)* for the *i*-th firm, with $0 \leq \theta \leq 1$ and where $\theta = 1$ indicates a point on the frontier and hence a technically efficient firm. The vector λ defines the projected point $(X\lambda, Y\lambda)$ against which the efficient score of the *i*-th firm is derived. The projected point is a linear combination of all the observed data points in the sample. The projected points obtained by each LP together form the feasible surface of the sample, known as the production technology (or frontier).

[56] This means that if (u^*, v^*) is a solution, then $(\alpha u^*, \alpha v^*)$is another solution

[57] Change in notation from u to μ is just to reflect the transformation made after imposing the constraint.

The original CCR DEA model was modified and extended by Banker, Charnes, and Cooper (1984) hereinafter referred to as BCC, by relaxing the restrictive assumption of CRS and assuming variable returns to scale (VRS).

An adjustment is made to equation (5.21) by the inclusion of a convexity constraint $I1'\lambda$ (where $I1$ is an I x 1 vector of ones) to give the BCC VRS DEA model as:

$$Minimise_{\theta,\lambda}\ \theta \tag{5.22}$$

$$subject\ to \quad -y_i + Y\lambda \geq 0$$

$$\theta x_i - X\lambda \geq 0$$

$$I1'\lambda = 1$$

$$\lambda \geq 0$$

The CCR–BCC DEA models have been used extensively and have also undergone various modifications and extensions to incorporate computation of other efficiency measures, including output-oriented efficiency, allocative efficiency, cost efficiency and profit efficiency.

For instance in the case of cost efficiency score computation using DEA, the following cost minimisation problem is solved:

$$Minimise_{\lambda,xi^*}\ w_i'x_i^*, \tag{5.23}$$

$$subject\ to \quad -y_i + Y\lambda \geq 0$$

$$x_i^* - X\lambda \geq 0$$

$$I1'\lambda = 1$$

$$\lambda \geq 0$$

where $w_i$is a N x 1 vector of input prices for the i-th firm and x_i^* (computed by the LP) is the cost-minimising vector of input quantities for the i-th firm, given the input prices $w_i$and output levels y_i.[58]

The modifications and extensions of the DEA notwithstanding, the underlying philosophy and fundamental characteristics of the DEA framework in constructing the frontier using non-parametric techniques and attributing all the deviation from the frontier as measures of inefficiencies are retained.

To conclude our review of the DEA technique, we note that the main merits attributable to the DEA framework are its simplicity, as well as the fact that it does not require any functional specification of the production process or any distributional assumptions regarding inefficiency. Accordingly, it avoids problems associated with model sensitivity and functional instability. However, DEA has a number of drawbacks. The major drawback lies in its deterministic nature that rules out the existence of noise and other measurement errors and attributes all deviations from the frontier to inefficiency. Second, and as a consequence of the first point, DEA is particularly sensitive to the presence of outliers in the data. Third, it precludes the possibility of performing direct statistical tests on the results. Fourth, the assumption of homogeneity of firms implicit in DEAs is often untenable.

5.3.2 Stochastic Frontier Approach (SFA)

The stochastic frontier model uses econometric methods for the measurement of efficiency. Being stochastic in nature, it distinguishes between deviations from the frontier that arise due to random errors and other deviations that measure 'true' inefficiency. Accordingly, the stochastic frontier model has a composite error term (ε_i) made up of a normally distributed random error term (v_i) and an inefficiency term (u_i) assumed to follow a one-sided distribution.

The origin of the SFA model is attributed to two papers simultaneously carried out by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) referred to subsequently as ALS and MB respectively. The stochastic production frontier, using cross-sectional data, is specified as follows:

$$lny_i = f(lnx_i;\ \beta) + \ \varepsilon_i \tag{5.24}$$

[58] We refer the reader to Coelli et al. (2005) and Ray (2004) who discuss in detail the applications of DEA.

where y is the output, $f(.)$ is the production function, x is the vector of inputs and β is the parameter vector to be estimated. ε_i is a composite error term defined as $\varepsilon_i = v_i - u_i$. v_i is a two-sided random error term capturing the effects of statistical noise, and follows the standard assumption of being independently and identically normally distributed with zero mean and constant variance (σ_v^2), that is, $v_i \sim iidN(0, \sigma_v^2)$. u_i is the inefficiency term capturing each firm's shortfall in output relative to the frontier.[59]

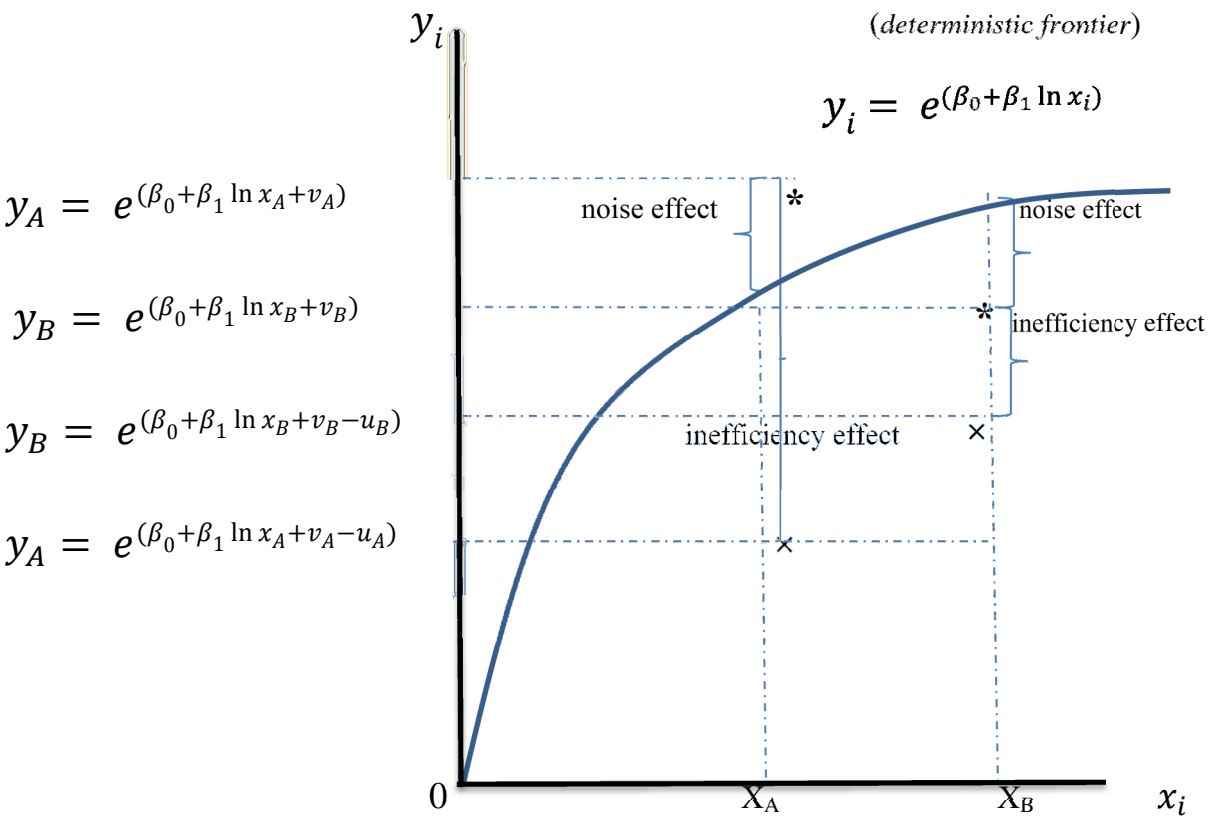
We highlight the features of the stochastic frontier model in Figure 5.5, where we have a stochastic production frontier model with a single output using a single input specified as:

$$\ln y_i = \beta_0 + \beta_1 \ln x_i + v_i - u_i \tag{5.25}$$

Equation 5.25 can be written as:

$$y_i = \underbrace{e^{(\beta_0 + \beta_1 \ln x_i)}}_{\textit{deterministic frontier}} \text{ x } \underbrace{e^{(v_i)}}_{\textit{noise}} \text{ x } \underbrace{e^{(-u_i)}}_{\textit{inefficiency}} \tag{5.26}$$

Figure 5.5 Stochastic production frontier



Source: Coelli et al. (2005)

[59] u_i is one-sided ($u_i \geq 0$) as inefficiency can only reduce output

Figure 5.5 shows the inputs and outputs of two firms, A and B, with Firm A using input X_A to produce output Y_A while Firm B uses input X_B to produce output Y_B (observed values are indicated by the points × while the frontier values are marked *).

The key point from Figure 5.5 is that the frontier output for Firm A lies above the deterministic part of the production frontier only because the noise effect is positive while the frontier output for Firm B lies below the deterministic part since the noise effect is negative.

Thus, noise can be positive or negative reflecting the assumption of v_i being a two-sided random error term. Inefficiency effect is one sided as the observed output is always below the deterministic part of the frontier.

The measure of technical efficiency can be computed as the ratio of observed output to the stochastic frontier output, that is:

$$TEi = \frac{y_i}{e^{(\beta_0+\beta_1 \ln x_i+v_i)}} = \frac{e^{(\beta_0+\beta_1 \ln x_i+ v_i-u_i)}}{e^{(\beta_0+\beta_1 \ln x_i +v_i)}} = e^{-u_i} \quad (5.27)$$

The estimation of stochastic frontier models is carried out by maximum likelihood, which requires distributional assumptions on the one-sided inefficiency term, u_i. The literature offers different specifications for the distribution of u_i, including (i) the half normal distribution with zero mean, that is, $u_i \sim N^+(0, \sigma_u^2)$ (ALS, 1997); (ii) the exponential distribution, $\mathcal{E}(\sigma_u)$ (MB, 1977); (iii) the truncated normal distribution with $u_i \sim N^+(\mu, \sigma_u^2)$ introduced by Stevenson (1980) and (iv) the gamma distribution proposed by Greene (1990).

In the original models by ALS (1977) and MB (1977), u_i is assumed to follow a half-normal and exponential distribution, respectively. These are considered as somewhat restrictive as the half-normal distribution with zero mean implicitly assumes that firms' average inefficiency is zero and suggests that most firms are operating near full efficiency. The development of the truncated normal and gamma distributions relaxes the restrictions of the earlier distributions and provides some flexibility. This flexibility however naturally increases the complication of estimation.

The estimation of efficiency using SFA involves two stages: first, estimation of the parameters of the underlying production technology and estimates of the parameters of the distributions of the error terms. The second stage involves obtaining estimates of firm-specific

inefficiency, which requires disentangling estimates of statistical noise and inefficiency from the estimated composite residuals. We explain the estimation techniques in both stages below.

Recalling our stochastic production function in equation (5.24) and assuming a half-normal inefficiency term, we have:[60]

$$lny_i = f(lnx_i;\ \beta) + \ v_i - u_i \tag{5.28}$$

$$v_i \sim iidN(0, \sigma_v^2)$$

$$u_i \ \sim iid\ N^+(0, \sigma_u^2)$$

where, as before, $\varepsilon_i = \ v_i - u_i$

Estimation of the parameters using maximum likelihood estimation requires derivation of the likelihood function. To form the density of lny_i underlying the likelihood function, the assumptions on the distribution of the error terms are used to derive the joint density function of u_i and v_i, which is given as:

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v}.\exp\{-\frac{u^2}{2{\sigma_u}^2} - \frac{v^2}{2{\sigma_v}^2}\} \tag{5.29}$$

Since $\varepsilon_i = v_i \ - u_i$, the joint density function for u_i and ε_i is

$$f(u_i, \varepsilon_i) = \frac{2}{2\pi\sigma_u\sigma_v}.\exp\{-\frac{u^2}{2{\sigma_u}^2} - \frac{(\varepsilon+u)^2}{2{\sigma_v}^2}\} \tag{5.30}$$

Finally, the marginal density function of ε can be obtained by integrating u out of the joint density $f(u, \varepsilon)$ which yields

$$f(\varepsilon_i) = \int_0^\infty f(u_i, \varepsilon_i)\, du$$

The marginal density function is accordingly given as (Greene, 2007):

$$= \frac{2}{\sigma\sqrt{2\pi}}.\phi\left(\frac{\varepsilon_i}{\sigma}\right).\left[\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)\right] \tag{5.31}$$

[60] We use the half-normal case for illustration purposes as it is the first to be developed under the SFA literature.

where $\sigma^2 = {\sigma_u}^2 + {\sigma_v}^2$ and $\lambda = {\sigma_u}/{\sigma_v}$ and $\Phi(.)$ and $\phi(.)$ are the standard normal cumulative distribution and density functions.

The parameterisation from ${\sigma_u}^2$ and ${\sigma_v}^2$ to σ^2 and λ is convenient, since λ provides an indication of the relative contributions of u and v to ε.[61] In particular, if $\lambda \to 0$ then ${\sigma_u}^2 \to 0$ so that there is no inefficiency in the disturbance, and we can estimate the model by OLS. On the other hand, if $\lambda \to +\infty$ then ${\sigma_v}^2 \to 0$ which represents a deterministic production frontier model with no noise.

The marginal density function $f(\varepsilon)$ is asymmetrically distributed with mean and variance, given by:

$$E(\varepsilon) = -E(\varepsilon) = -\sigma_u\sqrt{\frac{2}{\pi}} \qquad \text{and} \qquad V(\varepsilon) = \frac{\pi-2}{\pi}{\sigma_u}^2 + {\sigma_v}^2 \tag{5.32}$$

From the marginal density function in equation (5.31), the log likelihood function for the half-normal case based on a sample of N producers is given as:

$$\ln L = -N\ln\sigma - constant + \sum_i\{ln\,\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right) - \frac{1}{2}\left(\frac{\varepsilon_i}{\sigma}\right)^2\} \tag{5.33}$$

The log likelihood function in equation (5.33) can be maximised with respect to the parameters to obtain the maximum likelihood estimates of all parameters.

After obtaining the estimates of the technology parameters and the two parameters σ_u and σ_v (or σ and λ), the second stage is to obtain estimates of firm specific efficiencies. The estimated residuals of the model is the ε_i and not the u_i.

The standard estimator of u_i is given by the conditional mean function $E(u|\varepsilon)$. A technique developed by Jondrow, Lovell, Materov, and Schmidt (1982), hereafter referred to as JLMS which is based on the conditional distribution of u_i given ε_i is commonly is used. JLMS show that if $u_i \sim N^+(0, \sigma_u^2)$, the conditional distribution of u given ε is:

$$f(u|\varepsilon) = \frac{f(u,\varepsilon)}{f(\varepsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*}.\exp\{-(\frac{u-\mu_*)^2}{2{\sigma_*}^2}\}/[1-\Phi(-\frac{\mu_*}{\sigma_*})] \tag{5.34}$$

[61] An alternative parameterization sometimes used for some other forms of the model is $\gamma = {\sigma_u}^2/{\sigma^2}$.

where $\mu_* = -\varepsilon\sigma_u{}^2/\sigma^2$ and $\sigma_*{}^2 = \sigma_u{}^2\sigma_v{}^2/\sigma^2$

Further, since $f(u|\varepsilon)$ is distributed as $N^+(\mu_*, \sigma_*^2)$, the mean of the distribution can serve as a point estimator for u_i, given as[62]

$$E(u_i|\varepsilon_i) = \mu_{*i} + \sigma_* \left[\frac{\phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)}{1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)} \right]$$

$$= \sigma_* \left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{1-\Phi(\varepsilon_i\lambda/\sigma)} - \left(\frac{\varepsilon_i\lambda}{\sigma}\right) \right] \quad (5.35)$$

Finally, once we obtain the point estimators of u_i, the technical efficiency of each firm can be derived as as $TE_i = \exp(-\hat{u}_i)$, where $\hat{u}_i$ is either $E(u_i|\varepsilon_i)$ or $M(u_i|\varepsilon_i)$.

Battese and Coelli (1988) suggested the alternative point estimator for TE_i given by

$$TE_i = E(\exp\{-u_i\}|\varepsilon_i) = \left[\frac{1-\Phi\left(\sigma_* - \frac{\mu_{*i}}{\sigma_*}\right)}{1-\Phi\left(-\frac{\mu_{*i}}{\sigma_*}\right)} \right] . \exp\left\{-\mu_{*i} + \frac{1}{2}\sigma_*{}^2\right\} \quad (5.36)$$

The estimated technical efficiencies (TE_i) lie between 0 and 1, where $TE_i = 1$ is the most efficient situation. Usually $TE_i < 1$ and is the measure of the shortfall of observed output from the maximum feasible output.

To summarise the relative advantages and disadvantages of the DEA and SFA models, we note that the DEA technique is simple to apply, does not require any functional specification of the production process or any distributional assumptions regarding inefficiency. Its main challenge lies in its deterministic nature that rules out the existence of noise and other measurement errors and attributes all deviations from the frontier to inefficiency. Second, we noted that DEA is particularly sensitive to the presence of outliers in the data and third, that it precludes the possibility of performing direct statistical tests on the results.

[62] The mode of the distribution, $M(u_i|\varepsilon_i)$ can also serve as a point estimator but the mean is frequently used.

We deem SFA to have a stronger appeal than DEA in the light of the above drawbacks since SFA allows for statistical noise which addresses the random noise problem associated with the DEA method. This is crucial in the context of developing country studies where there could be measurement errors in data and variables. Second, SFA facilitates statistical tests of hypotheses to be conducted on the functional type of the production structure and the degree of inefficiency. Third, SFA is a powerful tool for incorporating and quantifying effects of policies as well as firm characteristics which influence inefficiency. While the main criticisms of SFA are that there is no *a priori* justification for the selection of both functional form of the production structure and particular distributional assumption of the inefficiency term, these can be addressed by the tests of hypotheses that are conducted on the empirical estimates.

Accordingly, in line with the objectives of the thesis, in analysing deregulation, and bank specific characteristics such as ownership and size on banking efficiency, this study adopts SFA for the empirical efficiency estimation. We therefore proceed to discuss some key issues relating to this approach and relevant to our study in the next section.

5.4 Key issues in SFA model estimation relevant to this study

Our research objectives require capturing deregulation, ownership and size variables in the stochastic frontier model as a way of assessing the impact of these variables on efficiency. The availability of a panel dataset for the study also requires a review of the approaches in modelling stochastic frontier within the panel data context. We discuss in this section these two related issues: first, the application of panel data techniques for stochastic cost frontier modelling, and secondly, the incorporation of exogenous variables as determinants of inefficiency in SFA models.

5.4.1 SFA panel data models

The issue of how to model inefficiency in a panel dataset context is a major point in the analysis of panel data. Panel data facilitates the estimation of bank-specific inefficiencies over time, and therefore makes it possible to examine the impact of reform policies on efficiency of banks overtime.

Stochastic frontier modelling using panel data was motivated by Schmidt and Sickles (1984) who outline how the use of panel data overcomes problems inherent in using cross-section data. First, estimated firm-specific inefficiency is not consistent in the cross-section case. Although the composite error term for a firm can be estimated consistently, it is made up of both random noise and inefficiency, and the variance of the distribution of inefficiency conditional on the composite error term does not vanish when the sample size increases. If one has sufficiently large time period, the inefficiency of a particular firm can be estimated consistently, as addition of more observations on the same firm yields information not attainable by adding more firms. Second, with a panel dataset one need not make the strong distributional assumptions on normality of the random noise term and the various assumptions on the inefficiency term made in a cross-section dataset for the estimation of the model. Third, estimates of the model parameters and of firms' inefficiencies can be obtained without assuming that inefficiency is uncorrelated with the regressors. This assumption is deemed not accurate as the knowledge of a firm's level of inefficiency could affect its input choices.

Based on the above Schmidt and Sickles (1984) proposed the following panel data stochastic frontier model (assuming *i* firms observed over *t* time periods):

$$y_{it} = \alpha + f(x_{it};\, \beta) + \epsilon_{it} \tag{5.37}$$

with ϵ_{it} defined as $\epsilon_{it} = v_{it} - u_i\,,\ u_i \geq 0,\ i = 1, \dots, N; t = 1, \dots, T$

The authors assume u_i to be fixed parameters which can be allowed to be correlated with the regressors. Moreover, the distributional assumptions on u_i can be dispensed with.

The inefficiency term u_i then is firm-specific but time-invariant and can be combined with the common intercept, α, that is, $\alpha_i = \alpha - u_i$ so that we can re-write the model in (5.37) as:

$$y_{it} = \alpha_i + f(x_{it};\, \beta) + v_{it} \tag{5.38}$$

Model (5.38) then reduces to a standard fixed effects model and conventional panel data estimations can be applied, for example, using the least squares dummy variable (LSDV) approach or within estimator to estimate the parameters.

Once the $\hat{\alpha}_i$s are estimated, the following transformation is used to recover the estimated value of $\hat{u}_i$:

$$\hat{u}_i = \max\{\hat{\alpha}_i\} - \hat{\alpha}_i \ \geq 0,\ 0, i, \dots, N \tag{5.39}$$

from which firm-specific inefficiency can be obtained in the usual way using

$$TE_i = \exp(-\hat{u}_i)$$

If instead u_i is assumed to be a random variable and uncorrelated with the regressors, then conventional random effects estimation using generalised least squares (GLS) technique provides more efficient estimates than the fixed effects model (Kumbhakar and Lovell, 2003).

Assume u_i is a random variable; and let $E(u_i) = \mu$ and $u_i^* = u_i - \mu$

Then model (5.37) can be re-written as:

$$y_{it} = \alpha^* + f(x_{it};\, \beta) + v_{it} - u_i^* \tag{5.40}$$

where $\alpha^* = \alpha - \mu$

The model in (5.39) is similar to a standard RE panel data model and the GLS estimator can be applied to estimate the parameters, from which the estimates of u_i can be estimated in a similar way as:

$$\hat{u}_i = \max\{\hat{u}_i^*\} - \hat{u}_i^*$$

The above fixed effects and random effects specifications are the first class of panel data stochastic frontier model estimators, which relax the distributional assumptions of the error terms and are estimated using the standard panel data estimation techniques.[63]

Notwithstanding the strong appeal for the fixed effects and random effects models in dispensing with the distributional assumptions in estimating inefficiency, a major concern is the assumption that inefficiency is time-invariant. As Greene (2007) notes, while intuition suggests that the longer the panel, the better the estimator of time-invariant inefficiency, at the same time, the longer the time period, the less tenable this assumption becomes. The assumption that inefficiency is bank-specific but time-invariant is therefore a strong one as it presupposes that banks neither learn over time nor competitive forces play no role in

[63] An alternative approach to the random effects model though was proposed by Pitt and Lee (1981) in which distributional assumptions on the random components of the model can be imposed and parameters estimated by maximum likelihood. Firm-specific inefficiency can then be estimated using JLMS in the usual way as discussed earlier.

enhancing efficiency. It is an assumption which is deemed untenable in a dynamic banking environment especially one which has undergone various reforms.

We therefore consider some time-varying inefficiency models which were subsequently developed to overcome the restrictiveness of the time-invariant model.

The development of time–varying efficiency models follows from the restrictiveness of the time-invariant models. Various time–varying efficiency models have been developed for panel data SFA estimation (see Kumbhakar and Lovell, 2003; Kumbhakar et al., 2015 for detailed reviews), but are of the general form:

$$y_{it} = \alpha + f(x_{it};\, \beta) + v_{it} - u_{it} \tag{5.41}$$

Using the Schmidt and Sickles (1984) fixed effects model, Cornwell, Schmidt, and Sickles (1990) proposed a time-varying inefficiency model by replacing α_i in equation (5.38) by α_{it} where

$$\alpha_{it} = \alpha_{0i} + \alpha_{1i}t + \alpha_{2i}t^2 \tag{5.42}$$

In this specification, the parameters α_{0i}, α_{1i} and $\alpha_{2i}$are firm-specific and t is the time trend, and it allows efficiency to vary through time and in a different manner for each firm. Inefficiency is still modelled using $\hat{u}_{it} = \max\{\hat{\alpha}_{it}\} - \hat{\alpha}_{it}$, and the most efficient firm can change from year to year.

The main challenge with this time-varying fixed effects model is that it can be over-parameterised in cases with large N and small T (typical of most panel data models) as there would be too many parameters (3N parameters in the α_{it} function in equation (5.42)) to be estimated.

A less heavily parameterised fixed effects frontier model proposed by Lee and Schmidt (1993) involves specifying u_{it} in equation (5.41) specified as:

$$u_{it} = \lambda(t)u_i \tag{5.43}$$

where both $\lambda(t)$ and u_i are deterministic, and so can be estimated without imposing any distribution on u_i. The problem with the Lee-Schmidt model is the presence of too many

parameters in $\lambda(t)$, which makes it unsuitable for cases of large T, and therefore appropriate only for short panels.

As seen, the above class of models are fixed effects time-varying models as the inefficiency term is non-stochastic but a parametric function of time. Accordingly, these models are estimated using distribution-free approaches.

An extension proposed by Kumbhakar (1990) is to add a time effect to the fixed effects model. This class of time-varying models accordingly have both deterministic and stochastic components, and thus, distributional assumptions can be made on the inefficiency term and estimated with maximum likelihood.

The general specification of this class of models is:

$$y_{it} = \alpha + f(x_{it};\, \beta) + \epsilon_{it} \tag{5.44}$$

$$\epsilon_{it} = v_{it} - u_{it}$$

where

$$u_{it} = G(t)u_i$$

$$v_{it} \sim N(0, \sigma_v^2)$$

$$u_i \sim N^+(\mu, \sigma_u^2)$$

where $G(t) > 0$ is a function of time (t).[64]

In this class of models, inefficiency is the product of a deterministic function of time $G(t)$ and a stochastic firm component, u_i. The inefficiency term is thus not fixed for a given firm, but changes over time as well as across firms.

Some of the specific models developed in this class of models include:

Kumbhakar (1990) who specifies

$$G(t) = [1 + \exp(\gamma_1 t + \gamma_2 t^2)]^{-1} \tag{5.45}$$

[64] Given that $u_i \geq 0$, to ensure that $u_{it} \geq 0$ requires $G(t) > 0$.

so that $G(t)$ is monotonically increasing or decreasing depending on the signs and magnitude of $\gamma_1$and γ_2.

Battese and Coelli (1992) proposed a simplified formulation by assuming

$$G(t) = \exp[-\gamma(t - T)] \tag{5.46}$$

where T is the terminal period. The model, referred to as the time decay model allows inefficiency to increase or decrease exponentially depending on the sign of γ.

Kumbhakar and Wang (2005) who proposed a similar model in which:

$$G(t) = \exp\left[\gamma\left(t - \underline{t}\right)\right] \tag{5.47}$$

and where $\underline{t}$ is the beginning period sample.

Notably, the above models of time-varying technical inefficiency models assume that inefficiency is driven by time and ignores other variables that can influence efficiency. This assumption of time variability being solely responsible for efficiency is therefore not plausible when efficiency is affected by exogenous or environmental variables. We therefore review the incorporation of exogenous or environmental variables in the stochastic frontier estimation as determinants of inefficiency.

5.4.2 Accounting for exogenous factors as determinants of inefficiency in SFA model

Exogenous or environmental variables are so called because they are strictly neither inputs nor outputs within the production process, but nonetheless exert some influence on efficiency. As discussed in Chapter 4, deregulation policies can create incentives for banks to enhance their efficiency, while differences in ownership status can also influence efficiency levels of banks. We denote the vector of exogenous variables as $z_i = (z_1, \dots, z_n)$. The incorporation of these exogenous factors as potential covariates of inefficiency in the stochastic frontier model has followed two main approaches: the two-step and the one-step approaches.[65]

[65] An earlier approach of incorporating z_i directly into the structure of the production process as $lny_i = ln\, f(x_i; z_i;\, \beta) + v_i - u_i$ has been discarded. The model assumes that z_i and u_i are uncorrelated, which means

(a) The two-step approach

The two-step approach assumes that z_i influences y_i indirectly through its effect on estimated efficiency and was extensively used in the earlier literature. Estimation of the parameters of the stochastic frontier model and firm-specific inefficiencies are obtained without controlling for the exogenous variables in the usual way in the first step. The second step then seeks to explore the determinants of inefficiency by regressing the (in)efficiency scores derived in the first step on these environmental variables.[66]

The two-step approach however has serious econometric anomalies. It must be assumed that the elements of z_i are uncorrelated with the elements of x_i. If they are correlated, then the model estimated in the first-step is 'mis-specified' and so the maximum likelihood estimates of the parameters in the first step are biased due to the omission of the relevant variables z_i. (Wang & Schmidt, 2002) further show that even if z_i and x_i are uncorrelated, ignoring the dependence of the inefficiency on z_i will cause the first stage technical efficiency indices to be serious under-dispersed, so that the results of the second-step regression are likely to be biased downwards, and they show that this is true regardless of whether z_i and x_i are correlated. Another criticism of the two-step approach is that in the first step it is assumed that inefficiencies are identically distributed but this assumption is contradicted in the second-step regression in which predicted efficiencies are now assumed to have a functional relationship with z_i. Due to these serious shortcomings of the two-step approach, recent empirical applications have been based on the one-step approach.

(b) The one-step approach

The one-step approach involves estimating the stochastic frontier model and the inefficiency term expressed as a function of the exogenous variables simultaneously using maximum likelihood in a single step. The one-step procedure accounts for the exogenous influences on inefficiency by parameterising the distribution of u_i as a function of the exogenous variables (z_i) that are potential correlates of inefficiency, although different approaches have been proposed based on the location of the distribution.

The first (and probably most widely used) approach is to parameterise the mean of the pre-truncated distribution as a way of studying the exogenous influences on inefficiency. Models

---

that variation in inefficiency is not explained by variation in the exogenous variables. It therefore provides little relevance to these exogenous variables as determinants of inefficiency.

[66] Since the dependent variable (efficiency) in the second step is usually bounded between zero and one, limited dependent variable estimation technique such as the Tobit model is usually employed at this stage.

based on this approach include: Kumbhakar, Ghosh, and McGuckin (1991), Huang and Liu (1994) and Battese and Coelli (1995), collectively referred to as KGMHLBC class of models. The models adopt the truncated-normal distribution of the inefficiency term but abandon the constant-mean assumption, μ, and assume that the mean of the distribution of the pre-truncated inefficiency term is a linear function of the exogenous variables, that is,

$$u_i \sim N^+(z_i\delta, \sigma_u^2) \tag{5.48}$$

where, as before, z_i is the vector of exogenous variables and δ is the corresponding coefficient vector to be estimated.

The second approach parameterises the *variance* of the pre-truncated distribution of the inefficiency term by the vector of exogenous variables. The class of models in this category include models by Caudill and Ford (1993), Caudill, Ford, and Gropper (1995), and Hadri (1999) (also collectively referred to as CFCFGH). The models were originally developed to account for heteroscedasticity in the inefficiency term. It accordingly parameterizes heteroscedasticity by a vector of observable variables such that

$$\sigma_u^2 = \exp\ (\ z_{u,i}; w_u) \tag{5.49}$$

where $z_{u,i}$ is the vector of exogenous variables and w_u is the parameter vector. Kumbhakar et al. (2015) point out that the CFCFGH model also addresses the issue of inefficiency effects of the exogenous variables by parameterising the variance of the inefficiency term by the exogenous variables.

The third approach by Wang and Schmidt (2002) parameterises both the mean and the variance of the inefficiency term as a function of the exogenous variables, with $u_{it} \sim N^+(\mu_{it}, \sigma_{it}^2)$ as before and where $\mu_{it} = z_{it}\delta$ and $\sigma_{it}^2 = \exp(z_{it}\gamma)$. Wang and Schmidt (2002) argues that either of the two earlier approaches, KGMHLBC and CFCFGH is plausible by showing that the first two moments equations of μ_{it} are functions of μ_{it} and, σ_{it}, and so there is no reasonable basis for preferring one model over the other. Wang however proposes a double parameterisation of both μ_{it} and σ_{it} by the same vector of exogenous variables which therefore encompasses both the KGMHLBC and CFCFGH as special cases.

A unique attribute of Wang's model is its ability to accommodate non-monotonic efficiency effects. A non-monotonic effect exists between two variables if their values are positively related in part of the parameter space, and negatively related in the rest. In particular, Wang and Schmidt (2002) demonstrates that z_{it} can have, within a sample, *both* positive and negative marginal effects on the production efficiency, and that the sign of the effect depends on the values of z_{it}.[67] The marginal effect in the class of KGMHLBC models is monotonic in that it is only *either* efficiency-enhancing *or* efficiency-impeding depending on the sign of the δ coefficient. The downside of Wang's model is that it is more complex and as a result convergence might be a problem.

From the proposed models of handling exogenous variables as determinants of inefficiency in the one-step approach, the KGMHLBC model has been widely used as it is deemed a sensible approach in investigating the exogenous influences on efficiencies. Another appeal of the KGMHLBC model is that it makes the distributional shape of u_i even more flexible as each observation has an observation-specific mean of the pre-truncated distribution, with the mean determined by observation-specific variables.

We briefly review the three models of KGMHLBC, in which as we noted, the mean of the pre-truncated distribution is parameterised as a linear function of the exogenous variables as specified in equation (5.48).[68]

Kumbhakar et al. (1991) model is a cross-sectional model of the form:

$$\ln y_i = ln(x_i\beta) + v_i - u_i \tag{5.50}$$

where $$u_i = \delta' z_i + \varepsilon_i \tag{5.51}$$

so that the one-step production frontier model is written as:

$$\ln y_i = ln(x_i\beta) + v_i - (\delta' z_i + \varepsilon_i) \tag{5.52}$$

[67] An illustration is the issue of age of a farmer, where a young farmer's efficiency may benefit from an increase in age as experiences accumulate, but an aged farmer, however, may likely suffer from efficiency loss, because of deteriorated mental and physical capability.

[68] We provide a review of these models as discussed in detail in Kumbhakar and Lovell (2003).

Since $u_i \geq 0$ it requires that $\varepsilon_i \geq -\delta' z_i$ Distributional assumptions can be imposed on v_i and ε_i, together with the restriction $\varepsilon_i \geq -\delta' z_i$ in order to derive the likelihood function.

KGM suggest the use of the model in equation (5.50), impose distributional assumptions on v_i and u_i and ignore ε_i. In that case, once it is assumed that $v_i \sim (0, \sigma_v^2)$ and $u_i \sim N^+(\delta' z_i, \sigma_u^2)$, and that v_i and u_i are distributed independently, the parameters in equation (5.50) can be estimated using maximum likelihood.

The log likelihood function is accordingly given as:

$$\ln L = constant - \frac{1}{2}\ln(\sigma_v^2 + \sigma_u^2) - \sum_i \ln \Phi\left(\frac{\delta' z_i}{\sigma_u}\right) + \sum_i \ln \Phi\left(\frac{\mu_{*i}}{\sigma_*}\right) - \frac{1}{2}\sum_i \frac{(e_i + \delta' z_i)^2}{\sigma_v^2 + \sigma_u^2} \quad (5.53)$$

where

$$\mu_{*i} = \frac{\sigma_v^2 \delta' z_i - \sigma_u^2 e_i}{\sigma_v^2 + \sigma_u^2}$$

$$\sigma_*{}^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$$

and $e_i = \ln y_i - ln(x_i\beta)$ are the residuals obtained from the estimation of equation (5.50).

From the maximisation of the log-likelihood function, we can obtain ML estimates of $(\beta, \delta, \sigma_v^2, \sigma_u^2)$, from which firm-specific estimates of technical inefficiency can be obtained using the JLMS decomposition.

Huang and Liu (1994)'s model is quite similar but with inefficiency specified in the form

$$u_i = g(z_i; \delta) + \varepsilon_i \quad (5.54)$$

so that the single-step production frontier model is:

$$\ln y_i = ln(x_i\beta) + v_i - [g(z_i; \delta) + \varepsilon_i] \quad (5.55)$$

To ensure that $u_i = [g(z_i; \delta) + \varepsilon_i] \geq 0$ means truncating ε_i from below such that $\varepsilon_i \geq -g(z_i; \gamma)$, and by assigning a distribution to ε_i such as $\varepsilon_i \sim (0, \sigma_\varepsilon^2)$. Thus instead of truncating a normal distribution with variable mode from below at zero (as in KGM), Huang and Liu truncate a normal distribution with zero mode from below at a variable truncation

point$[-g(z_i;\delta)]$. This allows $\varepsilon_i \leq 0$ but enforces $u_i \geq 0$. Estimation is also undertaken by maximum likelihood. A novelty of this model is that the function $g(z_i;\delta)$ is allowed to include interactions between elements of z_i and elements of x_i.

These two models were developed in the context of using cross-section data. The Battese-Coelli (1995) model (hereafter BC95) extends Huang and Liu's model but within the panel data context.[69]

The BC95 model of the stochastic production function is specified as follows:

$$\ln y_{it} = x_{it}\beta + v_{it} - u_{it} \tag{5.56}$$

$$u_{it} = z_{it}\delta + \varepsilon_{it}$$

The u_{it}s are assumed to be independently distributed, such that u_{it} is obtained by truncation at zero of the normal distribution such that $u_{it} \sim (z_{it}\delta, \sigma_u^2)$. ε_{it} is a random variable defined by the truncation of the normal distribution with zero mean and variance σ_ε^2 such that the truncated point is at $-z_{it}\delta$, that is, $\varepsilon_{it} \geq -z_{it}\delta$ which guarantees $u_{it} > 0$.

In other words, the non-negativity requirement $u_{it} = z_{it}\delta + \varepsilon_{it} \geq 0$ is modelled as $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$, with the distribution of ε_{it} being bounded below by the variable truncation point$-\delta' z_{it}$. Battese and Coelli note that this distribution assumption on ε_{it} is consistent with the distributional assumption on u_{it} that $u_{it} \sim N^+(\delta' z_{it}, \sigma_u^2)$.

The technical efficiency of the *i*th firm is given by:

$$TE_i = E(e^{-u_i}) = e^{(-\delta' z_i - \varepsilon_i)} \tag{5.57}$$

Battese and Coelli (1993) formulised the conditional expectation of the technical efficiency, $e^{(-u_{it})}$ as:

$$E[\exp\{-u_i\}\,|v_i - u_i]\Big| = \left[\exp\{-u_* + \tfrac{1}{2}{\sigma_*}^2\}\right].\left[\frac{\Phi\left(\frac{\mu_*}{\sigma_*} - \sigma_*\right)}{\Phi\left(\frac{\mu_*}{\sigma_*}\right)}\right] \tag{5.58}$$

[69] The BC95 model however does not include inputs in the specification of the inefficiency function

where

$$\mu_{*i} = \frac{\sigma_v^2(\delta^{'} z_i) - \sigma_u^2(\varepsilon_i)}{\sigma_v^2 + \sigma_u^2}$$

$$\sigma_*{}^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$$

Once the technical efficiency is estimated, the partial effect of each exogenous variable on efficiency can be derived.

5.5 Conclusions

This chapter has reviewed the conceptual framework of frontier efficiency and discussed the various techniques in measuring efficiency. It includes the structure of the production technology and its functional characterisation using production frontiers and distance functions in analysing technical efficiency. We then reviewed the concept of economic efficiency including cost, revenue and profit efficiency measures where behavioural objectives are imposed, and noted the dominant use of cost efficiency in most empirical efficiency analysis. We also noted that the parametric stochastic frontier model of measuring efficiency provides a stronger appeal than the deterministic non-parametric approach. We finally examined the main methods of incorporation inefficiency covariates in the one-step approach.

In view of the analysis in this chapter, and in line with the objectives of the thesis, we employ a time-varying stochastic cost frontier which incorporates exogenous factors such as deregulation reform indices, ownership and size variables within a panel data context. We accordingly adopt the BC95 model of estimating a one-step stochastic cost frontier model for the empirical analysis in Chapter 7.

CHAPTER 6 THE IMPACT OF DEREGULATION REFORMS ON COMPETITION IN GHANA'S BANKING SECTOR

6.1 Introduction

Although banking remains the dominant sector of Africa's financial system, competition in the sector remains low, notwithstanding the modest gains from the first wave of reforms during the 1990s. In the case of Ghana, existing evidence also suggests that the earlier reforms had a limited impact on banking competition due to lingering challenges such as high concentration, low penetration of new banks, fragmented banking system, and high borrowing by governments from the banking system as discussed in Chapter 2 (Antwi-Asare and Addison, 2000; Buchs and Mathisen, 2005; Biekpe, 2011).

The recent deregulation reforms implemented during 2003-2006 were aimed at enhancing competition in the sector by addressing those existing challenges. In particular, the relaxation of product and geographical restrictions on banking activities with the introduction of universal banking in 2003 was geared towards addressing the fragmented banking sector and creating a level playing field to facilitate competition. The policy of opening up of the industry through the licensing of new banks was also aimed at enhancing contestability and competition, and to address the problem of high concentration in the industry. The scrapping of the mandatory secondary reserve requirements was also intended to relax the credit control constraints imposed by the policy.

As expounded in Chapter 2, these deregulation reforms led to a significant structural transformation of the banking sector, including the licensing of new banks, the rapid growth in the industry's branch network, the introduction of new products and services, the decline in concentration levels, the emergence of more regional banks, the curtailment of the dominance of state banks, and greater opportunities for credit expansion and deeper financial intermediation. In spite of the notable structural changes accompanying these reforms, the impact of these reforms on banking competition has not been assessed.

The purpose of this chapter therefore is to contribute to the literature by empirically assessing the impact of these recent deregulation policies on the competitiveness of Ghana's banking sector. We contribute to the literature on competition in African banking markets from several respects. First, rather than examining competition in the entire banking market, our study

examines specifically competition in the loans market. The reasons for this focus are: (i) the loans market is the largest and most important segment of the banking sector in Ghana, accounting for about two-thirds of total investible funds, in line with the primary credit intermediation role of banks; (ii) it is also the most challenging and uncompetitive segment of the banking sector, characterised by high interest rates and limited access, and was the main target of the reforms; (iii) the money market's competitiveness is driven largely by government's fiscal operations and thus assessing overall banking competitiveness might be clouded by the competitiveness or otherwise of this market; (iv) other segments of the banking market are less developed in Ghana to warrant their study of competitiveness; and (iv) the reform policies, taken together, are expected to have a major impact on the credit market, which therefore deserves greater attention.

Second, we estimate two separate models of competition models, the persistence of profits (POP) and Boone Indicator (BI) models, which is also different in two respects. First, we do not use the Panzar–Rosse H-statistic (P-R) model which has been extensively used in Africa, and second, we use the BI, which to our knowledge, has not been used in any country-specific banking competition study in sub-Saharan Africa. Our choice is made for reasons of robustness (competition models can often lead to confliction results) as well as completeness (the models allow for a different analysis of the dynamics of competition). We provide the justification for the choice of these models in the methodology section below. We account for the pre- and post-reform competition measures by using a bilateral policy reform dummy variable in the POP model to capture the effects of the reforms on competition. The BI model estimates competition on an annual basis and therefore also facilitates analysis of the trend in competition measures before and after the reforms.

Third, our rich panel dataset of 25 banks over a relatively long period which captures both the pre- and post-reform periods, not only enriches the reliability of the model estimations and quality of empirical analysis, but also provides a more effective way of assessing policy impacts on competition. This enhances the quality of the resulting policy recommendations, and overcomes the limitations of previous research based on insufficient data.

The rest of the chapter is organised as follows. Section 6.2 details the methodology adopted for the empirical estimation, while Section 6.3 presents the results of the empirical models and a discussion of the results of the competition models. Section 6.4 concludes.

6.2 Methodology and data

As mentioned, we use the POP and BI models, and do not utilise the P-R model, which, despite its many shortcomings, has been heavily relied upon in most studies on banking competition in Africa. Almost all banking competition studies on African countries have employed the P-R model, and even the few recent studies employing multiple models use other models (such as the Lerner Index, Conjectural Variation and POP models) to complement the P-R in their empirical research work (see Kasekende et al., 2009; Mwenga, 2011; Poshakwale and Qian, 2011; Biekpe, 2011; and Simpasa, 2013).

Despite its popularity with some of the literature, and as we discussed in detail in Chapter 3, the P-R model suffers from some serious limitations that in some cases can entirely invalidate its results. First, the model is based on the assumption of long-run equilibrium which, if not met, makes the results inconclusive. Indeed it is argued that the coexistence of banks of different sizes within the banking sector, especially in the context of developing countries, is strong evidence of long-run disequilibrium, which undermines the reliability of competition results from this approach (Bikker et al., 2012). Second, the P-R model does not take into account banks' responses to reform policies and competitive pressures as it ignores interdependence of firms. Thus for policy-impact assessment studies such as ours, the model cannot be relied upon. While attempts at overcoming this challenge have included a dynamic specification of the P-R model, achieving long-run equilibrium on an annual basis is a more daunting challenge. Goddard and Wilson (2009) argue that the P-R model is based on a static equilibrium framework, but in practice the speed of adjustment towards equilibrium is less than instantaneous and markets might be out of equilibrium. Thus, the static equation used in the P-R model is mis-specified. Third, the P-R model is biased towards monopolistic competition. Another limitation is that the H-statistic index derived from the model is not continuous and hence it is difficult to meaningfully compare estimated H-statistic indices between two periods or on an annual basis so as to interpret trends in the evolution of competition using this measure. It is for the above reasons that, in spite of its wide application in earlier banking studies, recent studies have preferred alternative specifications.

The motivation for our selection of the POP and BI models is that they overcome the above limitations of the P-R model. First, they both assume interdependence among banks and that banks will react to competitive pressures induced by policy reform changes, and thus measure banks' responsiveness to reform policies (Delis, 2012; Schaeck and Cihák, 2014). They are

both dynamic models and can reliably measure competition levels over time and therefore facilitate impact analysis of policy reforms on competition.

In addition, the underlying theories of both models and the nature of the deregulation reforms make them particularly suited to our purposes. More specifically, the POP model examines competition in the light of the level of persistence of profits over a period of time. The underlying theory of the POP model is that in the absence of entry barriers, excess profits enjoyed by incumbent banks will attract new banks and this will lead to a gradual erosion of such excess profits, so that persistence of profits is expected to reduce to signal increasing competition. The model accordingly relates market contestability to competition and especially in this context where the reforms involve the relaxation of entry restrictions and activity restrictions which are expected to drive down the level of profit persistence that existed initially. This makes the POP model an appropriate choice for our study.

In the case of the BI, its underlying theory as expounded in Chapter 3 is that competition is expected to increase via two channels: (i) when bank products become closer substitutes and banks interact more aggressively; and (ii) when entry costs decline so as to facilitate the entry of new players. In both cases, the performance of more efficient banks should improve (Boone, 2008b). The reforms implemented in Ghana targeted exactly these two points. The universal banking policy is expected to foster stronger product substitution as it creates a common platform for all banks with similar products and services, thus enhancing the substitutability of products compared to the previous fragmented nature of the industry. Also, entry costs are expected to decline with the removal of entry restrictions that led to an increase in the number of banks in the sector. This makes the Boone indicator particularly suited to our case study. Finally, both models are well adapted for measuring competition in the loans market which is the focus of our study. It is on the basis of the above reasons that we deem the choice of the POP and BI as the most appropriate models to adequately address the research questions being investigated.

6.2.1 Empirical specification of the Persistence of Profits (POP) model

A review of the empirical work on the application of the POP model in banking competition studies shows that the POP model is usually formulated as a first-order autoregressive model of banks' profitability (Bektas, 2007; Kaplan and Çelik, 2008; Zhao et al. 2010; Goddard et al., 2011; Mwega, 2011; and Pervan et al., 2015). The definition of profitability depends on the type of competition being assessed. A broader view of competition that encompasses the

whole banking sector uses overall profitability or profit rate of the banks, measured by return on assets (ROA) or return on equity (ROE). For competition in the loans market, as in our case, loan 'profitability' is based on loan overcharge, which is defined as the ratio of the price of loans to the marginal cost of loans. A perfectly competitive loan market occurs where loan price equals their marginal cost so that the loan overcharge is unity.

We therefore formulate a partial adjustment model of loan price overcharge to examine the persistence of the loan overcharge over time. If the loan overcharge persists over time, then the level of competition is low; vice versa, a lower persistence of the loan overcharge reflects increasing competition. The pricing of loans reflects banks' perception of changes in competitive conditions in the credit market. Where competition in the loans market intensifies in a particular year, banks will seek to reduce the loan overcharge in order to remain competitive, and so the persistence of profits will fall.

We follow Zhao et al. (2010) and formulate our basic POP model as a partial adjustment model of loan overcharge (*LOC*) where the actual change in loan overcharge is expressed as a function of the change towards the optimum loan overcharge, as follows:

$$\ln LOC_{it} - \ln LOC_{it-1} = \lambda(\ln LOC^*_{it} - \ln LOC_{it-1}) + \varepsilon_{it} \quad (6.1)$$

In equation (6.1), LOC_{it} is the loan overcharge of bank *i* at time *t*, defined as the ratio of implicit loan price to marginal cost of loans; correspondingly.[70] We define LOC^*_{it} as the optimum loan overcharge, i.e. the value under perfect competition, which is unity.[71] The adjustment of the overcharge towards unity therefore reflects the dynamic evolution from imperfect to perfect loan market competition.

The parameter λ in equation (6.1) measures the speed of adjustment towards the optimum LOC^*_{it} and ε_{it} is the error term, with $\varepsilon_{it} \sim iid(0, {\sigma_\varepsilon}^2)$.

The intuition behind equation (6.1) is that we are looking at the change in the loan overcharge as a function of its distance from the desired level as λ measures the speed of adjustment

[70] We use the term implicit loan price as it is derived from the ratio of interest income on loans (from the income statement) and the volume or size of loans (from the balance sheet) due to the difficulties in obtaining explicit loan prices for each customer for each bank. The implicit loan price accordingly includes risk premium for credit risk and a normal profit margin.

[71] In this case loan price equals marginal cost, and thus $ln\ LOC^*_{it}$ equals zero.

towards the desired level and the extent to which deviations from the optimum loan overcharge feed into deviations from actual overcharge. A priori, we expect that $0 < \lambda < 1$. If $\lambda = 1$, the change in loan pricing reflects the optimum change and is therefore a case of perfect competition as the loan overcharge will equal the optimum loan overcharge. The larger is λ the faster is the adjustment speed towards perfectly competitive prices. A low λ on the other hand represents a less competitive environment as the adjustment towards perfect competition is slow.

Since the study seeks to investigate specifically the impact of the policy reforms on competition we introduce a bilateral policy reform dummy, R, to capture their effect. In recognition of the fact that the full implementation of the reforms was completed in 2006, our policy variable, R takes a value of 0 for the pre-reform period (2000–2006) and a value of 1 for the post-reform period (2007–2014). The dummy is interacted with the explanatory variable to measure the difference in the parameter λ between the two periods.[72]

Our policy-interacted model can accordingly be expressed as follows:

$$(\ln LOC_{it} - \ln LOC_{it-1}) = \lambda(\ln LOC^{*}_{it} - \ln LOC_{it-1}) + \delta R(\ln LOC^{*}_{it} - \ln LOC_{it-1}) + \varepsilon_{it} \quad (6.2)$$

where the parameter δ measures the change in the speed of adjustment during the post-reform period, with $0 < \lambda + \delta < 1$.

A positive and statistically significant δ implies a faster speed of adjustment towards a perfectly competitive situation and therefore an increase in the intensity of competition. If δ is negative and statistically significant the adjustment to the competitive level in the post-reform period is slower, suggesting a decline in competition. If δ is not statistically significant, we can infer then that there is no significant change in the level of competition between the two periods.

[72] In view of possible lagged effects of policy reforms on their intended outcomes we also estimate models in which the policy reform year is changed from 2006 to 2007 and 2008 respectively. The results of the alternative specifications are captured in Appendix 6.3 but did not show significant differences. Based on the tests of model adequacy using the r-squared and the F-tests, the model with policy reform at 2007 was chosen.

As defined earlier, the optimum loan overcharge $LOC_{i_{it}}^{*}$ is the loan overcharge of perfect competition where the loan price equals marginal cost, so that $LOC\ _{i,t}^{*}$ is unity, and thus ln LOC_{it}^{*} is zero. Substituting this in equation (6.2), and rearranging (6.2) we obtain:

$$\ln LOC_{it} = (1-\lambda)\ln LOC_{it-1} - \delta R * \ln LOC_{it-1} + \varepsilon_{it}$$

which can be re-written as:

$$\ln LOC_{it} = \alpha \ln LOC_{it-1} + \gamma R * \ln LOC_{it-1} + \varepsilon_{it} \quad (6.3)$$

where $\alpha = 1-\lambda$ and $\gamma = -\delta$

The interpretation of the parameter estimates from equation (6.3) is explained as follows. The parameter α is the persistence parameter as it measures the persistence of LOC_{it-1} into LOC_{it}. Given that $\alpha = 1-\lambda$, we can obtain the adjustment parameter, λ, from the estimated persistence coefficient as $\lambda = 1-\alpha$. A high value of α therefore means a low value of λ and consequently low competition, and vice versa. The parameter δ can also be derived as $-\gamma$ from equation (6.3). A significantly positive γ means a significantly negative δ and will indicate a slower speed of adjustment during the post-reform period and thus a decline in competition, and vice versa. Put differently, α measures the pre-reform persistence of loan overcharge and $\alpha+\gamma$ measures the post-reform persistence of loan overcharge.

To complete our model, we also account for macroeconomic factors, industry-specific variables and other exogenous factors that could potentially impact changes in the loan overcharge in the model.[73]

Our complete estimable POP model is accordingly specified as follows:

$$\ln LOC_{it} = \alpha \ln LOC_{it-1} + \gamma R * \ln LOC_{it-1} + \kappa_1 \ln RGDPG_t + \kappa_2 \ln RMPR_t + \kappa_3 \ln RTBR_t + \kappa_4 \ln FISCAL_t + \kappa_5 \ln RDEPN_t + \omega HHI_t + \tau CRISES + \varepsilon_{it} \quad (6.4)$$

[73] This can be done with time dummies (Zhao et al., 2010) or by the inclusion of specific macroeconomic and industry variables (Mwega, 2011; Poshakwale and Qian, 2011).

In equation (6.4), α is the persistence parameter of loan overcharge for the pre-reform period, while γ is the change in the persistence parameter between the pre- and post-reform period. The post-reform period's persistence parameter of loan overcharge is accordingly measured by $\alpha + \gamma$.

If γ is not significantly different from zero, then there is no change in the persistence and thus no change in competition between the two periods. If γ is significantly different from zero, and is positive then the post reform persistence parameter $(\alpha + \gamma)$ is bigger and reflects a reduction in competition. On the other hand, if γ is negative, then the post reform persistence parameter $(\alpha + \gamma)$ is smaller, and indicates an increase in competition during the post-reform period. In terms of the priors, $0 < \alpha < 1$ and $0 < \alpha + \gamma < 1$.

RGDPG is real GDP growth rate; RMPR is real monetary policy rate; RTBR represents the real Treasury bill rate; FISCAL is the fiscal balance to GDP ratio; and RDEPN is the real currency depreciation rate. These annual macroeconomic indicators enter the model in logarithmic form so as to facilitate interpretation of the coefficients as elasticities. We also include one industry-specific variable: the Herfindahl-Hirschman Index (HHI) of total assets as a measure of market concentration which could impact on the loan overcharge.[74] We finally introduce a dummy variable, CRISES, to capture any possible impact of the global financial and economic crises of 2008 on the loan overcharge and competitiveness of the banking sector, with value 1 for 2008 and beyond, and 0 otherwise. While the global financial crisis did not adversely affect African banking systems directly via the contagion effect, as a result of the low integration of Africa's financial system with the global financial system, the global economic recession that ensued indirectly had an adverse effect on domestic banking markets through higher interest rates in most African countries (IMF, 2009).[75]

As of the expected effects of the explanatory variables, for real GDP, it is anticipated that high economic growth enhances business opportunities, reduces business and credit risk and therefore leads to a lower loan overcharge. A slowdown in GDP growth limits business

[74]We also explored using time dummies to capture the exogenous macroeconomic and industry specific variables as was used in Zhao et al. (2009). The results of this model specification were however not better than our alternative specification. Including both time dummies and these variables resulted in most of them being dropped due to strong multicollinearity.

[75] The IMF (2009) report noted the spill-over of the global financial crisis to economic crisis, with depressed global demand for commodities resulting in weakening commodity prices, lower exports and lower government revenues, decline in capital flows and remittances, and a virtual dry-up of external financing resources. Under such conditions, there was increased pressure on domestic markets to accommodate demand for credit from both government and the private sector, and could result in high domestic interest rates in African economies.

opportunities, increases business and credit risks, and potentially increases default risks (Flamini, Schumacher, & McDonald, 2009). Accordingly, this will drive up loan overcharge. We therefore expect a negative relationship between real GDP growth and loan overcharge.

The monetary policy rate (MPR) is the policy rate used by the Bank of Ghana (BoG) in its conduct of monetary policy and is determined by the Monetary Policy Committee (MPC) of BoG.[76] The MPR is usually raised to curb higher inflationary pressures by increasing lending rates so as to curtail demand for loans as well as raise savings rates to increase savings. Changes in RMPR therefore have a direct impact on lending rates (Kovanen, 2011) and thus a positive relationship between RMPR and loan overcharge is expected.

The Treasury bill rate (TBR) is the rate on government securities which is an alternative investment instrument for banks. Increases in the TBR, reflecting higher domestic borrowing by the government, provide high returns to banks on investments in such risk-free treasury bills, and this could result in upward adjustments in the loan overcharge rates by banks. Real Treasury bill rates have been found to be positively related to interest rate spreads (Beck & Hesse, 2009) and thus a positive relationship between RTBR and loan overcharge is also anticipated.

Government fiscal operations also impact on the banking sector. Large fiscal deficits are expected to lead to high borrowing from the banking sector, which could feed into an increase in loan overcharge to the private sector. Fiscal surpluses vice versa will lead to reduced borrowing by government. A negative relationship is therefore expected between the FISCAL variable and loan overcharge.

We also account for the possible impact of changes in the exchange rate, currency depreciation, on loan overcharge. Changes in nominal exchange rates are expected to impact on interest rates (Beck and Hesse, 2009) and thus on loan overcharge, with currency depreciation leading to higher loan overcharge and vice versa.

On the industry specific variables, we examine the impact of concentration on loan pricing and the loan overcharge. It is argued that banks in more concentrated markets are able to collude and adjust loan rates, although this argument has been debated in both the theoretical

[76] The MPR is usually announced by the MPC at its quarterly meetings following an assessment of economic conditions and inflationary expectations.

and empirical literature. It is expected that the coefficient of the concentration measure, HHI will be positive.[77]

The CRISES dummy variable is expected to have an adverse impact by increasing LOC as explained above. The tightening of credit markets from the deleveraging that followed the crises seems to have had spill over effects on African banking systems. Accordingly, the coefficient of the CRISIS variable is expected to be positive.[78]

In summary, our primary focus and coefficients of interest in the POP model in equation (6.4) are the persistence parameter, α and changes in the persistence parameter, γ to assess both the competition level and changes in the intensity of competition between the pre- and post-reform periods. The macroeconomic and industry-specific variables provide insights into the determination of the loan overcharge and hence other determinants of competition while the crises dummy tests the impact of the global financial and economic crises on competitive conditions in Ghana's banking system.

6.2.2 Empirical modelling of the Boone Indicator (BI) model

As noted earlier, the BI model measures banking competition by relating bank performance to differences in efficiency. In terms of model specifications, a review of the empirical application of the Boone indicator in banking studies shows that the efficiency of banks is measured in terms of marginal costs while performance can be measured in terms of profits or market share of loans (see for instance, Schaeck and Čihák, 2008; Delis, 2012; Clerides et al., 2013; Van Leuvensteijn et al. 2013; Brissimis et al., 2014; and Xu et al., 2014). The established literature using the Boone model shows that bank performance is measured using profits when competition in the overall banking market is being analysed, while where competition in the loans market is being analysed, bank performance is usually measured using marginal cost of loans (Schaeck and Čihák, 2008; Van Leuvensteijn et al., 2011; Van Leuvensteijn et al., 2013).

Accordingly, since our focus is on analysing competition in the loans market as indicated earlier, in formulating our empirical Boone model, we follow Van Leuvensteijn et al., (2011)

[77] Here we include HHI as a proxy for the potential number of banks as the POP model is based on the contestable market theory which states that it is the potential and not the actual number of banks that matter.
[78] We introduce the CRISES dummy in 2008 but also test it in 2009 due to possible lagged effect of the global crises on African economies.

and adopt the measurement of bank performance using market share of loans, while efficiency in the loan market is proxied by the marginal cost of loans.

The basic Boone indicator model following from the above is specified as follows:

$$ln\ MS_{it} = \alpha_0 + \beta ln\ MC_{it} + \varepsilon_{i,t} \quad (6.5)$$

where MS_{it} is the market share of loans of bank *i* at time *t*; MC_{it} is the marginal cost of loans of bank *i* at time *t*, which is a proxy for efficiency; and β is the Boone measure of competition.

A priori, β is expected to be negative due to the inverse relationship between marginal cost and loan market share. Although there is no threshold of reference for the Boone indicator, the higher the absolute value of β the greater the level of competition (Boone, 2008b).

The estimated Boone indicator from the model specification in (6.5) gives the average competition measure for the period under study and is very useful for cross-country studies to analyse and compare average competitive conditions across banking sectors in different countries.

In this study, since our research interest is on the impact of the policy reforms on competition, we seek to investigate whether there has been any change in the competitive environments between the pre-reform and post-reform periods. We therefore specify a variant of the Boone model which enables us to examine the evolution of competition annually by computing the Boone measure on a yearly basis. This approach has been used in recent empirical studies on banking systems in developed countries by Schaeck and Čihák (2008), Delis (2012) and Van Leuvensteijn et al. (2011).

We follow Van Leuvensteijn et al. (2011) and specify our estimable model, tracking competition changes between 2000 and 2014, as follows:

$$ln\ MS_{it} = \alpha + \sum_{t=1}^{T} \beta_t\, D_t\ ln\ MC_{it} + \sum_{t=1}^{T-1} \Gamma_t D_t\ +\ \varepsilon_{it} \quad (6.6)$$

In equation (6.6) the vector of parameters β_t is the Boone measure of annual competition, tracking its evolution; D_t is a time dummy to control for factors common to all banks in the industry and specific to each year; and ε_{it} is the error term, with $\varepsilon_{it} \sim iid(0,\ {\sigma_\varepsilon}^2)$.

The dependent variable MS_{it} is the market share of loans of bank *i* in year *t*, computed as the share of each bank's net loans to the total industry loans On the right hand side, MC_{it} is the marginal cost of loans, which is not directly observable but will be derived following estimation of a total cost function.

6.2.4 Empirical modelling of the cost function

The specification of both the POP and Boone indicator models discussed above require the derivation of the marginal cost of loans, and hence the estimation of a cost function as marginal costs are not observed directly. The empirical estimation of a total cost function for the banking sector requires several considerations including the choice of an appropriate functional form of the cost function, the input-output composition of a bank's cost function, the measurement of quantities and prices of such inputs and outputs, the identification of other possible cost drivers, and post estimation tests of the regularity conditions to ensure that the cost function satisfies its theoretical properties as discussed in Chapter 5.

With regards to the choice of an appropriate functional form, we note that some of the frequently used functional forms in the literature include the Cobb-Douglas, the transcendental logarithmic (translog), and the Fourier functional forms. Given its flexibility, and in line with most of the empirical literature on cost functions, we choose the translog specification. The translog function provides a second order approximation without imposing a priori restrictions on the production technology, and its choice is based on the following reasons. First, as noted by Kumbhakar and Lovell (2003), the translog cost function is able to accommodate multiple outputs without necessarily violating curvature conditions, and is thus very apt for this study where multiple outputs are defined for the banking sector. The derivation of marginal costs from an estimated translog cost function is deemed more accurate and closely in line with economic theory, as the multi-product nature of the translog facilitates the derivation of marginal costs of different segments of the market, such as the loans market that this study is focusing on (Van Leuvensteijn et al. 2011). Second, it is a flexible functional form thus making fewer assumptions, than alternative forms, on the structure of the production process, in particular on production and substitution elasticities. Non-flexible functional forms such as the Cobb-Douglas, Leontief and linear functions are too restrictive as they place a priori restrictions on the substitution possibilities among the factors of

production.[79] It is however possible to impose restrictions on the parameters (homogeneity conditions) to ensure that the estimated model complies with theoretical properties of a cost function. The translog has received extensive attention in empirical estimation of cost efficiencies and is frequently used in application to the banking sector (Kumbhakar and Lovell (2003; Weill 2013).

The translog cost function requires the identification of inputs and their prices, outputs and total costs. The composition of inputs and outputs of banks, like most service sectors, is however not straightforward. The surveyed empirical literature distinguishes between the production approach and the intermediation approach in defining banks' input-output mix. The difference between the two approaches is therefore the treatment of deposits – whether as input or output. The production approach considers banks as producers, utilising inputs such as physical capital and labour to produce banking outputs such as deposits, loans and other banking services, and therefore treats deposits as output. The intermediation approach, as the name suggests, sees the production process in terms of the financial intermediation role of banks and therefore treats deposits as input (Sealey and Lindley, 1977).

In line with most of the empirical literature we adopt the intermediation approach; this is especially reasonable in our case given that we examine competition in the loans market, and in this context deposits can definitely be seen as one of the fundamental inputs. It is argued that the production approach is more suitable in cases such as the evaluation of bank branch performance, where deposits mobilisation is a key output required of branch managers. However, for overall performance assessments and analyses of banks, the intermediation approach is more favoured.

Based on the intermediation approach, our input variables are total customer deposits and other borrowed funds, capital and labour.[80] We define three output variables: performing loans (LOANS), other earning assets (OEA), and fee and commission income, used as a proxy for Off-Balance Sheet items (OBS). OBS items have grown in importance in most banking sectors and ignoring them would underestimate output levels of banks and could lead to incorrect implications for cost efficiency and productivity (Clark and Siems, 2002; Casu

[79] For instance, the Cobb-Douglas specification assumes that all firms have constant production elasticities and that substitution elasticities equal unity.

[80] Other borrowed funds relate to borrowings by banks from mainly development finance institutions (DFI) such as the African Development Bank, the Netherlands Development Finance Company (FMO), Proparco, among others, which are to support lending to small and medium sized businesses.

and Girardone, 2006).[81] We use commission income as a proxy due to its relatively ease of computation and follows other studies such as Zhao et al. (2010) and Van Leuvensteijn et al. (2011) that have used this proxy. It is lack of data that often leads empirical application to make use of proxies, as we do.

Based on the two inputs: loanable funds (deposits plus other borrowed funds) and capital (labour and physical capital and fixed assets), the respective input prices are calculated as the ratio of interest expense to loanable funds (price of loanable funds - P_{FUNDS}), and the ratio of operating costs to total assets (price of labour and capital to capture both labour-related and capital-related expenditure - P_{LK}). The non-separation of labour-related costs from capital-related expenditure is due to lack of specific data on staff costs distinct from other operational costs. Details of the variables used and their measurements are discussed in section 6.2.6.

The translog cost function is therefore expressed with 3 outputs (LOANS, OEA and OBS) and two input prices (P_{FUNDS} and P_{LK}). To capture the effect of technological changes on cost a time trend, T, is also included, in quadratic form. The trend variable is also interacted with the input price and output variables to model both non-neutral and scale augmenting technology changes respectively.

Our translog cost function incorporating 3 outputs (y_1, y_2, y_3) and 2 input prices (w_1, w_2) and our time trend and its interaction input prices and outputs can be specified as follows:[82]

$$ln\,TC_{i,t} = \alpha_0 + \sum_{j=1}^{2}\beta_j\ lnw_{jit} + \sum_{k=1}^{3}\gamma_k\ lny_{kit} + \sum_{j=1}^{2}\sum_{m=1}^{2}\beta_{jm}\,lnw_{jit}\,lnw_{mit} + \sum_{k=1}^{3}\sum_{l=1}^{3}\gamma_{kl}lny_{kit}\,lny_{lit} + \sum_{j=1}^{2}\sum_{k=1}^{3}\psi_{jk}lnw_{jit}\,lny_{kit} + \theta_1 T + \theta_{11}T^2 + \sum_{j=1}^{2}\eta_j\,Tlnw_{jit} + \sum_{k=1}^{3}\zeta_k\ Tlny_{kit} + \varepsilon_{i,t} \quad (6.7)$$

For the above specified cost function to correspond to a well-behaved production structure, the estimated cost function must satisfy these three key regularity conditions: (i) linear homogeneity in input prices; (ii) symmetry; and (iii) monotonicity in input prices and outputs (Kumbhakar et al., 2015). The first two regularity conditions are customarily imposed prior to estimation while the monotonicity condition is tested for afterwards.

[81] While studies have argued for inclusion of OBS activities as bank outputs, the different measures of computing such OBS activities all have drawbacks as enumerated in Casu and Girardone (2002). Fee and commission income is however used as an appropriate proxy due to its relative ease of computation compared to the others especially in the context of developing countries where relatively less financial data is provided on banks' activities.

[82] We use (y_1, y_2, y_3) to represent (LOANS, OEA and OBS) and (w_1, w_2) for (P_{FUNDS} and P_{LK}) for ease of expressing them in the equation format.

Linear homogeneity implies imposing the following restrictions on the parameters prior to estimation:[83]

(i) $\sum_{j=1}^{2} \beta_j = 1 \; or \; \beta_1 + \beta_2 = 1$;

(ii) $\sum_{j=1}^{2} \sum_{m=1}^{2} \beta_{jm} = 0 \; or \; \beta_{11} + \beta_{12} = 0; \; \beta_{21} + \; \beta_{22} = 0;$

(iii) $\sum_{j=1}^{2} \sum_{k=1}^{3} \delta_{jk} = 0 \; or \; \psi_{11} + \psi_{21} = 0; \psi_{12} + \; \psi_{22} = 0; \; \psi_{13} + \; \psi_{23} = 0.$

It can be demonstrated that this can be achieved by normalising total costs and prices by one of the input prices, that is, by dividing TC and all the terms involving w_1 by w_2. We therefore define $TC^* = TC/w_2$ and $w_1{}^* = w_1/w_2$.

Also since the translog cost function is continuous and twice differentiable, the second cross derivatives are symmetric; hence we impose the symmetry conditions as follows: $\beta_{jm} = \beta_{mj}$ and $\gamma_{kl} = \gamma_{lk}$ prior to estimation of the translog cost function.

Following the imposition of the above restrictions on the parameters, our estimable translog cost function becomes:[84]

$$ln\,(TC^*{}_{it}) = \alpha_0 + \beta_1 \ln\left(w_1{}^*{}_{it}\right) + \beta_{11} \ln\left(w_1{}^*{}_{it}\right)^2 + \gamma_1 lny_{1it} + \gamma_{11}(lny_{1it})^2 + \gamma_2 lny_{2it} + \gamma_{22}(lny_{2it})^2 + \gamma_3 lny_{3it} + \gamma_{33}(lny_{3it})^2 + \gamma_{12} lny_{1it} lny_{2it} + \gamma_{13} lny_{1it} lny_{3it} + \gamma_{23} lny_{2it} lny_{3it} + \psi_{11} ln\, w_{1it}{}^* lny_{1it} + \psi_{12} ln\, w_{1it}{}^* lny_{2it} + \psi_{13} ln\, w_{1it}{}^* lny_{3it} + \theta_1 T + \theta_{11} T^2 + \eta_1 l\, n\, w_{1it}{}^* T + \zeta_1 lny_{1it} T + \zeta_2 lny_{2it} T + \zeta_3 lny_{3it} T + \varepsilon_{it} \quad (6.8)$$

The monotonicity condition is derived from production theory and requires that total cost should be non-decreasing in input prices and output. Accordingly, monotonicity in outputs requires positive marginal costs and monotonicity in prices requires that total cost increases as input prices increase. This means that the partial derivative of total costs with respect to the three outputs and two input prices at each observation must be positive. In other words,

$\frac{\delta lnTC_{it}}{\delta ln\, y_{1it}} > 0$ $\frac{\delta lnTC_{it}}{\delta ln\, y_{2it}} > 0$ $\frac{\delta lnTC_{it}}{\delta ln\, y_{3it}} > 0$ $\frac{\delta lnTC_{it}}{\delta ln\, w_{1it}} > 0$ $\frac{\delta lnTC_{it}}{\delta ln\, w_{2it}} > 0$

[83] As seen in Chapter 5, linear homogeneity in input prices means that a proportional increase in all input prices must increase cost by the same proportion, holding output constant. The non-negative cost requirement was checked at the data input stage.

[84] The cost function is used as an intermediary variable to derive the marginal cost of loans for the computation of the loan overcharge for the POP model as well as the explanatory variable for use in the Boone indicator. An implicit assumption we make here is that banks do actually achieve the goal of cost minimisation. We therefore consider the translog cost function as a regression, and not a cost frontier, an assumption which will be relaxed in Chapter 7 which examines cost efficiency using stochastic cost frontier.

Where there are few violations of these conditions, the estimated cost function could still be said to satisfy these theoretical properties. However, major violations of these conditions are quite serious and warrant further investigation (Kumbhakar & Lovell, 2003). The estimated translog model and the parameter estimates might not be meaningful if the model fails to satisfy the regularity properties.

After estimating the cost function, and undertaking the regularity tests to ensure that it conforms to the theoretical properties, we then derive the observation-specific marginal costs with respect to loans since we are looking at competition in the loans market. Accordingly, following the estimation of (6.8), the marginal cost of loans (MC_{y_1}) is computed as:

$$MC_{y_1it=} \frac{\delta TC_{it}}{\delta y_{1it}} = \frac{TC_{it}}{y_{1it}} (\gamma_{1it} + 2\gamma_{11} lny_{1it} + \gamma_{12} lny_{2it} + \gamma_{13} lny_{3it} + \psi_{11} lnw_{1it}{}^{*} + \zeta_1 T) \quad (6.9)$$

6.2.5 Econometric estimation techniques and data issues

In this section we briefly present the general approach used in estimating our models. Specific techniques and tests are carried out in detail under each of the three models estimated in Section 6.3. In terms of data, we employ a panel data set consisting of annual bank-level data for all the banks in the industry during the period 2000-2014.[85] All banks in the industry in each year were used: from 16 in 2000, rising to 25 in 2014 (the increase due, as we know, to new entrants). We therefore utilise an unbalanced panel dataset structure and employ panel data estimation techniques for all the three models: translog cost function, POP model and the Boone indicator model. Panel data analysis enables us to analyse unobserved variation across time and banks and also helps deal with omitted variable bias if there are omitted variables which are bank-specific in any of the models. Panel data models assume that intercepts vary across individual banks due to unobserved bank-specific effects, whiles slope coefficients are identical and that regression coefficients do not vary over time (Cameron and Trivedi, 2005). If the bank-specific effects are assumed to be non-stochastic (i.e. fixed), then a fixed effects model is appropriate.

In our panel data formulation, we test for bank-specific effects and make a choice as to whether fixed effects or random effects modelling is appropriate, which depends on the

[85] These banks are all currently classified as universal banks following the abolition of the commercial, merchant and development banking classifications.

assumption regarding the correlation of the bank-specific effects and the explanatory variables each model. Random effects estimation assumes that none of the explanatory variables are correlated with the unobserved bank-specific effect. Fixed effects estimation assumes that there is correlation between bank-specific effect and any of the explanatory variables.

Thus, the key factor in distinguishing between the fixed effects and random effects methodology is the assumption regarding the correlation of the unobservable bank-specific effect with the explanatory variables. It is suggested that the choice of fixed effects or random effects model can be inferred from the data generation process. In particular, where the sample data is randomly drawn from a large population, one should use random effects estimation as a fixed effect model would lead to an enormous loss of degrees of freedom due to usually large number of cross-sectional units (individuals). On the other hand, the fixed effects model is deemed an appropriate specification if the sample is on a specific set of firms and inference is restricted to the behaviour of these firms (Baltagi, 2008; Dougherty, 2011; Brooks, 2014). It is worth pointing out however that this issue relates to the poor performance of the fixed effects estimator when it comes to making predictions and thus extending inference beyond the scope of the selected sample. As pointed out by Greene (2012), "the crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, and not whether these effects are stochastic or not" (Greene, 2012, p. 347).

For each of the three models, we accordingly test for the presence of fixed effects using the Chow test and random effects using the Breusch-Pagan Lagrange-Multiplier (LM) test. We formally test whether a random effects or fixed effects model is appropriate using the Hausman test. Details of these are captured in the empirical estimations in Section 6.3.

In terms of data sources, bank-level data was compiled from audited financial statements of all the banks in operation in each year for the period 2000–2014. The data sources for the bank-data were the banks' published annual reports and compilations compiled by the research units of Ecobank Ghana and the Ghana Bankers Association.[86] As a result of the entry of new banks, mergers and acquisitions, the number of banks increased over the sample period. Accordingly we have an unbalanced panel data set of 25 banks observed over 15 years

[86] The bank data and data sources are considered richer and more reliable than bank data from Bankscope. The Bankscope database on Ghana covers only a relatively shorter time period starting from 2006. In addition, some banks have been incorrectly classified.

(2000-2014).[87] Upon cleaning the data, and accounting for missing values we have a total of 321 observations. Macroeconomic data – including real GDP growth, monetary policy rate, Treasury bill rate, annual currency depreciation and fiscal deficit to GDP rates, were obtained from the Bank of Ghana annual reports and the IMF's World Economic Outlook database. All the bank-level data were adjusted to real values using the GDP deflator with 2006 as the base year.

We define the variables used in the estimation of the translog cost function, the POP model and the Boone model in Table 6.1, and discuss their measurement as follows. For the translog cost function, total cost was derived from the income statement of banks, and comprise both interest expenses and operational costs. Performing loans were extracted from the balance sheet of banks, which is defined as gross loans less non-performing loans.[88] Other earning assets consist principally of investments in treasury bills and placements with other banks. Fee and commission income is obtained from the income statements of banks. Loanable funds consist of total deposits and short-term borrowing and bonds, and the price of loanable funds is derived by dividing total interest expense by total loanable funds. Price of labour and capital (non-interest operating cost) is obtained by dividing operating costs by total assets. Descriptive statistics of the data for the estimation of the translog cost function are presented below in Section 6.3.1

For the POP model, we compute the loan overcharge by dividing the price of loans by the marginal costs of loans. The price of loans is computed as a ratio as follows. The reported interest income in the income statement is composed mainly of interest income on loans and interest income on government securities.[89]

[87] There was only a single merger of two banks in 2012. Their pre-merger financial data was treated separately, with the combined post-merger financial data for the merged entity.

[88] By using performing loans, we control for loan quality due to the heterogeneity in the quality of loans among banks, and high non-performing loans.

[89] Placements with/from other banks are mainly on overnight basis and do not represent a core investment decision. Such daily decisions are based on a Bank's cash-flow position at the end of each day. Due to their residual nature, they are excluded in the computations. Both placements with and from other banks are ignored in the pricing computations.

Table 6.1 Definition of variables for estimation of:

Translog Cost Function (TCF)

Variable	**Notation**	**Definition/Computation of Variable**
Total Cost	TC	Interest expenses + operating expenses
Outputs		
Performing loans (LOANS)	y_1	Gross Loans less non-performing loans
Other earning assets (OEA)	y_2	Investments in government securities and placements with other banks
Fee and commission income (OBS)	y_3	Proxy for Off-Balance Sheet activities
Input Prices		
Price of loanable funds (P_{FUNDS})	w_1	Interest expense/(deposits + other borrowed funds)
Price of labour and capital (P_{LK})	w_2	Non-interest operating costs/total assets

POP model

Loan overcharge	LOC	Loan price / marginal cost of loans
Real GDP growth	RGDPG	Annual GDP growth in real terms
Real Monetary Policy Rate	RMPR	Nominal MPR adjusted by inflation
Real Treasury Bill Rate	RTBR	Nominal TBR adjusted by inflation
Fiscal balance to GDP ratio	FISCAL_BAL	Government budget balance/GDP
Real currency depreciation	RDEPN	Nominal currency depreciation adjusted by inflation
Herfindahl-Hirschman Index	HHI_ASSETS	Sum of the squares of each bank's market share of assets

Boone Indicator (BI) model

Market share of loans	MS	Loan of Bank i in year t /total loans in year t
Marginal cost of loans	MC	Derived from translog cost function

Based on published monthly Treasury bill rates, we derive the interest income on government securities component using the average flow of investments in Treasury bills reported on the balance sheets. This is then deducted from the total interest income to obtain the interest income on loan component. This interest income on loans is then divided by the loan figure to obtain the price of loans.

Real GDP growth is obtained from the Central Bank of Ghana, as provided by the Ghana Statistical Service. The nominal monetary policy rate is determined and announced by the Monetary Policy Committee (MPC) at its quarterly review of the economy meetings. Thus, in a typical year, there are 4 or so MPR rates. The annual nominal MPR rate was computed as the average of those quarterly rates. Nominal Treasury bill rates are based on monthly 91-day bill rates published by the Bank of Ghana. The use of the 91-day bill rate is due to the high component of those bills in the overall Treasury bill portfolio (Bank of Ghana, 2014). The annual rate used was accordingly derived as the average of the monthly rates published by the central bank, and therefore gives a better approximation, than say, the year-end rates. Nominal currency depreciation is computed from the annual change in the year-end exchange rate of the Cedi to the US Dollar. The real depreciation rate is computed by adjusting the nominal exchange rate by the rate of inflation. The nominal monetary policy rates, Treasury bill rates and nominal currency depreciation are also adjusted for inflation to obtain the real values. The annual fiscal balance to GDP figures were extracted directly from the Bank of Ghana Annual reports. The HHI of assets was computed for each year as the sum of the squares of the market share of assets of the banks. We present descriptive statistics of the data for the estimation of the POP model in Section 6.3.2.

With regards to the Boone indicator, the market shares of loans were obtained for each bank in each year by dividing each bank's loan portfolio by the industry total loan portfolio for the year. The marginal cost of loans is obtained from the estimated translog cost function. Descriptive statistics of the data are also presented in Section 6.3.3.

6.3 Empirical estimation and discussion of results

We present in this section the results of the estimation of the models specified in the preceding section and a discussion of the empirical results of the competition models. Prior to the estimation of each model, we also provide statistics of the data used for the estimations. We start with the translog cost function in Section 6.3.1. Based on the derived marginal costs from the estimated translog cost function, we proceed with the estimation of the POP model in Section 6.3.2 and the Boone model in Section 6.3.3.

6.3.1 Estimation of translog cost function and derivation of marginal costs

Our translog cost function model for estimation from equation (6.8) is as follows:[90]

$$ln\,(TC^*_{it}) = \alpha_0 + \beta_1 \ln(w_1{}^*_{it}) + \beta_{11} \ln(w_1{}^*_{it})^2 + \gamma_1 lny_{1it} + \gamma_{11}(lny_{1it})^2 + \gamma_2 lny_{2it} + \gamma_{22}(lny_{2it})^2 + \gamma_3 lny_{3it} + \gamma_{33}(lny_{3it})^2 + \gamma_{12} lny_{1it} lny_{2it} + \gamma_{13} lny_{1it} lny_{3it} + \gamma_{23} lny_{2it} lny_{3it} + \psi_{11} ln\, w_{1it}{}^* lny_{1it} + \psi_{12} ln\, w_{1it}{}^* lny_{2it} + \psi_{13} ln\, w_{1it}{}^* lny_{3it} + \theta_1 T + \theta_{11} T^2 + \eta_1 l\, n\, w_{1it}{}^* T + \zeta_1 lny_{1it} T + \zeta_2 lny_{2it} T + \zeta_3 lny_{3it} T + \varepsilon_{it} \quad (6.10)$$

where $\varepsilon_{it} = u_i + v_{it}$, with u_i being the unobserved bank-specific effect and $v_{it} \sim IID(0, \sigma_v^2)$.

We first explore the data on the variables used for estimation by examining the descriptive statistics of total cost, the 3 output and 2 input price variables.

(1) Descriptive statistics of data for estimation of translog cost function

Tables 6.2a and 6.2b show the mean values of the main variables for the cost function, with the trends depicted in Figure 6.8.

Table 6.2a Mean values of TC, LOANS, OEA, OBS (Amounts in GHC Million)

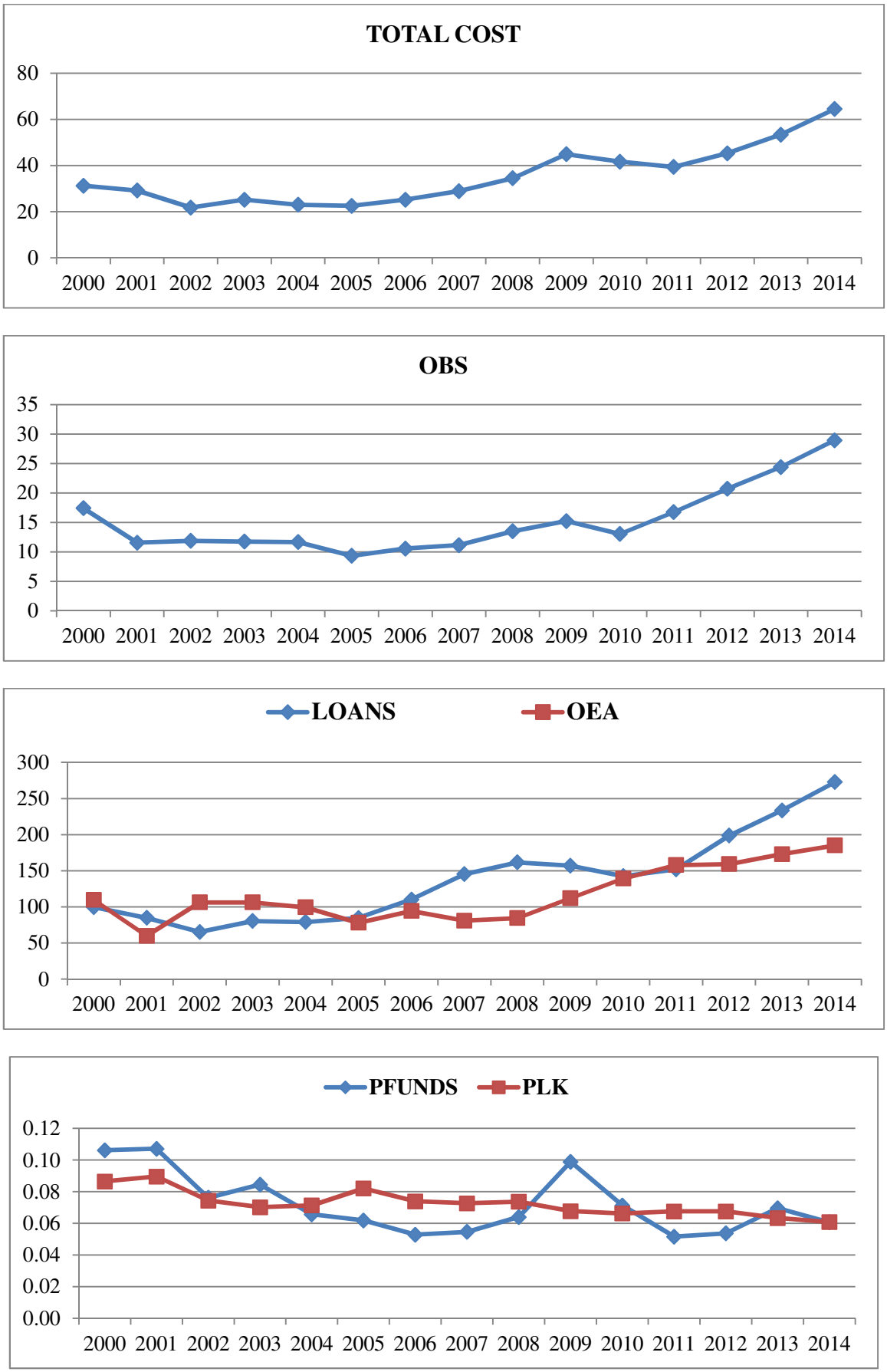
	2000	2002	2004	2006	2008	2010	2012	2014
TC	31	22	23	25	34	42	45	64
LOANS	100	65	79	110	162	143	199	273
OEA	110	106	99	94	85	139	159	185
OBS	17	12	12	11	13	13	21	29

Table 6.2b Mean values of input prices (P_{FUNDS} and P_{LK})

	2000	2002	2004	2006	2008	2010	2012	2014
P_{FUNDS} (unit price of funds)	0.106	0.076	0.066	0.053	0.064	0.071	0.054	0.061
P_{LK} (unit price of labour and physical capital)	0.086	0.074	0.071	0.074	0.074	0.066	0.068	0.061

[90] This is the specification of the estimated translog cost function. Alternative specifications including adding control variables such as total assets and equity/total assets did not yield sensible results with severe violations of the regularity conditions. In the end, we estimated the above model.

Figure 6.1 Trend in mean values of total cost, outputs and input prices (GHC Million)



The mean total cost of the banking industry shows an initial decline, inched up temporary in 2003 followed a marginal decline, and started increasing from 2006 until 2009 (Table 6.2 and Figure 6.1). Following further marginal declines in 2010 and 2011, the mean total costs increased thereafter.

The mean value of loans, after an initial decline during 2000-2002, picked up in 2003 and steadily increased to 2005, while OEA experienced a gradual reduction during that period. The growth in loans was more pronounced from 2006 to 2008 while OEA marginally declined between that period, the immediate response to the scrapping of the secondary reserves. Loans thereafter declined in 2009-2010 before rising from 2011 to 2014. OEA, after picking up in 2009 increased steadily thereafter. OBS was fairly stable for most part during 2002-2010, but picked up in 2011 and increased steadily afterwards.

For input prices, P_{FUNDS} declined for most part of 2000-2006 but increased sharply during 2008-2009. It declined in 2010-2011 but increased during 2012-2013 before declining in 2014. The trend broadly reflects developments in general interest rates. Changes in P_{LK} have been less erratic than P_{FUNDS} though with only marginal changes throughout the period.

In terms of variation of the panel distribution, we observe from Table 6.3 that there is significant variation both across banks (between variation) and over time (within variation) in all the variables. The only difference is that for the outputs and total costs, the variation across banks is higher than the within variation over the years, which is to be expected due to the relative differences in size of the banks and scale of operations. In respect of input prices, however, the within variation is greater than between variation, which suggests that banks' size does not seem to influence input prices.

(2) Estimation results of the translog cost function

We estimated equation (6.10) using panel data techniques and carried out various tests for the choice of the preferred model. The fixed effects model turns out to be the preferred model and the results are presented in Table 6.4. Details of the full results using pooled OLS, fixed effects and random effects estimations, and the results of the Hausman test for the selection of the fixed effects are shown in Appendix 6.1.

Table 6.3 Descriptive statistics of TCF variables in panel form

Variable		Mean	Median	Std. Dev.	Distribution of Variation	Min	Max
Total Cost	overall	36.486	28.434	31.472		0.549	156.366
	between			24.591	61%	1.011	113.104
	within			19.450	38%	- 0.035	128.939
P_{FUNDS}	overall	0.070	0.062	0.037		0.011	0.222
	between			0.026	48%	0.028	0.124
	within			0.026	52%	0.001	0.182
P_{LK}	overall	0.072	0.066	0.027		0.014	0.235
	between			0.018	43%	0.030	0.114
	within			0.022	67%	0.019	0.232
LOANS	overall	143.703	107.437	141.522		0.530	802.157
	between			99.578	50%	8.587	413.337
	within			100.112	50%	- 78.690	659.544
OEA	overall	118.951	67.226	132.206		0.953	669.051
	between			103.442	61%	17.638	400.945
	within			78.298	35%	- 147.131	455.249
OBS	overall	15.465	10.473	14.601		0.093	89.413
	between			11.336	60%	0.310	37.193
	within			9.003	38%	- 5.336	69.190

TC, LOANS, OEA and OBS are in GHC Million. P_{FUNDS} and P_{LK} are unit prices of funds and labour and physical capital respectively. No. of observations: 321.

Before interpreting the estimated parameters in Table 6.4, we discuss the results of various tests of hypotheses carried out to ensure model adequacy. The specific hypotheses we tested included adequacy of a more restrictive Cobb-Douglas functional form; existence of technological change; existence of non-neutral technological change; whether technology is homothetic; the relevance of deregulation in the cost frontier; and the overall significance of the inefficiency model. The results of the tests are summarised in Table 6.5.

Table 6.4 Fixed effects estimation of translog cost function (preferred model)

Ln(TC/P_{LK})	**Ln TC***	**Parameter**	**Coefficient**	**standard errors**	**t-statistic**
Ln (P_{FUNDS}/ P_{LK})	lnw1*	β_1	0.581***	0.098	5.92
Ln (LOANS)	lny1	γ_1	0.588***	0.114	5.14
Ln (OEA)	lny2	γ_2	0.282*	0.105	2.68
Ln (OBS)	lny3	γ_3	0.021	0.143	0.15
(Ln (P_{FUNDS}/ **P_{LK}**))^2	(lnw1)^2	β_{11}	0.096***	0.019	5.08
(Ln LOANS)^2	(lny1)^2	γ_{11}	0.065***	0.021	3.13
(Ln OEA)^2	(lny2)^2	γ_{22}	0.086***	0.016	5.29
(Ln OBS)^2	(lny3)^2	γ_{33}	0.033	0.025	1.31
(Ln P_{FUNDS}/ **P_{LK}**)(Ln LOANS)	(lnw1*)(lny1)	ψ_{11}	-0.005	0.030	-0.16
(Ln P_{FUNDS}/ P_{LK})(ln OEA)	(lnw1*)(lny2)	ψ_{12}	-0.038*	0.023	-1.67
(Ln P_{FUNDS}/ **P_{LK}**)(Ln OBS)	(lnw1*)(lny3)	ψ_{13}	0.042	0.035	1.22
(Ln LOANS)(Ln OEA)	(lny1)(lny2)	γ_{12}	-0.138***	0.028	-4.9
(Ln LOANS)(Ln OBS)	(lny1)(lny3)	γ_{13}	-0.004	0.040	-0.1
(Ln OEA)(Ln OBS)	(lny2)(lny3)	γ_{23}	-0.042	0.035	-1.2
T	T	θ_1	0.013	0.019	0.68
Tsqr	Tsqr	θ_{11}	0.001	0.001	1.23
T(Ln P_{FUNDS}/ P_{LK})	Tlnw1*	η_1	-0.003	0.004	-0.62
T(Ln LOANS)	Tlny1	ζ_1	-0.007	0.006	-1.11
T(Ln OEA)	Tlny2	ζ_2	-0.005	0.004	-1.02
T(Ln OBS)	Tlny3	ζ_3	0.014**	0.006	2.23
Intercept	_cons	α_0	2.025***	0.203	9.96

No. of observations: 321; R-squared: 0.979; F -test for overall significance: 2848 (Prob > F=0)

*** p<0.01, ** p<0.05, * p<0.1 (Significance at 1%, 5% and 10% level respectively)

Table 6.5 Tests of hypotheses on estimated parameters of the translog cost function

Null hypothesis (H0) defined by the following parameter restrictions. Full details provided below	Test Statistics	Decision (at 5% significant level)
(1) $\beta_{11} = \gamma_{11} = \gamma_{22} = \gamma_{33} = \psi_{11} = \psi_{12} = \psi_{13} = \gamma_{12} = \gamma_{13} = \gamma_{23} = \theta_{11} = \eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = 0$	F(15, 24) = 45.67	Reject
(2) $\eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = 0$	F(4,24) = 1.19	Do not reject
(3) $\theta_1 = \theta_{11} = \eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = 0$	F(6,24) = 2.70	Reject
(4) $\psi_{11} = \psi_{12} = \psi_{13} = 0$	F(3,24) = 1.50	Do not reject

Notes: Tests of hypotheses involves imposing restrictions on the parameters, (i.e. the parameters equal zero) of the estimated translog cost function. The tests are based on the F-test.

The null hypothesis of each of the above are summarised below:

- Hypothesis (1): the second-order coefficients in the translog function are zero, and so the Cobb-Douglas functional form is adequate to represent the data.
- Hypothesis (2): the interaction between the time trend (T) and the input prices and outputs are zero, that is, there is no non-neutral and scale augmenting technical change and there is no change in the output composition and input mix associated with the time trend.
- Hypothesis (3): all coefficients involving the time trend are zero, and therefore that there is no technical change or technological progress over time.
- Hypothesis (4): the production technology is homothetic, and so the interactive terms of input prices and output quantities are zero.

Based on the results of the various hypothesis tests, we reject the hypothesis that the second-order coefficients of the translog cost function are zero and confirm the suitability of the translog specification to represent the data rather than the Cobb-Douglas. We also reject the hypothesis that there is no technical change, and conclude that there is technological change associated with time. However, we fail to reject that there is no non-neutral and scale-

augmenting technical change present in the model. The production technology is however found to be homothetic.

We also carried out the regularity tests to check if our estimated translog cost function possesses the properties of a cost function.[91] The results of the test of the monotonicity conditions which involves checking the observation–specific cost elasticities of outputs and input prices (shown in Table 6.6) confirms that the cost elasticity with respect to both input prices as well as loans are all positive. We however note a few violations in respect of cost elasticity of other earning assets and fee-based income which are deemed minimal.[92] We therefore conclude that the estimated cost function seem to satisfy the properties of linear homogeneity, symmetry and monotonicity in input prices and outputs as required by theory.

Table 6.6 Monotonicity conditions of estimated translog cost function

	Mean	Std. Dev.	Minimum	Maximum	No. of violations of monotonicity conditions (i.e. cost elasticity < 0)
cost elasticity w.r.t. outputs					
Loans	0.505	0.103	0.085	0.877	0
Other earning assets	0.290	0.147	-0.144	0.826	6
Fee and commission income	0.080	0.082	-0.159	0.285	52
input price elasticities					
cost of funds	0.448	0.123	0.027	0.719	0
cost of labour and capital	0.552	0.123	0.281	0.973	0

Based on the estimated results of the translog cost function and the estimated average cost elasticities of outputs and input prices, we discuss below the results in some more detail. We observe that the estimated regression model shows that most of the regressors relating to LOANS and OEA as well as the normalised input price are significant at the 5% significance level. They also have the expected signs. In terms of model adequacy, our F-test of overall significance of the explanatory variables is high, and the very high R^2 also confirm the strong

[91] Having imposed linear homogeneity and symmetry conditions prior to estimation, we only test now for monotonicity conditions.

[92] These violations represent only 2% and 18% of the total observations. Such minimal violations of the monotonicity conditions are not uncommon in the literature. Large violations however raise serious concerns on the empirical estimates (Kumbhakar et al., 2015).

explanatory power of the translog model. OBS is the only output variable which is not statistically significant except for its interaction with the time trend, which is not uncommon due to the multicollinearity issues associated with such large number of explanatory variables. This suggests that the off-balance sheet activities do not seem to significantly influence total costs which further suggests limited investment in these non-core banking services. The lack of significance of OBS also explains the few monotonicity violations, which therefore do not cause any particular reason for concern.

The elasticity of loans is remarkably higher than that of other earning assets and fee-based income, underscoring the low cost associated with investments in treasury bills and generation of off balance sheet services despite the relative large share at least of the former on the balance sheets of banks. The sum of the cost elasticities of outputs also reveals scale economies in the banking industry.

The time trend (and its quadratic form) variables are not statistically significant, signifying that there have not been any technology-induced changes in the cost structure of banks. The interaction of the time trend with fee and commission income is however significant suggesting that technology-engineered changes in growing off-balance sheet activities could increase costs in the long run.

Finally, using the estimated parameters of the translog cost function, we obtain the observation-specific marginal cost of loans (MC_{y_1}), using equation (6.9) above. This is used with the derived price of loans for the calculation of the loan overcharge, which we defined as LOC = price of loans/marginal cost of loans for the estimation of the POP model, discussed in the next chapter.

6.3.2 Analysis of competition dynamics from the POP model

The first competition model we estimate is the POP model which in its estimable form was defined (in equation 6.4) as:

$$\ln LOC_{it} = \alpha \ln LOC_{it-1} + \gamma R * \ln LOC_{it-1} + \kappa_1 \ln RGDPG_t + \kappa_2 \ln RMPR_t + \kappa_3 \ln RTBR_t + \kappa_4 \ln FISCAL_t + \kappa_5 \ln RDEPN_t + \omega HHI_t + \tau CRISES + \varepsilon_{it} \quad (6.11)$$

We observed earlier that α is the pre-reform persistence of profit parameter; and γ, which measures the change in the persistence parameter for the post-reform period. As noted in Section 6.2.2, a high value of α (high persistence of profit) is indicative of a low level of competition; and vice versa. In addition, a significantly positive γ means a much higher persistence of profit ($\alpha + \gamma$) during the post-reform period and thus a decline in competition; and vice versa. The set of parameters κ capture the impact of the macroeconomic variables; ω measures the effect of concentration levels on competition and τ measures the impact of the global financial and economic crises on loan overcharge and competition.

(1) Descriptive Statistics of data on variables to be estimated in the POP model

The mean values of the banks' loan overcharge, loan price and marginal cost of loans are captured in Table 6.7 and Figure 6.2.

Table 6.7 Mean values of loan overcharge (LOC), loan price and marginal cost of loans

	2000	**2002**	**2004**	**2006**	**2008**	**2010**	**2012**	**2014**
LOC	1.376	1.555	1.775	1.615	1.450	1.974	1.952	2.146
Price of loans	0.365	0.303	0.300	0.249	0.222	0.289	0.202	0.243
MC of loans	0.292	0.210	0.175	0.152	0.157	0.150	0.119	0.122

The trend in the banking industry's mean price of loans and marginal cost of loans follow a similar pattern. There was a general decline in MC of loans during 2001-2007 except for a temporary increase in 2005. It increased during 2008-2009, declined in 2010-2011 and thereafter remained fairly stable during 2012-2014. The loan price reflects a similar pattern although it experienced an increase earlier in 2004 before declining until 2007. Unlike MC which increased only during 2008-2009, the loan price rise in 2008 continued until 2010 before the brief decline in 2011-2012 and subsequent rise during 2013-2014 (Figure 6.2a).

Reflecting the pattern of changes in the mean values of loan price and MC of loans, the mean values of loan overcharge (LOC) show marginal changes except for 2004 and 2009 where the increases were relatively bigger. Following the higher loan price and declining MC of loans in 2010, LOC shot up significantly in 2010 and steadily increased during the last three years.

Figure 6.2a Mean values of banks' loan price and marginal cost of loans

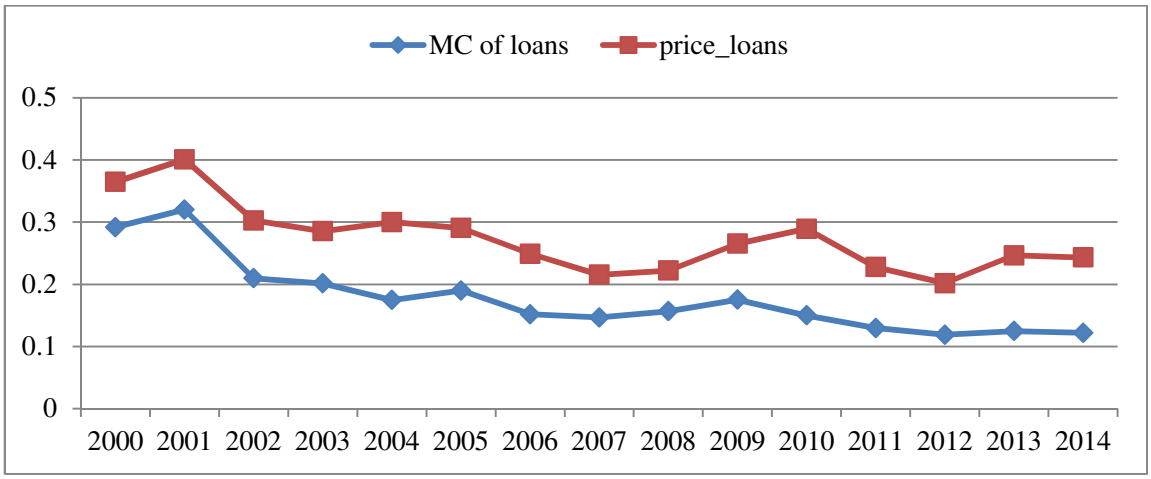
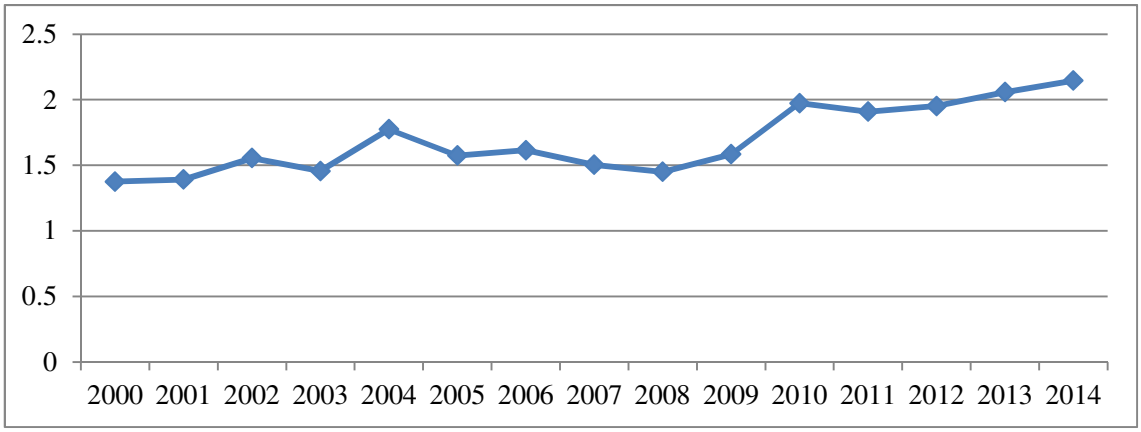


Figure 6.2b Mean values of banks' loan overcharge (LOC)



We next examine descriptive statistics of the macroeconomic and industry-specific variables used in the estimation of the POP model as captured in Table 6.8a. The descriptive statistics points to a relatively wide variability in the macroeconomic indicators in most of the variables, reflecting the relatively unstable macroeconomic environment of the country.[93]

[93] Detailed discussion of the trends in these variables and graphs are discussed in Appendix 6.2.

Table 6.8a Descriptive statistics of macroeconomic and industry variables

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
RGDP	6.32	5.7	2.59	3.70	14.00
MPR	0.19	0.17	0.05	0.13	0.27
TBR	0.20	0.19	0.09	0.10	0.42
DEPN	0.18	0.05	0.24	0.01	0.94
FISCAL	-7.35	-7.55	2.77	-11.50	-2.30
HHI_ASSETS	0.092	0.088	0.0301	0.058	0.147

Macroeconomic variables are in annual percentages (%).

We also show the correlation matrix of the explanatory variables in the POP model in Table 6.8b to analyse any potential multicollinearity problems associated with the regression arising from strongly correlated explanatory variables. From the table however, we do not note any significant correlation between any two variables except for currency depreciation and Treasury bill rates with a correlation coefficient of 0.6. Accordingly, we do not envisage any multicollinearity problems with the regression.

Table 6.8b Correlation matrix of macroeconomic and industry variables

	lnrgdp	lnrmpr	lnrtbr	lnfiscal	lnrdepn	hhi_assets
lnrgdp	1					
lnrmpr	0.2978	1				
lnrtbr	-0.5894	-0.2231	1			
lnfiscal	-0.0899	-0.118	0.4991	1		
lnrdepn	-0.1847	0.2168	0.605	0.5081	1	
hhi_assets	-0.5741	-0.1967	0.5287	-0.0612	-0.0584	1

(2) Estimation results of the POP Model

We carried out six different specifications of the POP model in which we had the policy dummy variable, R, moved from 2006 to 2007 and 2008 due to the possible lagged effects of the impact of the reforms, and interacted each of these policy reform years with the LOC. For each of these scenarios, we also explored two specifications of the financial/economic crisis

dummy variable, CRISES, in 2008 and 2009 respectively. Appendix 6.3 summarises the results of all the estimation of the six specifications of the POP model.[94]

On the bases of the overall R^2 and the F-statistic of overall significance, our preferred model is the fixed effects specification with the policy reform variable set in 2007 and the CRISES variable in 2008. The results are reported in Table 6.9.[95]

Table 6.9 Estimated results of POP model with fixed effects regression

	coefficient	*robust standard errors*	*t-statistics*	*p-value*
LnLOC_1	0.324***	0.056	5.750	0.000
R*LnLOC_1	0.185***	0.057	3.220	0.004
Ln RGDP	-0.253***	0.057	-4.400	0.000
Ln RMPR	3.101***	0.593	5.240	0.000
Ln RTBR	-1.217*	0.651	-1.870	0.074
Ln FISCAL	0.084**	0.038	2.240	0.035
Ln RDEPN	-0.674**	0.320	-2.100	0.046
HHI_ASSETS	1.035	1.774	0.580	0.565
CRISES	0.183**	0.086	2.130	0.044
Intercept	0.525***	0.183	2.870	0.008

Observations: 287; *R-squared:* 0.51; F(9,24) = 23.7 test for overall significance (Prob > F = 0);
*** p<0.01, ** p<0.05, * p<0.1 (Significance at 1%, 5% and 10% level respectively)

Our F test confirms joint significance of all the parameters and our overall R^2 of 50.1% shows that the model is adequate. We discuss the results in more detail below. From Table 6.9, the estimated parameter of the lag of loan overcharge α, is 0.324 and is statistically significant at the 1% level. The estimated value of γ is 0.185 and it is statistically significant at the 1% level. It is positive which means the persistence parameter increases during the post-reform period. Hence, the post-reform period has a higher loan overcharge persistence of 0.51 (0.324 + 0.185), which is suggestive that average competitive conditions declined during the post-reform period. Competitive conditions were thus on average relatively stronger during the

[94] All the six estimated models were carried out using pooled OLS, fixed effects and random effects estimations, and the tests for fixed effects and random effects led to choosing the fixed effects specification.
[95] Due to the high level of government spending and macroeconomic challenges during election cycles, an election dummy variable was included in the model and estimated. The election dummy variable was however found to be non-significant. It was accordingly dropped from the model.

pre-reform period, but declined during the post-reform period.[96] Thus, the policy reforms did not seem to have resulted in a strengthening of competitive conditions in the banking sector.

The pre-reform period's low persistence parameter of 0.324 seems to suggest that the industry was reasonably competitive to begin with at least in the context of African banking sectors. For instance, in a study of competition in Kenya for the period 1998–2007, Mwega (2011) found a persistence parameter of 0.51, while Egypt had a high persistence parameter of 0.81 in a study by Poshakwale and Qian (2011) covering 1992-2007. Our results on the impact of reforms on competition however differ from those two studies as competition increased during the post-reform period in the case of Kenya, while there was no significant change in competitive conditions in Egypt following implementation of the reforms.

We discuss below the impact of the other explanatory variables on loan overcharge and hence on competition conditions.

Real GDP growth enters significantly at the 1% level and is negative, in line with theoretical expectations, with an increase in real GDP growth leading to a reduction in loan overcharge as we noted earlier. The coefficient of -0.25 suggests that a 1% increase in real GDP growth will lead to a 0.25% reduction in loan overcharge.

The real monetary policy rate (RMPR) exerts a positive impact on loan overcharge as expected and is statistically significant at the 1% level. The estimated coefficient of 3.1 means that a 1% increase in RMPR induces a 3.1% increase in loan overcharge. The high size of the coefficient of RMPR (largest among the explanatory variables) shows that RMPR has an economically significant impact on loan overcharge and banking sector competition. This is not surprising due to the direct feeding of changes in MPR into bank rates, but the elastic nature shows how detrimental increases in RMPR has on loan overcharge.

The real Treasury bill rate (RTBR) is not statistically significant at the 5% level but only significant at the 10% level. It's coefficient of -1.22 however suggests that a 1% increase in RTBR will induce a 1.22% reduction in loan overcharge. This seems surprisingly at variance

[96] Alternatively, the persistence of loan overcharge of 0.324 corresponds to a relatively high adjustment parameter (λ) of 0.676 during the pre-reform period, with γ = 0.185 implying δ = -0.185. Hence there is a reduction in the speed of adjustment to competitive levels from 0.676 during the pre-reform period to 0.49 during the post-reform period, which shows a reduction in competition during the post-reform period.

with the a-prior expectation as discussed earlier. A possible explanation could be that increases in RTBR attract huge investments by banks in treasury bills and this reduces loanable funds available for private sector lending. Potential loan supply to the private sector is curtailed and banks selectively restrict lending creditworthy large corporates, reputable medium-sized companies, and high net-worth individuals at lower loan overcharge rates. The reduction in loan overcharge is deemed feasible as the interest income from investment treasury bills might more than compensate for the lower loan overcharge to these clients.

The fiscal balance to GDP ratio is statistically significant at the 5% level with the expected sign as a high fiscal deficit will lead to an increase in loan overcharge arising from increased domestic borrowing by the government, *ceteris paribus*. A 1% increase in budget deficit to GDP ratio induces a 0.08% increase in loan overcharge. The magnitude of the coefficient reveals a smaller economic impact of the variable.

The HHI is found not to be statistically significant in influencing loan overcharge. Thus market concentration, or the lack of it, seems to play little role in influencing loan overcharge and hence competition in the sector. This reinforces the theoretical underpinnings of the POP model which is based on the contestable markets theory and argues in favour of potential entry rather than actual entry and concentration levels as impacting competition.

The real currency depreciation is also statistically significant at the 5% level, meaning currency depreciation impacts loan overcharge but surprisingly in the opposite direction. The estimated coefficient shows that a 1% depreciation leads to a 0.67% reduction in loan overcharge.

The CRISES dummy variable is significant at the 5% level and positive, suggesting a 0.18% increase in loan overcharge as a result of the adverse effects of the financial crisis, The tightening of credit market conditions and increased pressure on domestic banking markets resulting from the slump in export revenues and fiscals revenues seem to have contributed to higher interest rates and loan overcharge during 2009-2014 as seen in the trend in interest rates, and seem to have contributed in constraining competitive conditions post-2008.

To sum up the results of the POP model, the persistence of profits in the banking sector was low during the pre-reform period, compared to the post-reform period. The ushering in of the reforms was expected to have reduced the level of persistence of profits due to the expected competitive conditions arising from the deregulation reforms. It is observed that the reduction

in the competitive conditions in the banking sector during the post reform period could be traced to the strong effects of the adverse macroeconomic developments post-2008, contributed in part by global financial/economic crises. These developments exerted an adverse effect on banking competition, overshadowing the intended benefits of the reform policies. To double check the robustness of these results, we estimate as an alternative model, the Boone indicator in the next section.

6.3.3 Evolution of competition from the Boone model

We recall from Section 6.2.3 that the Boone model estimates a relationship between performance (in our case, the market share of loans) and efficiency (measured by the marginal cost of loans). We adopted a variant of the Boone model that provides the annual competition measure to track evolution of competition from 2000 to 2014, specified as follows:

$$ln\ MS_{i,t} = \alpha + \sum_{t=1}^{T} \beta_t \, D_t \, ln\ MC_{i,t} + \sum_{t=1}^{T-1} \gamma_t D_t \ + \ \varepsilon_{i,t} \qquad (6.12)$$

where MS is the market share of loans, MC is the marginal cost of loans, β_t is the Boone measure of competition, in this case the year measures of competition to track the evolution of competition during the sample period; D_t is a time dummy to control for other factors common to all banks in the industry and specific to each year; and $\varepsilon_{i,t}$ is the error term, with $\varepsilon_{i,t} \sim iid(0,\ {\sigma_\varepsilon}^2)$.

(1) Descriptive statistics of data on variables to be estimated in the Boone model

Having already explored descriptive statistics on the marginal cost of loans in Section 6.3.2, we show below the descriptive statistics of the computed market share of loans in Table 6.10a. The decline in the mean values and standard deviation of the market share of loans from 2000 to 2009 reflects reduction in the concentration in the industry and confirmed by the decline in the HHI observed earlier. The decline, however, is not only due to losing of market share by the large banks as the maximum market share after a sharp decline in 2001 has remained fairly unchanged but even increased during 2007 to 2010. The sharp drop in the maximum market share since 2010 seems to reflect the shift from loans to government securities even by big banks.

Table 6.10a Descriptive statistics of market share of loans – annual basis

	Mean	Median	Std. Dev.	Minimum	Maximum
2000	6.25	2.07	8.31	0.03	27.65
2001	5.88	2.05	8.78	0.18	34.10
2002	5.56	3.14	5.95	0.29	17.46
2003	5.56	2.43	6.33	0.34	19.75
2004	5.56	2.66	5.93	0.39	18.68
2005	4.76	2.50	5.30	0.05	17.25
2006	4.76	3.12	4.70	0.14	15.75
2007	4.35	2.72	4.82	0.26	18.95
2008	4.17	2.91	4.57	0.02	19.85
2009	4.00	2.87	3.95	0.07	19.90
2010	4.80	3.81	4.41	0.27	20.00
2011	4.00	4.14	2.64	0.20	14.41
2012	4.17	3.95	2.71	0.25	11.69
2013	4.55	4.69	3.01	0.60	14.06
2014	4.17	3.93	2.70	0.37	12.26

Values are in percentages (%)

Table 6.10b Descriptive statistics of market share of loans in panel form

Variable		Mean	Median	Std. Dev.	Min	Max
Market share of loans	overall	4.7	3.2	0.049	0.001	0.341
	between			0.040	0.002	0.161
	within			0.027	-0.058	0.227

Table 6.10b shows the wide overall variation in the market share of loans. As expected, the between variation (variation across banks) in the market share of loans is higher than the within variation (variation across years).

(2) Estimation results of the Boone model

The results of the estimation of equation 6.12 are presented in Table 6.11. The model was estimated using panel data techniques with the random effects model selected as the preferred model, based on the Hausman test (the detailed results are reported in the Appendix 6.4).

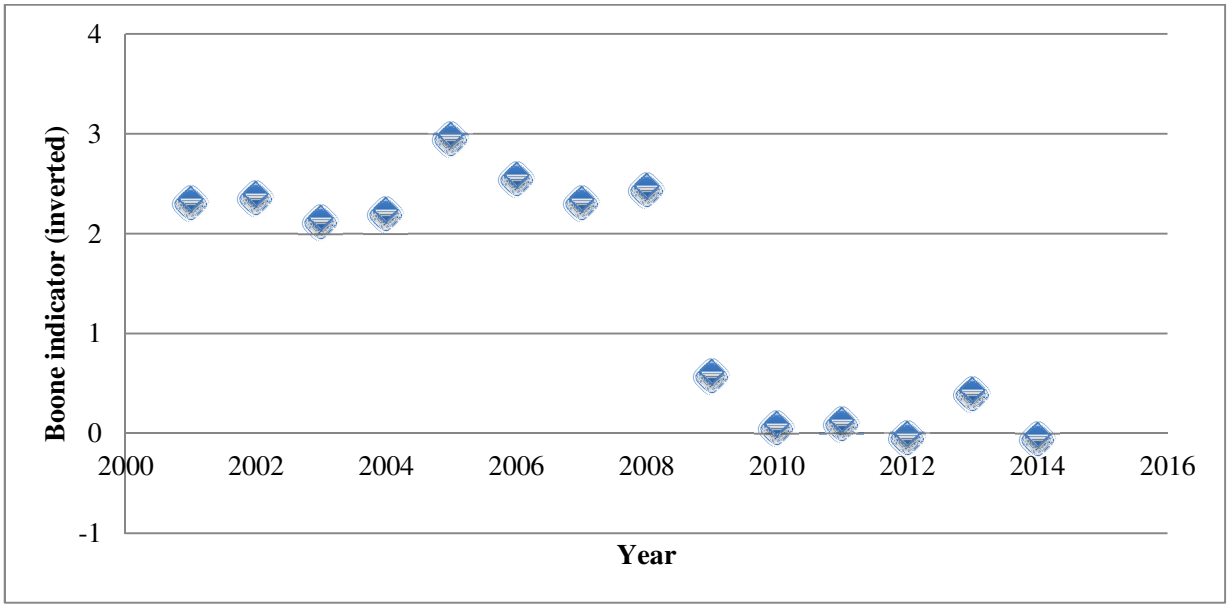
Table 6.11 Estimates of annual Boone indicators for 2001–2014

Year	Coefficient (stdev)	Year	Coefficient (stdev)
2001	-2.349***	2008	-2.482***
	[0.299]		[0.376]
2002	-2.400***	2009	-0.613
	[0.323]		[0.987]
2003	-2.160***	2010	-0.097
	[0.253]		[0.896]
2004	-2.237***	2011	-0.13
	[0.462]		[0.611]
2005	-2.992***	2012	0.008
	[0.523]		[0.638]
2006	-2.589***	2013	-0.434
	[0.396]		[0.674]
2007	-2.345***	2014	0.014
	[0.343]		[0.537]

As shown in the table, the estimated annual Boone indicator parameters are all negative and significant at the 1% level, until 2009. However, none of the coefficients for 2009-2013 is significant. The lack of significance of the estimated annual Boone measures indicates woefully low level of competitiveness in the loans market beyond 2008. It is worth noting that the non-significance of the yearly Boone estimates for the post-2008 era is not unique. Van Leuvensteijn et al. (2011) found, in a study of competition in a cross-country study of five the major countries in the euro area, the UK, US and Japan for the period 1994-2004, that in most of the countries, not all the estimated annual Boone were statistically significant. For instance, for the annual Boone estimates for the UK for 1994-2004, the estimates for 1994-1997 and 2002-2004 were not statistically significant meaning only estimates for 4 years were statistically significant during the 11-year period. Similar high numbers of insignificant annual estimates were recorded for Spain and the Netherlands.

To clearly analyse the trend in competition, we show in graphical form the annual beta estimates. For ease of representation (since the Boone indicator is usually negative), we plot the inverted beta estimates on the vertical axis in Figure 6.3 to track the evolution of competition over 2001-2014.

Figure 6.3 Trend in annual Boone competition measures, 2001-2014



The trend in annual competition estimates from the Boone model shown in Figure 6.3 reveals significant differences in competition on a yearly basis. It shows that competitive conditions were strong and broadly stable during 2001-2004, with the average (inverted) Boone measure of 2.27 for the 4-year period. This marginally strengthened during 2005–2008, with an average (inverted) Boone measure of 2.60. The stronger competitive condition during 2005-2008 was however not sustained as competition significantly declined as evidenced by the low Boone estimates of 2009–2014 which are not even significant, and as explained above are indicative of very low competition. The interpretation of the results suggests that the banking industry's competitiveness was fairly stable during 2001-2004, but increased during 2005-2008, the increase arguably due to the impact of the reforms.

Per Boone's intuition, reallocation of market share of loans from the less efficient banks to the more efficient banks was stronger during this period, compared to the pre-reform period. Strategies by banks to grow loans and maintain a competitive edge in the loan market during this period was stronger. The sharp drop in the beta coefficient from 2008 and 2009 shows that the loan reallocation efforts were significantly low in 2009; and this is due to the fact that the loans market was wholly unattractive compared to the alternative governments securities market during this period. The unattractiveness of the loans market was due to two reasons. First, the sharp increases in interest rates from 2008 made government securities more profitable and attractive hence banks began to re-allocate investible funds from the credits market into government securities. Second, the excessive growth in loan portfolio witnessed in 2007-8 in response to the scrapping of secondary reserves had triggered high loan default

rates in 2009 as interest rates soared. This made banks feel "they had bitten more than they could chew" in the loans market, and so decided to scale back loan growth. Thus, the intensity with which efficient banks sought to win market share in loans from less efficient ones dwindled resulting in the low beta coefficients.

6.4 Conclusions

This chapter empirically investigates the impact of recent deregulation reforms on competition in Ghana's banking sector using panel data spanning 2000-2014. The study focused on competition in the loans market as it is the largest and most important, and yet the most uncompetitive segment of the banking sector, and for which the reforms were designed for. Two dynamic competition measurement models, the POP and BI models, which facilitate impact assessments of reforms on competition in the loans market, were used.[97]

In the POP model, competitive conditions in the loans market is measured by the loan overcharge and the results show that persistence of loan overcharge was relatively low, at 0.324 during the pre-reform and to some extent the transition period. The persistence parameter however increased to 0.51 during the post-reform period, suggesting that on average, competitive conditions declined during the post-reform period. The decline in competitive conditions was very probably driven by the adverse macroeconomic environment, with rising interest rates, which was contributed in part by the effects of the global financial and economic crises as shown above.

The empirical results from the alternative competition model, the estimated annual Boone indices, yield similar results and shed further light on the evolution of the competitive environment during the period. The annual Boone estimates show an immediate increase in competition following the reforms, during 2005-2008 but this was not sustained post-2008, as competition declined sharply thereafter. The attractiveness of the government securities market and emerging loan defaults made the credits market unattractive.

[97] The formulation of both models requires the derivation of the marginal cost of loans, and this was done via the estimation of a translog cost function. The estimated translog cost model satisfied both model adequacy and met the regularity conditions of a cost function.

To summarise, both the POP model and the Boone indicator conclusively show that notwithstanding the deregulation reforms implemented by the Central Bank, the gains made in stimulating competition in the banking sector was not sustained for most part of the post–reform period. The results of the two models taken together also suggest that the initial increase in competition arising from implementation of the reforms was supported by the relatively strong macroeconomic environment for most of the pre-reform and transition period. The gains in competition were not sustained as macroeconomic weaknesses during post-2008 undermined banking competition.

On institutional development, the passage of the Credit Reporting Act in 2007 and the subsequent licensing of credit reference institutions should have preceded the scrapping of the secondary reserves. This would have at least helped strengthen the credit environment before banks are allowed the liberty to grow their loan portfolio. Accordingly, an important consideration in the implementation of reforms is an appropriate sequencing of reforms.

In the broader context of the literature on banking reforms and competition, the results of this study seem to corroborate findings of other studies on developing countries. Moyo et al. (2014) in a cross-county study of African banking sectors deregulation reforms enhanced competition, but that increased competition and stability hinged on a stable economic environment and strong institutions. Delis (2010) also found that financial liberalisation policies enhanced competition in developed countries with strong institutions, whereas competition is sluggish in developing countries with weak macroeconomic fundamentals and low institutional development.

In conclusion, we must mention that the two competition models used are not without limitations. Both models do not provide threshold values that can be described as the optimum level of competition. In other words, the estimated persistence parameter and the Boone values do not have specific benchmarks that could be described as the optimum level of competition. This limitation does not however affect our results and outcomes as our study looks at the relative measures over time and thus easy to assess policy impacts. Another limitation is that these models ignore differences in bank product quality, loan processing speed and delivery of service and attractiveness of product innovations which though seems to be a generic limitation for all competition models. These limitations notwithstanding, the two competition measures and the specification of the models seem to corroborate the results on competition and adequately address the research questions that this study seeks to answer.

Appendix 6.1 Estimation of translog cost function using POLS, FE, RE

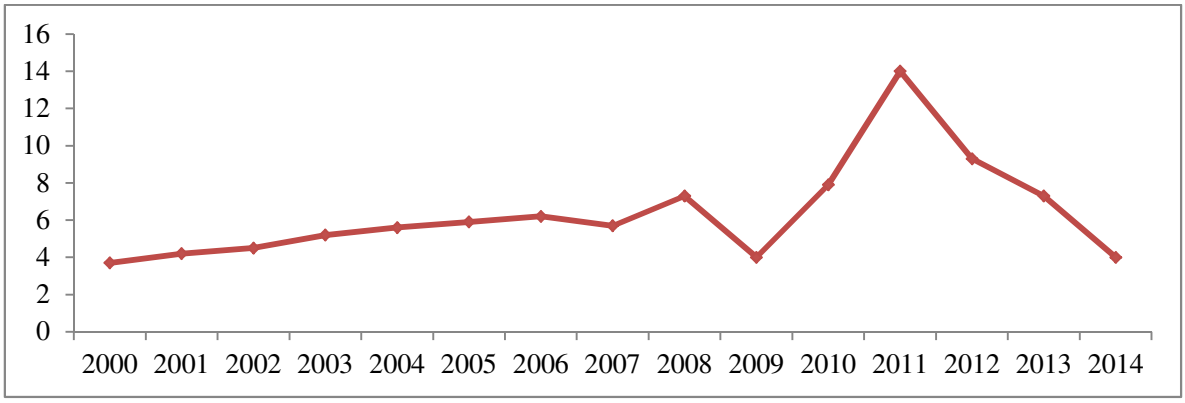
VARIABLES		POLS	FE	RE
Ln (P_{FUNDS}/ $\mathbf{P_{LK}}$)	lnw1*	0.532***	0.581***	0.533***
		[0.116]	[0.121]	[0.116]
Ln (LOANS)	lny1	0.491***	0.588***	0.492***
		[0.151]	[0.154]	[0.151]
Ln (OEA)	lny2	0.168	0.282*	0.168
		[0.143]	[0.142]	[0.143]
Ln (OBS)	lny3	0.172	0.021	0.171
		[0.255]	[0.253]	[0.255]
(Ln (P_{FUNDS}/ $\mathbf{P_{LK}}$))^2	$(lnw1)^2$	0.101***	0.096***	0.101***
		[0.022]	[0.021]	[0.022]
$(Ln\ LOANS)^2$	$(lny1)^2$	0.088***	0.065***	0.088***
		[0.021]	[0.018]	[0.021]
$(Ln\ OEA)^2$	$(lny2)^2$	0.118***	0.086***	0.118***
		[0.019]	[0.021]	[0.019]
$(Ln\ OBS)^2$	$(lny3)^2$	0.046	0.033	0.046
		[0.046]	[0.042]	[0.045]
(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)(Ln LOANS)	(lnw1*)(lny1)	-0.019	-0.005	-0.019
		[0.036]	[0.039]	[0.036]
(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)(ln OEA)	(lnw1*)(lny2)	-0.034	-0.038*	-0.034*
		[0.020]	[0.022]	[0.020]
(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)(Ln OBS)	(lnw1*)(lny3)	0.057	0.042	0.057
		[0.048]	[0.049]	[0.048]
(Ln LOANS)(Ln OEA)	(lny1)(lny2)	-0.144***	-0.138***	-0.144***
		[0.037]	[0.032]	[0.037]
(Ln LOANS)(Ln OBS)	(lny1)(lny3)	-0.02	-0.004	-0.02
		[0.056]	[0.052]	[0.056]
(Ln OEA)(Ln OBS)	(lny2)(lny3)	-0.075	-0.042	-0.075
		[0.054]	[0.046]	[0.054]
T	T	0.028	0.013	0.028
		[0.025]	[0.029]	[0.025]
Tsqr	Tsqr	0	0.001	0
		[0.001]	[0.001]	[0.001]
T(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)	Tlnw1*	-0.001	-0.003	-0.001
		[0.005]	[0.005]	[0.005]
T(Ln LOANS)	Tlny1	-0.013**	-0.007	-0.013**
		[0.006]	[0.006]	[0.006]
T(Ln OEA)	Tlny2	-0.004	-0.005	-0.004
		[0.006]	[0.007]	[0.006]
T(Ln OBS)	Tlny3	0.021***	0.014*	0.020***
		[0.007]	[0.008]	[0.007]
Intercept	_cons	2.117***	2.025***	2.116***
		[0.308]	[0.303]	[0.308]

Observations: 321; Robust standard errors in brackets; Hausman test for fixed effects vs. random effects: chi^2 (20) = 62.44 (hence FE is preferred) *** p<0.01, ** p<0.05, * p<0.1

Appendix 6.2 Trends in key macroeconomic and industry variables

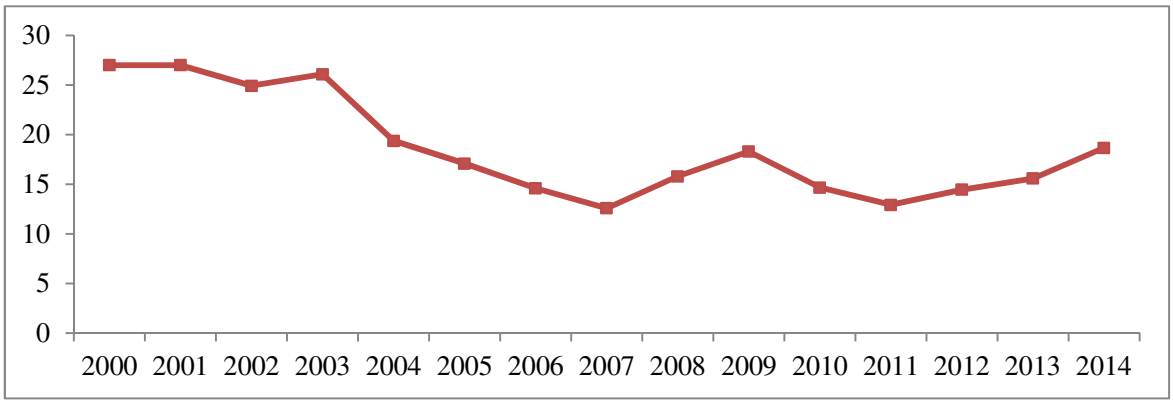
We capture in Figure A6.1 the trends in the key macro-economic variables during the period.

Figure A6.1a Real GDP growth



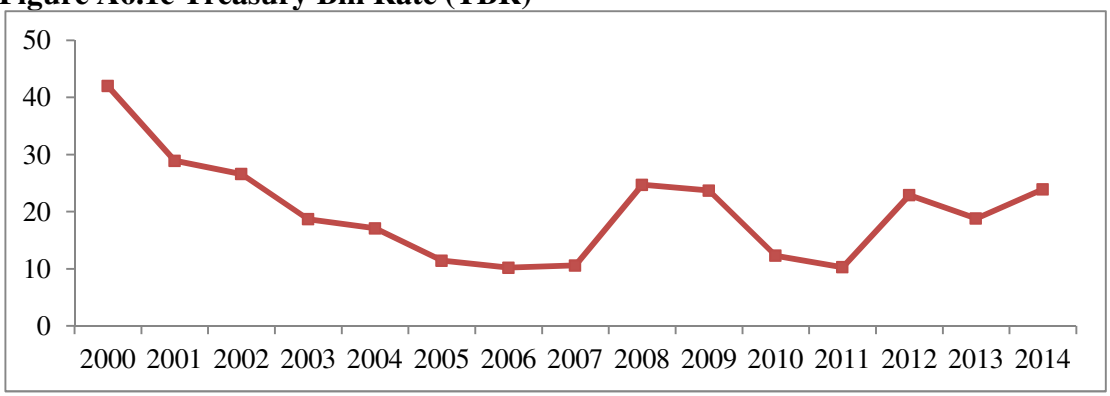
Real GDP growth rate increased consistently during 2000-2008 but fell in 2009 due to the slump in gold and cocoa prices following the economic downturn caused by the global financial crisis, macroeconomic difficulties with huge fiscal and balance of payments deficits in 2008-2009. Growth rebounded in 2010 with the sharp rise seen in 2011 driven by the oil sector on account of production of crude oil in commercial quantities in 2011, but subsequently declined.

Figure A6.1b Monetary Policy Rate (MPR)



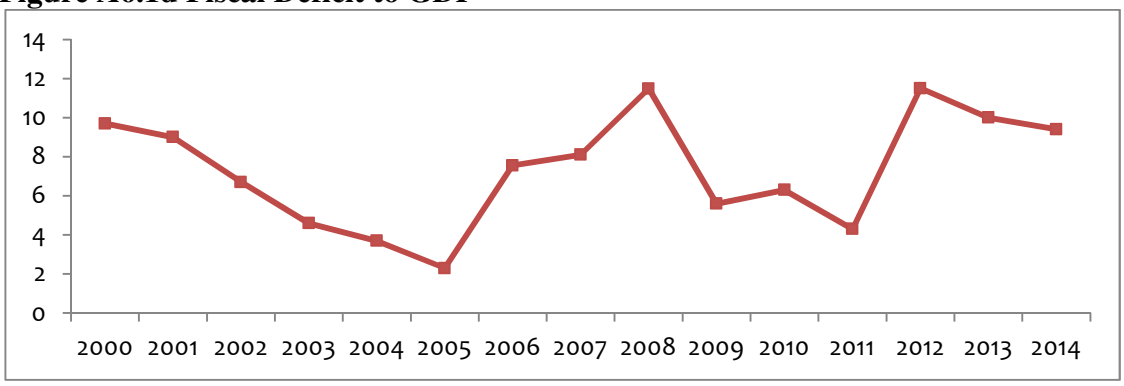
The persistent reduction in MPR by the Central Bank during 2000–2007 was in line with declining inflation, but was raised in 2008-2009 due to an upturn in inflationary pressures resulting from a combination of high crude oil prices, food crisis, election-related spending, and worsening terms of trade. The rate was marginally reduced during 2010 and 2011 but picked up from 2012 to 2014.

Figure A6.1c Treasury Bill Rate (TBR)



TBR follows developments in fiscal deficit to GDP. The gradual reduction in fiscal deficit to GDP ratio from 10% in 2010 to 2.3% in 2005 meant lower borrowing from the banking sector which led to a continual decline in the TBR from a high of 42% in 2000 to a low of 10.2% in 2007 (Figure A6.1c and Figure A6.1d).

Figure A6.1d Fiscal Deficit to GDP



Fiscal deficits increased sharply in 2006 through to 2008, and the resulting government borrowing via Treasury bills pushed TBR to more than double to around 24.5% during 2008-2009. TBR rates thereafter declined during 2010- 2011 in response to lower budget deficits, but shot up in 2012 and broadly stayed at relatively high levels during 2012-2014 as deficits widened during 2012-2014.[98]

In terms of currency depreciation, after an exceptionally 100% currency depreciation in 2000, the exchange rate was fairly stable during 2001-2007, with an average annual depreciation rate of 5%.[99] The high fiscal and external deficits of 2008-2009 translated into high currency

[98] Fiscal deficits are chronic in Ghana, with no surpluses recorded throughout the study period. It's for ease of presentation that we use the absolute values in the graph in Figure A6.1d.

[99] 2000 was an exceptionally challenging year as low prices of gold and cocoa, reduction in external financing, high external debt and huge government spending resulted in the worst currency crisis in recent times with the Cedi depreciating by almost 100%. The currency stability experienced during 2001-2007 was driven by strong

depreciation, before the currency stabilised momentarily in 2010-2011, and have increased since then.

Figure A6.1e Annual Depreciation (USD)

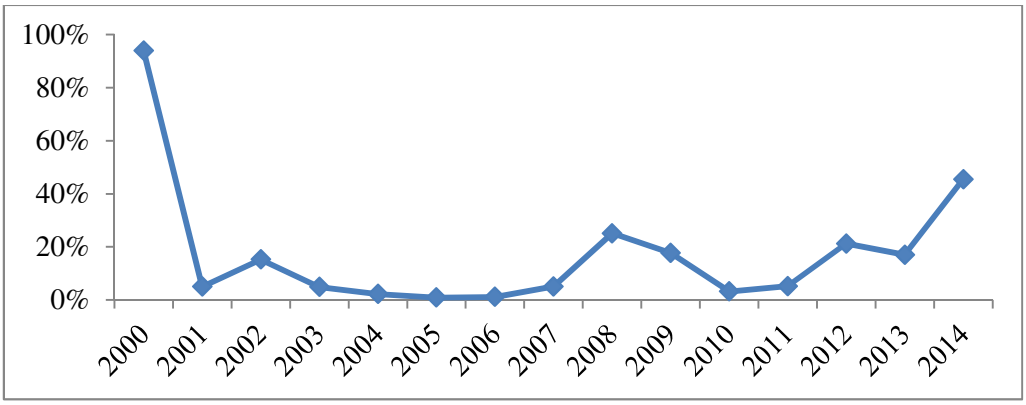
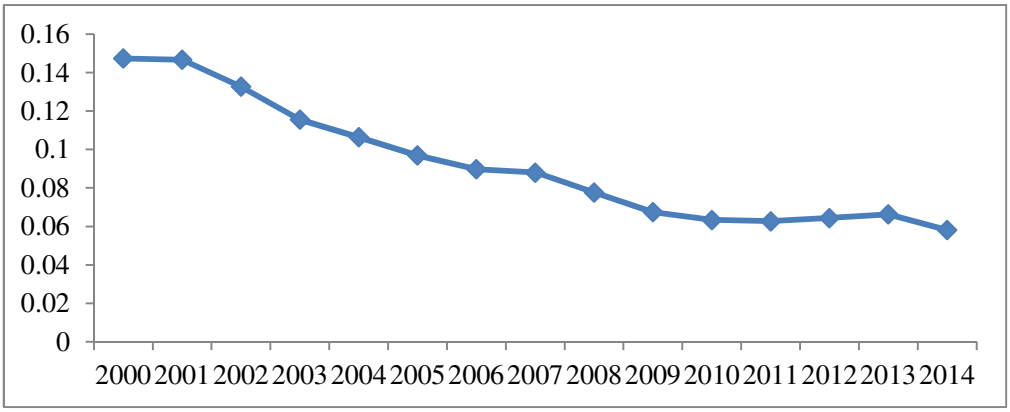


Figure A6.1f Herfindahl-Hirschman Index (HHI)



The HHI pattern reveals a continuous decline in banking concentration throughout the study period.

exports volumes and good export prices, increased external financing, and foreign currency savings arising from external debt cancelations under the Heavily Indebted Poor Countries (HIPC) initiative.

Appendix 6.3 Estimated results of alternative specifications of POP model

	POLICY VARIABLE IN 2006		POLICY VARIABLE IN 2007		POLICY VARIABLE IN 2008	
	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
	CRISES = 2008	CRISES = 2009	CRISES = 2008	CRISES = 2009	CRISES = 2008	CRISES = 2009
VARIABLES						
LnLOC_1	0.334***	0.330***	0.324***	0.304***	0.318***	0.315***
	[0.062]	[0.061]	[0.056]	[0.053]	[0.053]	[0.055]
R*LnLOC_1	0.096	0.07	0.185***	0.181***	0.173***	0.151**
	[0.058]	[0.063]	[0.057]	[0.060]	[0.057]	[0.072]
Ln RGDP	-0.026***	-0.018**	-0.253***	-0.022***	-0.025***	-0.019***
	[0.007]	[0.006]	[0.057]	[0.007]	[0.006]	[0.006]
Ln RMPR	4.738***	3.226***	3.101***	3.530***	4.128***	3.044***
	[0.858]	[0.590]	[0.593]	[0.651]	[0.855]	[0.577]
Ln RTBR	-1.973***	-1.161**	-1.217***	-1.252**	-1.620***	-0.962**
	[0.563]	[0.467]	[0.651]	[0.469]	[0.563]	[0.450]
Ln FISCAL	-0.016*	-0.015*	-0.084**	-0.016**	-0.017**	-0.015**
	[0.008]	[0.007]	[0.038]	[0.007]	[0.008]	[0.007]
Ln RDEPN	-0.398	-0.297	-0.674**	-0.432	-0.453	-0.394
	[0.298]	[0.289]	[0.320]	[0.297]	[0.297]	[0.289]
HHI_ASSETS	0.474	-0.27	1.035	-0.614	-0.133	-0.932
	[1.181]	[1.122]	[1.774]	[1.101]	[1.127]	[1.095]
CRISIS2008	0.233***		0.183**		0.158**	
	[0.063]		[0.086]		[0.064]	
CRISIS2009		0.146***		0.106**		0.079
		[0.046]		[0.044]		[0.057]
Constant	0.146	0.24	0.525***	0.285*	0.218	0.322*
	[0.181]	[0.165]	[0.183]	[0.160]	[0.176]	[0.160]
Observations	287	287	287	287	287	287
F test	17.02	16.29	23.67	23.71	21.65	19.81
Prob > F	0	0	0	0	0	0
r-squared						
Number of banks	25	25	25	25	25	25
r-squared overall	0.493	0.485	0.514	0.512	0.514	0.505
Robust standard errors in brackets						
*** p<0.01, ** p<0.05, * p<0.1						

On the bases of the overall R^2 and the F-statistic of overall significance, we selected model 3, which is the model with the policy variable set in 2007 and the CRISES variable in 2008.

Appendix 6.4 Estimates of annual Boone indicators for 2001–2014

VARIABLES	Pooled OLS	Fixed Effects	Random Effects
2001	-3.366***	-1.919***	-2.349***
	[0.311]	[0.292]	[0.299]
2002	-3.208***	-2.016***	-2.400***
	[0.488]	[0.297]	[0.323]
2003	-3.041***	-1.747***	-2.160***
	[0.275]	[0.292]	[0.253]
2004	-3.225***	-1.774***	-2.237***
	[0.441]	[0.490]	[0.462]
2005	-3.628***	-2.641***	-2.992***
	[0.408]	[0.616]	[0.523]
2006	-3.084***	-2.274***	-2.589***
	[0.293]	[0.492]	[0.396]
2007	-3.025***	-1.931***	-2.345***
	[0.405]	[0.377]	[0.343]
2008	-3.452***	-1.722***	-2.482***
	[0.658]	[0.197]	[0.376]
2009	-1.084	-0.572	-0.613
	[1.428]	[0.397]	[0.987]
2010	-0.063	-0.466	-0.097
	[1.431]	[0.336]	[0.896]
2011	0.173	-0.611*	-0.13
	[0.971]	[0.330]	[0.611]
2012	0.327	-0.559*	0.008
	[1.065]	[0.317]	[0.638]
2013	-1.562*	0.227	-0.434
	[0.833]	[0.668]	[0.674]
2014	0.69	-0.773	0.014
	[0.800]	[0.477]	[0.537]

*** p<0.01, ** p<0.05, * p<0.1 (Significance at 1%, 5% and 10% level respectively)

CHAPTER 7 IMPACT OF DEREGULATION, OWNERSHIP AND SIZE ON BANKING EFFICIENCY IN GHANA

7.1 Introduction

Deregulation policies in the banking sector are expected to induce operational efficiency among banks through the creation of a more competitive environment, as discussed in Chapter 4. The competition-efficiency hypothesis also established the efficiency-enhancing role of increased banking competition as we saw in Chapter 3. Our empirical study of the impact of deregulation on competition in Ghana's banking sector in Chapter 6 shows that the reforms had an initial positive effect on stimulating banking competition, but this was not sustained in the later years due to macroeconomic weaknesses. It therefore may seem to suggest that efficiency levels of banks might have also increased initially in line with the higher competition but may have subsequently slackened as competition weakened in the later period. This however requires an explicit examination of the effect of the deregulation policies directly on bank-level efficiency, since the reduction of inefficiency levels is an important objective of such banking reforms.

This chapter accordingly examines empirically the impact of deregulation reforms on banking efficiency during the study period in line with our primary research objective. Our second research objective is to analyse to what extent bank ownership and bank size impact on efficiency. As discussed in Chapter 2, the ownership dynamics in Ghana's banking sector following the reforms make it imperative for a study to account for the relative efficiencies of these diverse banking groupings of state-owned banks, private domestic banks, regional banks and foreign banks.[100] The issue of bank size in influencing efficiency is relevant in view of concerns that the Central Bank should introduce policies to foster consolidation in the industry in view of the growing number of banks. We examine efficiency levels of large, medium and small banks. Large banks are banks with total assets above the upper quartile of total assets of the industry; small banks are banks with assets size below the lower quartile; while banks in-between are classified as medium-sized for each year. The aim of this chapter therefore is to provide empirical evidence on the effect of deregulation reforms, ownership status and bank size on bank-level efficiencies in Ghana.

[100] It is worth noting that all these different classes of banks operate on a level playing field, that is, within the same market and under the same regulatory requirements.

Our study is unique compared to other studies on African banking efficiency in several respects. Rather than treating deregulation as a single category, we examine the potential different efficiency impacts of different deregulation policies implemented, and the extent to which each of those policies affect efficiency?[101] We do this by using the informative and internationally recognised survey data and database on banking regulation and reforms by Barth et al. (2001, 2003, 2007, 2012) and Abiad et al. (2010), which have not been applied in Africa banking markets. Second, the unique bank-level data over a sufficiently long period facilitates effective assessment of bank-level efficiency and efficiency levels by ownership and size categories. This helps in the design of different policies for different bank classes if need be.

We also use stochastic frontier analysis (SFA) in measuring cost efficiency, in preference to the widely used DEA in the African literature for reasons detailed in Chapter 5. Finally, we adopt the one-step Battese-Coelli (1995) (BC95) in modelling cost efficiency of banks, which overcomes the econometric challenges associated with the two-step approach as discussed in Chapter 5. The BC95 one-step model simultaneously estimates the stochastic cost frontier with an inefficiency term expressed as a function of exogenous factors and bank–specific characteristics in the context of panel data. This methodology facilitates computation and analysis of the efficiency impacts of deregulation, ownership and bank size using panel data covering 2000–2014.

The rest of the chapter is arranged as follows. Section 7.2 discusses the methodology and data issues, including model specification of the cost frontier and the inefficiency model. It also discusses in detail the derivation of deregulation policy variables inefficiency, and measurement of the variables of the model. Section 7.3 details the empirical estimation of the model, hypothesis tests and analysis of the results. We conclude the chapter in Section 7.4.

7.2 Methodology and data

7.2.1 Methodology

We use the stochastic cost frontier to model a bank's cost characteristics, in which a bank is inefficient if its total costs are higher than those predicted for a fully efficient bank using the

[101] That is, whether the different deregulation policies had a uniform impact or different impacts on banking efficiency.

same inputs and outputs combination, and such differences are not explained by statistical noise. As discussed earlier, the general form of the stochastic cost frontier (in panel form) is specified as follows:

$$\ln C_{it} = C^*(\ln y_{it}, ln\, w_{it};\, \beta) + \varepsilon_{it} \tag{7.1}$$

where C_{it} is total observed cost; y_{it} is a vector of outputs; w_{it} is a vector of input prices; $C^*(.)$ is the underlying deterministic cost frontier representing the best or fully efficient firm operating in that environment; and β is the vector of parameters to be estimated. Like all stochastic frontier models, equation (7.1) has a composite error term $\varepsilon_{it} = v_{it} + u_{it}$; v_{it} represents a symmetric pure random error term capturing effects of statistical noise such as measurement errors and omitted variables, while u_{it} is an asymmetric non-negative term representing bank-level inefficiency. The subscripts denote bank *i* at time *t*.

In the BC95 model, the inefficiency term is parameterised as a function of exogenous factors and bank characteristics, Z'_{it} , that is,

$$u_{it} = Z'_{it}\,\delta + W_{it} \tag{7.2}$$

where u_{it} is the non-negative inefficiency term assumed to be independently but not identically distributed and follows a truncated normal distribution with mean $Z'_{it}\delta$ and constant variance, σ_u^2, that is, $u_{it} \sim N^+(z_{it}\delta, \sigma_u^2)$. The error term in equation (7.2), W_{it}, is a random variable defined by the truncation of the normal distribution with zero mean and variance σ_W^2 such that the truncated point is at $-z_{it}\delta$, that is, $W_{it} \geq -z_{it}\delta$ which guarantees the non-negativity of u_{it} (Battese and Coelli, 1995). In other words, the non-negativity requirement of u_{it} means that $W_{it} \geq -z_{it}\delta$ and so we assume $W_{it} \sim (0, \sigma_W^2)$, with the distribution of W_{it} being bounded below by the variable truncation point $-z_{it}\delta$.[102]

The BC95 model specified by equations (7.1) and (7.2) is estimated simultaneously in a single-step using maximum likelihood (ML) techniques based on the distributional assumptions, $v_{it} \sim N(0, \sigma_v^2)$, and $u_{it} \sim N^+(\delta' z_{it}, \sigma_u^2)$. ML estimation involves maximising the log-likelihood function to derive the parameter estimates.

[102] The distributional assumption on W_{it} is consistent with the distributional assumption on u_{it}, that is, $u_{it} \sim N^+(\delta' z_{it}, \sigma_u^2)$. If all the elements of the parameter vector δ are not significantly different from zero, then the inefficiency effects are not related to the z-variables, and so we obtain the half-normal distribution originally proposed by Aigner et al. (1977), while if only one of them is significantly different from zero, then we have the truncated normal with constant mean distribution proposed by Stevenson (1980) (Battese and Coelli, 1993).

The log-likelihood function of the BC95 model is given as (Kumbhakar et al, 2015):[103]

$$\ln L = constant - \frac{1}{2}\ln(\sigma_v^2 + \sigma_u^2) - \sum_i \ln \Phi(\frac{\delta' z_{it}}{\sigma_u}) + \sum_i \ln \Phi(\frac{\mu_{*it}}{\sigma_*}) + \frac{1}{2}\sum_i \frac{(\varepsilon_{it} + \delta' z_{it})^2}{\sigma_v^2 + \sigma_u^2} \quad (7.3)$$

where $\mu_{*it} = \frac{\sigma_v^2\, \delta' z_{it} + \sigma_u^2 \varepsilon_{it}}{\sigma_v^2 + \sigma_u^2}$; $\sigma_*{}^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$; $\varepsilon_{it} = v_{it} + u_{it}$

The maximisation of the above log-likelihood function produces the estimates of the parameters of the model $(\beta, \delta, \sigma_v^2, \sigma_u^2)$.

As discussed in Chapter 5, the conditional expectation of cost efficiency using the JLMS (1982) technique can be formulated as (Battese and Coelli, 1993):

$$E[\exp\{-u_{it}\}\,|\varepsilon_{it}]\Big| = \left[\exp\{-u_* + \frac{1}{2}\sigma_*{}^2\}\right].\left[\frac{\Phi\left(\frac{\mu_*}{\sigma_*} - \sigma_*\right)}{\Phi\left(\frac{\mu_*}{\sigma_*}\right)}\right] \quad (7.4)$$

where, as defined already,

$$\mu_{*it} = \frac{\sigma_v^2(\delta' z_{it}) + \sigma_u^2(\varepsilon_{it})}{\sigma_v^2 + \sigma_u^2}; \qquad \sigma_*{}^2 = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2}$$

The maximum likelihood estimates of the parameters are substituted into equation (7.4) to obtain the bank-specific cost efficiency estimates.

The cost efficiency of bank *i* at time *t* is given by:

$$CE_{it} = E(e^{-u_{it}}) = e^{(-\delta' z_{it} - W_{it})} \quad (7.5)$$

The cost efficiency estimates range between 0 and 1 with 1 representing full cost efficiency.

In the BC95 model, the estimated parameters of the inefficiency model show the direction of the impact of the covariates on inefficiency. They however do not represent the marginal effects. The marginal effects of the exogenous variables on cost inefficiency can however be obtained using the formula (see Battese and Coelli, 1993):

$$\frac{\partial \mathrm{CI}}{\partial \mathrm{z}_{ik}} = \left(\frac{\Phi\left(\frac{\mu_*}{\sigma_*} - \sigma_*\right) \mathrm{e}^{-u_* + \frac{1}{2}\sigma_*{}^2}}{\sigma_* \Phi\left(\frac{\mu_*}{\sigma_*}\right)} - \frac{\Phi\left(\frac{\mu_*}{\sigma_*} - \sigma_*\right) \phi\left(\frac{\mu_*}{\sigma_*}\right) \mathrm{e}^{-u_* + \frac{1}{2}\sigma_*{}^2}}{\sigma_* \left(\Phi\left(\frac{\mu_*}{\sigma_*}\right)\right)^2} - \frac{\Phi\left(\frac{\mu_*}{\sigma_*} - \sigma_*\right) \mathrm{e}^{-u_* + \frac{1}{2}\sigma_*{}^2}}{\Phi\left(\frac{\mu_*}{\sigma_*}\right)} \right) \frac{\partial \mathrm{u}}{\partial \mathrm{z}_{ik}} \quad (7.6)$$

[103] This is a generalisation of the truncated normal distribution, but with the constant mean μ replaced by the variable mean, $z_{it}\delta$,

where $\phi(.)$ is the probability density function of the standard normal distribution, and $\frac{\partial u}{\partial z_{ik}}$ is the partial effect of the k-th z-variables on u_{it} defined in equation 7.2.

7.2.2 Specification of empirical cost frontier and determinants of inefficiency

The empirical specification of our cost frontier derives much from our discussion on the estimation of cost function (for purposes of generating marginal costs) in Section 6.2.4 of Chapter 6 regarding the choice of an appropriate functional form, definition of inputs and outputs, and the measurement of quantities and prices of such inputs and outputs. As discussed in detail in that section, we choose the translog cost functional form with three outputs and two inputs following the intermediation approach. The only difference here is that we now specify a composite error term in the cost frontier, and introduce a deregulation variable.

To capture the impact of deregulation reforms in our cost frontier, we introduce a 'combined' deregulatory reform variable (DEREG) in the cost function which may affect bank cost directly through shifting the cost frontier. DEREG is a measure of the extent of banking sector liberalisation and is derived as the average of three individual deregulation variables that we capture separately in the inefficiency function (details follow shortly).[104]

We also assume that deregulation reforms will influence banks' production technology over time. Accordingly, we include the interaction of DEREG with the time trend (TDEREG), so that the evolving impact of deregulation is allowed to vary over time. Thus, we are able to assess the response of efficient practices in respect of cost minimisation to be associated with the evolution of the policy reforms rather than with the simple passing of time. (Zhao et al., 2010).

[104] While it is more customary to include deregulatory variables as the determinants of inefficiency as we do, the rationale for including DEREG in the cost function is to allow for the deregulated environment to influence cost technology. The individual deregulation variables and DEREG are normalised between 0 and 1 as explained later.

Our empirical specification of the cost frontier based on the translog cost function incorporating 3 outputs (y_1, y_2, y_3) and 2 input prices (w_1, w_2), time trend and its interaction with outputs and input prices, and the deregulation reform variables is as follows:[105]

$$lnTC_{i,t} = \alpha_0 + \sum_{j=1}^{2} \beta_j \, lnw_{jit} + \sum_{k=1}^{3} \gamma_k \, lny_{kit} + \sum_{j=1}^{2} \sum_{m=1}^{2} \beta_{jm} lnw_{jit} \, lnw_{mit} + \sum_{k=1}^{3} \sum_{l=1}^{3} \gamma_{kl} lny_{kit} \, lny_{lit} + \sum_{j=1}^{2} \sum_{k=1}^{3} \psi_{jk} lnw_{jit} \, lny_{kit} + \theta_1 T + \theta_{11} T^2 + \sum_{j=1}^{2} \eta_j \, Tlnw_{jit} + \sum_{k=1}^{3} \zeta_k \, Tlny_{kit} + \tau_1 DEREG + \tau_2 TDEREG + v_{it} + u_{it} \quad (7.7)$$

Following our discussion on the imposition of parameter restrictions in line with the linear homogeneity of input prices and symmetric properties of the second cross derivatives, we can write the estimable stochastic cost frontier as follows:

Stochastic cost frontier:

$$ln\,(TC_{it}*) = \alpha_0 + \beta_1 \ln\left(w_1{}^*{}_{it}\right) + \beta_{11} \ln\left(w_1{}^*{}_{it}\right)^2 + \gamma_1 lny_{1it} + \gamma_{11}(lny_{1it})^2 + \gamma_2 lny_{2it} + \gamma_{22}(lny_{2it})^2 + \gamma_3 lny_{3it} + \gamma_{33}(lny_{3it})^2 + \gamma_{12} lny_{1it} lny_{2it} + \gamma_{13} lny_{1it} lny_{3it} + \gamma_{23} lny_{2it} lny_{3it} + \psi_{11} ln\, w_{1it}{}^* lny_{1it} + \psi_{12} ln\, w_{1it}{}^* lny_{2it} + \psi_{13} ln\, w_{1it}{}^* lny_{3it} + \theta_1 T + \theta_{11} T^2 + \eta_1 l\, n\, w_{1it}{}^* T + \zeta_1 lny_{1it} T + \zeta_2 lny_{2it} T + \zeta_3 lny_{3it} T + \tau_1 DEREG + \tau_2 TDEREG + v_{it} + u_{it} \quad (7.8)$$

where TC is total cost, comprising interest cost and operating cost;

w_1 *and* w_2 are input prices: price of loanable funds and price of labour and capital respectively;

y_1 , y_2 *and* y_3 are outputs: loans, other earning assets and fee and commission income;

T is a time trend;

$Tlnw_{jit}$ is the interaction of time trend (T) with input prices;

$Tlny_{kit}$ is the interaction of time trend (T) with outputs;

DEREG is an average of the three deregulation policy indices; and
TDEREG is the interaction of time trend with the DEREG policy reform index.[106]

---

[105] Equation (7.6) is similar to equation (6.7); the only difference here is that equation (7.6) is a cost frontier with the one-sided inefficiency term, u_{it}, and the introduction of the deregulation variables in the cost frontier. The other variables are as defined in equation (6.7) in Chapter 6.

We also specify our inefficiency model as follows:

Inefficiency model:

$$u_{it} = \delta_0 + \delta_1 ACTV\_R + \delta_2 ENTRY\_R + \delta_3 CCRR + \delta_4 DOMESTIC\_D + \delta_5 REGIONAL\_D + \delta_6 FOREIGN\_D + \delta_7 lnTA + \delta_8 CRISIS + W_{i,t} \quad (7.9)$$

where

ACTV_R is a deregulation index on banking activity restrictiveness or openness; it ranges from 0 to1 increasing in liberalisation or openness;

ENTRY_R is a deregulation index measuring banking entry restrictiveness or openness; it also ranges from 0 to1, increasing in liberalisation or more openness;

CCRR is a deregulation index measuring the extent of credit controls through reserve requirements; it ranges from 0 to1 with higher values associated with more liberalisation or openness;

DOMESTIC_D, REGIONAL_D and FOREIGN_D are all ownership dummy variables;
Ln TA is the natural log of total assets, a proxy for bank size; and
CRISIS is a dummy variable with 1 for post crisis period, 0 otherwise.

As the focus of this study lies in the impact of deregulation, ownership and size on efficiency, we explain the determinants of the inefficiency model specified in equation (7.9) in more detail. The deregulation indices (ACTV_R, ENTRY_R, and CCRR) measure the extent of liberalisation in respect of activity restrictions, bank entry and credit controls respectively, corresponding to the three main deregulation policies implemented. Specifically, the ACTV_R index captures relaxation on banking activity restrictions following the introduction of universal banking. The ENTRY_R index measures the openness of the banking sector to both foreign and domestic private sector participation and signals relaxation of banking entry restrictions, following the new licencing policy introduced by the Bank of Ghana. The CCRR index measures the relaxation of credit control and reserve requirements following the abolition of secondary reserve requirements. The derivation and computation of these indices

[106] The DEREG variable is explained in more detail together with the individual deregulation indices later in Section 7.2.3

and the average deregulation reform index (DEREG) are based on surveys on measuring banking regulation by Barth et al. (2001, 2003, 2007, 2012), and guidelines on financial reform database by Abiad et al. (2010), and discussed further in detail below in Section 7.2.3 Each of these indices range from 1 to 4 with higher values representing more liberalisation, and are then normalised to lie between 0 and 1. Details of the coding rules are also presented in below.

To assess the impact of ownership on efficiency, we also include bank-group specific ownership dummy variables, DOMESTIC_D, REGIONAL_D and FOREIGN_D to capture private domestic, regional and foreign banks respectively.[107]

We use total assets as the measure for bank size in the inefficiency model.

Finally, since the study period straddles the global financial and economic crises, we also examine any possible effect of the crises on inefficiency, by including a CRISIS dummy variable which takes of 0 for 2000-2007 and 1for 2008-2014.

In terms of the priors of the coefficients in the inefficiency model, the expected impact of the deregulatory variables on cost inefficiency is dependent on the underlying theory. As discussed in Section 4.2 of Chapter 4, the public interest theory predicts an adverse effect of deregulation on efficiency while the private interest theory sees deregulation as enhancing efficiency. Accordingly, the parameters of each of the three deregulatory variables are expected to be positive as per the public interest theory since deregulation is seen as detrimental to efficiency and hence will increase inefficiency, while the parameters are expected to be negative as per the private interest theory, as it sees deregulation as reducing inefficiency.

Regarding the likely impact of the ownership dummies on inefficiency, using state banks as the ownership base, it is expected that the ownership dummies will be negative. As discussed in section in Section 4.3, while there are different theoretical predictions, the empirical literature surveyed seems to suggest that in most cases, foreign banks tend to be more efficient than domestic banks, while private domestic banks tend to be relatively more efficient than state-owned banks. In terms of the inefficiency function therefore, we expect the

[107] We omit the state-owned bank dummy as that serves as the base ownership category to avoid the dummy variable trap and multi-collinearity. Further, the inclusion of ownership dummies in the inefficiency function is to reflect managerial inefficiency. Altunbas et al. (2001) argues for such an approach rather than inclusion in the cost frontier to control for variations in technology as they argue that a common technology of intermediation is assumed to be available to all types of banks so the choice of technology is purely a managerial decision.

parameters of the ownership dummies to be negative, as we anticipate each of the ownership classes to reduce cost inefficiency in comparison with state-owned banks. In respect of bank size on efficiency, we expect that a negative relationship between bank size and inefficiency to the extent that large banks are expected to be more efficient than small as we discussed in Section 4.4. We expect that the crisis will have an adverse effect on efficiency and so there should be a positive coefficient between the CRISIS variable and inefficiency.

Table 7.1 Definition of variables of the Inefficiency function

Deregulation variables	ACTV_R Range: (0,1)	Activity restriction measures the extent to which banks are allowed to engage in other non-core banking activities such as insurance, securities and real estate, the degree of restrictiveness on mixing of banking and commerce, as well as the extent to which bank branching is allowed for all banks. The level of restrictiveness of each activity is coded on a scale of 1 to 4, where 1=prohibited; 2=restricted; 3=permitted; and 4=unrestricted. Thus, higher values denote less restriction and more liberalisation. The ACTV_R index is the average of these 3 activities. It is then normalised to lie between 0 and 1. **Source**: Barth et al. (various issues).
	ENTRY_R Range: (0,1)	Bank entry restriction is a measure of the restrictiveness of bank entry and is based on the ratio of new bank applications denied (D). It is also coded 1 – 4, where 1 represents D=100%; 2 represents (50% < D < 100%; 3 represents (10% < D < 50%; and 4 represents (D < 10%). Higher values denote less % denials, and hence less restriction on bank entry, and more liberalisation. It is also normalised to lie between 0 and 1. **Sources**: Barth et al. (various issues), Abiad et al. (2010).
	CCRR Range: (0,1)	Credit Control and Reserve Requirement measures the level of credit restriction via reserve requirements. It is normalised to be between 0 and 1.**Source**: Abiad et al. (2010).
Ownership dummies	STATE_D	1 if banks are majority owned by the State, 0 otherwise
	DOMESTIC_D	1 if banks are private domestic banks, and 0 otherwise
	REGIONAL_D	1 if banks are regional banks, and 0 otherwise
	FOREIGN_D	1 if banks are pure foreign banks, and 0 otherwise
Total Assets	TA	Fixed assets + loans + other earning assets + other assets
Financial Crisis dummy	CRISIS	As a proxy for the global financial/economic crises, taking the value 1 for 2008 and after, and 0 before.

Notes: (i) Definition of deregulation variables and data for computation are based on Barth et al. (various issues) and Abiad et al. (2008). Details of the computation of the indices are given in Table 7.2
(ii) The STATE_D variable is omitted in the inefficiency model equation to avoid the dummy variable trap and multicollinearity.

The definition of variables in the translog cost frontier is the same as in the cost function discussed in Section 6.2.5 in Chapter 6. Table 7.1 therefore captures the definition of the variables of the inefficiency function, together with their measurement. Details on the derivation of the deregulation indices are discussed below.

7.2.3 Derivation and measurement of deregulation reform policy variables

To account for the multi-faced nature of the banking reform policies and the different timing of the reforms, we introduce three distinct deregulation reform variables to measure the three main deregulation policies implemented: removal of activity and branch restrictions through universal banking; relaxation of entry restrictions through the new licensing policy; and removal of credit restrictions through the abolition of high secondary reserve requirements.[108],[109]

The three deregulation variables ACTV_R, ENTRY_R and CCRR accordingly correspond to banking activity restrictions, bank entry restrictions and credit control and reserve requirements. ACTV_R, which represents the universal banking model introduced, measures the removal of both activity and branch restrictions in Ghana's banking model prior to 2004. The ENTRY_R index reflects the open licencing policy by the Central Bank in 2006 and captures the relaxation of banking entry restrictions. The CCRR signals the abolition of the secondary reserve requirements in 2006 and thus seen as a relaxation of restrictions on loan growth. Since our study is on the impact of the deregulation policies implemented by the Bank of Ghana, we focus on only those three reforms rather than all the reform components proposed in Abiad et al. (2010) and Barth et al. (2012) databases.

The derivation of the ACTV_R and ENTRY_R indices is based on the surveys on measuring banking regulation by Barth et al. (2001, 2003, 2007 and 2012), Based on survey responses from regulators/central banks which are coded by the authors ranging between 1 and 4, we utilise these indices to generate policy liberalisation scores for each of these two variables.

[108] The open licensing policy was not due to a 'relaxation in the requirements for new banks' to be licensed per se. It was more of a change in policy to attract new banks (without relaxing the requirements) after the regulator had the autonomy to do that under the new Bank of Ghana Act (2002). Prior to that, the regulator was accountable to government for such decisions.

[109] Although a fourth reform policy, passage of the Credit Reporting Act was passed in 2007, its operationalization in terms of licensing credit reference bureaus (CRBs) and commencement of activities by these CRBs did not take place until 2011-12. Accordingly, we do not include this reform policy.

The ACTV_R variable is derived from the activity restrictions and mixing banking/commerce as defined in the Barth et al. (2012) database.[110] The ENTRY_R index is derived from the ratio of banking applications denied in the Barth et al. (2012) database, and coded 1-4 using a coding structure proposed by Abiad et al. (2010). The derivation of the CCRR index is also based on guidelines on the financial reform database by Abiad et al. (2010).[111] Coding ranges between 1 and 4, with higher values reflecting more liberalisation.

Table 7.2 shows the derivation of the deregulation indices and their sub-components. The sum of the indices for each category is normalised to lie between 0 and 1 to facilitate comparison and aggregation into the combined deregulation index (DEREG).

Table 7.2 Derivation of deregulation policy indices

A. ACTV_R - Relaxation of activity restrictions

Banks undertaking non-core banking activities	**2000**	**2002**	**2004**	**2006**	**2008**	**2010**	**2012**	**2014**
Securities	2	2	1	1	2	2	2	2
Insurance	0	0	1	1	1	1	1	1
Real Estates	0	0	2	1	1	1	1	1
Mixing Banking/Commerce	2	2	2	2	2	2	2	2
Branch banking restrictions	1	1	2	2	3	3	3	3
Total score	5	5	8	7	9	9	9	9
ACTV_R index	0.3	0.3	0.5	0.5	0.6	0.6	0.6	0.6

*Additional line created by author to reflect relaxation of branch restrictions. Restrictions were imposed on merchant banks, but not on commercial and development banks. Gradual relaxation during the transition period of 2003-2006.

Sources: 1, 2) Barth et al. (2012); 3) Author

[110] A major regulatory activity which was relaxed in the case of Ghana was bank branching, which was partially restricted but abolished as part of the universal banking model introduced. An additional category on bank branching is accordingly created with the same coding structure. As noted by Barth et al. due to the usual challenges in information retrieval especially from developing countries, the authors suggest that these databases remain work in progress, and so could be combined with country knowledge information to explore such reform policies. It is in this context that we added this sub-component on branch liberalisation.

[111] We note that the database covers the period up to 2005. However, this is a quantitative measure based on percentage reserve requirements and thus the indices can easily be computed for the post-2005 period.

B. ENTRY_R – Relaxation of bank entry restrictions

	2000	2002	2004	2006	2008	2010	2012	2014
Based on % of entry applications denied [1]	50%	50%	27%	27%	0%	0%	0%	0%
Score [2]	1.0	1.0	2.0	2.0	3.0	3.0	3.0	3.0
ENTRY_R index	**0.333**	**0.333**	**0.667**	**0.667**	**1**	**1**	**1**	**1**

Policy indicator is based on the Competition Regulatory Variables, based on the World Bank database on Bank regulation by Barth et al. (2012). Codes (based on % of banking entry applications denied (D))

Data is for previous five years, so we assign the average percentage of the previous 5 years for each of the years.

* Although authors suggest percentage of applications denied can be used, we follow a similar coding by Abiad et al. (2010) to categorise this into an appropriate coding. This facilitates both the interpretation of the policy variable and also the aggregation into the combined policy reform index (DEREG).

Sources: 1) Barth et al. (2001, 2003, 2007, 2012) 2) Author's based on Abiad et al. (2010)

C. CCRR – Relaxation of credit controls

	2000	2002	2004	2006	2008	2010	2012	2014
Reserve requirements restrictive*	0	0	0	1	2	2	2	2
Mandatory credit allocations**	1	1	1	1	1	1	1	1
Subsidized credit allocations**	1	1	1	1	1	1	1	1
Total Score (out of 4)	2	2	2	3	4	4	4	4
CRR index	0.333	0.333	0.333	0.5	0.667	0.667	0.667	0.667

*Coded 0 if RR >20%; 1 if 10% < RR < 20%; 2 if RR < 10%): ** Coded 0 if they exist; 1 if they do not exist

35% secondary reserves reduced to 15% in 2005 and abolished end of 2006.

Mandatory credit allocations (rationing) and credits at subsidized rates were not in place during the sample period.

Source: Abiad et al. (2010)

D. Summary and aggregation of deregulation reform indices

	2000	2002	2004	2006	2008	2010	2012	2014
ACTV_R	0.3	0.3	0.5	0.5	0.6	0.6	0.6	0.6
ENTRY_R	0.3	0.3	0.7	0.7	1.0	1.0	1.0	1.0
CCRR	0.3	0.3	0.3	0.5	0.7	0.7	0.7	0.7
DEREG (Average of the 3)	0.3	0.3	0.5	0.5	0.8	0.8	0.8	0.8

7.2.4 Data and descriptive statistics

The dataset for the stochastic cost frontier is the same as used for the translog cost function in Chapter 6. Detailed descriptive statistics of the data are provided in Section 6.3.1. We provide below descriptive statistics of the deregulation indices used in the inefficiency function, and the aggregate index used in the cost frontier in Table 7.3.

Table 7.3 Descriptive statistics of deregulation indices

	Mean	**Median**	**Std. Dev.**	**Minimum**	**Maximum**
ACTV_R	0.511	0.6	0.108	0.333	0.600
ENTRY_R	0.756	1	0.285	0.333	1.000
CCRR	0.533	0.667	0.152	0.333	0.667
DEREG	0.600	0.755	0.179	0.333	0.756

We observe an increase in the level of liberalisation in each of the three components, although at different paces. Bank entry experienced the highest level of liberalisation, followed by liberalisation in credit control and reserve requirements. Activity restrictions recorded the lowest form of liberalisation. On account of these developments, the average deregulation index shows an increase in the overall level of liberalisation in the banking sector.

In terms of bank size, we classify banks into 3 categories: large, medium and small banks. Large banks are banks with total assets above the upper quartile of total assets of the industry; small banks are banks with assets size below the lower quartile; while banks in-between are classified as medium-sized for each year. Based on these criteria, the number of banks and the average total assets for each category is captured below in Table 7.4

Table 7.4 Bank classification by assets size

	2000		**2014**	
	No of banks	average asset size	No of banks	average assets size
Large banks	4	721	6	1150
Medium banks	8	156	12	573
Small banks	4	17	6	190

Assets size in GHC Million

7.3 Empirical results and analysis

The estimation of equations (7.7) and (7.8) was performed via a one-step maximum-likelihood method using Stata, and the **sfpanel** stochastic cost frontier code developed by Belotti, Daidone, Ilardi, and Ate (2013).

7.3.1 Analysis of the stochastic cost frontier

Table 7.5 Maximum likelihood estimates of stochastic cost frontier and inefficiency model

Coefficients of frontier model

Ln(TC/P_{LK})	**Ln TC***	**Parameter**	**Coefficient**	**Standard Error**	**t-ratio**
Ln (P_{FUNDS}/ P_{LK})	lnw1*	β_1	0.6266***	0.0863	7.2600
Ln (LOANS)	lny1	γ_1	0.4621***	0.1134	4.0800
Ln (OEA)	lny2	γ_2	0.0061	0.1042	0.0600
Ln (OBS)	lny3	γ_3	0.1202	0.1378	0.8700
(Ln (P_{FUNDS}/ $\mathbf{P_{LK}}$))^2	(lnw1)^2	β_{11}	0.0991***	0.0151	6.5800
(Ln LOANS)^2	(lny1)^2	γ_{11}	0.0646***	0.0192	3.3700
(Ln OEA)^2	(lny2)^2	γ_{22}	0.1154***	0.0144	8.0000
(Ln OBS)^2	(lny3)^2	γ_{33}	0.0545**	0.0234	2.3300
(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)(Ln LOANS)	(lnw1*)(lny1)	ψ_{11}	-0.0548**	0.0280	-1.9600
(Ln P_{FUNDS}/ P_{LK})(ln OEA)	(lnw1*)(lny2)	ψ_{12}	-0.0483**	0.0208	-2.3200
(Ln P_{FUNDS}/ $\mathbf{P_{LK}}$)(Ln OBS)	(lnw1*)(lny3)	ψ_{13}	0.0882***	0.0322	2.7400
(Ln LOANS)(Ln OEA)	(lny1)(lny2)	γ_{12}	-0.1019***	0.0260	-3.9100
(Ln LOANS)(Ln OBS)	(lny1)(lny3)	γ_{13}	-0.0023	0.0367	-0.0600
(Ln OEA)(Ln OBS)	(lny2)(lny3)	γ_{23}	-0.0948***	0.0336	-2.8200
T	T	θ_1	-0.0595**	0.0298	-1.9900
Tsqr	Tsqr	θ_{11}	-0.0068***	0.0018	-3.8000
T(Ln P_{FUNDS}/ P_{LK})	Tlnw1*	η_1	0.0055	0.0043	1.3000
T(Ln LOANS)	Tlny1	ζ_1	-0.0133**	0.0052	-2.5400
T(Ln OEA)	Tlny2	ζ_2	-0.0048	0.0042	-1.1300
T(Ln OBS)	Tlny3	ζ_3	0.0216***	0.0052	4.1300
DEREG	DEREG	τ_1	-2.2184***	0.5193	-4.2700
TDEREG	TDEREG	τ_2	0.3275***	0.0629	5.2000
Intercept	Intercept	α_0	3.5947***	0.3049	11.7900

Coefficients of inefficiency model

ACTV_R	δ_1	0.5878	0.8571	0.6900
ENTRY_R	δ_2	-4.4812***	0.8905	-5.0300
CCRR	δ_3	7.6924***	1.9254	4.0000
DOMESTIC_D	δ_4	-0.1146*	0.0603	-1.9000
REGIONAL_D	δ_5	-0.2708***	0.0875	-3.0900
FOREIGN_D	δ_6	-0.2050**	0.0846	-2.4200
Ln TA	δ_7	-0.2232***	0.0525	-4.2500
CRISIS	δ_8	0.3577***	0.1122	3.1900

Log-Likelihood = 204 Wald (chi^2) = 7227 $\sigma_u = 0.055$*** $\sigma_v = 0.124$*** $\lambda = 0.44$***

*** p<0.01, ** p<0.05, * p<0.1 (Significance at 1%, 5% and 10% level respectively)

The estimated parameters of the stochastic cost frontier and the inefficiency function, together with the variance parameters using the Battese-Coelli (1995) one-step estimation are reported in Table 7.5.

Table 7.6 Marginal effects of the determinants of cost inefficiency

ACTV_R	δ_1	0.2158
ENTRY_R	δ_2	-1.645***
CCRR	δ_3	2.8239***
DOMESTIC_D	δ_4	-0.0421*
REGIONAL_D	δ_5	-0.0994***
FOREIGN_D	δ_6	-0.0753**
Ln TA	δ_7	-0.082***
CRISIS	δ_8	0.0742***

The marginal effect is interpreted as the as the % change of bank cost inefficiency due to a per unit change in any of the z-variables. A positive sign indicates that the z-variables positively impacts cost inefficiency or reduces cost efficiency. Results computed in Stata, using the sfpanel developed by Belotti et al. (2012)

Before interpreting the estimated parameters of the cost frontier, we again analyse the various tests of hypotheses of the model. We tested the suitability of the translog functional form; existence of technological change, and the existence of non-neutral and scale-augmenting technological change; whether technology is homothetic; the relevance of deregulation in the cost frontier; and the overall significance of the inefficiency model. The results of the tests are summarised in Table 7.7.

Table 7.7 Tests of hypotheses: stochastic cost frontier and inefficiency model

Null hypothesis (H0) defined by the following parameter restrictions. Full details provided below	Test Statistics (χ^2)	Decision (at the 1% significant level)
(1) $\beta_{11} = \gamma_{11} = \gamma_{22} = \gamma_{33} = \psi_{11} = \psi_{12} = \psi_{13} = \gamma_{12} = \gamma_{13} = \gamma_{23} = \theta_{11} = \eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = 0$	439.12	Reject
(2) $\eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = 0$	17.69	Reject
(3) $\theta_1 = \theta_{11} = \eta_1 = \zeta_1 = \zeta_2 = \zeta_3 = \tau_2 = 0$	40.60	Reject
(4) $\psi_{11} = \psi_{12} = \psi_{13} = 0$	9.10	Accept
(5) $\tau_1 = \tau_2 = 0$	30.39	Reject
(6) $\delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = \delta_7 = \delta_8 = 0$	61.73	Reject

Notes: Tests of hypotheses based on the Wald test which involves imposing restrictions on the parameters, (i.e. the parameters equal zero) of the estimated stochastic cost frontier and inefficiency model. The tests are chi-squared distributed with the degrees of freedom equal to the number of restrictions.

- Hypothesis (1): the second-order coefficients in the translog function are zero, and so the Cobb-Douglas functional form is adequate to represent the data.
- Hypothesis (2): the interaction between the time trend (T) and the input prices and outputs are zero, that is, there is no non-neutral and scale augmenting technical change and there is no change in the output composition and input mix associated with the time trend.
- Hypothesis (3): all coefficients involving the time trend are zero, and therefore that there is no technical change or technological progress over time.
- Hypothesis (4): the production technology is homothetic, and so the interactive terms of input prices and output quantities are zero.
- Hypothesis (5): the combined deregulation variable and its interaction with time trend are zero, and therefore deregulation is not relevant in explaining cost, and there is no deregulation-induced technical change.
- Hypothesis (6): all the coefficients of the inefficiency model are jointly zero, and so the variables are not relevant in explaining cost inefficiency.

With regards to the functional form of the cost frontier, we reject the Cobb-Douglas and conclude that the translog fits the data more appropriately. We also note that there is technological change associated with time, and that there is both non-neutral and scale-augmenting technical change present in the model. Deregulation was also found to be relevant in affecting costs and cost technology. The production technology is however found to be homothetic. We also reject the null hypothesis that all the coefficients of the inefficiency model are jointly zero, and conclude that the inefficiency covariates are relevant in explaining cost inefficiency.

The regularity tests to check if the estimated translog cost function possesses the properties of a cost function yielded similar results as those reported in section 6.3.1, with most of the observation-specific cost elasticities of outputs and input prices being positive and a few violations in respect of the cost elasticity of OEA and OBS.[112] We therefore conclude that the estimated cost function satisfies the regularity conditions of linear homogeneity, symmetry and monotonicity in input prices and outputs as required by theory.[113]

Regarding the estimated coefficients of the cost frontier (Table 7.5), we observe that most of the regressors are statistically significant at the 5% significance level, with the expected signs. In particular 18 (out of the 23) are statistically significant at the 5% level. The Wald test on joint significance or overall significance of the explanatory variables of the model is high. Here again, a few variables (the first-order levels of OEA and OBS) are not statistically significant, which as explained in Section 6.3.1 is normal in the translog due to the collinearity it has, and also explains the few monotonicity violations observed above.

Both the time trend, T (and its quadratic form, T-sqr) variables are statistically significant, suggesting technology-induced changes in the cost structure of banks. The negative coefficient of the time trend signifies pure technological progress, which however declines over the time after the fourth year.[114] We also note that the coefficient of the variable DEREG is negative and statistically significant, which suggests that the deregulation reforms taken together helped to reduce cost or achieved some cost efficiency gains during the period, although its effect tends to taper down over time since the TDEREG coefficient is positive.

[112] These violations represent only 2% (OEA) and 18% (OBS) of the total observations.

[113] As observed earlier, homogeneity in input prices was imposed prior to estimation while symmetry condition was also imposed. Non-negativity of costs had already been checked during data input.

[114] The effect of pure technological change is derived as the derivative of equation (7.6) with respect to the time trend (T) (excluding variables with time interactions) which is $\theta_1 + 2\theta_{11}T$. Based on the values of $\theta_1$and θ_{11}, and setting the derivative to zero, we note that -0.06 + 2(-0.0068)T = 0, which gives T=4.4.

Cost efficiencies are influenced not only by technological progress but also by scale economies. Economies of scale (EOS) measure the relative change in a bank's total cost for a given proportional change in all its outputs. We accordingly measure scale efficiency in the banking industry with reference to the cost elasticities of outputs evaluated at the means, shown in Table 7.8. We observe that the banking industry exhibits scale economies as the sum of the cost elasticity w.r.t to all outputs is less than one.[115] The implication of scale economies is Ghana's banking sector is that large banks are able to make cost efficiency gains. Scale economies are also found across all the ownership class of banks, with private domestic banks exhibiting the highest scale economies, followed by regional banks and foreign banks, with state banks exhibiting the lowest scale economies.

Table 7.8 Output elasticities and scale economies – all banks and by ownership class

	All banks	State-owned	Domestic	Regional	Foreign
CE_{y1}	0.476	0.532	0.473	0.445	0.464
CE_{y2}	0.302	0.277	0.243	0.321	0.391
CE_{y3}	0.118	0.127	0.140	0.125	0.066
Returns to scale	1.12	1.07	1.17	1.12	1.09

7.3.2 Analysis of the determinants of inefficiency

The test on the joint significance of the inefficiency effects also confirms that the inefficiency covariates are relevant in explaining cost inefficiency. This is also confirmed by the statistically significant value of lambda, $\lambda = \frac{\sigma_u}{\sigma_v} = 0.44$ which indicates that inefficiency significantly contributes to the distance from the frontier and is a robust check on the model specification test earlier carried out.[116]

In respect of the deregulation indices, the coefficient of ACTV_R is not statistically significant, suggesting that the relaxation of activity restrictions did not influence cost inefficiency. We interpret this to be due to the somewhat limited scope of the implementation of this policy. While universal banking in principle broadened the definition of banking to

[115] The sum of the cost elasticities of input prices is unity as expected due to the linear homogeneity assumption. As shown earlier, we can make inferences about economies of scale (EOS) based on the cost elasticity of output (CE_{yi}). EOS is given by $1/(CE_{y1} + CE_{y2} + CE_{y3})$.

[116] A low and insignificant λ suggests that inefficiency is not significant in contributing to the distance of the frontier and so all variation is due to pure randomness and thus OLS could be applied to the estimation.

include leasing, mortgage financing and insurance business, these non-core banking activities do not seem to have really taken off. This is backed up by the trend in the level of openness in the ACTV_R deregulation index from the descriptive statistics which shows a relatively small level of liberalisation in activity restrictions. The survey results on Ghana from Barth et al. (2012) suggest restricted activity with regards to the new business lines, and thus the expected synergies and associated cost efficiency to be gained from product diversification into non-core banking business are yet to be realised. Banks seem to focus on their traditional banking business model of short-term lending and investment in government securities and this culture might take a long time to change. This is also confirmed by the improvement in their technology of production which gets cheaper over time as suggested by the significant negative interaction terms of LOANS and OEA with T in the frontier estimations. Non-core banking activities should which reflects in OBS is also the output whose cost of production increases over time (positive interaction with T in the frontier).

There could be several reasons for this phenomenon. First, the lack of a unified regulatory and supervisory body for the financial services industry makes it difficult for banks to enter in particular into the insurance and securities markets as they would have to acquire licensing from the regulatory bodies overseeing the insurance and securities markets.[117] Second, some attempts by banks to collaborate with insurance companies to provide joint products, say banc assurance have also not been successful due to different requirements from the two different regulatory bodies for such joint-product offerings. Third, the medium- to long-term nature of these businesses (insurance, securities and mortgage) seem unattractive to banks who are culturally used to the short-term banking business model, and especially as they continue to earn attractive returns from such traditional banking business. Thus, in spite of the introduction of universal banking, there are operational barriers which constrain the full realisation of this policy.

The ENTRY_R variable has a statistically significant and negative coefficient, which suggests that the relaxation of entry restrictions contributed in reducing cost inefficiency. The implementation of the policy reform seems to have impacted on competitive pressures and facilitated the enhancement of efficiency among banks. To remain competitive, banks have to adopt cost-efficiency measures including deploying electronic banking technology, process

[117] There is no single financial services regulator as the insurance and securities industries have separate regulatory bodies, the National Insurance Commission and the Securities and Exchange Commission respectively, whose mandates are distinct from the Central Bank.

re-engineering to cut waste, and other cost-minimisation strategies to enhance efficiency and remain competitive.

The CCRR variable is also statistically significant but positive. This suggests that the removal of credit restrictions through the scrapping of reserve requirements is associated with an increase in cost inefficiency. We interpret this outcome to reflect the risks inherent in increasing credit portfolio in a deregulated environment where the credit environment is not strong. It could also be that banks were not yet "good enough" at efficiently managing the extra liquidity that came from the scrapping of the reserves, as discussed below. Strong credit growth is usually accompanied by high non-performing loans and higher loan loss provisions, especially where the credit environment is not supportive of such loan growth.

The banking industry's loan book increased significantly with loan-to-deposit ratio of 53% in 2004 (prior to removal of the secondary reserves), rising to 67% in 2006 (the year the reserves were fully abolished), and further to 79% in 2008. This deepening of financial intermediation immediately following the scrapping of the secondary reserves subsequently affected loan quality as non-performing loans (NPLs) of the banking industry shot up from 8% in 2008 to 16.2% and 17.6% in 2009 and 2010 respectively.[118] High NPLs have been found to be associated with decline in cost efficiency (Berger and DeYoung, 1997; Girardone et al., 2004). Accordingly, the significant growth in loans following the relaxation of credit constraints and the attendant high NPL accounted for the increase in cost inefficiency in Ghana's banking sector.

On the effect of the ownership dummies on inefficiency, we observe that the ownership dummy variables are all negative and statistically significant, but only REGIONAL_D and FOREIGN_D are statistically significant at the 5% level. A test on the equality of all the three ownership dummies is rejected at the 5% level of significance. Further tests confirm that the difference in efficiency between state-owned and private domestic banks is not statistically significant, while regional and foreign banks efficiency levels are statistically significant from state-owned banks at the 5% significance level.[119] We accordingly conclude that on average regional and foreign banks are relatively more cost efficient than state-owned and private domestic banks.

[118] NPLs usually lag behind loan growth; hence the high NPL ratio in 2009 and 2010 reflects the large loan build up from 2006 to 2008. Loan growth was subsequently scaled down thereafter with a corresponding fall in NPL ratio subsequently.

[119] That is, tests that the coefficients of the ownership dummies are significantly different from zero.

The total assets variable is also found to be statistically significant, and the negative coefficient indicates that large banks are able to reduce cost inefficiency than smaller banks. In other words, large banks seem to be able to gain cost advantages through economies of large scale operations over small banks. This confirms the scale economies found to exist within the banking industry, and across all the ownership categories.

The CRISIS variable is also statistically significant but positive, suggesting an adverse impact of the global financial crisis on cost efficiency, as expected. As shown in Figures 7.1 and 7.2 (below), cost efficiency declined in 2008 across all ownership and bank size groups. Banks however recovered the year after and were able to improve on efficiency levels thereafter.

The estimated coefficients of the determinants of cost inefficiency in Table 7.5 only indicate the direction of the effect of the inefficiency covariates on the inefficiency term and not on the magnitude of the impact on cost inefficiency.[120] A quantification of the marginal effects of each of the determinants on cost inefficiency is given in Table 7.6. The marginal effects show that for a per unit reduction in entry restrictions, cost inefficiency decreases by 1.6% while a unit reduction in credit constraints leads to a 2.8% increase in inefficiency.

We complete our analysis by showing the estimated average cost efficiencies for the banking industry, across the ownership groupings, and across different bank sizes over the study period. Table 7.9 and Figure 7.1 show the trend in the industry and per ownership class. It shows that the average cost efficiency of Ghana's banking industry for the sample period is 91.88%.

Banking industry efficiency increases gradually during 2000 to 2003 and then experiences a sharp increase in 2004 with technological progress contributing to the efficiency enhancement as seen earlier. Efficiency marginally declines in 2005 and thereafter increases steadily until 2007 before experiencing a slump in 2008, which could be attributed to the global financial crisis. Efficiency rebounds in 2009, increases marginally but consistently throughout the rest of the study period.

[120] As mentioned in Chapter 5, the efficiency effects in the BC (95) model, like the whole KGMHLBC class of models are monotonic.

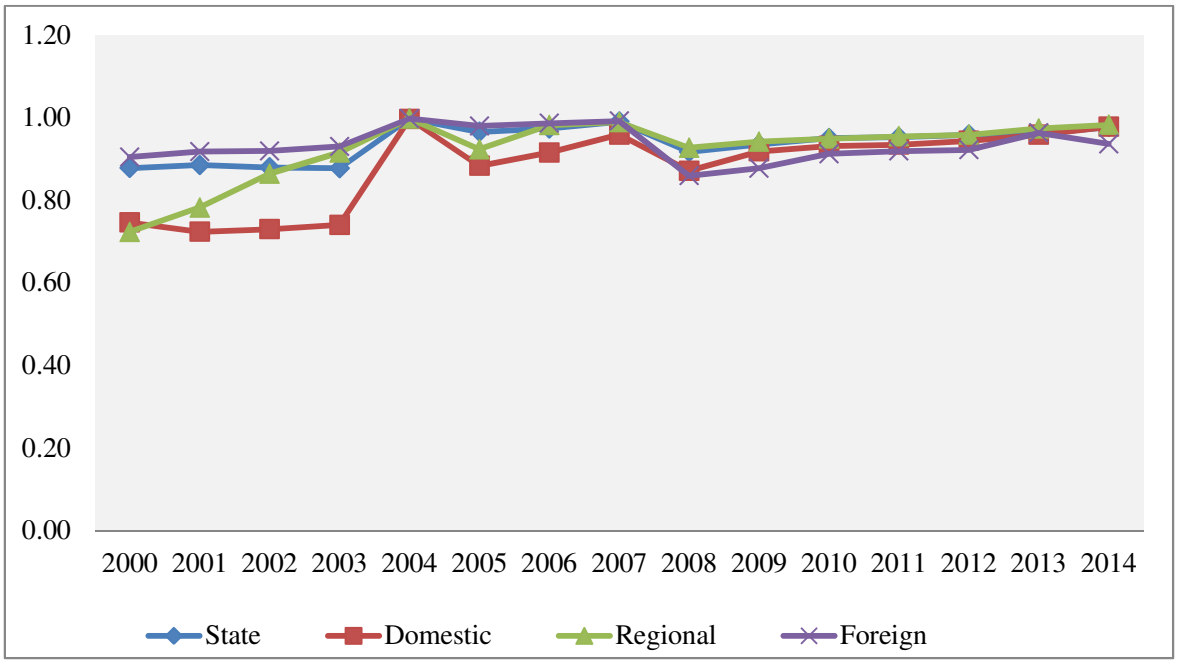
Table 7.9 Annual mean efficiencies of banks and across ownership groups

	All banks	State-owned	Domestic	Regional	Foreign
2000	0.8227	0.8781	0.7461	0.7231	0.9047
2001	0.8272	0.8852	0.7236	0.7823	0.9178
2002	0.8360	0.8794	0.7300	0.8644	0.9194
2003	0.8499	0.8776	0.7402	0.9157	0.9303
2004	0.9976	0.9977	0.9972	0.9979	0.9980
2005	0.9309	0.9657	0.8833	0.9234	0.9799
2006	0.9588	0.9738	0.9157	0.9818	0.9866
2007	0.9809	0.9904	0.9593	0.9894	0.9922
2008	0.8952	0.9172	0.8714	0.9274	0.8595
2009	0.9211	0.9349	0.9185	0.9419	0.8775
2010	0.9371	0.9506	0.9314	0.9492	0.9123
2011	0.9417	0.9530	0.9347	0.9548	0.9193
2012	0.9469	0.9587	0.9441	0.9587	0.9224
2013	0.9654	0.9620	0.9586	0.9736	0.9632
2014	0.9709	0.9771	0.9780	0.9829	0.9366

A similar pattern in the industry trend is exhibited by all the 4 different banks by ownership status. Figure 7.1 shows the evolution of average annual efficiencies of banks per ownership class.

Prior to the implementation of the reforms, the existing and incumbent state banks and foreign banks operated at higher levels of efficiency compared to the newly-established private domestic banks and regional banks. However, while those state and foreign banks increased their efficiency levels gradually, the domestic and regional banks experienced sharp increases in efficiency to catch up with the other two ownership bank categories in 2004. After some form of convergence in 2004, the domestic and regional banks experienced a relatively steeper dip in efficiency levels in 2005 before catching up again thereafter until 2007.

Figure 7.1 Trend in annual mean efficiencies of banks by ownership class



The decline in efficiency witnessed in 2008 affected all categories of banks but more especially private domestic and foreign banks. Post-2008, all banks witnessed a steady rise in cost efficiency, with regional banks experiencing the highest rise thereafter.

An interesting observation is the higher efficiency levels of regional banks from 2007. One possible reason is that the regional banks leverage on their strengths as foreign banks as well as their understanding of the African banking terrain. They combine superior management expertise, training and technological innovation, among others as quasi-foreign banks with the local knowledge and experience in doing business in Africa. Some of them have also centralised back office operations such as loan processing at their regional head offices to make significant cost savings, in addition to leverage their regional bank networking platform to take advantage of regional cross-border products such as trade finance and remittance businesses. In general, regional and foreign banks seem to enjoy cost advantages over the entire due to superior management and technology-driven products.

We now analyse the pattern of efficiency levels of different bank sizes. We noted that bank size (proxied by total assets) was significant in reducing bank inefficiency. Furthermore, the presence and significance of scale economies attest to the positive impact of size on enhancing cost efficiency. This is confirmed by the pattern of the average efficiencies of

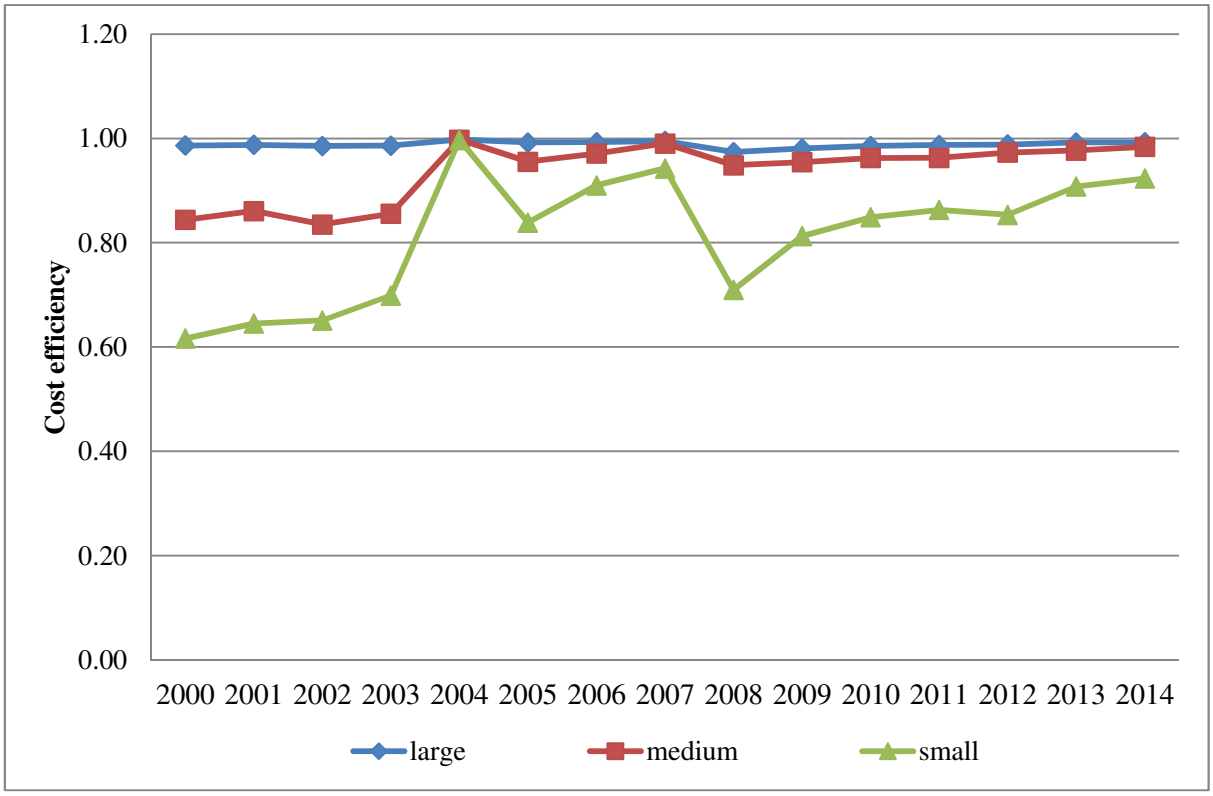
large, medium and small banks. As mentioned earlier, we define large banks as banks whose assets size in each year is above the upper quartile in that year; while small banks are those with assets size within the lower quartile. Banks within the lower and upper quartile range are classified as medium-sized banks in each year. The average cost efficiencies of these different bank sizes are captured in Table 7.10 and Figure 7.2.

Table 7.10 Annual mean efficiencies of banks by size

	Large Banks	Medium Banks	Small Banks
2000	0.9865	0.8440	0.6162
2001	0.9876	0.8608	0.6451
2002	0.9856	0.8351	0.6509
2003	0.9862	0.8557	0.6990
2004	0.9982	0.9977	0.9971
2005	0.9922	0.9555	0.8388
2006	0.9927	0.9711	0.9100
2007	0.9948	0.9904	0.9424
2008	0.9738	0.9487	0.7097
2009	0.9806	0.9545	0.8127
2010	0.9854	0.9624	0.8490
2011	0.9872	0.9627	0.8631
2012	0.9883	0.9728	0.8536
2013	0.9922	0.9769	0.9079
2014	0.9926	0.9838	0.9232

As shown, large banks exhibited the highest efficiency level; followed by medium-sized banks; with small banks having the lowest cost efficiency throughout the study period. Here again, the incumbent large banks (mostly foreign and state-banks) exhibited higher efficiency levels, with the new small and medium-sized banks (characterised by mostly regional and private domestic banks), catching up initially. Small banks were also hit most by the indirect effects of the global financial crisis, but managed to thereafter enhance their efficiencies, while there is a sort of convergence in efficiency levels of large and medium-sized banks.

Figure 7.2 Bank efficiency levels for different bank sizes



7.4 Conclusions

This chapter has examined empirically the impact of deregulation on banking efficiency, and the role of bank ownership and size in influencing efficiency. The study uses the stochastic cost frontier approach and specifically the BC95 model, which estimates the cost frontier with an inefficiency model in a one-step simultaneous estimation using maximum likelihood techniques. The inefficiency model captured deregulation variables, ownership, assets size and a crisis variable. Deregulation policy variables were constructed using data, information and coding rules from international surveys and databases on banking regulations and reforms from Barth et al. (2012) and Abiad et al. (2010).

The empirical results show non-uniform impacts of the three deregulation policies on banking efficiency. The relaxation of activity restrictions through universal banking had no effect on enhancing efficiency; relaxation of bank entry restrictions had a positive effect in enhancing efficiency; while removal of credit constraints had a negative effect on cost efficiency. In terms of ownership, we observe that the new regional and private domestic banks managed to increase their efficiency levels initially to catch up those of the incumbent foreign and state-

owned banks. Over the study period though, regional and foreign banks are found to be more efficient than state-owned and private domestic banks. Of particular interest in the African context is the emergence of these regional pan-African banks which this study shows have high efficiency levels. We also find that bank size is significant in positively influencing cost efficiency, confirmed by the economies of scale as well as the variation in efficiency across different bank sizes. The study also finds evidence to suggest that the global financial crisis had an adverse impact on banking efficiency.

The results of this study in general compare with those of other studies on African countries such as Hauner and Peiris (2008) on Uganda; Zhao and Murinde (2011) on Nigeria; Mwega (2011) on Kenya and Mwenda and Mutoti (2011) on Zambia, who also find the efficiency-enhancing impact of deregulation reforms. However, as our study shows, there is the need to separately analyse different deregulatory policies due to the non-uniformity of effects each policy could have on efficiency. In terms of ownership-efficiency, our findings are similar to those of Hauner and Peiris (2008) and Mwega (2011) who find that foreign banks are generally more efficient than domestic banks. Our results differ from Poshakwale and Qian (2011) who find domestic banks more efficient than foreign banks in their study on Egypt, although we observe similarity as they do not find any efficiency differential between state-owned banks and private domestic banks just as we do. In terms of size-efficiency, our results are similar to Hauner and Peiris (2008) and Zhao and Murinde (2011) who find large banks to be more efficient than small banks.

The policy recommendations are that deregulation reforms could enhance banking efficiency in the context of African countries, but specific reforms should take into account country peculiarities in terms of the financial infrastructure and development. More specifically, opening up of banking system to attract new banks could enhance efficiency; especially to regional bank and foreign banks, as this study shows that that they are the most efficient. Also, state-owned banks may not always need to be privatised for them to be efficient. With the right reforms and operating environment, state banks can be motivated to enhance their efficiency levels. The policy of relaxation of activity restrictions must be backed by further regulatory reforms if banks are to take advantage of synergies in entering the insurance and securities markets. The scrapping of secondary reserves, while it is welcome to deepen financial intermediation, efforts at improving the credit environment must be intensified so as to minimise the risks inherent in loan growth in a weak credit environment.

CHAPTER 8 SUMMARY AND CONCLUSIONS

8.1 Research summary and findings

Financial deregulation policies seem to have gained prominence more recently in most developing economies since the 1970s, with African countries experiencing some form of deregulation reforms during the 1980s and 1990s as part of broader economic reform programs supported by the World Bank and the IMF. Those initial reforms did help transform most banking sectors from a controlled system to a more market-based system following with liberalisation of interest rates and exchange rates, abolition of directed credit and sectoral credit ceilings, strengthening the regulatory framework and undertaking some form of privatisation of state-owned banks. In addition, measures to develop other components of the financial system were implemented with the establishment of new stock exchanges in eight African countries during 1980-1999 including the Ghana Stock Exchange.

The reforms however did not achieve much in terms of facilitating competition and efficiency in the banking sector, which remains the dominant sector of Africa's financial system. Understandably, being the maiden liberalisation across the continent, a somewhat gradual approach was taken as experiences in other developing regions showed that sweeping financial deregulation reforms could have detrimental consequences especially in environments characterised by macroeconomic uncertainties, ineffective supervision and inadequate market-support infrastructure (Nissanke and Aryeetey, 1998).

Opening up of the banking sector to domestic private and foreign participation was limited, the banking systems remained fragmented as different classes of banks were licensed to operate in different segments, governments relied on extensive borrowing from the banking sector to finance budget deficits through imposition of high reserve requirements, and the regulatory environment somehow remained under government control. These lingering effects constrained competition and efficiency in the banking sector.

The crucial role of competition and efficiency in banking cannot however be over-emphasized. They serve as key drivers of financial and economic development through their effect on lowering the cost of financial intermediation, improving access to banking services, deepening financial intermediation, facilitating technological progress and innovation, and contributing to the overall growth of the economy. To benefit from the effects of competition

and efficiency in Africa's banking sector therefore requires the pursuit of further reforms to address those inhibitions.

It is against this background that a few African countries have since the turn of the 2000s initiated some reforms aimed at stimulating banking competition and efficiency. Ghana pursued deregulatory reforms under its Financial Sector Strategic Plan (FINSSIP) during 2003-2006, specifically aimed at addressing lingering challenges to the development of the banking sector, and foster competition and efficiency in the sector.

The thesis accordingly sought to evaluate the impact of such deregulatory reforms on the competitiveness and efficiency of Ghana's banking sector, in line with the objectives for which they were implemented.

Specifically, the thesis examined the following three key research questions:

- What has been the impact of the deregulation reforms on the competitiveness of the banking sector in Ghana?

- To what extent have banks' efficiency levels been impacted by the deregulatory reforms?

- Do bank ownership and bank size play any role in influencing bank-level efficiency?

Given that these reforms were implemented during 2003-2006, we choose the sample period of 2000-2014, which we believe is adequate to analyse the impacts of the reforms due to the usual lagged impact of reforms on their intended objectives.[121] Our chosen sample period covers both the pre- and post-reform periods and therefore makes it adequate to undertake our research.

The focus on Africa is due to the gradually changing banking landscape on the continent since the 2000s, while the choice of Ghana for our empirical investigation is because of its comprehensive reform experience which is particularly suitable for addressing our research questions, and makes a useful case study to draw lessons for other African countries. A

[121] We were constrained by data for the pre-2000 period to expand our study period. We however deem the sample period adequate to address the research questions as it fully covers the pre- and post- reform period for our study.

politically stable democracy, a more reasonably autonomous central bank, systematically implemented deregulation reforms, a more unified banking system, and a well-diversified ownership base provide a more naturally controlled empirical condition to address the research questions.

The thesis is also innovative from several respects. First, in our analysis of competition, we focused on the loans market rather than on the general banking industry. The loans market is the largest and most important segment of African banking and indeed the main target of the reforms. Second, we estimated two separate models of competition: the persistence of profits (POP) model and the Boone indicator, which is a relatively new approach never applied to African banking markets. This approach enhances the robustness and completeness of our competition results. Third, unlike most of the literature, rather treating deregulation policies as a general category, we examined the (potentially different) impacts of specific deregulation policies on banking efficiency, using indexes produced by Barth et al. (2001, 2003, 2007, 2012) and Abiad et al. (2010). The different efficiency impacts from the different policies underscores the uniqueness of this approach, and clearly set this study from other studies on Africa which have not explored the informative power of this approach. Finally, our rich bank-specific data covering the pre- and post-reform facilitated the analysis of competition and efficiency effects over time and more importantly provided an effective mean of analysing policy impacts and efficiency of different classes of banks.

Two empirical studies were undertaken to examine the impact of the reforms on competitive conditions in the banking sector and efficiency levels of banks. The empirical paper on competition focused primarily on the loans market as it represents the most important, yet uncompetitive segment of the financial sector, and for which the banking reforms were specifically aimed at. Two credible and well-established dynamic competition models, the Persistence of Profits (POP) and Boone indicator models, which facilitate the assessment of policy impacts on competition in the loans market, were estimated.

The persistence of profits (POP) model shows that the level of persistence of loan overcharge was low during the pre-reform period, but increased during the post-reform period. The 0.324 persistence parameter of loan overcharge during the pre-reform period of 2000-2007, suggests that on average 32.4% of loan overcharge in one year feeds into the following year during this period. The increase in the persistence parameter to 0.5 during the post-reform period shows that on average 50% of a year's loan overcharge feeds into the next year's overcharge. To the extent that loan overcharge reductions are associated with increased competition, the increase

in the persistence parameter suggests that average competitive conditions declined during the post-reform period. Other macroeconomic variables, in particular, the real monetary policy rate was found to be a key variable positively impacting loan overcharge. Banking concentration was however not found to be significant, while the global financial crisis was found to have an adverse impact on competition. The decline in competitive conditions was driven by the adverse macroeconomic environment which was contributed in part by the effects of the global financial and economic crises.

The Boone model also examined competition in the loans market, using the relationship between banks' market share of loans and the marginal cost of loans (a proxy for inefficiency in the loans market). A high value of the Boone coefficient is indicative of a stronger competitive environment as it shows stronger reallocation of market share from inefficient banks to more efficient banks. The computed annual Boone competition measures showed that competition was relatively stable during 2001-2004, but intensified during 2005-2008 and reflected the immediate positive impact of the reforms in facilitating competition. The annual Boone measures for the subsequent years declined significantly, suggesting that competition in the loans market declined considerably during the post-2008 period.

Taken together, the results from the two models are consistent, and suggest that the increased competition impacted initially by the reforms was not sustained due to macroeconomic weaknesses due in part by the effects of the global financial crisis. The adverse macroeconomic environment post-2008 was characterised by high interest rates which led to portfolio re-allocation by banks from loans to government securities, due to the higher yields on risk-free government securities and the high credit risks on loans, contributing to the weakening in competition in the loans market.

The second empirical paper examined the impact of deregulation reforms on efficiency in the banking sector, and the role of bank ownership and assets size in influencing bank-level efficiency. The study adopted stochastic frontier analysis and used the Battese-Coelli (1995) model, which combines estimation of the stochastic cost frontier with the covariates of inefficiency in a one-step maximum likelihood method. To account for the three main deregulation policies, three deregulation reform indices were constructed and measured using survey data, information, and coding rules from two cross-country studies and databases on banking regulations and reforms. These, together with ownership dummies, assets size were captured as inefficiency covariates and estimated with the cost frontier.

The empirical results showed that the reforms taken together impacted on enhancing cost efficiency of banks. The deregulation-efficiency dynamics however show a non-uniform impact of the different deregulation policies on banks' efficiency. While relaxation of entry restrictions had a positive effect in enhancing efficiency, removal of credit constraints had a negative effect on cost efficiency. The lifting of activity restrictions had no effect on enhancing cost efficiency. In terms of ownership effects, regional and foreign banks were found to be relatively more cost efficient compared to private domestic banks and state-owned banks. The efficiency differences between the classes of banks are however marginal, with all of them experiencing improved efficiency over the period. Bank size was found to positively influence cost efficiency, confirmed by economies of scale observed in the banking sector as a whole and across all ownership groups. The study also showed that the global financial crisis had an adverse impact on banking efficiency as efficiency levels across all ownership groups declined during 2008. Efficiency levels rebounded thereafter and increased for the rest of the period.

In summary, the study seems to provide evidence that the deregulation reforms targeted at the credits market did increase competition in the loans market initially but this was somehow short-lived. Macroeconomic weaknesses in increase in interest rates experienced in 2008 and after seem to have not only heightened credit risks but also made the alternative government securities a more attractive market for the banks. These influences accordingly weakened competition in the loans market. Competition was not affected by the level of market concentration. Bank-level efficiencies of the mostly incumbent foreign banks and state-owned banks experienced marginally changes while those of regional and private domestic banks increased during the first four years. Banks accordingly seem to have braced themselves for the reforms by improving on efficiency levels. Technological progress also contributed to the initial increased cost efficiency, but at a declining rate up to 2004. Efficiency levels rebounded from the dip in 2005 and steadily increased until 2007 as the banking industry became more competitive in response to the reforms. The results however point out that deregulation reforms are not enough to sustain banking competition and that, macroeconomic weaknesses and external shocks could interfere with the competitive environment. Macroeconomic weaknesses and external shocks also have adverse effects both on competition and efficiency as the study found, although recovery from the shocks was slower with industry competition than with bank-level efficiency. This suggests that once banks have responded to a more competitive environment by enhancing their efficiency levels, they do not relax their efficiency even if competition declines. Put differently, banks recovered

quickly from the external shocks so that they could respond appropriately to future competitive conditions. The results also suggest that different deregulation policies might have different impacts on efficiency, and so each policy must be analysed in the context of the country's financial development and institutional context.

8.2 Policy implications and recommendations

We summarise below the policy implications of the study and make a couple of policy recommendations. First, banking reforms could potentially drive competition in the banking and also facilitate efficiency improvements. This however requires an assessment of the type of policies to introduce and which are likely to make a greater impact based on the macroeconomic situation, the soundness of the credit environment and the necessary operational changes required to make the policy reforms work. Second, opening up of banking systems to both domestic and foreign participation seem to greatly increase competition directly and enhance efficiency. Banking penetration by the regional pan-African banks as well as foreign banks should be encouraged as the study shows their relatively higher efficiency levels. Third, strong state-owned banks may not necessarily have to be privatised for privatisation sake for them to be efficient. With the right reforms and operating environment, state banks could be strengthened to compete effectively with other classes of banks.

Universal banking may hold promise for economies of scope and efficiency enhancement but institutional changes and other regulations that facilitate incorporation of non-core banking business such as an insurance and mortgage financing with traditional banking must also be pursued. In addition, there is the need for strong incentives that facilitate medium- to long-term financing for banks to enter such businesses. Furthermore, continued reliance of short-term financing government through the banking system serves as a major dis-incentive for the banks to explore other investment avenues, and undermines the policy of abolishing reserve requirements. Alternative sources of financing government deficits must be explored and a limit placed on the level of financing government can access through the domestic banking system.

Bank size matters for efficiency. It may therefore be argued that consolidation policies might be effective in the promotion of large banks to ensure greater efficiency. Any such policy

must however be weighed against any potential challenges of increasing concentration and market power in the banking industry.

The scrapping of reserve requirements, though to deepen financial intermediation, should be weighed against the credit environment in which banks operate. The adverse impact of this policy on efficiency through increased risk behaviour of banks raises the question of the industry's preparedness to deal with rapid credit expansion in a weak credit environment. Measures to improve the credit environment such as credit referencing should be put in place to ensure a sound credit environment before loosening tight credit limits so as to minimise the risks inherent in loan growth in a weak credit environment experienced in most African countries. Although the Credit Reporting Act was passed in 2007 to facilitate the licensing of credit reference bureaus so as to help improve the credit environment, the policy could have been implemented to ensure that the credit environment is improved before the abolition of the secondary reserves. Thus, policy sequencing of reforms is very important.

8.3 Limitations of the research and suggestions for future research

While the thesis provides an empirical assessment of deregulation impacts on banking competition and efficiency in Ghana, there are some limitations which we acknowledge. These limitations are described below some of which could serve as the basis for future research.

The study focused on primarily on the impact of deregulation reforms on competition and efficiency, and incorporated the macroeconomic environment in the competition measure and ownership and size in the efficiency analysis. The study does not take into account other factors which could drive competition and efficiency, and suggests that such factors could be incorporated in future research.

The use of a deregulation policy reform dummy variable in the POP competition model as a proxy to distinguish the pre- and post-reform periods suggests that policy reforms are either there or not, and thus neither captures the different deregulation policies nor the different timing of these policies. While this was somehow effectively handled in the efficiency study by analysing the impact of each deregulation policy on efficiency, the formulation of both the POP and BI models makes it unsuitable to account for the individual deregulation policies.

Future research could look at models that facilitate assessment of different deregulation reforms on competition.

Although the thesis represents a classic single country case study which adequately addresses our research objectives, the research could have benefitted from a cross-country study with two or more African countries with similar reforms and outlooks. This would have provided an opportunity for results to be more robustly generalised. While data constraints prevented us from carrying out such cross-country studies, future research can look into such studies.

In the broader context of Africa's financial sector, microfinance and shadow banking issues have potential relevance in financial development. No consideration was however given to these, especially microfinance. This is a choice made at the start of the research when we decided to focus on the main banking sector. It would be useful for a wider consideration of financial sector issues in Africa to include them in future work.

REFERENCES

Abiad, A., Detragiache, E., Tressel, T. (2010). "A New Database of Financial Reforms". *IMF Staff Papers, 57*(2), 281-302.

African Development Bank (2016), African Economic Outlook. Available at *https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/AEO_2016_Report_Full_English.pdf*

Adjei-Frimpong, K., Gan, C., Hu, B. (2014). "Cost Efficiency of Ghana's Banking Industry: A Panel Data Analysis". *The International Journal of Business and Finance Research, Vol. 8 (2) pp. 69-86, 2014*

Afriat, S. N. (1972). "Efficiency estimation of production functions". *International Economic Review*, 568-598.

Aigner, D. J., Chu, S.-F. (1968). "On estimating the industry production function". *The American Economic Review, 58*(4), 826-839.

Alchian, A. A. (1965). "Some economics of property rights". *Il politico*, 816-829.

Allen, F., Gale, D. (2000). *Comparing financial systems*. Cambridge, Mass; London: MIT Press.

Allen, F., Otchere, I., Senbet, L. W. (2011). "African financial systems: A Review". *Review of Development Finance, 1*(2), 79-113.

Amidu, M., Wilson, J. O. (2014). "Competition in African Banking: Do Globalization and Institutional Quality Matter?"

Angelini, P., Cetorelli, N. (2003). "The Effects of Regulatory Reform on Competition in the Banking Industry". *Journal of Money, Credit and Banking, 35*(5), 663-684.

Anginer, D., Demirgüç-Kunt, A., Zhu, M. (2013). "How does bank competition affect systemic stability?"

Antwi-Asare, T. O., Addison, E. K. Y. (2000). *Financial sector reforms and bank performance in Ghana*. Overseas Development Institute, London.

Ariss, R. T. (2010). "On the implications of market power in banking: Evidence from developing countries". *Journal of Banking & Finance, 34*(4), 765-775.

Aryeetey, E., Kanbur, S. M. R. (2006). *Ghana's economy at half century : an overview of stability, growth and poverty*. Ithaca, N.Y.: Cornell University.

Baltagi, B. (2008). *Econometric analysis of panel data*: John Wiley & Sons.

Bank of Ghana, Annual Reports, Banking Laws and Press Releases (various issues)

Banker, R. D., Charnes, A., Cooper, W. W. (1984). "Some models for estimating technical and scale inefficiencies in data envelopment analysis". *Management Science, 30*(9), 1078-1092.

Barros, C. P., Wanke, P. (2014). "Banking efficiency in Brazil". *Journal of International Financial Markets Institutions & Money, 28*(1), 54-65.

Barros, C. P., Williams, J. (2013). "The random parameters stochastic frontier cost function and the effectiveness of public policy: Evidence from bank restructuring in Mexico". *International Review of Financial Analysis, 30*, 98-108.

Barth, Caprio, G., Levine, R. (2012). "The evolution and impact of bank regulations".

Barth, J., Caprio, G., Levine, R. (2006). *Rethinking bank regulation: till angels govern*. Cambridge: Cambridge University Press.

Barth, J., Caprio Jr, G., Levine, R. (2013). "Bank Regulation and Supervision in 180 Countries from 1999 to 2011". *Journal of Financial Economic Policy, 5*(2), 111-219.

Barth, J. R., Caprio, G., Jr., Levine, R. (2001). "The regulation and supervision of banks around the world: a new database". *Brookings-Wharton Papers on Financial Services*, 230.

Battese, G., Coelli, T. J. (1995). "A model for technical inefficiency effects in a stochastic frontier production function for panel data". *Empirical Economics, 20*(2), 325-332.

Battese, G. E., Coelli, T. J. (1988). "Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data". *Journal of Econometrics, 38*(3), 387-399.

Battese, G. E., Coelli, T. J. (1992). "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India". *International applications of productivity and efficiency analysis* (pp. 149-165): Springer.

Bauer, P. W., Berger, A. N., Ferrier, G. D., Humphrey, D. B. (1998). "Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods". *Journal of Economics and Business, 50*(2), 85-114.

Bawumia, M. (2010). *Monetary policy and financial sector reform in Africa: Ghana's experience*.

Beck, T. (2008). "Bank Competition and Financial Stability: Friends or Foes?"

Beck, T., Cull, R. (2014). "Banking in Africa (Vol. 1): Oxford University Press".

Beck, T., De Jonghe, O., Schepens, G. (2013). "Bank competition and stability: cross-country heterogeneity". *Journal of financial Intermediation, 22*(2), 218-244.

Beck, T., Demirgüç-Kunt, A., Levine, R. (2006). "Bank concentration, competition, and crises: First results". *Journal of Banking and Finance, 30*(5), 1581-1603.

Beck, T., Demirgüç-Kunt, A., Maksimovic, V. (2004). "Bank Competition and Access to Finance: International Evidence". *Journal of Money, Credit and Banking, 36*(3), 627-648.

Beck, T., Demirgüç-Kunt, A., Pería, M. S. M. (2011). "Bank Financing for SMEs: Evidence Across Countries and Bank Ownership Types". *Journal of Financial Services Research, 39*(1), 35-54.

Beck, T., Hesse, H. (2009). "Why are interest spreads so high in Uganda?" *Journal of Development Economics, 88*(2), 192-204.

Bektas, E. (2007). "The persistence of profits in the Turkish banking system". *Applied Economics Letters, 14*(3), 187-190.

Belotti, F., Daidone, S., Ilardi, G., Ate, V. (2013). "Stochastic frontier analysis using Stata". *Stata Journal, 13*(4), 719-758.

Berg, S. A., Førsund, F. R., Jansen, E. S. (1992). "Malmquist indices of productivity growth during the deregulation of Norwegian banking". *Scandinavian Journal of Economics, 94*, 211-228.

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*(2), 404-431.

Berger, A. N., Demirguc-Kunt, A., Levine, R., Haubrich, J. G. (2004). "Bank concentration and competition: An evolution in the making". *Journal of Money, Credit, and Banking, 36*(3), 433-451.

Berger, A. N., DeYoung, R. (1997). "Problem loans and cost efficiency in commercial banks". *Journal of Banking and Finance, 21*(6), 849-870.

Berger, A. N., Hannan, T. H. (1998). "The Efficiency Cost of Market Power in the Banking Industry: A Test of the "Quiet Life" and Related Hypotheses". *The Review of Economics and Statistics, 80*(3), 454-465.

Berger, A. N., Humphrey, D. B. (1997). "Efficiency of financial institutions: International survey and directions for future research". *European Journal of Operational Research, 98*(2), 175-212.

Berger, A. N., Mester, L. J. (1997). "Inside the black box: What explains differences in the efficiencies of financial institutions?" *Journal of Banking and Finance, 21*(7), 895-947.

Berger, A. N., Mester, L. J. (2003). "Explaining the dramatic changes in performance of US banks: technological change, deregulation, and dynamic changes in competition". *Journal of financial Intermediation, 12*(1), 57-95.

Biekpe, N. (2011). "The Competitiveness of Commercial Banks in Ghana". *African Development Review, 23*(1), 75-87.

Bikker, J. A., Shaffer, S., Spierdijk, L. (2012). "Assessing competition with the Panzar-Rosse model: The role of scale, costs, and equilibrium". *Review of economics and statistics, 94*(4), 1025-1044.

Boone, J. (2008a). "Competition: Theoretical Parameterizations and Empirical Measures". *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft, 164*(4), 587-611.

Boone, J. (2008b). "A New Way to Measure Competition". *The Economic Journal, 118*(531), 1245-1261.

Boot, A. W., Thakor, A. V. (1993). "Self-interested bank regulation". *The American Economic Review*, 206-212.

Boyd, J. H., Nicoló, G. D. (2005). "The Theory of Bank Risk Taking and Competition Revisited". *The Journal of Finance, 60*(3), 1329-1343.

Bresnahan, T. F. (1982). "The oligopoly solution concept is identified". *ECONOMICS LETTERS, 10*(1-2), 87-92.

Brissimis, S., Iosifidi, M., Delis, M. D. (2014). "Bank market power and monetary policy transmission".

Brissimis, S. N., Delis, M. D., Papanikolaou, N. I. (2008). "Exploring the nexus between banking sector reform and performance: Evidence from newly acceded EU countries". *Journal of Banking and Finance, 32*(12), 2674-2683.

Brooks, C. (2014). *Introductory econometrics for finance*: Cambridge university press.

Brownbridge, M., Gockel, A. F. (1996). *The impact of financial sector policies on banking in Ghana*: Institute of Development Studies.

Bruhn, M., Farazi, S., Kanz, M. (2013). "Bank Competition, Concentration, and Credit Reporting".

Buchs, T. D., Mathisen, J. (2005). *Competition and efficiency in banking : behavioral evidence from Ghana*. Washington, D.C.: International Monetary Fund, African Dept.

Cameron, A. C., Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*: Cambridge University Press.

Carletti, E., Hartmann, P. (2003). "Competition and Financial Stability: What's Special about Banking?", In Monetary History, Exchange Rates and Financial Markets: Essays in Honour of Charles Goodhart: Cheltenham, UK: Edward Elgar.

Casu, B., Girardone, C. (2006). "Bank competition, concentration and efficiency in the single european market". *The Manchester School, 74*(4), 441-468.

Casu, B., Girardone, C., Molyneux, P. (2015). *Introduction to Banking 2nd edition* (GB): Pearson Education.

Caudill, S. B., Ford, J. M. (1993). "Biases in frontier estimation due to heteroscedasticity". *Econometric Letters, 41*(1), 17-20.

Caudill, S. B., Ford, J. M., Gropper, D. M. (1995). "Frontier estimation and firm-specific inefficiency measures in the presence of heteroscedasticity". *Journal of Business & Economic Statistics, 13*(1), 105-111.

CEPA. (2003). *The poverty reduction and growth facility (PRGF) arrangement : May 1999-November 2002*. Accra: Centre for Policy Analysis.

Cetorelli, N. (2001). "Competition among banks: Good or bad?" *Economic Perspectives, 25*(2), 38.

Cetorelli, N. (2002). "Entry and competition in highly concentrated banking markets". *Economic Perspectives, 26*(4), 18.

Charnes, A., Cooper, W. W., Rhodes, E. (1978). "A Data Envelopment Analysis Approach to Evaluation of the Program Follow through Experiment in US Public School Education: DTIC Document".

Chen, X., Skully, M., Brown, K. (2005). "Banking efficiency in China: Application of DEA to pre- and post-deregulation eras: 1993–2000." *China Economic Review, 16*(3), 229-245.

Claessens, S. (2009). *Competition in the Financial Sector : Overview of Competition Policies*. Washington: International Monetary Fund.

Claessens, S., Demirgüç-Kunt, A., Huizinga, H. (2001). "How does foreign entry affect domestic banking markets?" *Journal of Banking and Finance, 25*(5), 891-911.

Claessens, S., Laeven, L. (2004). "What Drives Bank Competition? Some International Evidence". *Journal of Money, Credit and Banking, 36*(3), 563-583.

Claessens, S., Van Horen, N. (2012). "Being a foreigner among domestic banks: Asset or liability?" *Journal of Banking & Finance, 36*(5), 1276-1290.

Clark, J. A., Siems, T. F. (2002). "X-efficiency in banking: Looking beyond the balance sheet". *Journal of Money, Credit and Banking*, 987-1013.

Clarke, G. R. G., Cull, R., Shirley, M. M. (2005). "Bank privatization in developing countries: A summary of lessons and findings". *Journal of Banking and Finance, 29*(8), 1905-1930.

Clerides, S., Delis, M. D., Kokas, S. (2015). "A new data set on competition in national banking markets". *Financial Markets, Institutions & Instruments, 24*(2-3), 267-311.

Coelli, T., Coelli, T. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). New York: Springer.

Coelli, T. J. (1995). "Recent developments in frontier modelling and efficiency measurement." *Australian Journal of Agricultural and Resource Economics, 39*(3), 219-245.

Cornwell, C., Schmidt, P., Sickles, R. C. (1990). "Production frontiers with cross-sectional and time-series variation in efficiency levels". *Journal of Econometrics, 46*(1-2), 185-200.

Delis, M. D. (2010). "Competitive conditions in the Central and Eastern European banking systems". *Omega, 38*(5), 268-274.

Delis, M. D. (2012). "Bank competition, financial reform, and institutions: The importance of being developed". *Journal of Development Economics, 97*(2), 450-465.

Demirgüç-Kunt, A., Huizinga, H., Claessens, S. (1998). "How does foreign entry affect the domestic banking market?"

Demirgüç-Kunt, A., Martínez Pería, M. S. (2010). "A framework for analyzing competition in the banking sector: an application to the case of Jordan".

Demsetz, H. (1973). "Industry Structure, Market Rivalry, and Public Policy". *Journal of Law and Economics, 16*(1), 1-9.

Dick, A. A., Hannan, T. H. (2010). "Competition and antitrust policy in banking".

Domanski, D. (2005)."Foreign banks in emerging market economies: changing players, changing issues".

Dong, Y. (2010). *Cost efficiency in the Chinese banking sector: A comparison of parametric and non-parametric methodologies.* © Yizhe Dong.

Dougherty, C. (2011). *Introduction to econometrics*: Oxford University Press.

Elyasiani, E., Mehdian, S. (1995). "The comparative efficiency performance of small and large US commercial banks in the pre- and post-deregulation eras". *Applied Economics, 27*(11), 1069.

Fang, Y. W., Hasan, I., Marton, K. (2011). "Bank efficiency in South-Eastern Europe The role of ownership, market power and institutional development". *Economics of Transition, 19*(3), 495-520.

Ferrier, G. D., Lovell, C. K. (1990). "Measuring cost efficiency in banking: econometric and linear programming evidence". *Journal of Econometrics, 46*(1-2), 229-245.

Fischer, K.-H. (2000). "Acquisition of information in loan markets and bank market power-An empirical investigation".

Flamini, V., Schumacher, M. L., McDonald, M. C. A. (2009). *The determinants of commercial bank profitability in Sub-Saharan Africa*: International Monetary Fund.

Førsund, F. R., Lovell, C. K., Schmidt, P. (1980). "A survey of frontier production functions and of their relationship to efficiency measurement". *Journal of Econometrics, 13*(1), 5-25.

Fries, S., Taci, A. (2005). "Cost efficiency of banks in transition: Evidence from 289 banks in 15 post-communist countries". *Journal of Banking and Finance, 29*(1), 55-81.

Genay, H., Udell, G. F., DeYoung, R., Berger, A. N. (2000). "Globalization of Financial Institutions: Evidence from Cross-Border Banking Performance". *Brookings-Wharton Papers on Financial Services, 2000*(1), 23-120.

Gerschenkron, A. (1962). *Economic backwardness in historical perspective: a book of essays*: Belknap Press of Harvard University Press Cambridge, MA.

Girardone, C., Molyneux, P., Gardener, E. P. M. (2004). "Analysing the determinants of bank efficiency: the case of Italian banks". *Applied Economics, 36*(3), 215-227.

Goddard, J., Molyneux, P., Wilson, J. O. (2004). "The profitability of European banks: a cross-sectional and dynamic panel analysis". *The Manchester School, 72*(3), 363-381.

Goddard, J., Wilson, J. O. S. (2009). "Competition in banking: A disequilibrium approach". *Journal of Banking and Finance, 33*(12), 2282-2292.

Hadri, K. (1999). "Estimation of a doubly heteroscedastic stochastic frontier cost function". *Journal of Business & Economic Statistics, 17*(3), 359-363.

Hauner, D., Peiris, S. J. (2008). "Banking efficiency and competition in low income countries: the case of Uganda". *Applied Economics, 40*(21), 2703-2720.

Hawkins, J., Mihaljek, D. (2001). "The banking industry in the emerging market economies: competition, consolidation and systemic stability: an overview". *BIS papers, 4*, 1-44.

Hellmann, T. F., Murdock, K. C., Stiglitz, J. E. (2000). "Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?" *The American Economic Review, 90*(1), 147.

Huang, C. J., Liu, J.-T. (1994). "Estimation of a Non-Neutral Stochastic Frontier Production Function". *Journal of Productivity Analysis, 5*(2), 171-180.

Hughes, J. P., Lang, W., Mester, L. J., Moon, C.-G. (1996). "Efficient Banking under Interstate Branching". *Journal of Money, Credit and Banking, 28*(4), 1045-1071.

Humphrey, D. B., Pulley, L. B. (1997). "Banks' Responses to Deregulation: Profits, Technology, and Efficiency". *Journal of Money, Credit and Banking, 29*(1), 73-93.

IMF (2011). "Ghana : 2011 Article IV Consultation" from http://site.ebrary.com/id/10557058

IMF. (2013). *Ghana : 2013 Article IV Consultation*. Washington: International Monetary Fund.

Isshaq, Z., Bokpin, G. A. (2012). "Expansion and Efficiency in Banking: Evidence from Ghana". *Managerial and Decision Economics, 33*(1), 19-28.

Iwata, G. (1974). "Measurement of conjectural variations in oligopoly". *Econometrica: Journal of the Econometric Society*, 947-966.

Jondrow, J., Lovell, C. K., Materov, I. S., Schmidt, P. (1982). "On the estimation of technical inefficiency in the stochastic frontier production function model". *Journal of Econometrics, 19*(2-3), 233-238.

Kaplan, M., Çelik, T. (2008). "The impact of tourism on economic performance: the case of Turkey". *The International Journal of Applied Economics and Finance, 2*(1), 13-18.

Kasekende, L., Mlambo, K., Murinde, V., Zhao, T. (2009). "Restructuring for competitiveness: The financial services sector in Africa's four largest economies". *Word Economic Reform: Africa Competitiveness Report, 209*, 49-81.

Keeley, M. C. (1990). "Deposit Insurance, Risk, and Market Power in Banking". *The American Economic Review, 80*(5), 1183-1200.

Koetter, M., Meesters, A. (2013). "Effects of specification choices on efficiency in DEA and SFA". *Efficiency and Productivity Growth: Modelling in the Financial Services Industry*, 215-236.

Kovanen, A. (2011). "Monetary Policy Transmission in Ghana: Does the Interest Rate Channel Work?" IMF Working Paper

Kroszner, R. S., Strahan, P. E. (1999). "What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions". *The Quarterly Journal of Economics, 114*(4), 1437-1467.

Kumbhakar, S. C. (1990). "Production frontiers, panel data, and time-varying technical inefficiency". *Journal of Econometrics, 46*(1-2), 201-211.

Kumbhakar, S. C., Ghosh, S., McGuckin, J. T. (1991). "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms". *Journal of Business & Economic Statistics, 9*(3), 279-286.

Kumbhakar, S. C., Lovell, C. K. (2003). *Stochastic frontier analysis*: Cambridge university press.

Kumbhakar, S. C., Lozano-Vivas, A., Lovell, C. A. K., Hasan, I. (2001). "The Effects of Deregulation on the Performance of Financial Institutions: The Case of Spanish Savings Banks". *Journal of Money, Credit and Banking, 33*(1), 101-120.

Kumbhakar, S. C., Wang, H.-J. (2005). "Estimation of growth convergence using a stochastic production frontier approach". *Economics Letters, 88*(3), 300-305.

Kumbhakar, S. C., Wang, H., Horncastle, A. P. (2015). *A practitioner's guide to stochastic frontier analysis using Stata*: Cambridge University Press.

Lau, L. J. (1982). "On identifying the degree of competitiveness from industry price and output data". *Economics Letters, 10*(1-2), 93-99.

Lee, Y. H., Schmidt, P. (1993). "A production frontier model with flexible temporal variation in technical efficiency". *The measurement of productive efficiency: Techniques and applications*, 237-255.

Leon, F. (2015). "Measuring competition in banking: A critical review of methods".

Levine, R., Barth, J. R., Caprio, G. (2001). "The Regulation and Supervision of Banks around the World: A New Database". *Brookings-Wharton Papers on Financial Services, 2001*(1), 183-240.

Liu, H., Molyneux, P., Wilson, J. (2013). "Competition in banking: Measurement and interpretation". *Handbook of Research Methods and Applications in Empirical Finance*, 197.

Love, I., Peria, M. (2012). "How bank competition affect firms' access to finance". The World Bank. *Development Research Group, Policy Research Paper (6163)*.

Mahesh, H. P., Bhide, S. (2008). "Do Financial Sector Reforms Make Commercial Banks More Efficient?" A Parametric Exploration of the Indian Case. *Margin: The Journal of Applied Economic Research, 2*(4), 415-441.

Maudos, J., Pastor, J. M., Serrano, L. (2000). "Convergence in OECD countries: technical change, efficiency and productivity". *Applied Economics, 32*(6), 757-765.

Moyo, J., Nandwa, B., Council, D. E., Oduor, J., Simpasa, A. (2014). "Financial sector reforms, competition and banking system stability in Sub-Saharan Africa". *New Perspectives*.

Mwega, F. (2011). "The Competitiveness and Efficiency of the Financial Services Sector in Africa: A Case Study of Kenya". *African Development Review, 23*(1), 44-59.

Mwenda, A., Mutoti, N. (2011). "Financial Sector Reforms, Bank Performance and Economic Growth: Evidence from Zambia". *African Development Review, 23*(1), 60-74.

Nissanke, M., Aryeetey, E. (1998). *Financial integration and development : liberalization and reform in Sub-Saharan Africa*. London ; New York: Routledge.

Panzar, J. C., Rosse, J. N. (1987). "Testing for "Monopoly" Equilibrium". *The Journal of Industrial Economics, 35*(4), 443.

Pasadilla, G., Milo, M. (2005). *Effect of liberalization on banking competition*: Philippine Institute for Development Studies.

Petersen, M. A., Rajan, R. G. (1995). "The effect of credit market competition on lending relationships". *The Quarterly Journal of Economics, 110*(2), 407-443.

Porta, R. L., Lopez-de-Silanes, F., Shleifer, A. (2002). "Government Ownership of Banks". *The Journal of Finance, 57*(1), 265-301.

Porter, N., Takáts, E., Feyzioglu, T. (2009). "Interest rate liberalization in China".

Poshakwale, S. S., Qian, B. (2011). "Competitiveness and Efficiency of the Banking Sector and Economic Growth in Egypt: Competitiveness of the Banking Sector in Egypt". *African Development Review, 23*(1), 99-120.

Schaeck, K., Cihák, M. (2014). "Competition, Efficiency, and Stability in Banking". *Financial Management, 43*(1), 215-241.

Schaeck, K., Čihák, M. (2008). "How does competition affect efficiency and soundness in banking?" New empirical evidence.

Schaeck, K., Cihak, M., Wolfe, S. (2009). "Are Competitive Banking Systems More Stable?" *Journal of Money, Credit and Banking, 41*(4), 711-734.

Schmidt, P., Sickles, R. C. (1984). "Production Frontiers and Panel Data". *Journal of Business & Economic Statistics, 2*(4), 367-374.

Sealey, C. W. J., Lindley, J. T. (1977). "Inputs, outputs and a theory of production and cost at depository financial institutions". *The Journal of Finance, 32*(4), 1251.

Semih Yildirim, H., Philippatos, G. C. (2007). "Efficiency of Banks: Recent Evidence from the Transition Economies of Europe, 1993-2000". *The European Journal of Finance, 13*(2), 123-143.

Shirley, M. M., Walsh, P. P. (2000). *Public vs. private ownership: the current state of the debate*: World Bank, Development Research Group, Regulation and Competition Policy.

Simpasa, A. M. (2013). "Increased foreign bank presence, privatisation and competition in the Zambian banking sector". *Managerial Finance, 39*(8), 787-808.

Sowa, N. K. (2002). *An assessment of poverty reducing policies and programmes in Ghana*: Centre for Policy Analysis.

Stevenson, R. E. (1980). "Likelihood functions for generalized stochastic frontier estimation". *Journal of Econometrics, 13*(1), 57-66.

Sturm, J.-E., Williams, B. (2004). "Foreign bank entry, deregulation and bank efficiency: Lessons from the Australian experience". *Journal of Banking and Finance, 28*(7), 1775-1799.

Uchida, H., Tsutsui, Y. (2005). "Has competition in the Japanese banking sector improved?" *Journal of Banking and Finance, 29*(2), 419-439.

van Leuvensteijn, M., Bikker, J. A., van Rixtel, A. A. R. J. M., Sørensen, C. K. (2011). "A new approach to measuring competition in the loan markets of the euro area". *Applied Economics, 43*(23), 3155-3167.

van Leuvensteijn, M., Sørensen, C. K., Bikker, J. A., van Rixtel, A. A. R. J. M. (2013). "Impact of bank competition on the interest rate pass-through in the euro area". *Applied Economics, 45*(11), 1359-1380.

Wang, H.-J., Schmidt, P. (2002). "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels". *Journal of Productivity Analysis, 18*(2), 129-144.

Weill, L. (2013). "Bank competition in the EU: How has it evolved?" *Journal of International Financial Markets, Institutions and Money, 26*, 100-112.

Xu, B., Rixtel, A. V., Leuvensteijn, M. v. (2014). "Measuring bank competition in China: A comparison of new versus conventional approaches applied to loan markets".

Yuan, Y., Phillips, R. D. (2008). "Financial integration and scope efficiency in US financial services post Gramm-Leach-Bliley".

Zhao, T., Casu, B., Ferrari, A. (2010). "The impact of regulatory reforms on cost structure, ownership and competition in Indian banking". *Journal of Banking and Finance, 34*(1), 246-254.

Zhao, T., Murinde, V. (2011). "Bank Deregulation and Performance in Nigeria". *African Development Review, 23*(1), 30-43.