DOCTOR OF PHILOSOPHY

The economics of GCC banking efficiency

Shams, Khalid

Award date:
2003

Awarding institution:
Bangor University

The Economics of GCC Banking Efficiency

A Thesis Submitted to the University of Wales
in Fulfilment for the Requirements
for the Degree of Doctor of Philosophy

By

Khalid Shams

School for Business and Regional Development
University of Wales, Bangor

- February 2003 -

Acknowledgements

This thesis owes its existence, after almighty ALLAH, to the help, support, and inspiration of many people. In the first place, I would like to express my sincere appreciation and gratitude to Prof. Phillip Molyneux for providing me the support and encouragements to accomplish this work. I thank him for his many constructive comments and discussions at any time of the day, and for always believing in my research and providing me with confidence to continue during the most difficult times. I ask ALLAH to give him the best in this worldly life and in the hereafter.

I am also proud to thank Dr. Yener Altunbas and Dr. Jonathan Williams for constructive discussions and insights on theoretical and methodological issues. I also extend my appreciation to all staff members and fellow students for their assistance, support, and ecouragment.

I am also indebted to Dr. Khaled Hussein (of Islamic Development Bank) and all the staff members in the University of Kent's Department of Economics where I started this research.

I am very grateful to Mr. Ebrahim Abdul-Hadi of the UAE Central Bank for his admirable service to provding me copies of Financial Position of Commerical Banks in the UAE. I am also thankful to my cousine Mr. Saeed Al-Banna for his contiuous assistance.

I would like also to thank Shikh Abdulla Bin Saud Al-Thani, the former deputy governer of Qatar Central Bank, Dr. Abudulla Al-Kholaifi, the president of the University of Qatar, Dr. Saif Al-Sowaidi, Dr. Nasser Abdul-Ghani, and Mr. Basheer Al-Kahloot for their ecouragment and support.

I owe special gratitude to my wife Ebtesam Al-Mulla for her continuous patienece, support, and prayers.

Finally, I am delighted to acknowledge all that I owe to my parents, and the best I can say to them is "My Lord! Bestow on them Your Mercy as they did bring me up when I was small" (The Nobel Quraan, Al-Isra, Verse 24).

ABSTRACT

This thesis analyses the cost and profit efficiencies of the GCC banking sector over the period 1995-2000. Efficiencies are estimated using the most recent frontier technique, the Fourier Flexible form. The thesis also uses a logistic regression model to estimate the determinants of GCC banking efficiency. The findings show that the level of inefficiencies in the GCC banking industry ranges between 8 and 10% for costs, and 30 and 32 % for profits. There are no major differences in banks inefficiency levels among GCC countries. Moreover, inefficiencies show almost stable trends over 1995-2000. Comparisons of inefficiency levels across bank ownership type and assets size reveal that national banks are more cost efficient but less profit efficient than foreign banks. In terms of bank size, large banks are found to be more cost efficient but less profit efficient than other sized banks. The results also indicate that foreign banks have on average been operating with higher scale diseconomies than national banks. Moreover, scale diseconomies decline as the assets sizes of both national and foreign banks increase. The main results from our logistic regression are that the strengthening of financial capital is a central element explaining bank efficiency in the GCC region; however, the erosion in loan quality reduces banking sector efficiency. The main policy conclusion from this thesis is that GCC governments need to continue to implement financial reform packages that strengthen banking system soundness, foster banking competition, and also devise incentive schemes to improve managerial efficiency in order that GCC banks are better placed to meet the challenges of greater openness.

TABLE OF CONTENTS

4 FINANCIAL SYSTEM EFFICIENCY: AN OVERVIEW

5 MEASURING BANKING SECTOR EFFICIENCY: THEORY AND EMPIRICAL EVIDENCE

6 METHODOLOGY AND DATA

7 EFFICIENCY IN GCC BANKING: EMPIRICAL EVIDENCE

LIST OF TABLES

2 GCC ECONOMIES AND BANKING SYSTEMS DEVELOPMENT

6 METHODOLOGY AND DATA

7 EFFICIENCY IN GCC BANKING: EMPIRICAL EVIDENCE

LIST OF FIGURES

2 GCC ECONOMIES AND BANKING SYSTEMS DEVELOPMENT

4 FINANCIAL SYSTEM EFFICIENCY: AN OVERVIEW

5 MEASURING BANKING SECTOR EFFICIENCY: THEORY AND EMPIRICAL EVIDENCE

LIST OF ACRONYMS

ATM	-	Automated Teller Machine
BCCI	-	Bank of Credit and Commerce International
BMA	-	Bahrain Monetary Currency
CBO	-	Central Bank of Oman
CN	-	Cost inefficiency half-normal
DEA	-	Data Envelopment Analysis
DFA	-	Distribution Free Approach
E	-	Total equity
FCA	-	Functional Cost Analysis
FDH	-	Free Disposal Hull
FDI	-	Foreign direct investment
FIX	-	Fixed assets
FOREIGN	-	Foreign banks – dummy variable
GATT	-	General Agreement on Tariffs and Trade
GCC	-	Gulf Cooperation Council
GDP	-	Gross Domestic Products
IMF	-	International Monetary Fund
IT	-	Information Technology
KD	-	Kuwaiti Dinar
LM	-	Lagrange Multiplier
LTA	-	Ratio of loans to total assets
M & A	-	Merger and Acquisition
MLE	-	Maximum likelihood Estimation
OBU	-	Offshore Banking Units
OLS	-	Ordinary Least Squares
P1	-	Price of deposits
P2	-	Price of labour
G1	-	Price of loans
G2	-	Price of other earning assets
PROF	-	Profits

PROV	-	Total provisions
Q1	-	Total loans
Q2	-	Other earning assets
QCB	-	Qatar Central Bank
QMA	-	Qatar Monetary Agency
QR	-	Qatari Riyal
R	-	Spearman Rank Correlation Coefficient
Repos	-	Repurchase Agreements
RO	-	Riyal Omani
ROA	-	Return on Assets
ROE	-	Return on Equity
S & L	-	Savings and Loans
SAMA	-	Saudi Arabia Monetary Agency
SDRs	-	Special Drawing Rights
SFA	-	Stochastic Frontier Approach
SN	-	Standard profit inefficiency half-normal
SR	-	Saudi Riyal
T	-	Time trend
TA	-	Total assets
TBGDP	-	Total bank assets as a ratio to GDP
TC	-	Total cost
TCTA	-	total cost to total assets
TFA	-	Thick Frontier Approach
UAE	-	United Arab Emirates
US	-	United States (of America)
USD	-	United States Dollar currency

CHAPTER 1

BACKGROUND, AIMS, METHODOLY, AND THE STRUCTURE PLAN

1.1 Introduction

This thesis examines the efficiency of GCC banking system between 1995-2000.[1] Given the ongoing deregulation process, it is important to have an indication of the efficiency features of GCC banks in order to evaluate the influence of financial reforms that aim to improve the soundness and enhance competitiveness of the GCC financial systems overall.

The aim of this chapter is to provide the background to the thesis. It includes the aims of the study, its importance, and the methodology followed. The chapter also provides the structure and the plan of the study.

1.2 Background and aims of the study

Over the last decade, GCC countries' banking systems have experienced many regulatory changes. The most important of these has been the gradual removal of interest rate ceilings on loans and deposits, which commenced from the mid 1990s onwards. The aim of these regulatory changes was to bring about a more competitive environment and to foster improved efficiency in the banking system. GCC banking systems will also be exposed to even more competition by the time they become more integrated within the recently announced GCC economic and monetary union or when

[1] GCC refers to Gulf Cooperation Council consisting of the Kingdom of Saudi Arabia, the United Arab of Emirates (the UAE), the Kingdom of Bahrain, the Sultanate of Oman, the State of Qatar, and the State of Kuwait.

the GATT's agreement (which all GCC countries have joined except Saudi Arabia) will come into effect. In essence, GCC banking industries are expected to be more prepared for regional and international competition since each GCC country will have to open its doors to foreign banks. As competition increases, the issue of cost efficiency becomes an interesting area for investigation since improvements in cost efficiency are expected to be brought about by the more competitive environment. In addition, analysing the profit efficiency of Gulf banks is also important as this helps to inform us about the ability of banks to maximise revenues through various strategies relating to their production and diversification features.

This thesis aims to examine cost and profit inefficiencies in GCC banking sectors, where these inefficiencies are depicted as the deviation of actual cost and profit from the optimal banking industry's cost and profit functions. This deviation is known as X-inefficiency, an important feature of operational inefficiency. The measurement of this deviation enables us to know the status of GCC banking efficiency and how it is compared to banking sector efficiency in other countries. The interest in measuring X-inefficiency in banking has increased over the last decade as commentators have sort to examine the impact of increased competition on banking sector costs. While an extensive literature has developed to examine banking sector efficiency in the US and Europe (see Berger and Humphrey, 1997; Goddard et al., 2001), there is only a limited literature on developing countries (Bhattacharyya, Lovell, and Sahay, 1997; Isik and Hassan, 2002; and Al-Jarrah, 2002).

The second aim of this thesis is to compare the efficiency of both national and foreign banks. Generally, most bank efficiency studies to date have been undertaken on national banks although a handful of studies do make comparison between the efficiency of national and foreign banks in countries (including Srivastava, 1999; Intarachote, 2000; Isik and Hassan, 2002). The interest in comparing the efficiency of national and foreign bank comes from the fact that (particularly in developing countries) it is generally expected that foreign banks are more cost and profit efficient because they are driven into the local markets by various comparative advantages such as experience, a higher quality of management, and better access to international financial markets. Comparisons between the efficiency of national and foreign banks therefore can tell us whether these advantages are exploited or not. National banks, for instance, might have certain advantages over foreign banks (they may be favoured in financing government's projects) that may preserve their competitive status. These factors may affect the efficiency of foreign banks and distort the ability of national banks to compete within an environment of greater financial system openness.

In addition to these two broad aims, this thesis also takes into consideration the influence both risk and asset quality factors have on the levels of measured inefficiency in GCC banking markets. Generally, there is evidence that both risk and asset quality factors can influence both cost and profit efficiencies (Mester, 1996; Berger and Mester, 1997; Altunbas et al., 2000). These factors are typically closely monitored by regulatory authorities so as to ensure that banks keep adequate levels of capital and have acceptable quality of loan portfolios. The links between efficiency, risk and asset quality may therefore be important from a policy-makers perspective. Especially, for instance if we find that efficient banks have high asset quality and are less risky.

The thesis also investigates the extent to which GCC banks exploit economies of scale in conducting banking operations. Knowledge of optimal bank size provides more information about the competitive status of GCC banking. Finally, we investigate the main determinants of efficiency in Gulf banking.

1.3 Data and methodology

The empirical part of this study is based on a sample of 93 banks operating in the GCC banking markets over the period 1995-2000. The data used in this sample are obtained from the London based IBCA bank credit rating agency's Bankscope data (January, 2002), as well as audited annual reports obtained from the UAE and Qatar foreign banks. The definition of bank inputs and outputs are based on the intermediation approach, in which banks are considered as financial firms using inputs such as deposits, labour, and fixed assets to produce a range of outputs, such as loans and securities' investments. This approach also suggests the consideration of interest on deposits as a component of bank total costs. Moreover, our analysis considers the nature of our data set, which suggests the use of a panel data approach. We use the constructed sample to estimate cost and profit efficiency, where profit efficiency estimation consists of standard profit and alternative profits as proposed in Berger and Mester (1997).

Cost and profit X-efficiencies can be measured using two different methodologies – the parametric, stochastic cost frontier and non-parametric, linear programming approach. This thesis adopts the stochastic frontier model as a parametric approach to measure the efficiency of GCC banking systems. The stochastic frontier model is estimated using the most recent frontier technique, the Fourier Flexible form. In order to obtain robust findings we set two distributional assumptions (half-normal and exponential) on the inefficiency term. Moreover, the thesis also uses the distribution-free approach to see how inefficiency scores obtained under a different approach vary. This is undertaken to investigate the consistency of the main results. Moreover, our model estimating efficiency is based on two main specifications - with and without the consideration of risk and asset quality factors. These specifications are used in order to explore how the inefficiency scores may be influenced when controlling for these aforementioned factors in our model. (For further robustness checks, we utilize some of the Bauer et al. (1998) consistency tests and also employ rank correlation tests to investigate the reliability of the obtained efficiency measures).

Estimates of economies of scale are obtained using the preferred Fourier Flexible specification, which includes risk and quality factors. Moreover, scale efficiencies are derived using the methodology proposed in Evanoff and Israilevich (1995).

The last task in the empirical work investigates the determinants of inefficiency. Using the estimates of both cost and profits inefficiencies as regressands, we estimate a logistic function model to find how inefficiency measures are correlated with a range of factors that reflect various banking business and market specifics such as bank size, risk preferences, loan quality, ownership, and market size.

1.4 The thesis structure plan

The thesis is divided into eight chapters as follows:

Chapter 1 Introduction

Chapter 2 GCC Economies and Banking Systems Development

This chapter describes the overall perspective of the GCC economies. It outlines their main economy characteristics, developments, and challenges.

The chapter also focuses on the GCC financial sectors' structure and development. It highlights the main characteristics and then provides a review of recent changes in banking sector structure and regulation in each country. Moreover, the chapter provides an overview of GCC banking systems' soundness and performance and highlights the progress undertaken by GCC countries to form monetary and economic union.

Chapter 3 The Role of Foreign Banks – A Literature Review

This chapter provides an overview of the reasons why foreign banks exist and the role they play in financial system stability and economic growth. The chapter provides an extensive review on studies on foreign bank performance and efficiency and shows how the role of foreign banks compares to those of national banks. This is undertaken so as to provide an insight into the main issues surrounding the behaviour of foreign banks. This helps to inform our empirical analysis on the efficiency of GCC banks that is conducted later in this thesis.

Chapter 4 The Efficiency of the Financial System - An Overview

Chapter 4 examines the efficiency of financial systems and its importance in the economic growth process. The chapter defines financial system efficiency and classifies efficiency of the financial system into four categories: operational, informational, regulatory, and structural efficiency. The chapter also provides an analysis of how the existence of efficiency in the financial system can lead to greater stability and growth in the economy.

Based on the efficiency framework described in this chapter, the thesis will mainly focus on the operational dimension of financial system efficiency, which particularly covers the issue of bank efficiency.

Chapter 5 Measuring Banking Sector Efficiency – Theory and Empirical Evidence

This chapter provides a survey on bank efficiency studies. This includes the study of the nature of banking firms and how their inputs and outputs are treated. The chapter also explores both the theoretical and empirical literature concerning the main aspects of operational efficiency that are extensively studied in the banking literature; these aspects relate to economies of scale, economies of scope, and X-efficiency. The chapter continues by reviewing the efficiency measurement techniques and discusses the functional forms used to estimate banking efficiency.

Chapter 6 Methodology and Data

This chapter explains the econometric approach which will principally be used to estimate banking X-inefficiency in GCC banking markets: cost, standard profit, and alternative profit inefficiencies. The chapter describes the functional forms, the data, and the variables used in the efficiency estimation. The chapter also illustrates the methods used to measure both scale economies and scale inefficiency, which are features mainly related to the cost characteristics of banks. To explore the possible variables determining inefficiency in the GCC banking industry, the chapter outlines a logistic regression model approach that will be used to evaluate efficiency determinants.

Chapter 7 Empirical Results - Efficiency in GCC Banking Systems

This chapter empirically measures the efficiency of banks operating in the GCC banking sectors over the period 1995-2000. It presents and compares the results using three efficiency concepts: cost efficiency, standard profit efficiency, and alternative profit efficiency. In order to arrive at more reliable efficiency measures, two main specifications are examined: the traditional and the preferred specification that controls for risk and quality factors. These are based on the parametric approach using the Fourier Flexible form model. As this form is estimated using the half-normal and exponential distribution that are set on the inefficiency term, the chapter also estimates the model using the distribution-free approach for robustness purposes. In addition, the inefficiency analysis is extended to evaluate both scale economies and scale inefficiencies. In order to investigate the determinants of inefficiency in GCC banking systems, we use a logistic model approach to see if banking sector inefficiency is related to various bank-specific and market-specific features.

The main results found are as follows. The mean cost inefficiency from the preferred model is about 9 per cent (from the half-normal and distribution-free estimates). In both standard and alternative profit functions, the inefficiency results indicate that nearly one third of the profits that could be earned by the best-practice bank are lost to inefficiency. This evidence shows that there are higher levels of profit inefficiency than of cost inefficiency and this result supports the importance of inefficiencies on the revenue side, either due to the wrong choice of output or to the mispricing of output.

In order to arrive at a consistent result, we compare efficiency estimates derived using the distributional-free approach with our cost inefficiency measures. As they tend to be almost identical, this suggests that the half-normal cost inefficiency measures are more likely to concur with 'actual' cost inefficiencies in the system.

The findings also show that the risk and quality factors provide information influencing bank inefficiency levels when we use either the cost or profit function models. When risk and quality factors are considered, the mean inefficiency measures show a slight decrease.

The rank-order coefficients show a close association between inefficiency estimated across different specifications. The association is also close for inefficiency estimates using different distributional assumptions. Rank correlation is also undertaken to investigate the relationship between estimated inefficiency results and financial ratios. The relationship between inefficiency and profitability is almost consistent. That is, profit inefficiency and profitability are negatively related. In the same way, the relationship between costs and inefficiency measures is consistent because the positive rank correlation suggests that the more cost- and profit-inefficient banks incur higher costs.

Foreign banks are found to be less cost efficient, but more profit efficient than national banks. This suggests that foreign banks focus more on revenue generating than do their national counterparts. As foreign banks tend to have a different business mix (high end retail clients, large corporate banking services, and so on), it is perhaps not surprising that they are found to be less cost efficient but more profit efficient. Moreover, large banks are more cost efficient than medium and small banks, and small banks are more profit efficient than medium and large banks.

The sample shows that the GCC banking industry has been exhibiting scale diseconomies driven mostly by banks that belong to the GCC countries' exhibiting scale diseconomies (namely banks in the UAE, Oman, and Qatar). Scale economies of large GCC banks are much closer to unity than are those of small and medium banks. Moreover, small banks show more scale diseconomies than do medium-size banks. Scale economies have also been calculated for the foreign banks operating in the GCC countries. The results show that foreign banks have been operating with, on average, higher scale diseconomies than national banks over the six-years study period.

In comparing between X-inefficiency and scale inefficiency, X-inefficiencies are consistently larger than scale inefficiencies during the study period. This result also suggests that banks need to improve their managerial practices as a priority in order to increase the efficiency of their performance.

When looking at scale inefficiency in terms of the size of banks, the results indicate that large banks are the most scale efficient in the GCC banking industry. The results also show that medium-size banks are more scale efficient than their smaller counterparts. With regard to scale inefficiency comparisons between foreign and national banks, GCC national banks are found to be much more scale efficient than foreign banks.

Lastly, in the logistic regression, cost and profit inefficiencies are found to be negatively related with the risk variable. There is also evidence that inefficiency is. positively related to loan quality variables, suggesting that banks with enhanced financial capital and high loan quality are more efficient.

Chapter 8: Conclusions

This chapter provides a summary of the main findings, policy implications, and recommendations. It also provides limitations of this research and proposes areas for future research.

CHAPTER 2

GCC ECONOMIES AND BANKING SYSTEMS DEVELOPMENT

2.1 Introduction

The first part of this chapter outlines the main characteristics of GCC economies, and the second part provides an overview of their financial systems. Section 2.2 presents a background to the GCC countries' economies. This includes their history, the size of their economies (in terms of GDP), their demography, and various indicators relating to recent economic performance. Section 2.3 provides an overview of the GCC financial systems. This covers the development of individual GCC countries' banking systems, evaluates the performance of the GCC banks, and describes recent moves to create a GCC economic and financial union. Section 2.4 contains our conclusions.

2.2 Background to the GCC countries' economies

The GCC, i.e. the Gulf Cooperation Council, was founded in 1981 with the aim of coordinating policies in various political, economic, and social matters across the Gulf region.[1] The GCC countries consist of six Arab Gulf states: the Kingdom of Saudi Arabia, the United Arab Emirates (UAE), the Kingdom of Bahrain, the Sultanate of Oman, the State of Qatar, and the State of Kuwait (see Figure 2.1).

[1] GCC Secretariat General (http://www.gcc-sg.org/Foundations.html).

Figure 2.1 Map indicating the location of GCC countries

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

..._gulf_wars_2/maps/mideast.gif

The GCC countries stand out as one of the most important economic regions of the world. In particular, the capability of the region to meet the world demand for hydrocarbon consumption has driven the region's strategic economic significance in the global economy (Crystal, 1990). GCC countries were responsible for about 18 per cent of total world oil production in 1999, and they account for around 45 per cent of the world's proven crude oil reserves and 15 per cent of total world's proven natural gas reserves.[2]

The consequent importance of the Gulf region to the global oil market and economy lies in the fact that any interruption in Gulf oil production can destabilize the world economy, especially through deliberate limiting of the supply of oil. GCC countries can also make up any shortages in the world oil supply when oil production is interrupted elsewhere.

Table 2.1 shows the amount of oil and gas production and reserves in each GCC as well as each country's share in the total GCC production and reserves of these products. The

[2] GCC Secretariat General's Economic Bulletin, 2001, p. 12-13.

largest oil producer in the GCC is Saudi Arabia with a share of 57 per cent in 2000. Saudi Arabia also has the largest proportion of oil reserves (56 per cent). Although Saudi Arabia's natural gas production comprises the largest share among GCC countries, Qatar's natural gas reserves are the largest.

Table 2.1 Oil and Gas production and reserves, 1999

	Oil production (1000 barrel per day)		Oil reserve (billion barrel)		Natural gas production** (million cubic meters/year)		Natural gas reserves (million cubic meters/year)	
	Country total	% in GCC total	Country total	% in GCC total	Country total	% in GCC total	Country total	% in GCC total
Kuwait	1882.9	14.3%	96.5	21%	10860	6.9%	1480	6.5%
Qatar	632.5	4.8%	4.5	1%	26200	16.5%	8500	37.5%
Oman	895	6.8%	5.7	1%	11565	7.3%	805	3.6%
Saudi Arabia	7560	57.2%	263.5	56%	49780	31.4%	5777	25.5%
Bahrain	179.8	1.4%	0.15	0%	11030	7.0%	110	0.5%
UAE	2060	15.6%	98.1	21%	48980	30.9%	6003	26.5%

Source: Secretariat General's Economic Bulletin, 2001. Percentages are the authors' own calculations based on this source.
** Data for this item correspond to the year 1998.

History

From the sixteenth to the nineteenth century, the Arabian Peninsula was under the control of the Ottoman Empire (Savory, 1980). By the second half of the nineteenth century, increased pressure from local Arab tribes in the Peninsula along with British naval domination in Gulf waterways and coasts caused the Ottoman influence in the Arabian Peninsula gradually to weaken. In the meantime, local Arab rulers embraced the strong British presence in the region with protectorate arrangements. This connection with the British resulted in the presence of governmental institutions established by the British, who also constructed strong commercial and military ties with the indigenous inhabitants (Anthony, 1975). Before their withdrawal in 1968, the British defined most of the borders of the Gulf states. During the 1960s and early 1970s, most of the GCC countries declared their independence (Kuwait had declared its independence in 1961. Bahrain, Qatar, and the UAE declared independence in 1971.

However, Saudi Arabia's unification had already taken place in 1932, and Oman's declaration of independence came in 1950.)

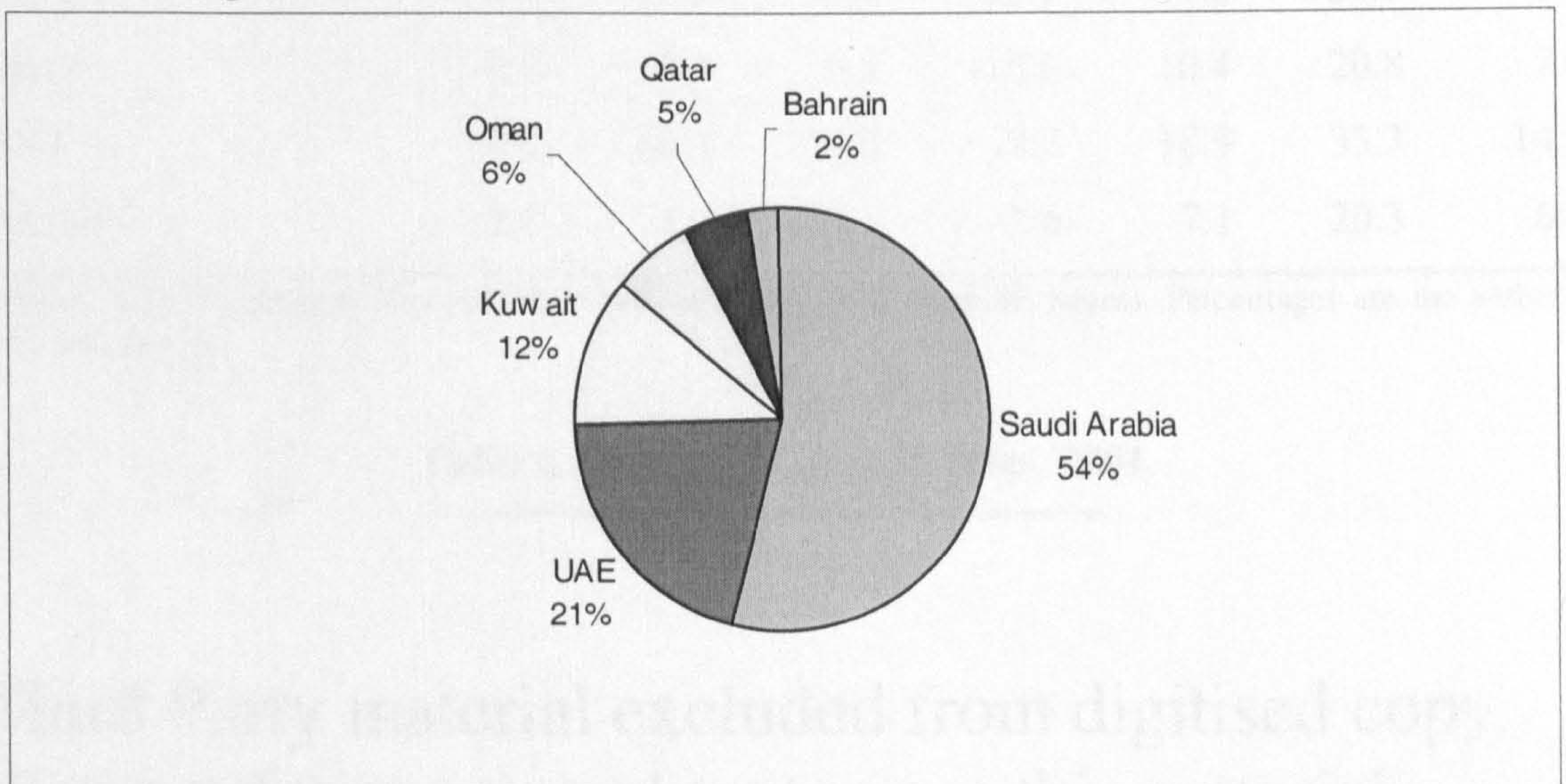
With the withdrawal of British protection from the region, the prevailing political environment became relatively unstable, so much so that on occasions the security of the Gulf region became threatened, especially after the Iranian Revolution of 1979 and the Iran-Iraq war in 1980 (Peterson, 1988). These security fears led to the formation of the Gulf Cooperation Council in May 1981. However, political differences have continued to affect the security and the relationships between countries in the Gulf region. The Iraqi invasion of Kuwait, tensions between Iran and some Gulf states - as well as some tensions within the GCC stemming from undefined borders - have dominated the political environment over the last twenty years or so. Recently, the level of overall political tension has significantly reduced, resulting in improved relationships between most of the GCC countries and Iran, accompanied by a resolving of almost all undefined intra-GCC states' borders. In addition, the strong presence of US and British bases in the GCC states has been viewed as a major factor safeguarding security and stability in the region.

Prior to the commercial production of oil in the 1950s, the economic status of the Arab Gulf region was that of an underdeveloped region. The main source of income had been generated by professions such as trade, fishing, shipbuilding, pearling, and cattle raising, as well as limited farming activities (Anthony, 1975; Crystal, 1990). Although pearling was well known to be one of the main economic activities in the Gulf region, the Japanese development and the commercial production of artificially planted pearls affected the competitiveness of Gulf pearls. Moreover, the exploration for oil (first found in Bahrain in 1932) and the growth of the oil industry in the region lessened the importance of these activities' contributions to GDP. With slow but unremitting growth in oil revenues, the economies of the Arab Gulf states significantly improved, especially when world oil prices soared in 1973 causing a rapid increase in living standards that at times reached first world levels. This also led to a rapid expansion in urbanization and growth in population (mostly by inflows of foreign labour).

Recent economic growth in the region

GDP is widely used as an indicator to measure economic development in a country. Using this reference, it is clear that the GCC countries achieved significant economic development throughout the 1990s. According to the GCC Secretariat General's report (2001, p. 15-16), the GDP of GCC countries grew by 78 per cent from $180 billion in 1990 to $321 billion in 2000. The size of the GCC economy as of 2000 bounded between the largest, Saudi Arabia, accounting for 54 per cent of total GCC GDP, and the smallest, Bahrain (see Figure 2.2). In 2000, Saudi Arabia's GDP amounted to some $173.3 billion, followed by $66.1 billion for the UAE, $37.7 billion for Kuwait, $19.7 billion for Oman, $16.4 billion for Qatar, and $7.9 billion for Bahrain. In fact, the differences in oil production quantities (and their revenues) are the main determinant of the respective economies' sizes (as shown above in Table 2.1).

Figure 2.2 The distribution of GCC GDP at current prices, 2000



Source: Percentages are the author's own calculations based on the GCC Secretariat General's Economic Bulletin, 2001

The significant income generated from the wealth of the hydrocarbon resources, accompanied by relative small population (see the demography subsection), led to high records of per capita incomes. For instance, the average per capita income in the GCC countries stood at around $10,362 in 2000, up from $8,144 in 1990 and $8,653 in 1995. Individually, Qatar had the highest per capita income in the GCC region in 2000, standing at around $29,000, followed by the UAE with a per capita income of around $21,500 (GCC Secretariat General's Economic Bulletin, 2001).

In addition, over the period 1995-2000, most GCC countries achieved positive nominal GDP average annual growth rates (see Table 2.2). This indicator averaged 6.4 per cent in Saudi Arabia and Bahrain and reached 14.9 per cent in Qatar, the latter being mainly explained by the fact that this country undertook substantial capital expenditure on developing its gas sector during this period (Gulf Business, July 2002). In real terms, GCC economic rates of growth exceeded world levels in 2001 (see Table 2.3).

Table 2.2 GCC GDP annual growth over the period 1995-2000 (current prices)

	1995	1996	1997	1998	1999	2000	AVG
Saudi Arabia	6.4	10.6	3.7	-12.4	8.4	21.8	6.4
UAE	11.9	12.1	5.0	-6.0	10.1	20.4	8.9
Kuwait	7.1	17.0	-3.4	-15.7	17.2	26.9	8.2
Oman	6.8	10.7	3.7	-10.6	10.4	20.8	7.0
Qatar	10.4	11.3	24.7	-9.2	18.9	33.3	14.9
Bahrain	5.1	4.3	4.1	-2.6	7.1	20.3	6.4

Source: GCC Secretariat General's Economic Bulletin, 2001 (various pages). Percentages are the author's own calculations.

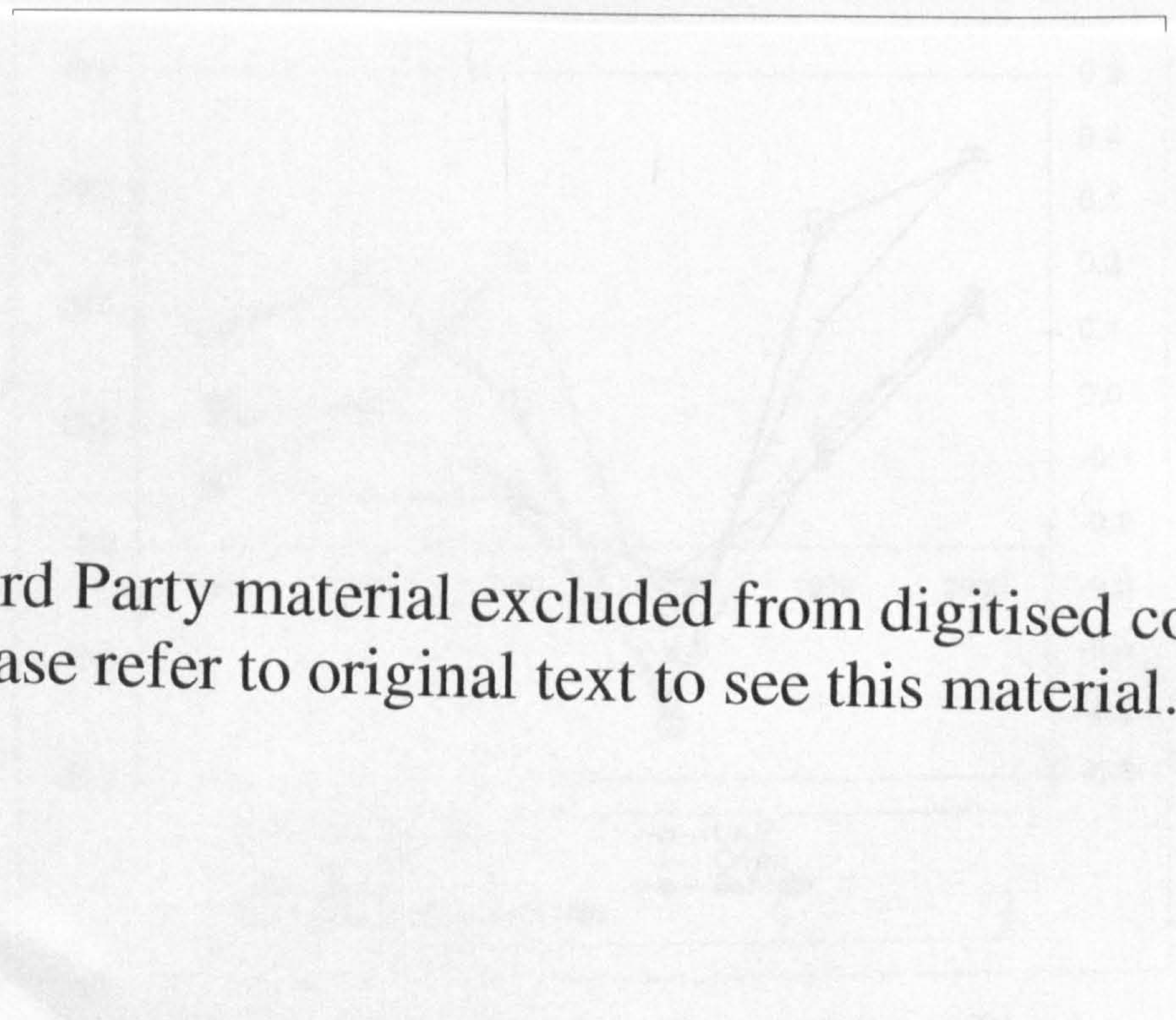
Table 2.3 Real GDP growth rates, 2001

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Sources: The World Factbook, 2002 (http://www.cia.gov/cia/publications/factbook).

Generally, the GCC economies are vulnerable to international price conditions surrounding their primary export product, oil. For instance, during 1995-2000, GCC GDP performance fluctuated mainly on account of the vulnerability of the oil sector. Thus, as Figure 2.3 shows, both GDP and oil sector growth rates exhibit similar patterns. The figure also indicates that all GCC countries experienced negative GDP growth in 1998 because of the crash in oil prices, so that the average oil price stood at \$12.60 a barrel for Brent in this year compared with \$19.12 in 1997 (MEED, 25 June 1999). The decrease in the oil price came after a huge excess supply in the oil market, mainly due to reduction in oil demand by the countries affected by the Asian financial crisis.[3] The recovery in oil demand and the success of the OPEC cartel to limit oil supply resulted in an increase in economic growth after 1998. The figure shows also that the least affected country in the GCC during 1998 was Bahrain, mainly because of the more diversified nature of its economy and the country's low dependence on oil income. As mentioned earlier, the strong growth of the Qatari economy shown in the figure was mainly due to the large capital expenditures on gas projects that were undertaken over the period.

Figure 2.3 GDP and oil growth rates of individual GCC countries over the period 1995-2000



Third Party material excluded from digitised copy. Please refer to original text to see this material.

[3] Source: Qatar Central Bank, Annual reports, 1998 and 1999.

Demography

The Gulf region experienced a rapid growth in population during the 1990s. Between 1990 and 2000, the population of the GCC countries increased by 40.3 per cent. In 2000, the population reached 31 million, distributed between 71 per cent in Saudi Arabia, 10 per cent in the UAE, 7.7 per cent in Oman, 7.2 per cent in Kuwait, and 1.9 per cent in Qatar. Moreover, data for 2000 reveal that GCC populations are very young since about 45 per cent are under 20 years old.[4]

In fact, non-citizen labourers dominate GCC populations. These grew in number because of the need to cover the shortage of labourers among the indigenous population and to catch up with economic growth in the last three decades of the twentieth century. Despite the measures taken by local governments to replace non-national workers by nationals, the latter seem unlikely or unwilling to take up jobs in areas such as construction and other low-paid tasks, which are largely executed by expatriates (Gulf Business, August 2002).

Inflation

Another characteristic of GCC countries' economies is that all experienced low levels of inflation (generally less than 5 percent) throughout the 1990s.[5] For example, in 2001, the inflation rate ranged form 1 per cent in Oman to 4.4 per cent in the UAE (see Table 2.4). Most of these inflation rates are similar to these experienced in the developed countries, which range between 1 and 4 per cent for the same year.[6] The most likely source of GCC countries' inflation levels is to imports, as more than 90 per cent of GCC countries' imports are supplied by non-GCC countries.[7] Nevertheless, increased competition and substitutes for imported goods probably helped moderate the inflation level. Moreover, since most of the imports of GCC economies are Dollar denominated, GCC economies can face price inflation due to unstable exchange rates against the US

[4] GCC Secretariat General's Economic Bulletin, 2001, p. 39-40.
[5] Annual reports of 2001 of Central Banks in each GCC country.
[6] The World Factbook, 2002 (http://www.cia.gov/cia/publications/factbook).
[7] This percentage is calculated from the GCC Secretariat General's Economic Bulletin, 2001, p. 28.

Dollar. However, on average, the Dollar was relatively stable over the 1990s against major international currencies and this helped dampen potential inflationary pressures (Qatar Central Bank, 2000).

In addition, the use of appropriate monetary and fiscal policies to control liquidity and finance budget deficits helped, to some extent, in keeping pace with changes in oil prices and achieving stability in average general prices. The low interest rate/inflation climate in the global economy throughout the 1990s must also have been an important factor in limiting inflationary forces. The broad effect of this low-inflation environment has clearly been seen in the maintenance of a stable macroeconomic climate.

Table 2.4 Inflation rates in the GCC countries, 2001

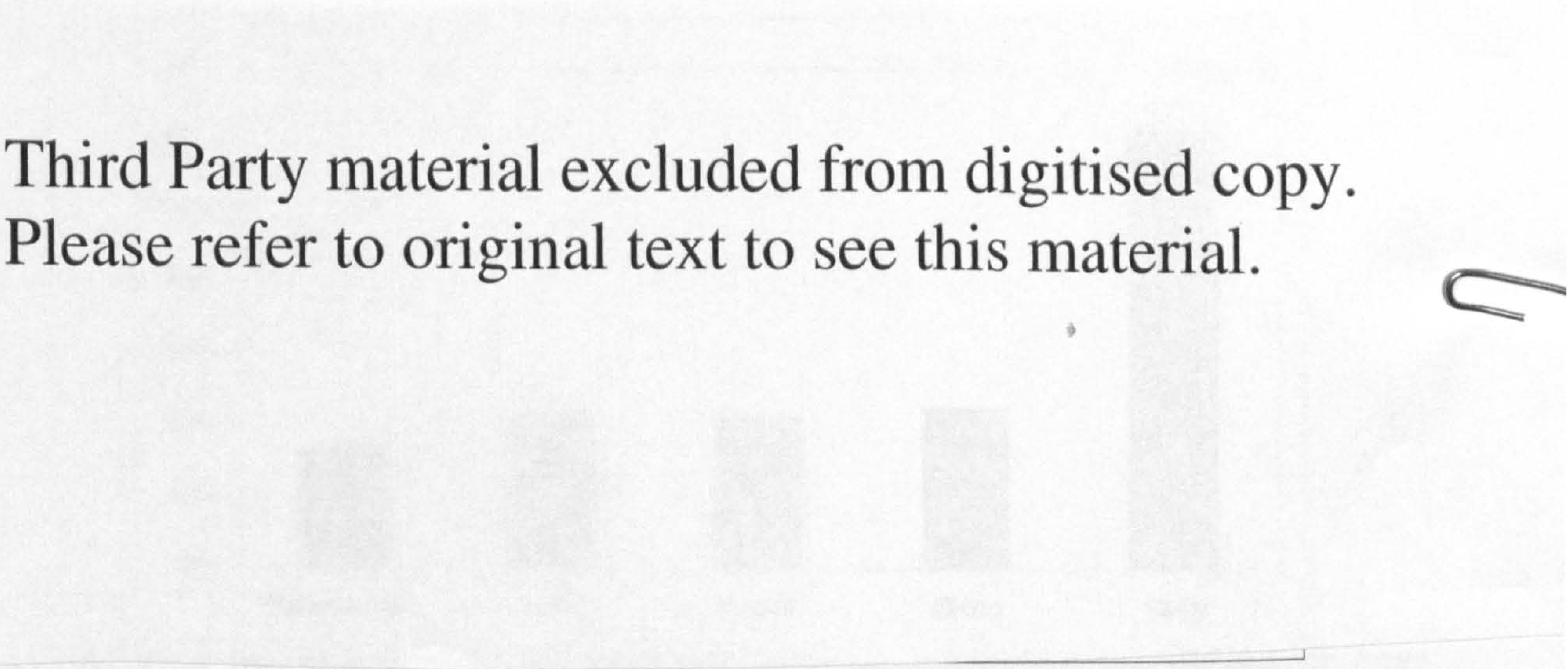
Source: The World Factbook, 2002 (http://www.cia.gov/cia/publications/factbook).

External indebtedness

Most of the GCC countries were able to settle the high levels of external debt that were generated from the financial burden associated with the military operation that terminated the Iraqi invasion of Kuwait. The current external indebtedness of GCC economies reflects, more or less, the extent to which they have financed their development projects as well as public deficits. For example, governments like Qatar and Oman have tapped international markets and sold bonds to finance government projects in the gas and petrochemicals area. Various budget deficits have also been run to bolster domestic government policy (Gulf Business, January and February 2000). Because of low oil prices in 1998, GCC countries were downgraded by the international credit rating agencies, compelling them to offer more attractive payment of interest (or in certain cases to postpone bonds debuts) until improved oil prices prevailed. However, oil prices had increased by mid 1999 and these countries gained improved ratings allowing them to make various successful international bond issues.

Saudi Arabia has the highest external debt, amounting to $28.8 billion in 2000, followed by the UAE ($14.1 billion), Qatar ($10.1 billion), Kuwait ($7.9 billion), and Oman ($4.4 billion). Relative to GDP, most Gulf countries' external debts are modest, except for those of Qatar whose external debts in 2000 amounted to 60 per cent of GDP (see Figure 2.4), having actually declined from about 80 per cent of GDP in 1998. The large Qatari external debt was mainly due to the country's determined plan to complete the construction of its huge gas field project. Repayments of these debts are expected to be arranged from the sales of gas (Gulf Business, July 2002).

Figure 2.4 External debts as a share of GDP, 2000



Third Party material excluded from digitised copy. Please refer to original text to see this material.

Source: Bank for International Settlements, 2002 (http://www.bis.org).

Exchange rates

GCC countries' exchange rate policies aim at providing stability and convertibility, and maintaining the value of the national currencies against major international currencies. This goal has been pursued for some time given that the majority of GCC countries have maintained fixed exchange rates against the US Dollar since the early 1970s. Kuwait is the main exception as it has tied its currency with a basket of major international currencies, although in January 2003 it started to peg its currency to the US Dollar as part of the GCC policy to introduce a single currency by 2010.

The adoption of the fixed exchange regime, as well as the choice of the US currency to which most GCC currencies have been pegged, is a result of the fact that most GCC income is in US Dollars generated from oil exports, in addition to the fact that the US currency is an internationally accepted medium of payment in world trade. In essence, GCC economies might be more exposed to the risk of currency fluctuation if they floated against non-US Dollar currencies. GCC countries back their pegged currencies to the Dollar by using Dollar reserves generated from oil revenues. However, in certain cases, it has cost these countries their reserves in order to keep their exchange rates as officially specified at the pegged rate. For example, during the first days of the Iraqi invasion of Kuwait, most GCC governments intervened with their reserves in order to maintain their exchange rates. Moreover, in late 1998 and early 1999, the Saudi authorities strengthened the position of the riyal against speculation activities in the international currency market, especially in the wake of the Asian financial crisis (Gulf Business, August 1999).

Fiscal status

Although GCC countries are believed to run surpluses because of their abundant wealth of hydrocarbon resources, most of these countries actually ran fiscal deficits in the 1990s. This was mainly because their governments shouldered the burden of huge expenditures associated with maintaining and providing public services as well as financing various state projects in the petrochemical industry, expansions in water and electricity facilities, and in oil and gas field developments. The deficits are also due to low diversification of government income, mainly derived from oil revenues that amount to no less than 70 per cent of total government revenues. Although none of the GCC governments levy personal income taxes, many of the GCC countries have introduced varieties of indirect taxes. These include fees on expatriates visa extensions and renewals, fees on medical services, airport tax, fuel price increases, and electricity tariffs. Many of these were introduced with the aim of diversifying government income but are subject to increase or decrease according to oil market conditions and government financial status.

From a record 30 per cent of GDP in 1991, the fiscal deficits of GCC countries have been dramatically reduced over time to around 5 per cent in 1999. In that year, the GCC countries operated with deficits amounting to $15.5 billion. The GCC budget, however, returned a surplus of $0.94 billion in 2000 owing to sustained strong oil prices and the continuation of tight fiscal policies.[8] In fact, the adverse impact of the 1998 economic downturn is believed to have had some positive outcomes. It has accelerated financial and economic reforms and privatisation policies, and paved the way for the private sector to play a greater role in economic development. This subsequently may have helped reduce the burden of government finance in various areas.

Diversification policy

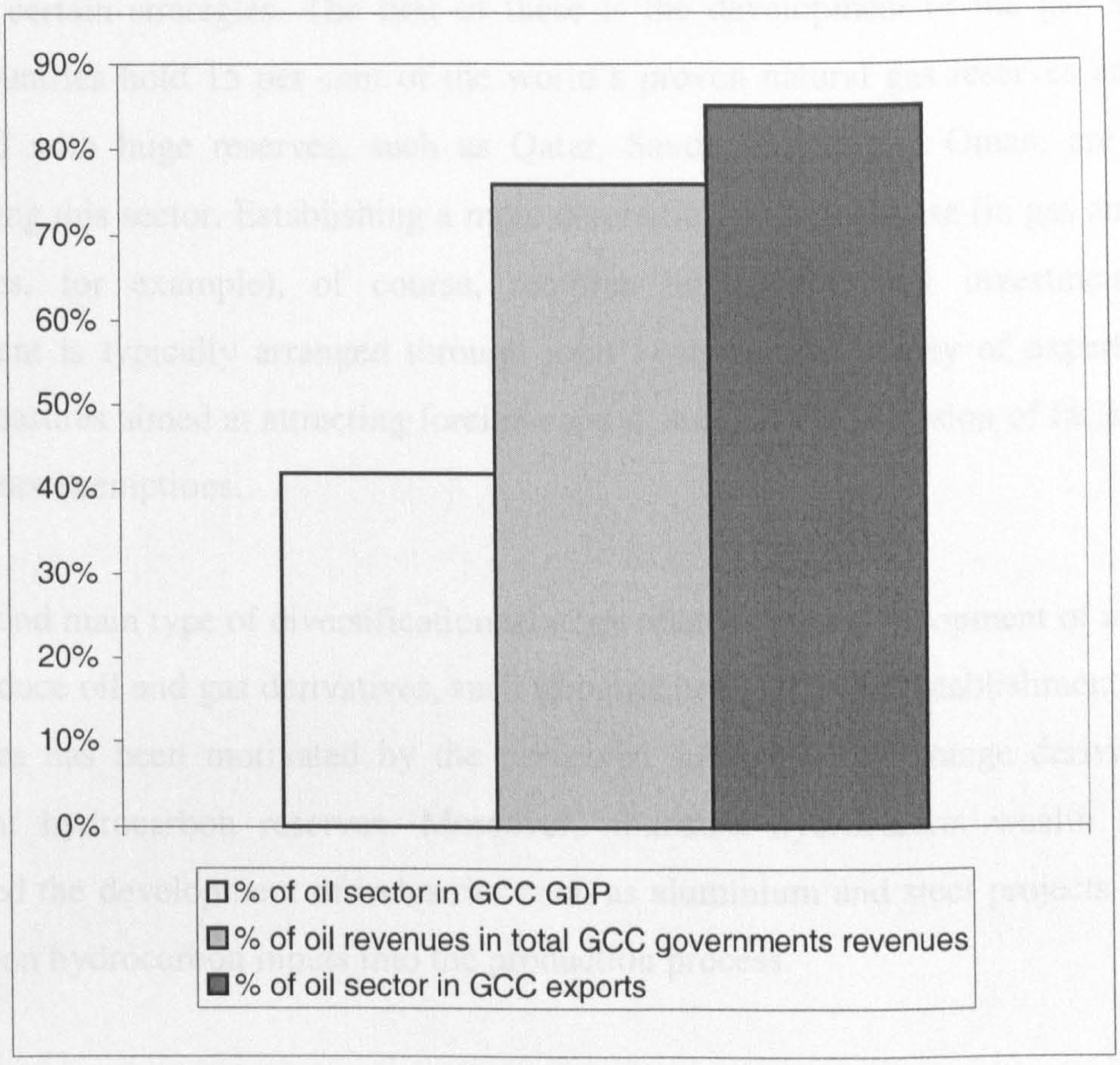
The GCC countries' economies' dependence on oil has reached substantial levels, as can be seen from the contribution of the oil sector to GDP, government revenues, and export earnings (see Figure 2.5 a and b). For example, across the GCC countries, the oil sector's contribution to GDP ranges between about 28 per cent in the UAE and 58 per cent in Qatar. Revenues of oil sales in total government revenues were no less than 70 per cent for each of the GCC countries (excluding Qatar). Moreover, oil exports comprised no less than 80 per cent of total GCC exports. Hence, the vulnerability of the overall economy to international oil prices should not be understated.

---

[8] GCC Secretariat General's Economic Bulletin, 2001, p. 20.

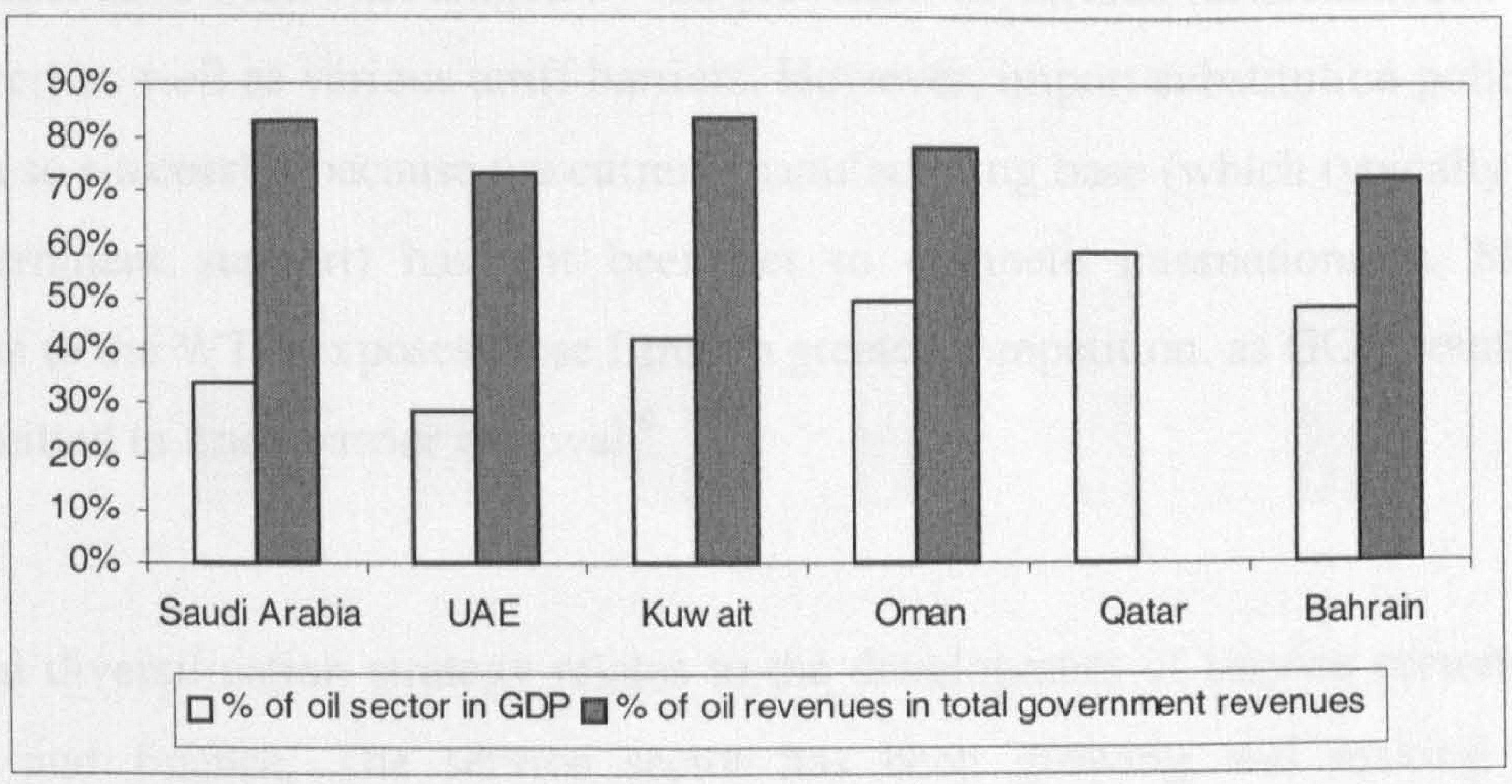
Figure 2.5 a and b The share of oil revenues in GCC GDP, exports, and total government revenues, 2000

(a)



Source: Percentage calculations are based on GCC Secretariat General's Economic Bulletin, 2001 (various pages).
Note: Data on Qatar's oil revenues are not available.

(b)



Sources: Percentage calculations are based on GCC Secretariat General's Economic Bulletin, 2001.
Note: Oil revenues as a share in total government revenues for Qatar are not available.

In order to develop a more diversified economic environment, the GCC countries have adopted certain strategies. The first of these is the development of the gas industry. GCC countries hold 15 per cent of the world's proven natural gas reserves and those endowed with huge reserves, such as Qatar, Saudi Arabia, and Oman, are rapidly developing this sector. Establishing a more diversified industrial base (in gas and heavy industries, for example), of course, requires intensive capital investment. Such investment is typically arranged through joint ventures, the supply of expertise, and other measures aimed at attracting foreign capital, such as the provision of facilities and various tax exemptions.

The second main type of diversification strategy relates to the development of industries that produce oil and gas derivatives, such as petrochemicals. The establishment of these industries has been motivated by the perceived substantial advantage deriving from abundant hydrocarbon reserves. Moreover, abundant hydrocarbon wealth has also permitted the development of industries such as aluminium and steel projects that rely heavily on hydrocarbon inputs into the production process.

The third strategy relates to promoting import substitution and, with the aim of promoting import substitution industries, many small to medium-size light manufacturing firms have been established in the Gulf (Gulf Business, August 2002). These firms have been encouraged by the provision of various facilities (such as water and power) as well as various tariff barriers. However, import substitution policies have not been so successful because the current manufacturing base (which typically depends on government support) has not been set to compete internationally. Moreover, accession to the WTO exposes these firms to greater competition, as GCC countries will be committed to trade barrier removal.[9]

The final diversification strategy relates to the development of service sectors such as tourism and finance. The service sector has been growing and making a larger contribution to GCC GDP. For example, the service sector in the GCC countries grew

[9] All GCC countries are members of the Word Trade Organisation except Saudi Arabia (which is in process of joining).

by more than 63 per cent over the period 1990-2000.[10] Some GCC countries have undertaken substantial service sector development in various areas, including tourism, the financial sector, information technology, education, and promoting exhibition and conference activities. For example, the UAE (in particular, the Emirate of Dubai) has a well-established and a premier re-export centre equipped with modern facilities that attract local and international firms. Further, Bahrain has focused on developing banking services that are mainly aimed at attracting offshore banking units. Moreover, Bahrain announced the establishment of an International Islamic Financial Market in August 2002. Dubai has a well-developed 'Internet City' and is competing hard with Bahrain to develop a major offshore financial sector. It also has the most developed tourism industry in the Gulf. Qatar is aiming to follow such steps by promoting tourism and conferencing. These developments in the service sector are aimed at achieving greater economic diversification and creating opportunities for the private sector that result in greater foreign investment.

Overall, the GCC countries' economies have been growing and powered mostly by their main source of income from oil production. However, these economies still remain exposed to fluctuations in international oil prices. This indeed suggests an increased need for reforms and greater economic diversification.

The above provides a broad insight into the main economic features of the GCC countries' economies. The following sections present the main features of financial and banking system developments in the region.

2.3 Overview of GCC financial systems

This section of the chapter outlines the development of individual GCC financial systems focusing mainly on the banking sector. We then examine the overall GCC banking sector performance. Finally, the section concludes by noting recent efforts aimed at shaping a more integrated GCC financial system.

[10] This includes wholesale, retail, hotels, restaurants, transport, finance, insurance, real estate, government services, and other services (GCC Secretariat General's Economic Bulletin, 2001, p.11).

2.3.1 Background to GCC financial systems

Banking sector development in Qatar

Prior to commercial export of Qatar's oil, Qatar did not have any banking entities practising banking activities (Qatar Monetary Agency, 1992). The first ever bank in Qatar was established in 1950, when the Eastern Bank (known today as the ANZ Standard Chartered Bank) established its Qatar branch after Qatar's oil exports commenced in December 1949. In 1954 and 1956, the British Bank of the Middle East (known today as the HSBC bank) and the Ottoman Bank (currently known as the Grindlays Bank) respectively opened their Qatar branches. Two Arab banks were also established later: the Arab Bank Limited in 1957 and the Intra Bank (known later as Almashreq Bank) in 1960. Until the mid 1960s, foreign bank branches dominated banking activities, until Qatar established its first national bank (know as the Qatar National Bank) in 1965 with joint venture capital shared equally between the Government of Qatar and the public. The economic expansion in Qatar attracted more foreign banks; thus, in the second half of the 1960s, the government authorised four new foreign banks.

Because of the strong presence of the British administration in the Gulf region, the dominant currencies formerly in circulation were either the Pound Sterling or the currencies that were linked to it, such as the Indian Rupee and the Gulf Rupee (the Gulf Rupee was issued in India and used especially for the Gulf region's cash transactions) (see Qatar Monetary Agency, 1992; Bahrain Monetary Agency, 2002). While these currencies were considered to be the main media of exchange to obtain goods and services in the Gulf region, negotiations between Qatar, Bahrain, Dubai, and Abu Dhabi had been taking place in order to create a common Gulf currency that would replace the aforementioned currencies. However, these negotiations failed to achieve this goal, but they did, at least, lead to a successful agreement, reached in 1966 between Qatar and Dubai, to create one currency to circulate within these two Gulf emirates. The responsibility for issuing and managing this currency was vested in the Qatar–Dubai

Currency Board. Prior to the circulation of this new currency (called the Qatar-Dubai Riyal), the Indian Government devalued the Indian Rupee by 35 per cent, which was followed by a parallel depreciation of the Gulf Rupee. To ensure a successful debut of the Qatar-Dubai Riyal, the two governments asked existing banks to exchange the Gulf Rupee with this new currency at the pre-devaluation rate. However, for technical reasons, the two governments decided to circulate the Saudi Riyal and withdraw the Gulf Rupee. This was followed by the issuance of the Qatar-Dubai Riyal in the last quarter of 1966 with a value equal to the pre-devaluation Gulf Rupee. The Qatar-Dubai Riyal was also covered by the Pound Sterling; however, when the Pound Sterling was devalued in 1967, the two emirates agreed to maintain the value of the new currency against gold.

According to the Qatar Monetary Agency (1992), the Qatar-Dubai currency circulated in Qatar until 1972, the year in which Dubai merged in the United Arab Emirates (UAE) and issued its own currency. After gaining independence in 1972, Qatar became a member of the International Monetary Fund (IMF) and in 1973 introduced its own currency (the Qatari Riyal), which was pegged to the IMF's special drawing rights, and then pegged to the US Dollar at a rate of QR3.64 per $1, which is in effect till today.

Qatar established in 1973 the country's central bank known as the Qatar Monetary Agency (QMA, later called the Qatar Central Bank, QCB). The QMA regulates banking credit and finances, issues currency, and manages the foreign reserves necessary to support the Qatari Riyal. One of the first steps taken by the QMA was to restrict the licensing of new bank establishments or branch openings of foreign banks. The oil boom started in 1973, promoting economic growth, and this resulted in an expansion of the banking sector as three national banks were established during the latter part of the 1970s. Furthermore, another two national banks were added to the banking structure during the 1980s. However, one foreign bank, the Qatar branch of Al-Mashrek Bank - headquartered in Beirut- was closed and put into liquidation in 1989 (Qatar Monetary Agency, 1992).

As a result of the Iraqi invasion of Kuwait, banks in Qatar lost an estimated 15 to 30 per cent of deposits in late 1990, while QMA (with its ready reserves) left banks free to accept or reject the withdrawal of deposits before their maturity but in accordance with their liquidity status.[11] Moreover, QMA directed money exchangers to sell the Dollar at the official rate, with penalties to be set for any reported violation. These measures adopted during the Gulf crisis maintained confidence and soundness in the financial system that resulted throughout the 1990s.

According to Gulf Business (August, 2002), one important banking problem occurred in 2000 when one of Qatar's national banks (Al-Ahli Bank of Qatar) was hit by a severe loan problem caused by one of its major corporate clients' defaulting. Al-Ahli Bank's credit risk exposure to this corporate was discovered to approach 40 per cent of the total bank loan portfolio. To bail out the bank, QCB rescued the bank on an agreement providing a 10-year guarantee with an amount close to the amount of the bank's non-performing loan ($28 million). QCB has also changed the bank's management and required significant bank restructuring. It has been argued that confidence in Qatar's banking sector would have been harmed if the QCB let this bank fail. Moreover, one of the major weaknesses that appeared to have led to this problem was that the bank's management generally remained hostage to the key shareholders and political influence. This necessitated moves to enhance the management and monitoring systems in order to reduce the likelihood of conflicts of interest in the future.

The current regulations indicate that banks' credits are limited to 95 per cent of their total deposits. In addition, banks must maintain a ratio of no less than 6 per cent of their capital to total assets at all times. Moreover, capital adequacy must maintain a minimum of 8 per cent, in line with the Basel 1988 recommendations. Nevertheless, it should be noted that, starting from the mid 1990s, QCB has gradually lifted the restriction on deposit rates and, currently, all deposit rates are set according to market forces.[12] Banks are also permitted to offer interest on demand deposit accounts with balances exceeding

---

[11] Qatar Monetary Agency, 1992.
[12] Qatar Central Bank Guidelines to banking institutions (http://www.qcb.gov.qa/pages/English_Site/intro.html).

QR2 million. QCB amended the reserve requirements from 19 per cent on demand deposits to 2.75 per cent effective on the total of all deposit accounts.

Within the period 1990-2000, the level of credit in the economy increased by 188 per cent, progressing by an average annual rate of 13 per cent and reaching $7.6 billion by 2000. Deposits increased by 136 per cent with an annual growth of 11 per cent totalling $9.9 billion by the end of the decade. Bank capital and reserves grew by 61 per cent, achieving an average annual growth of 6 per cent and reaching some $1.7 billion by 2000.[13] Moreover, the level of assets stood at $14.8 billion. The Qatari banking system currently includes fourteen commercial banks, seven national and seven foreign, as well as one specialised bank (Qatar Central Bank, 2000).[14]

Overall, the Qatari banking sector has substantially developed during the 1990s. The authorities continue to strengthen supervision of the banking system in order to ensure improved soundness and to comply with various international standards. Moreover, the relaxation of various barriers, such as interest rate ceilings, should help facilitate greater competition in the banking system.

The development of banking business in the UAE

According to the UAE Central Bank (2001), the British Bank of the Middle East started as the first bank in the UAE in 1946, taking a location in Dubai. This bank opened its second branch in Abu Dhabi following the discovery of oil. Later, the Eastern Bank and the Ottoman Bank opened their branches in Abu Dhabi in 1961 and 1962 respectively. The year 1963 witnessed the establishment of the first national bank, the National Bank of Dubai, followed in 1968 by the opening of Abu Dhabi National Bank in 1968. Obviously, the attractiveness of these two cities in the UAE derives mainly from the acceleration of trade activities (primarily in Dubai) and oil exports (largely in Abu Dhabi).

[13] Author's own calculation based on the GCC Secretariat General's Economic Bulletin, 2001.
[14] Among the national banks, two are Islamic. One of the foreign banks, the Grindlays bank Ltd., changed into a national bank by 1st August 2000. The specialized bank is Qatar Industrial Development Bank, initiated in 1997 to provide loans to small and medium-sized manufacturing firms.

The UAE Central Bank (2001) notes that after the formation of the federation which resulted in the establishment of the state of UAE in 1972 (consisting of seven emirates), the rush to open national and foreign branches accelerated. In 1972, the Currency Board was established to issue the UAE national currency, the Dirham, and to supervise and regulate the banking system. In the same year, the number of commercial banks increased to six domestic and fifteen foreign banks, most of them concentrated in Abu-Dhabi and Dubai and a few in the third largest emirate, Sharjah. Following the dramatic increase in international oil prices, the number of banks reached thirteen nationals and twenty-eight foreign banks by 1975. After 1975, the Currency Board realised that the economy needed more banking institutions to help with financing associated with the economic boom; thus, more bank licenses were issued and by 1977 there were twenty national and thirty-four foreign banks operating throughout the Emirates. In 1980, the UAE issued a Federal Law establishing the Central Bank of the UAE, with extensive powers to operate as the country's central bank. The central bank was formally in charge of issuing and controlling the supply of the Dirham and maintaining gold and foreign currencies to support its value. In 1981, the UAE Central Bank lifted the freeze on new bank establishments; however, it imposed it again specifically on the licensing of new foreign banks. It also instructed the existing foreign banks that from 1984 they would not be allowed to have more than eight branches throughout the UAE.

In the early 1980s, the UAE Central Bank adopted in the early 1980s several measures to strengthen the banking system (UAE Central Bank, 2001). It set minimum capital requirements, enhanced audit and reporting requirements, increased inspection, established a department dedicated to oversee bank loan risks, and set regulations that limited the amount of loans that could be given to the board of directors. In 1983, one bank failure resulted from the violation of the loan limit to the board of directors. This caused the UAE Central Bank to appoint administrators to this bank and, in essence, the central bank and the government of Dubai bailed out the bank with an amount of $380 million.

As the UAE Central Bank (2001) notes, the oil price fell below $10 per barrel in 1986 and this led to a sharp decline in federal revenues. Consequently, contractions in government expenditure slowed down economic activities and, as a result, the banking sector experienced loan problems arising from accelerated loan losses. This led to a restructuring of the banking sector when three banks in Dubai merged, as did another three in Abu Dhabi. This resulted in banking sector numbers falling to nineteen national and twenty-nine foreign banks. Another threat emerged in the wake of Iraq's 1990 invasion of Kuwait, when between 15 and 30 per cent of customer bank deposits were transferred out of the UAE. At this time, the UAE Central Bank injected funds into at least two banks in order to strengthen their liquidity and restore confidence in the banking system as a whole.[15]

During the 1990s, the UAE Central Bank introduced various regulations aimed at improving banking sector soundness (UAE Central Bank, Annual Report, 2001). By 1993, banks were subjected to a capital to assets ratio of 10 per cent. Moreover, banks were required to accumulate reserves by shifting 10 per cent of their annual net profits to the reserve accounts until the latter equalled 50 per cent of their paid-up capital. In 1994, banks were urged to move toward adopting International Accounting Standards. These directions enhanced, to some extent, the capitalisation of the UAE banking system. For example, in 1997, the average ratio of capital to risk-weighted assets for all banks was 21 per cent, which was well above the Basel 1988 recommendations. Recently, the UAE Central Bank has raised the capital reserve ratio to 14 per cent. This move came after one banker fled the UAE with an estimated quarter billion Dollars of customers' funds (The Banker, Sept., 2000).

Commercial banks in the UAE made significant developments during the 1990s. Commercial credit to different economic sectors grew by 169 per cent over the period 1990-1999, with an average annual growth of 12 per cent.[16] These credits amounted to $37.6 billion in 2000. Deposits in the commercial banks grew by 72 per cent with an annual growth of 8 per cent, and total deposits reached $36.8 billion in 2000. Moreover, bank capital and reserves amounted to some $9.3 billion by 2000, having experienced

[15] Economist Intelligence Unit, 1991.

[16] Author's own calculation based on the GCC Secretariat General's Economic Bulletin, 2001.

annual average growth of 9 per cent throughout the 1990s. Total banking sector assets amounted to $75.5 billion in 2000. Over the decade of the 1990s, only small changes in the number of banks occurred and by the end of 2000, the number of national banks had reached twenty while foreign banks stood at twenty-six banks.

Overall, UAE banks operate in a relatively healthy financial system. The banking system development over the last twenty years or so reflects the system's ability to cope with minor crises as well as the changing demands of clients and the economy.

Banking sector development in Kuwait

According to Al-Sharrah (1999), the first attempt to establish a bank in Kuwait was in 1935 when both the Ottoman Bank and the British Bank of the Middle East competed to establish a branch for their banks; however, neither succeeded because of the hesitant Kuwaiti rulers. In 1941, the British Bank of the Middle East was permitted to set up a branch in Kuwait. Many banks tried later to enter the Kuwaiti banking market, but the authorities prohibited foreign banks from conducting banking business in the country. When the British Bank's concession ended in 1971, this bank changed its name to the Kuwait Bank for the Middle East and Kuwaitis purchased 60 per cent of the bank's capital.

Foreign currencies - largely the Indian Rupee and then the Gulf Rupee - circulated in Kuwait between 1930 and 1961. However, in May 1961, Kuwait issued its own currency, called the Dinar (Al-Sharrah, 1999). The Kuwaiti Dinar (KD) came into existence after the Kuwaiti economy was strengthened, primarily through increased revenues from oil, which led to the development of the financial and other economic sectors. Moreover, the need for its own currency came because Kuwait wanted a stronger and stable currency, hoping to avoid the fluctuations associated with the Gulf Rupee. Prior to the establishment of the Kuwait Central Bank, a currency board was in charge of issuing the Kuwaiti Dinar and administering money exchange. In 1959, the Central Bank of Kuwait was created and took over the functions of the currency board.

In 1952, a group of Kuwaiti families founded the First National Bank in Kuwait, known as the National Bank of Kuwait, which is currently the largest commercial bank in the country. In fact, after Kuwait gained its independence in 1961, the establishment of several other banks, all under Kuwaiti ownership, followed. Moreover, some specialized financial institutions also emerged: the Credit and Savings Bank, was established in 1965 by the government to channel funds into domestic projects, agriculture, and housing; the Industrial Bank of Kuwait, established in 1973, aimed to fill the gap in medium- and long-term industrial financing; and the private Real Estate Bank of Kuwait emerged in 1973 as a financier of property developments in the country. By 1978, the number of commercial banks operating in Kuwait amounted to seven, the same as today (2003).

The huge revenues generated from oil production that coincided with the rise in oil prices after 1973 resulted in a substantial increase in the wealth of Kuwait and its inhabitants. Some of the increased prosperity was channelled into speculative activities on the Kuwaiti stock market and this resulted in a small stock market crash in 1977.[17] As a response to these difficulties, the government provided compensation for certain investors and also introduced reforms and stricter regulations. The introduction of tougher capital market regulations unintentionally contributed to the creation of an illegal stock market, known as the Suq al-Manakh. The Suq al-Manakh emerged as an unofficial stock market operating alongside the official one and its stocks were mainly traded by wealthy families trading in large amounts. Because deals were undertaken using post-dated cheques, this created a huge demand for credit, and when stock prices fell in 1982, the Suq al-Manakh crashed creating a severe shake-out of the Kuwaiti financial sector and the entire economy. The officials revealed that total outstanding cheques amounted to $94 billion from about 6,000 investors. The debts from the crash left all but one bank in Kuwait technically insolvent. Only the National Bank of Kuwait, the largest commercial bank, survived the crisis. In response, the government devised a complicated set of policies, embodied in the Difficult Credit Facilities Resettlement Program, to bail out banks and investors.

[17] Economist Intelligence Unit, 1992.

During the Iraqi invasion of Kuwait in 1990, the largest commercial bank in Kuwait (National Bank of Kuwait) was the bank least affected by the Iraqi invasion thanks to its substantial international funds.[18] It controlled the exiled government's finances during the invasion. However, over the 1990-1994 period in the aftermath of the Iraqi invasion, the annual decline in the Kuwaiti banks' assets reached 6.5 per cent, and the decline in these banks' foreign assets reached 13.4 per cent as the Kuwait government directed these banks to fulfil their international liabilities so as to maintain international confidence in the institutions (Al-Sharrah, 1999).

Since April 1993, the domestic interest rate structure has been linked to the KD discount rate and banks have been permitted to set their interest charges with a margin (not to exceed a certain level) set with reference to the Central Bank of Kuwait's rate (Central Bank of Kuwait, 2000). However, since January 1995, all ceiling rates on deposit's were lifted and are now determined according to the market mechanism.

Figure 2.6 Credits and deposits of Kuwaiti commercial banks, 1990-2000

Third Party material excluded from digitised copy. Please refer to original text to see this material.

GCC Secretariat General's Economic Bulletin, 2001.

In sum, the Kuwaiti banking sector has been restoring its pre-invasion position. If we look at banking credit over the last decade (see Figure 2.6), we notice that Kuwaiti banking credits were severely affected in the years after the Iraqi invasion. However,

[18] Economist Intelligence Unit, 1992; Central Bank of Kuwait, 2002.

banking credit recovered and reached $17.1 billion by 2000, showing more confidence. Total deposits in the banking sector have reached $25.8 billion by the end of the decade. Moreover, the level of financial capital and reserves of the banking system reached $5.7 billion, increasing by 50 per cent over the decade and suggesting a strengthened banking environment.

The banking sector in Kuwait has experienced difficult periods over the last twenty years or so resulting from the Suq al-Manakh crash and the Iraqi invasion. However the Kuwaiti banking system has illustrated its resilience in these difficult periods and has emerged as a solid banking system.

The development of banking business in Bahrain

According to the Bahrain Monetary Agency (1994), banking business in Bahrain started when a branch of the Eastern Bank opened in 1921. This bank was the only one operating in Bahrain until the British Bank of the Middle East opened its branch two decades later in 1944. The National Bank of Bahrain, the first local bank, opened its doors in 1957 followed in 1960 by the establishment of a Jordanian bank, the Arab Bank Limited.

Within this emerging financial system, the Gulf Rupee was the main currency enjoying locally a common acceptance.[19] In the mean time Bahrain engaged in negotiations with neighbouring emirates (Qatar, Dubai, and Abu Dhabi) to issue a common Gulf currency. However, because the negotiations failed, Bahrain continued to deal with the Gulf Rupee until it replaced it with the Bahraini Dinar, which was introduced in 1965 (Bahrain Monetary Agency, 1994). The Bahrain Currency Board was in charge of issuing and managing the supply of the Bahraini Dinar. The expansion of the financial system, powered by the rapid increase in oil revenues, brought to the surface the need to direct, supervise, and control the financial system by a well-equipped institution. Consequently, the Bahrain Monetary Agency (BMA) was created in 1973 to take over

[19] Discussion on currency development is based on Bahrain Monetary Agency, 2002.

the work of the Currency Board and practise extensive central banking powers. The Bahraini Dinar was linked to the British Pound Sterling and then to the US Dollar. The devaluation of these two currencies (the Pound Sterling in 1967 and the US Dollar in 1971 and then in 1973) created losses to the value of the Dinar held by the Currency Board, banks, and the public. As a result, the Dinar was linked to the SDR and, in the mean time, the reserves of the US Dollar were used as an interventional currency serving to stabilise the international value of the Dinar with a margin set at certain limits. The establishment of the BMA, the issuance of the national currency, and the progressive strengthening of the economy after the 1973 oil boom made Bahrain's financial system increasingly attractive to banking business. As a result, the number of commercial banks reached fifteen by 1977. Moreover, two specialised banks were added to the Bahraini banking structure, namely the Housing Bank and the Bahrain Development Bank.

Because the Government of Bahrain was conscious about its declining oil reserves, Bahrain was among the first GCC countries to undertake initiatives aimed at diversifying its economy away from oil (Bahrain Monetary Agency, 1994). Bahrain focused on developing itself as a centre for financial services in the Gulf region with the aim of attracting oil revenues from the neighbouring Gulf countries. In fact, Bahrain has successfully attracted offshore banking units (OBUs) and has developed the main offshore financial centre in the Gulf region. Offshore banks located in Bahrain are not required to pay income taxes. Moreover, they are exempted from foreign exchange controls and cash reserve requirements. On the other hand, OBUs must not accept deposits from citizens and residents of Bahrain, and must refrain from transactions involving Bahraini Dinars. In return, Bahrain benefits from employment opportunities for its national labour force and collects annual license fees. The first OBUs to operate in Bahrain were Citibank and Algemene Bank Nederland (opened in 1975). One of the main factors that induced the fast growth of Bahrain's OBUs market was the shift of OBUs located in Lebanon to Bahrain. The number of OBUs in Bahrain reached a maximum of seventy-six in 1984. However, owing to the dramatic decline in oil prices in the mid 1980s, many OBUs contracted their business, resulting in non-renewal of various licenses. Moreover, trends towards consolidation within and between banking groups increased. As a result, the number of OBUs in Bahrain declined and, by 2002, around forty-eight were active in the country. According to the Bahrain Monetary Agency (2001), around 32.9 per cent of the assets of OBUs are from Arab countries (mostly from other GCC countries). Western European banks account for 32 per cent, American banks 21.3 per cent, and Asian banks 11.3 per cent of total OBU banking sector assets.

In 1977, Bahrain also introduced a third category of banking licences, called Investment Banking licences (IBs), for banks intending to carry out investment business (Bahrain Monetary Agency, 1994). The first of these banks was Bahrain Investment Bank (in 1977). The number of these types of banks increased from a handful in the late 1970s to thirty-four by 2001.

Bahrain also aims to establish itself as a centre for Islamic banking and finance. Early on Bahrain took the lead in introducing a comprehensive prudential set of regulations for Islamic banks, which follow guidelines from the Bahrain-based Accounting and Auditing Organization for Islamic Financial Institutions and the Basel Committee on Banking Supervision, as well as guidelines from the accounting firm Ernst & Young. These regulations aim mainly to cover regulatory issues concerning capital adequacy, asset quality, and liquidity management. These regulations may give Bahrain-based Islamic banks a competitive edge and may create interest among other countries to adopt Islamic banking regulations similar to those developed by Bahrain (Standard and Poor's Creditweek, October 16, 2002).[20]

In 2000, total banking sector assets amounted to around $106 billion, a GDP multiple of about fourteen, with OBUs' assets occupying the largest stake (87.4 per cent), followed by commercial bank assets (9.4 per cent) and investment banks (3.2 per cent). By the end of 2000, Bahrain's banking sector comprised fifty-one OBUs and nineteen commercial banks, of which two were Islamic and thirteen were investment banks (Bahrain Monetary Agency, 2000).

[20] Bahrain hosted a first Islamic bank in 1975 and currently there are two commercial Islamic banks as well as a number of Islamic banks operating on the basis of OBUs and investment banks. According to the BMA (2001), the consolidated assets of Islamic banks stood at $6,051 million.

Although the Bahraini commercial banking sector is the smallest in the GCC region, Bahrain commercial banks have achieved significant growth over the last decade or so. Commercial banking credit experienced a growth of 112 per cent from the year 1990 to 2000, increasing annually by an average of 9 per cent and totalling $3.7 billion by 2000.[21] Over the same period, deposits increased by 70 per cent, with an annual growth rate of 7 per cent. These deposits totalled $6.5 billion in 2000. In addition, capital and reserves of the banking sector amounted to $0.6 billion by 2000. The assets size of Bahrain commercial banks reached $7.9 billion by the year 2000.

Overall, Bahraini banking sector development reflects its special position as a major financial centre in the Gulf region. The country constantly aims to provide an environment conducive to banking and financial activity, and has recently made various moves to establish itself as the major Islamic finance centre in the region. While there is increasing competitive pressure from Dubai, Bahrain still remains one of the world's premier financial centres. Given its role as an offshore centre, the domestic banking sector remains relatively small, in fact the smallest in the GCC; nevertheless, the domestic banks continue to provide an important role in mobilising domestic savings and financing economic development within the country.

The development of the banking sector in Oman

Banking activities in the Sultanate of Oman commenced a few years following the end of World War II, when the British Bank of the Middle East was the only available bank in the Sultanate, starting its operations in 1948. Although banking activities were relatively limited until the exploration of oil in 1967, the situation changed just three years after the start of the commercial exporting of oil when the Omani banking structure extended with the opening of three new banks (Oman Central Bank, 1996; Al-Sharrah, 1999).

[21] Author's own calculation based on the GCC Secretariat General's Economic Bulletin, 2001.

Prior to 1970, Oman did not have its own national currency. In 1970, Oman announced a decree establishing the Muscat Currency Authority to act as an official entity issuing Oman's currency (called the Riyal Omani, RO), managing Oman's foreign assets, and accepting deposits from banks in Oman. The British Bank of the Middle East (its Omani branch) was entrusted with administering this entity. In 1972, Oman established the Muscat Currency Authority to issue the national currency, manage government accounts, and to execute banking transactions with commercial banks and international institutions. Moreover, all banks were asked to acquire licenses from this entity in order to practise banking business (Al-Sharrah, 1999).

Together, the Muscat Currency Authority and the Oman Currency Board were the first steps taken towards the creation of the Oman Central Bank (Al-Sharrah, 1999; Bahrain Monetary Agency, 2002). In November 1974, a banking law established the Central Bank of Oman (CBO) that began operation in April 1975. The CBO is empowered to make advances to the government to cover temporary deficiencies in current revenues; to purchase government treasury notes and securities with a maximum maturity of ten years; to make advances to commercial banks; and to buy, sell, discount, and rediscount commercial paper.

According to the Central Bank of Oman (1996), the CBO establishment law also facilitated the entry of foreign-owned banks and permitted an increase in the number of local banks in the Sultanate. During the seventies (the period that witnessed an oil price boom), the number of banks operating in Oman increased, reaching twenty by the end of the decade. In addition, three specialized development banks were established: the Oman Development Bank (1977), the Oman Housing Bank (1977), and the Oman Bank for Agriculture and Fisheries (1981). Although the increase in the number of banks facilitated an inflow of foreign capital and increased funds to the development process, during the early 1980s the CBO froze new bank licensing, fearing that the available number of banks might lead to excess capacity in the Omani banking system. Moreover, the steep fall in oil prices in the mid 1980s exposed the Omani banking system to pressures that led to a rationalisation of various lending schemes and forced the authorities to encourage banks to strengthen their capital and to make adequate provisions and reserves.

Bank licensing was relaxed from the mid 1980s onwards and the number of banks increased to twenty-two by the end of 1980s, with nine national and thirteen foreign banks. In 1991, the CBO was given increased powers allowing the central bank to suspend or withdraw licenses of banks violating regulatory rules. In fact, the CBO exercised its new power on the Bank of Credit and Commerce International (BCCI) because of the institution's engagement in illegal practices such as weapon finances.[22] The CBO liquidated the BCCI branch in Oman and offered it to a national bank (Bank Dhofar al-Omani al-Fransi), which agreed to take over the BCCI branch in 1992. After this event, Bank Dhofar al-Omani al-Fransi became the second largest bank in the Sultanate after the National Bank of Oman. The restructuring trend in the Omani banking system had already started in January 1989 when the Bank of Muscat purchased the assets and liabilities of the Oman Banking Corporation. Moreover, the first half of the 1990s witnessed a decrease in the number of Omani national banks, falling to only seven banks as a result of various mergers, while the number of foreign banks fell to eleven (Al-Sharrah, 1999).

During the 1990s, certain banking regulations were put in force in order to advance the soundness of the Omani banking system.[23] In 1991, the CBO amended the ceiling on the amount banks could lend to their directors from a maximum of 20 per cent to 15 per cent of their capital. Moreover, although banks in Oman had been in full compliance with the Basel capital adequacy minimum requirement of 8 per cent since 1992, the CBO wanted to further enhance the capital cushion, and thus it asked banks in Oman to achieve a minimum ratio of 12 per cent by 1998 (Oman Central Bank, 2000). This led all banks in Oman to achieve a ratio even higher than the 12 per cent target. Moreover, an expansion in personal lending in 1997 and 1998 induced the CBO to put a ceiling of 30 per cent on the proportion of personal loans in total private sector lending. However, this limit was relaxed in 2000 as the ceiling increased to 35 per cent (owing to the improved macroeconomic climate). The loan to deposit or lending ratio is currently set at 87.5 per cent. The minimum reserve requirement for banks is set at 5 per cent of total deposits. Until 1993, the authorities set ceilings on the interest rates commercial banks

[22] See Oman Central Bank (1996).
[23] Oman Central Bank, Annual Reports (various years).

could charge on both deposits and loans. In a move toward deregulation, the authorities decided to gradually prepare the banking market for market-determined interest rates. Oman freed up the ceiling imposed on deposits of Riyal Omanis in the last quarter of 1993. In mid 1994, the authorities also deregulated interest rates on consumer loans of RO 9,000 or less. By January 1999, consumer loans were fully deregulated (Oman Central Bank, 2000).

Over the period 1990-2000, Omani banking credit grew by 198 per cent, increasing annually by an average of 13 per cent over the period, and totalling $7.7 billion by 2000.[24] Total deposits in the banking sector stood at $6.8 billion. In addition, capital and reserves of the banking sector reached $1.1 billion in 2000, reflecting an average annual growth of 15 per cent. Total commercial bank assets reached $15.2 billion in 2000. Overall, these indicators show that, as in other GCC markets, the Omani financial sector has expanded substantially over the last decade. Following a series of mergers during the 1990s, the number of commercial banks at the end of 2000 stood at fifteen, of which six are locally incorporated and nine are branches of foreign banks (Oman Central Bank, 2000).

Banking sector development in Saudi Arabia

Early banking activities in Saudi Arabia were limited to the presence of a handful of foreign-based trading houses, such as the branch of Algemene Bank Nederland, and of various money changers (Al-Jarrah, 2002). Their main business was to provide financial services for locals and pilgrims. The more formal and organised form of banking system emerged after the exploration of oil in 1939 and, as soon as World War II ended, the Saudi market attracted leading foreign banks to open branches. Hence, the French Banque de l'Indochine and Arab Bank Limited opened their branches in Jeddah in 1948; while in 1950, three international banks opened their braches, namely the British Bank of the Middle East, the National Bank of Pakistan, and Bank Misr (of Egypt).

[24] Author's own calculation based on the GCC Secretariat General's Economic Bulletin, 2001.

Moreover, Saudi Arabia did not have a national currency until 1952, a year that witnessed the establishment of the Saudi Arabia Monetary Agency (SAMA) (Al-Sahlawi, 1997; Al-Jarrah, 2002). However, over the years 1950 and 1956, SAMA introduced a paper money in the form of pilgrim receipts, which was covered by foreign currencies and precious metals. The introduction of the Saudi national currency, called the Riyal, came in 1960.

SAMA was (and continues to be) responsible for issuing and preserving the value of the Saudi Riyal, and for supervising and setting regulations governing the banking sector. At the time of SAMA's establishment, the Saudi government continued to use the Al-Kake and Bin Mahfouz Money Changer Company as its agent to undertake its payment services. In 1953, this company was permitted by the government to be transformed into a bank known as the National Commercial Bank, the first ever Saudi bank. By the end of the 1950s, the Saudi banking system witnessed the opening of an additional three foreign banks and two domestic banks. However, the two newly established Saudi banks, namely the Riyad Bank and Al-Watani Bank that started in 1957 and 1959 respectively, faced financial difficulties due to various liquidity problems. These were mainly caused by poor governance as board members of the two banks borrowed heavily, exposing the banks to various default problems. Being unable to meet depositors' claims, Al-Watani Bank became insolvent and was liquidated, ending up by merging with the Riyad Bank (Al-Suhaimi, 2001). As a result, in 1966, a banking law provided SAMA with broader supervisory powers that made banks subject to various liquidity, capital adequacy, lending, and reserve requirements.

By the early 1970s, other banks had entered the Saudi banking system, attracted by the opportunities brought about by the boom in the economy resulting from the increased oil revenues, especially from 1973 onwards. The strong presence of foreign banks, of which there were ten by the mid 1970s, encouraged the Saudi authorities to introduce a policy encouraging foreign banks to be converted into publicly traded companies with the participation of Saudi nationals. The legislation introduced in 1975 aimed to preserve the rights and interests of foreign banks' positions as partners in the newly incorporated banks. In order to maintain the performance and stability of the banking sector, foreign banks were allowed to hold up to 50 per cent ownership and include the name of their origins in the bank title.[25] They could also maintain management responsibilities and were allowed to enjoy treatment equal to that of national banks (Saudi Monetary Agency, 1998).

During the 1970s, five major specialised lending institutions were also established: namely, the Saudi Credit Bank, Saudi Agricultural Bank, Public Investment Funds, Saudi Industrial Development Fund, and the Real Estate Fund (Al-Sahlawi, 1997). These banks were established by the government to provide funds for specific sectors. The loans offered by these banks typically financed mid- to long-term development projects at subsidised rates.

In the 1980s, the Saudi economy experienced two major incidents. One was the sharp rise in oil prices during 1979-1981 due to the Iran-Iraq war, and the second was the severe decline in oil prices in 1986 (Al-Suhaimi, 2001). These incidents affected the Saudi banking system in that Saudi banks substantially extended their lending in the early 1980s, backed by the increase in their balance sheets after the oil price hike. Many of these loans were made without adequate assessment and monitoring procedures. Consequently, when oil prices fell in 1986, many banks faced difficulties recovering their loans owing to the severe contraction in the domestic economy, mainly because of declining government revenues. (For instance, government revenues fell from SR333 billion in 1981 to SR74 billion in 1987). As a result, non-performing loans in the banking system increased sharply, amounting to 20 per cent of total loans by 1986. This, understandably, depressed bank profits on account of the substantial rise in loan loss provisions. However, these incidents helped discipline banks' lending activities and, by 1988, most banks had adequate provisions for doubtful loans, with average loan provisions increasing to more than 12 per cent of total lending (Banks for International Settlements, 2001).

Another noteworthy event during the 1980s was the near failure of the Saudi Cairo Bank resulting from unauthorised bullion trading during 1979 and 1981. Accumulated losses exceeded the bank's capital, forcing the authorities to intervene (Al-Suhaimi,

[25] For example, the Saudi British Bank and the Saudi American Bank.

2001). In response, SAMA directed the bank to issue new shares and double its capital by 1986, and the increase in capital was undertaken by the Saudi Public Investment Fund.

During the 1980s, various other national banks were established, including Al-Rajhi Banking and Investment Corporation (the largest money exchanger licensed as a full commercial bank), Saudi Investment Bank (authorised as a full commercial bank with foreign ownership reduced to 25 per cent and the remaining shares sold to the public), and the United Saudi Bank (formed after the take-over of three foreign banks). These banks contributed to the restructuring of the Saudi banking sector. Meanwhile, SAMA encouraged banks to strengthen their capital positions so as to improve the soundness of the system (Al-Sahlawi, 1997; Al-Jarrah, 2002).

Another major development during the 1980s was the introduction of government bonds that helped strengthen bank's investment portfolios. In addition, automated teller machines were introduced in order to advance the quality of banks' services to the public, and debit and credit card services became more widely available.

The decade of the 1990s commenced with a serious test to the Saudi banking system after the Iraqi invasion of Kuwait (Al-Sahlawi, 1997; Al-Jarrah, 2002). Banks faced substantial deposit withdrawals in August 1990, accounting for 11 per cent of total banking sector deposits and these were exchanged into foreign currencies. By the end of 1990 the withdrawals eased (declining to 1.1 per cent of total deposits) owing to the intervention by SAMA. The authorities provided the banking system with substantial liquidity in Saudi Riyal and foreign currencies through greater use of repo arrangements. This helped to stabilise the system and maintain a healthy banking system during these turbulent times.

From 1991 to 1995, domestic loans and advances increased by 90 per cent, and banking profitability indicators continued to show sustained improvement (Al-Sahlawi, 1997; Al-Jarrah, 2002). The second half of the 1990s witnessed a merger between the United

Saudi Commercial Bank and the Saudi Cairo Bank, to form the United Saudi Bank. The United Bank also merged with the Saudi American Bank in 1998. Moreover, Saudi banks continued to embrace operational development by investing in new technologies such as electronic funds transfer systems and by setting up widespread point-of-sale terminals.

The Saudi banking sector expanded during the 1990s. Banking credit grew by 147 per cent with an annual average growth rate of 11 per cent, amounting to $46.2 billion by 2000. Also, deposits rose by 73 per cent, reaching some $71.2 billion.[26] Moreover, the level of financial capital and reserves of the banking system reached $11.6 billion, mirroring an annual growth of 10 per cent over the 1990-2000 period. By 2000, total banking assets amounted to some $121.1 billion, when there were eleven commercial banks operating in Saudi Arabia, of which four were joint ventures with foreign banks. From mid 1975, no new foreign bank entities have been allowed to enter the Saudi banking system. However, in the move towards GCC financial sector integration, the International Gulf Bank of Bahrain and the Abu Dhabi National Bank of the UAE have been lately granted licenses to open branches on Saudi soil.

Overall, the Saudi financial system, the largest in the Gulf region, has witnessed a remarkable expansion in banking accompanied by ongoing updating and revision of its regulatory framework to ensure increased soundness and prudence in the banking system.

To summarise, before 1950, business in the GCC countries was dependent on self-financing and wealthy families controlled and financed most of the economic activities (e.g. pearling, cattle, and food trade). In the late 1950s and in the following decade, most GCC countries were under British protection agreements, and the region witnessed the establishment of foreign banks, of which many were of British origin. By the early 1960s, the structure of the GCC financial system started to grow through the establishment of national banks. The financial systems also set up currency boards responsible for the control of money supply, aiming to replace the British Pound

[26] Author's own calculation based on the GCC Secretariat General's Economic Bulletin, 2001.

Sterling and the Gulf Rupees that had been circulating in the region. In general, the 1960s and 1970s were characterized by the establishment of central banks as well as the issuance of national currencies. During the 1980s and 1990s, the number of banks operating in the region increased, improving the sophistication of financial activity. In addition, regulatory authorities started to place greater emphasis on banking sector soundness and prudential regulation. As the competitive environment heightened, this was accompanied by consolidation and gradual deregulation of various banking systems, a process which continues.

Finally, it is important to note the role played by foreign banks in GCC countries. This can be summarised as follows:

- Foreign banks were the first well-established institutions to conduct, facilitate, and provide various financial services in the region.
- Foreign banks helped facilitate foreign direct investment (for example, in oil companies) and channelled these flows throughout the region and also in the respective countries.
- Using their relatively more advanced technologies and banking experience, foreign banks assisted host country governments in administering their finances, helping them, at least in certain cases, to issue national currencies (as, for example, in Oman).
- Most importantly, foreign banks provided assistance and the impetus for host countries to establish their own indigenous local banking systems. Their continuous existence in the GCC region has been an important factor in maintaining the confidence of infant financial sectors. Moreover, although the foreign banks in the Saudi banking sector were converted to reflect a dominant share of Saudi ownership, the legislation granted foreign banks ownership and management over these converted banks and this helped maintain their performance and banking system stability.

The following subsection outlines the recent performance features of GCC countries' banking systems.

2.3.2 Recent performance and soundness of GCC banking systems

In this section we analyse a range of indicators reflecting the main financial features of GCC countries' banking systems. One broad type of indicator that can be used to judge the development of the banking system relates to the ratio of bank deposits to the size of the economy. Typically, one can examine the relationship between various monetary aggregates, such as M1 and M2, to GDP (see Figure 2.7).[27] [28] Total GCC countries' money supply in its narrower definition (M1) varied around 20 per cent of GDP over the period 1995-2000. The ratio M2/GDP, which also measures financial deepening, is relatively high, averaging around 50 and 60 per cent over the period. These measures reflect how the GCC banking sector is able to attract deposits, and this degree of monetisation reflects the high use of money (cash and banks accounts) in preference to other means of exchange. This also reflects increased confidence in the banking system and suggests a readiness to use technology to serve customer financial needs (Jbili, Galbis and Bisat, 1997).

[27] M1 consists of currency in circulation and demand deposits in local currency. M2 consists of M1 plus time deposits and deposits in foreign currencies (Qatar Central Bank, 2001).
[28] Each M1, M2, and GDP are summed across GCC countries.

Figure 2.7 Degree of monetisation in the GCC countries' banking systems, 1995-2000

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Source: GCC Secretariat General's Economic Bulletin, 2001.

Moreover, the contribution of the financial sector to GCC countries GDP increased from 4.5 per cent in 1990 to 6.6 per cent in 2000, reflecting an increase in importance of the sector. In 2000, the ratio of commercial bank assets to GDP in GCC countries suggested that the banking sector was relatively important since this ratio ranged from about 70 per cent in Saudi Arabia to 125 per cent in Bahrain (excluding OBUs assets) and these levels appear relatively high by international standards (Jbili, Galbis and Bisat, 1997) (see Figure 2.8).

Figure 2.8 GCC commercial bank assets relative to GCC countries' GDPs, 2000

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Sources: GCC Secretariat General's Economic Bulletin, 2001; Annual reports of GCC central banks, 2000.

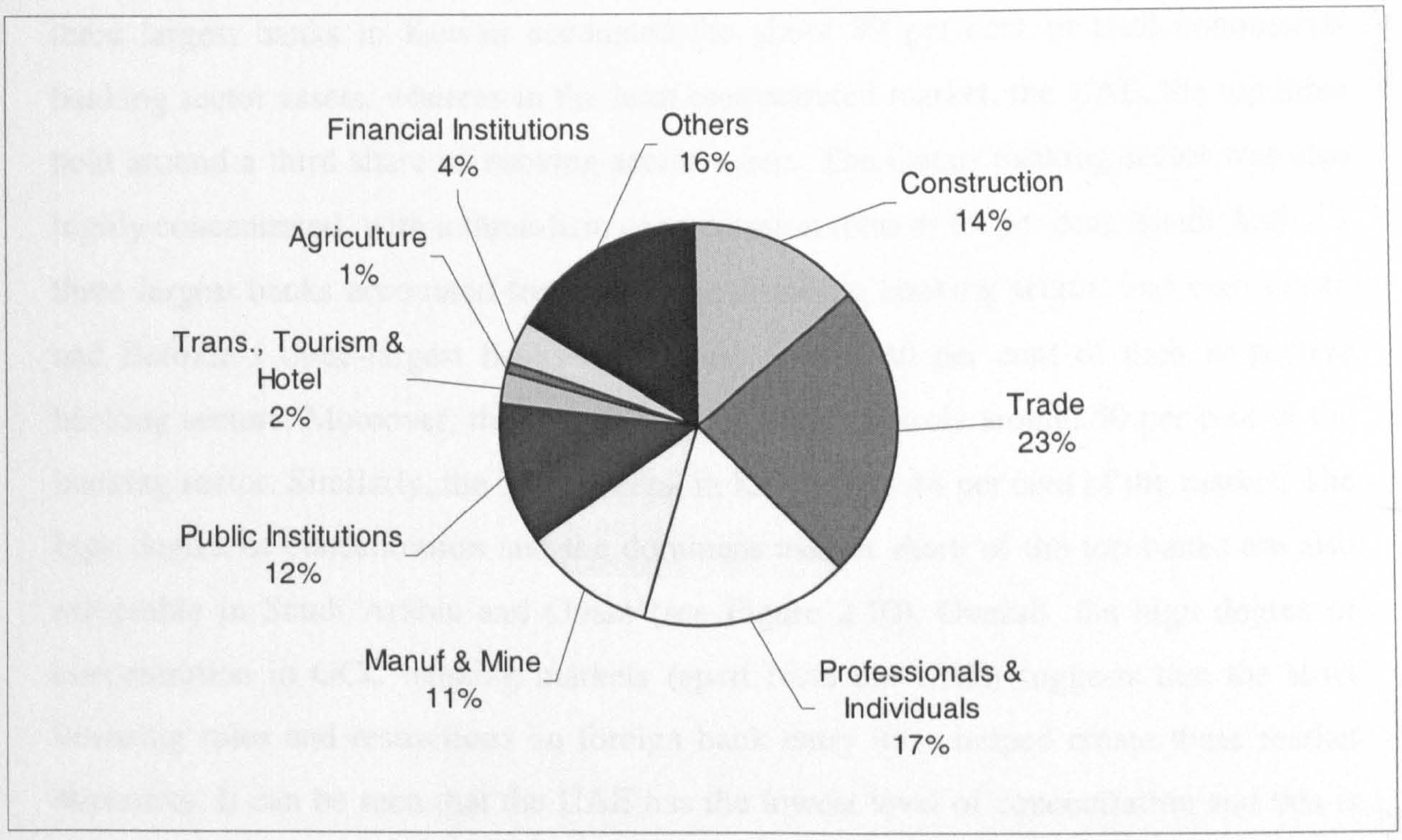
In terms of size, Table 2.5 shows that the Saudi commercial banking sector was the largest among those of the GCC countries. As of 2000, the figures (shown in Table 2.5) express the size of the GCC banking market in terms of assets, loans, and deposits.

Table 2.5 The size of the commercial banking sector across GCC states, 2000 ($ billion)

Sources: GCC Secretariat General's Economic Bulletin, 2001; Annual Reports of GCC central banks, 2000.

Bank lending was mainly concentrated in two economic sectors, trade and construction, that jointly occupied a stake of 37.4 per cent of total loans in 2000 (see Figure 2.9 and Table 2.6). Over the period 1995-2000, the GCC banks witnessed a remarkable average annual growth of 18.2 per cent in loans to the personal sector, the second position after the trade sector, with 17.1 per cent of total bank loans. This suggests the growing importance of retail banking in the Gulf (Jbili, Galbis and Bisat, 1997).

Figure 2.9 The share of commercial banks' credits to the economic sectors, 2000



Source: Author's own calculations based on GCC Secretariat General's Economic Bulletin, 2001.

Table 2.6 GCC Commercial Banks: Credit Structure (1995-2000)

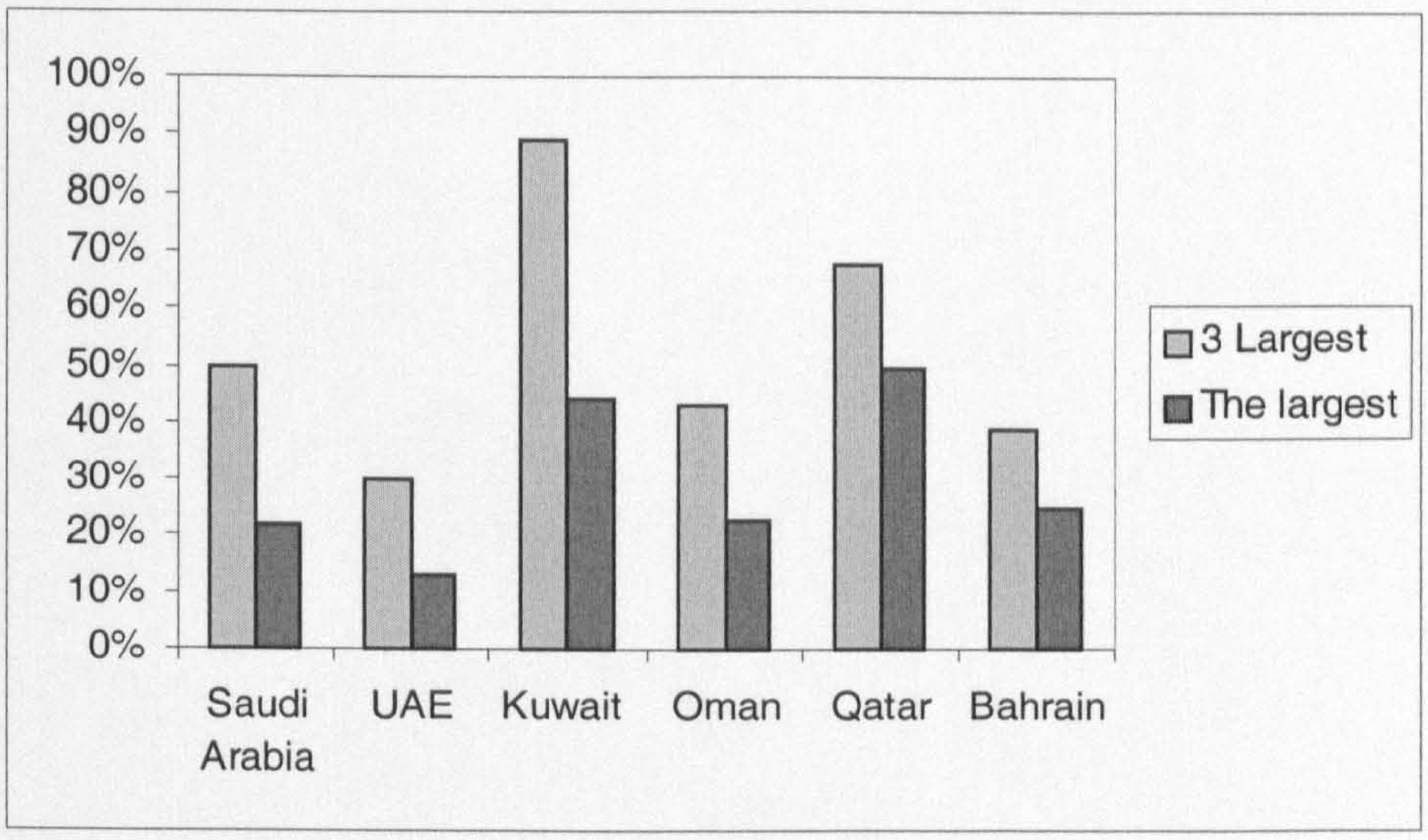
Year	Construction	Trade	Professionals & Individuals	Manuf. & Mining	Public Institutions	Trans., Tourism & Hotels	Agriculture	Financial Institutions	Others
1995	13.2%	23.3%	8.4%	9.3%	18.3%	2.6%	0.9%	3.9%	20.1%
1996	12.7%	24.2%	9.1%	8.9%	16.8%	2.0%	0.9%	4.4%	20.9%
1997	13.7%	25.3%	10.6%	9.2%	14.8%	1.9%	0.6%	3.4%	20.5%
1998	13.5%	25.0%	9.7%	10.1%	11.8%	1.6%	0.6%	3.3%	24.5%
1999	14.2%	23.6%	16.7%	11.0%	13.4%	2.6%	0.8%	3.4%	14.2%
2000	13.8%	23.6%	17.1%	10.6%	11.6%	2.5%	0.8%	3.8%	16.3%
Average annual growth (1995-2000)	1.1%	0.3%	18.2%	2.8%	-8.1%	2.9%	-2.2%	0.1%	-1.0%

Source: Author's own calculations based on GCC Secretariat General's Economic Bulletin, 2001.

The main source of deposits generated by commercial banks is interest-bearing accounts (mainly consisting of savings, time deposits, and foreign currency deposits), which account for 65 per cent of total commercial bank deposits, followed by demand deposits (27 per cent) and government deposits (7 per cent) as of 2000.

GCC banking industries are characterised by high market concentration. In 2000, the three largest banks in Kuwait accounted for about 89 per cent of total commercial banking sector assets, whereas in the least concentrated market, the UAE, the top three held around a third share of banking sector assets. The Qatari banking sector was also highly concentrated, with a three-firm concentration ratio of 69 per cent. Saudi Arabia's three largest banks accounted for half of the domestic banking sector, and both Oman and Bahrain's three-largest banks constituted around 40 per cent of their respective banking sectors. Moreover, the largest bank in Qatar controls around 50 per cent of the banking sector. Similarly, the biggest bank in Kuwait has 44 per cent of the market. The high degree of concentration and the dominant market share of the top banks are also noticeable in Saudi Arabia and Oman (see Figure 2.10). Overall, the high degree of concentration in GCC banking markets (apart from the UAE) suggests that the strict licensing rules and restrictions on foreign bank entry have helped create these market structures. It can be seen that the UAE has the lowest level of concentration and this is almost a consequence of laxity in restrictions on the licensing of domestic and foreign banks that increased the number of such institutions, especially in the late 1970s and 1980s, as was mentioned earlier (section 2.3.1).

Figure 2.10 The three largest and the largest banks' assets relative to total banking sector assets, 2000



Sources: Author's estimates calculated using Bankscope (January, 2002) and annual financial reports of foreign banks in Qatar and the UAE (2000).

The following indicators analyse the overall performance and soundness of GCC banks, including the foreign banks that operate in both Qatar and the UAE. This analysis is based mainly on the data sample that is used later in this thesis for the empirical evaluation of domestic and foreign bank efficiency comparisons detailed in Chapter 7. Table 2.7 summarises these indicators that reflect the growth, profitability, capital strength, and asset quality of GCC banks between 1995 and 2000.

Table 2.7 Indicators of GCC countries' commercial banks performance and soundness

The indicators	Qatar			UAE			Saudi	Kuwait	Bahrain	Oman	GCC
	Foreign	National	Overall	Foreign	National	Overall					
Growth indicators											
Total assets annual growth (1995-2000)	4.8%	5.6%	5.5%	9.0%	6.1%	6.7%	6.5%	3.5%	4.4%	16.9%	6.2%
Total deposits annual growth (1995-2000)	5.0%	5.7%	5.6%	8.5%	5.1%	8.2%	6.4%	3.2%	3.7%	14.9%	5.5%
Total loans annual growth (1995-2000)	5.6%	4.7%	4.8%	8.0%	5.7%	6.2%	6.1%	6.8%	4.4%	19.0%	6.9%
Equity annual growth (1995-2000)	5.7%	7.9%	7.6%	6.6%	7.7%	7.5%	5.7%	5.3%	5.0%	21.4%	6.7%
Profitability Indicators											
ROE (average 1995-2000)	45.0%	34.0%	39.5%	38.0%	24.1%	31.1%	19.4%	18.4%	21.1%	32.3%	35.4%
ROA (average 1995-2000)	10.0%	3.1%	6.6%	6.6%	4.6%	5.6%	1.9%	2.8%	3.4%	3.7%	4.8%
Capital strength											
Equity to total loans (average 1995-2000)	16.2%	15.9%	16.1%	20.0%	23.0%	21.5%	18.6%	20.8%	22.0%	16.9%	19.3%
Equity to total assets (average 1995-2000)	10.6%	10.1%	10.3%	15.3%	18.1%	16.7%	10.1%	15.4%	17.8%	11.9%	14.8%
Liquid assets to total assets (average 1995-2000)	7.6%	6.7%	6.8%	9.4%	8.6%	8.7%	6.1%	6.8%	4.4%	19.0%	6.9%
Asset Quality											
Loan Provisions annual growth (1995-2000)	38.6%	31.1%	34.9%	36.7%	5.3%	21.0%	42.9%	0.6%	37.3%	54.4%	31.8%
Loan to deposit ratio (average 1995-2000)	-	-	-	-	-	-	-	-	-	-	77%
Number of banks	8	6	14	24	19	43	9	10	11	6	93

Sources: Authors calculations using Bankscope (January, 2002) and annual financial reports of foreign banks in Qatar and the UAE (1995-2000).
Note: Annual growth (1995-2000) is the average of the rate of change from year to year taken for the whole period.
Average (1995-2000) is the average (or the mean) of the period.
Note: 'Overall' is the average for both national and foreign banks.

As mentioned earlier, GCC countries commercial banks experienced substantial growth during the 1990s. In general, whether one considers assets, loans, or deposits growth, annual growth rates range around the ten per cent level. In addition, the equity of the banking systems has increased at an annual rate of around 6 per cent. The growth of the banking system overall is also mirrored by the increased importance of foreign banks in the UAE. For instance, foreign banks in the UAE occupied 26 per cent of the total commercial banking sector in 2000, a percentage that increased from the 22 per cent in 1995. In this country, foreign bank assets have grown faster than national banks, experiencing an annual average growth rate of 9 per cent over the period (UAE national bank assets grew by 6.1 per cent annually over the same period).

In contrast, the share of foreign bank assets in Qatar's banking sector remained unchanged over the 1995-2000 period. While they accounted for 14 per cent of total commercial bank assets in 2000, their assets have grown at a slower rate compared with national banks. Foreign bank assets increased by 50 per cent with an annual growth of 4.8 per cent between 1995 and 2000, while national bank assets grew by 54 per cent, achieving an average annual growth rate of 5.6 per cent. Foreign bank deposits also increased by a lower percentage (51.2 per cent) and with lower average annual growth (5 per cent) compared with national banks. However, loans made by foreign banks increased at a faster rate than those of domestic banks, presumably a reflection of the involvement of these banks in financing major gas industry projects.

The widely used financial accounting profits indicators, return-on-equity (ROE) and return-on-assets (ROA), show that on average GCC banks performed very well in the second half of the 1990s. This was clearly reflected in the average ROEs of 35 per cent and ROAs of 4.5 per cent, which were high figures by international standards. Across GCC countries, banks in Qatar, the UAE, and Oman have generated the highest returns, although profitability elsewhere has been high. Moreover, ROE and ROA indicators show that foreign banks in both Qatar and the UAE generate higher profits than do their domestic peers. (In the empirical analysis in Chapter 7, the results indicate that foreign banks are more profit efficient than national banks).

The capital ratios (capital-to-loan and capital-to-assets) for Qatari and Omani banks are slightly lower than those of other GCC countries, but they are still high relative to those of many banks in the developed world. Overall, the equity to loans and the equity to assets ratios for the GCC countries (19.3 and 14.8 per cent respectively), show a strong capital position that maintains soundness in accordance with local and international guidelines.

In terms of banking sector liquidity, the ratio of liquid assets to total assets, which includes cash and due from banks, measures the ability of banks to meet unplanned withdrawals. This ratio shows a very high level of liquidity in the Omani banking sector, as it averaged 19 per cent over the period 1995-2000. Liquidity ratios vary between 6 and 10 per cent in most other GCC banking systems, with Bahraini banks having the lowest level of liquidity at 4.4 per cent.[29]

The average annual growth of loan loss provisions shows high levels for GCC banks; and the annual growth of loan provisions of foreign banks operating in Qatar and the UAE is even greater than that of their national peers. Generally, this indicates the presence of loan problems facing GCC banks during this period, probably a reflection of non-performing loans resulting from the substantial economic downturn and contraction of expenditure in 1998. Moreover, in Qatar, the default of a major borrower firm of one of Qatar's banks in the late 1990s induced the central bank to take measures assuring the adequacy of provisions to meet possible future defaults in the banking sector.

Despite the high level of loan loss provisioning, the GCC banking sectors have expanded and shown positive performance during the second half of the 1990s. However, an ongoing regulatory process aimed at strengthening prudent regulation and improving loan assessment methods continues to take place in order to ensure that GCC

[29] In general, these levels of liquidity are higher than those of the average US bank, typically around 5 per cent in 2000, but lower than those of European banks, where ratios vary around 20 per cent plus (Bankscope, 2003 estimates).

banks comply with both domestic and international safety and soundness standards. The following section sheds light on the various achievements of GCC countries in realising the aim of economic and monetary union.

2.3.3 Financial integration of the GCC countries

The formation of the GCC by Arab Gulf countries aims to establish a foundation for cooperation between countries that will lead to greater economic convergence and a more unified and integrated market. From the date the GCC agreement was signed (1981), negotiations have commenced aimed at increasing the free flow of products and factors of production within GCC countries. The council's negotiations yielded an agreement, signed in 1999, aimed at unifying trade customs charges. This agreement took effect from January 2003, and under it all products entering the GCC zone will face a unified customs rate. This is expected to increase non-price competition within the GCC zone by encouraging each country to improve their ports and alter trade facilities (by attracting higher volumes, cheaper warehousing services, and so on) so that the cost of imports may foster re-exporting business.

A major part of the GCC economic integration programme focuses on the creation of an economic and monetary union. In achieving this goal, the GCC has agreed to introduce a unified currency by 2010. Certain steps have clearly been accomplished that will gradually help pave the way to establishing the unified currency. For example, GCC countries have completed a project linking all ATM networks throughout the region. In essence, residents within the GCC countries will be able to obtain money from their own bank accounts at the same cost they pay in their own countries and at the same official exchange rate of GCC countries' currencies. Moreover, the GCC countries agreed at their last Omani summit (in 2000) to establish a timetable that enables them to adopt the Dollar as a currency to which all current GCC countries' currencies' will be pegged. (This was in place at start of 2003 as Kuwait, the only GCC country adopting a basket of currency, pegged its Dinar currency to the US Dollar commencing January 2003).

The committee of GCC central bank governors is also currently studying ways in which to develop GCC capital markets and especially bond markets, because of their expected positive effect on attracting investment and enhancing monetary policy tools. Moreover, with the aim of encouraging GCC banks to expand regionally, the GCC summit of 2000 issued a resolution urging central banks to allow banks from GCC countries to open branches throughout the region. This calls for GCC countries to change their local laws in order to permit greater bank entry. The impact of this resolution is already bearing fruit. For example, branches of a bank from the UAE have been opened in Saudi Arabia and Bahrain. Bahrain has also permitted the establishment of branches from Oman and the UAE. Qatar also has one branch of the UAE bank. In fact, the phased opening up of GCC banking markets should foster greater competitiveness and this may encourage increased mergers and alliances between banks within the region.

2.4 Conclusions

This chapter has provided an overview of GCC countries economies and their banking systems. The first part of the chapter has outlined the history of GCC economies and banking systems, their characteristics and various recent developments. The chapter has also analysed the performance and soundness of banks in different GCC countries. Finally, the chapter has outlined various important developments that are aimed at achieving GCC economic and monetary union. While foreign banks were instrumental in establishing domestic banking systems within the Gulf, their role has been somewhat curtailed as domestic governments have sought to nurture their own indigenous banking systems. The recent moves to create economic and monetary union within the Gulf are likely to encourage a greater role for foreign banks (from different GCC countries at least). This being so, the following chapter addresses the issue of the role of overseas banks, and the reasons explaining their performance and growth across borders.

CHAPTER 3

THE ROLE OF FOREIGN BANKS: A LITERATURE REVIEW

3.1 Introduction

Over the last two decades, the presence of foreign banks has been increasing in most financial systems in the world. For example, from 1995 to 2000, foreign bank assets increased to more than 40% of the total assets of a number of Latin American banking systems, for instance, in Argentina, Mexico, Peru, and Chile (Crystal et al., 2001). In transition economies, over 50% of banking industries' assets are controlled by foreign banks, operating in Hungary, the Czech Republic, and Poland (Clarke et al., 2001). In other economies, such as those of Middle Eastern, Asian, and African countries, the advent of foreign banks has been slower, but their growth has been experiencing an upward trend (Clarke et al., 2001).

This chapter attempts to shed light on the reasons for the growth of foreign banks' presence in many banking sectors around the world. To this end, the chapter reviews the literature containing studies undertaken on foreign banks. Overall, the main issues examined in the foreign bank literature are: the reasons behind foreign banks' existence and entry into foreign economies, the determinants of foreign banks' performance and growth, and comparisons between foreign and domestic banks' performance from the experience of both developed and developing countries.

The chapter is organized as follows. Section 3.2 provides a brief definition of both the theory of banks and international banking. Section 3.3 addresses the reasons why banks establish entities beyond their home countries. Section 3.4 discusses the different forms of the presence of foreign banks. Section 3.5 explains the types of foreign bank activities in host countries. Section 3.6 presents several reasons why foreign banks' entry is either favoured or restricted. Section 3.7 sheds light on the issue of the performance and efficiency of foreign banks in developed and developing countries. Section 3.8 contains the conclusions.

3.2 The theory of banks and international banking

As the consideration is of basic relevance to the banking literature, we start with the theory that explains why banks exist (see Heffernan, 1996). It is believed that the chief driving force promoting the establishment of banks is their owners' recognition of the profits that are to be made from resolving the difficulties of meeting the needs of lenders and borrowers when conducting credit business directly. To explain, lenders want to increase their income by channelling their excess money into activities that increase their wealth, but they face problems such as loan default, lack of information, search costs, and monitoring and screening of credit recipients' activities. Likewise, these recipients, or borrowers, want funds to meet their current consumption, to finance projects, or to exploit investment opportunities that generate profits. However, borrowers also face difficulties such as search costs, lack of lenders' confidence, and difficulties of pooling funds to meet needed large amounts. The difficulties burdening both lenders and borrowers render the credit market imperfection, which may lead to deceleration in the growth of the economy. These difficulties encourage banks to exist and make profits through meeting the needs of both lenders and borrowers by engaging in the business of intermediation.

In the intermediation[1] business, banks assure lenders of safety and returns on their funds. By means of this assurance, banks pool funds from lenders and lend these to borrowers whom banks charge them a rate that covers banks' costs such as interest paid to lenders (depositors), information costs (monitoring and screening), costs arising from the risks of default and inflation, as a well as a margin for bank profits. Since banks deal

[1] Chapter 4 elaborates the intermediation role of banks and relates this to the efficiency of the financial system.

with a high volume of lenders and borrowers, they are able to minimize these costs through economies of scale, the employment of expertise, and the use of technology, making banks' intermediation business more attractive to society than business carried out directly by individuals.

Therefore, where banks exist, although borrowers and lenders do not know each other, banks' intermediation between them alleviates various difficulties that would otherwise create market imperfections and hinder the smooth channelling of funds. Thus, the existence of banks will benefit the economy by inducing more funds to be channelled into the economy, fostering its growth.

In moving to the theory that explains the international presence of banks, it is much simpler if we first define international banking. This is because there are many aspects of banking activities that take on an international dimension and may complicate our task of establishing a theory accounting for the international presence of banks. Aliber (1984) defines international banking using three criteria: location, currency denomination of loans and deposits, and the nationality of both customers and banks. Under the first criterion, the location, a bank is considered international if it uses its foreign entities to conduct banking activities in the host country. Under the second criterion, the currency of denomination, a bank is said to be international when its deposits and loans are denominated in foreign currency and not in the currency of the country the bank is locating in. As for the third criterion, the nationality, when a bank engages in activities with customers who do not belong to the same country of origin as the bank, then the bank in this case would be called international.

Heffernan (1996) has attempted a much clearer definition of international banking. She argues that previous definitions do not explain the dividing line between international banks' activities and international banks' presence; it is, she asserts, important to know what theory to employ when explaining these aspects of international banking. Heffernan therefore argues that the theory of international banking should cover two areas: international banking services, which are explained by the theory of international trade; and multinational banking, which is explained by the theory of multinational enterprise. International banking services include all sorts of banking engagements in transactions that involve foreign countries. For example, dealing with currency exchange business, individuals' use of credit cards issued by their home banks to make purchases in foreign countries, and banks' use of the interbank market as a window for borrowing from banks located within or across national boundaries are all aspects of international banking services undertaken by local banks. International banking services are, then, better explained in the theory of international trade as, on account of their ability to collect information and reduce transaction costs through overcoming market imperfections, banks have comparative and competitive advantages over individuals, and sometimes also other banks, in functioning as intermediaries between trans-national borrowers and lenders, in trading in foreign exchange, and in engaging in international portfolio business.

A more specific sort of international banking, which is the focus of this thesis, is when a local bank extends its activities across borders by providing banking services through a physical representation, leading to such a bank becoming known as a multinational firm. Multinational banking is, then, better explained by the theory of multinational enterprise, which focuses on the entry determinants, such as entry barriers and location, which motivate banks to establish entities to undertake banking business overseas. Therefore, a bank is defined as multinational (hereafter a foreign bank) when it owns and controls a banking entity in more than one country (Casson, 1989). In practice, banks are known as foreign banks when non-citizens own 51 per cent or more of the shares of such banks located in host countries. [2]

Historically, foreign banking coexisted with the colonial age, developing rapidly from the mid nineteenth century onwards (Heffernan, 1996). In early days, foreign branches functioned to facilitate the finance of imports and exports between colonies and the mother country. They also existed in order that foreign branches could help the parent banks to obtain information on their international customers' financial status and projects, as well as the provision of payment facilities. In fact, foreign banks operating in the colonies contributed to the formation of the financial systems of many colonies,

[2] This definition is a default in the Bankscope definition of a foreign bank (Bankscope, January 2002).

since most colonies did not have their own financial systems. Most of these banks remained in the colonies after independence, but the major growth of foreign banks occurred in the post-World-War-II period and onwards.

In more recent times, it has been observed that advances in information and communication technologies have enabled banks to perform international banking without really having physical representation in the target country. However, many determinants lead banks to seek a physical existence in the international banking industries. These are discussed in the next section.

3.3 Why do foreign banks exist?

As previously noted, the theory of multinational banking focuses on the determinants/factors motivating banks to extend their activities across national borders. In order to provide a deeper understanding of the factors stimulating banks to tap into international banking business, an overview of the reasons provided in the literature will be presented here. These fall into two main categories: (1) reasons specific to changes in the international financial environment; and (2) reasons specific to banks and financial systems.

3.3.1 Entry reasons specific to changes in the international financial environment

The international financial environment can influence the evolution and development of a country's financial and banking system. Environmental forces, such as political and social factors, as well as a county's aptitude to embrace international trade, can all be factors fostering foreign banks' presence. Other factors, such as aspects of globalization, rate of return divergence, and geographical distance, also influence foreign banks' entry.

As an example, the political environment is an important element influencing economic relations between nations; when political relations improve, trade and market openness should increase and a bilateral bank presence may also become more apparent. Moreover, political stability promotes stability and certainty in financial systems, which also encourage banks' presence.[3] Similarly, social factors such as culture, language, and religion can affect the extent of closeness between countries and can facilitate a greater foreign bank presence. For instance, Islamic banks have a greater presence in Islamic rather than non-Islamic countries. This may be due to the easy understanding of social and cultural needs of countries with similar cultures. Thus, the ability to overcome cultural differences is an advantage that stimulates the foreign bank entry option (Bain, 1956). Understanding of the culture and social environment and an appreciation of societies' credit demands may, then, reduce the cost of making non-socially desirable financial products which may conflict with social tastes.

Greater political and economic certainty and a high degree of social similarities tend to facilitate trade between countries, which can be a source of attraction for foreign banks. Increased trade between countries induces the establishment of foreign banks to provide trade credit, foreign exchange, and payment facilities, reducing transaction costs and increasing the level of trade. The studies by Goldberg and Saunders (1981a and 1981b), for instance, find a positive and statistically significant relationship between the presence of foreign banks in the US and bilateral trade between the US and the foreign banks' countries of origin. Fisher and Molyneux (1996) have also found that banks' presence in London between 1980 and 1989 was greater in the case of the main trading partners of UK in international trade. This indicates that banks' presence in overseas countries could be a sign of an existing, and indeed increasing trade relationship.

Related to the trade argument is the issue of globalization, which leads to increased market interdependencies and linkages (Rybcynski, 1988). Aspects of globalization, such as advances in technology and communications, as well as changes in international financial regulations, have helped foster foreign banks' presence (Berger, DeYoung, and Udell, 2000). Advances in technology and means of communication have enabled

[3] Fisher and Molyneux (1996) stress that a more stable banking industry plays a greater role in attracting foreign banks. Moreover, Grosse and Goldberg (1991) find evidence from the US financial system supporting the hypothesis that the more risk there is available in the foreign bank's country of origin, the more the foreign bank's presence will be witnessed in low-risk financial systems.

banks to pursue financial deals and activities promptly and with low costs through advanced financial networks. The changes in international financial regulations towards more financial system openness have been promoted by the World Trade Organisation (WTO). Members of the WTO are expected to remove all barriers on banks' entry in order to establish equal opportunities and level bases of competition between foreign and domestic banks (Krugman and Obstfeld, 1997, p. 240). Under this globalized environment, banks find themselves facing fewer restrictions on their activities and more attractions for establishing a presence in foreign countries.

The variance of the interest rate margin across countries is another reason for banks to expand across national boundaries. In general, when the interest rate margin narrows within a country, local banks seek overseas banking industries that have more attractive interest margins and profit opportunities. Conversely, banks with greater ability to offer a narrower interest margin in a country that has a wider interest margin may enter this country's banking sector and offer a more competitive interest rate margin that extracts profits and market shares. For example, Darby (1986) finds that among the reasons explaining foreign banks' entry into the US banking industry is the interest rate differential of Dollar-denominated assets that attract banks to offer cheaper rates compared to the available higher rates in the banking industry. However, equalization or convergence of the rate of return across international markets may reduce the incentives for entry; nevertheless, market segmentation, exchange rate variation, and regulatory asymmetries across countries can generate variations in the rate of return, which can also stimulate banks' entry.

The geographic location is also viewed as an important factor promoting foreign banks' entry. It can be argued that countries located in one region and close to each other may show more progress in the direction of market openness and integration. In essence, various studies have tested whether geographical distance has an influence on determining foreign banks' presence. Goldberg and Saunders (1981a, 1981b), Hultman and McGee (1989), and Grosse and Goldberg (1991), for instance, find a positive correlation between foreign banks' presence and their countries' geographical distance from the host countries. That is, the shorter the geographic distances are between countries, the greater degree of foreign banks' presence is found.

The environmental factors influencing financial systems are important for stimulating banks' presence. In addition, the next section reviews another range of factors, belonging to bank and financial system specifics, which also have an influence on foreign banks' decision to enter overseas financial systems.

3.3.2 Entry reasons specific to banks and financial systems

There are several factors that are specific to banks and financial systems which encourage banks to launch into multinational business. These can be addressed as follows.

Following customers

Banks may be motivated to target their customers who are home countries' citizens or immigrants based in foreign countries. For example, the existence of foreign banks of home countries of a significant number of the workforce will probably help in attracting the payments and remittance business (Casson, 1989). In addition, the need of multinational corporations for smooth banking facilities in a host country may attract home country banks to open entities capable of providing such corporations with access to banking services and finance better than can be found elsewhere. Therefore, the opportunity for banks to follow their customers abroad will protect their assets and induce them to expand, as the growth of the activities of their customers based in host countries can increase the foreign bank's range of operations.

One of the important advantages that foreign banks obtain by following their customers overseas connects to information gathering relating to home-country customers. Foreign banks have better knowledge of their customers' status and needs than do domestic banks (Coulbeck, 1984; Hempel & Simenson, 1999). Because of a lack of information, domestic banks might be hesitant to lend to foreign customers and this may create an incentive for overseas banks to follow their clients in the new markets.

Another advantage of banks' following their customers overseas is the reduced cost of collecting information on customers, projects, and macroeconomic conditions of the host country. The cost of collecting this information and monitoring customers' credit allocation is expected to be far less if a bank has a foreign entity than if this information has to be collected from the home country (Heffernan, 1996).

From a psychological viewpoint, bank customers who have multinational operations may feel more comfortable in dealing with their home banks operating in the same host country because of transaction benefits associated with the bank-customer relationships. Moreover, it may be easier for clients to access account information and track financial records. This may make it more attractive for customers to deal with their home bank entities based overseas, as most records needed by them can be obtained from the parent bank. By contrast, domestic banks, which have no records relating to new foreign entrant customers, may find it necessary to ask foreign customers for more formal references and personal information, making it less attractive to deal with overseas customers, in the short run at least. Thus, both customers and banks of the same origin already enjoy existing relationships that may have been fostered over a long period. Such relationships will be reflected in the prompt extension of facilities to customers based overseas.

It can, however, be argued that foreign banks are keen to build bank-customer relationships with domestic customers as well. The expansion of their banking market share to include non-home country borrowers may be another purpose, as foreign banks' expansion of their business beyond their home-country customers gives them enhanced ability to exploit potential economies of scale and to obtain more long-term bank-customer relationships.

On the other hand, foreign banks may have little information on which to base their decision, especially when a financial system is underdeveloped and information is not readily available at a public level. Information about the industry and economy of a country and what success may be anticipated of projects in a foreign country is generally more accessible to domestic banks than to their foreign peers (Hymer, 1970). Moreover, foreign banks may incur greater costs than local banks in studying project feasibility. In this case, domestic banks may enjoy an information advantage and become the providers of banking services to overseas customers when lack of information or high costs of information gathering stand as a barrier in the way of foreign banks' entry.

Various empirical studies have sought to investigate whether the hypothesis that foreign banks follow home customers overseas is substantiated or not. Usually, the test of this hypothesis is based on an examination of elements such as bilateral trade and foreign direct investment. However, in most cases, the 'following-customer' hypothesis is linked to the relationship between foreign bank entry and foreign direct investment (FDI). For instance, in the investigation of Cho, Krishnan, and Nigh (1989) of the determinants of US banks' branching presence internationally, their results are mainly supportive of the hypothesis that the presence of US bank branches overseas is motivated by the presence of US businesses. Moreover, Grosse and Goldberg (1991) find a positive relationship between foreign investment and the presence of foreign banks in the US over the period 1989-1987. Fisher and Molyneux (1996) study FDI from Europe and Japan and find that FDI from these countries is positively correlated with the size of the presence of foreign banks located in London. A study by Miller and Parkhe (1998) of US banks active in thirty-two countries during 1987-1995 finds that foreign bank entry increased as FDI increased. However, their findings do not hold for tests including developing countries. Moreover, some studies have used more precise measures, such as lending patterns, to test the 'following-customer' hypothesis. For example, Seth, Nolle, and Mohanty (1998) analyse the lending activities of non-US banks (from Japan, the UK, Canada, the Netherlands, Germany, and France) that are located in the US banking industry. They find that foreign banks from four out of the six countries (Japan, the UK, Canada, and the Netherlands) allocated most of their lending to non-home country borrowers during the period 1981-1992. This evidence leads the authors to conclude that the hypothesis of foreign banks' following their customers overseas is not so apparent in the US market.

Overall, the majority of the aforementioned findings relatively show strong evidence of an existing relationship between FDI and foreign bank entry. It may be concluded that, by taking FDI as a proxy to test the 'following-customer' hypothesis, this hypothesis seems to hold and is most applicable to foreign banks' entry in developed countries.[4]

Regulatory asymmetries

Generally, regulation could have an impact on a foreign banks' decision regarding foreign entry. For example, banks tend to locate in financial systems with less tight regulations. The study of Goldberg and Grosse (1994) across US states leads to the conclusion that foreign banks' presence grew more in states with less strict regulations on foreign activities. Therefore, regulatory asymmetries may encourage banks to shift operations from restricted local banking activities to a more flexible regulatory environment in which banks can extend their banking activities (Fieleke, 1977; Goldberg and Saunders, 1988 a & b). For instance, an important policy tool to control the supply of loans is the discount rate. When this rate is high in one country, a bank's short-run cost of loans from the central bank to meet short-run liquidity will also be high. The regulatory environment in which high discount rates prevail may be a constraint on bank lending (Ramchander et al., 1999), which may increase the incentive to search for other, less restricted markets.

Regulatory asymmetries can encourage banks to circumvent restrictions they face in their domestic banking system. The US and the British governments, for instance, impose reserve requirements only on their banks' domestic currency business (Krugman and Obtsfeld, 1997). This, therefore, can encourage local banks to attract deposits denominated in widely internationally demanded currencies and then they can use these for funding purposes on more attractive terms. More importantly, since the British government does not impose reserve requirements on Dollar deposits within UK borders, Dollar deposits of US bank branches in London are not subject to reserve requirements by the British government. Because the Dollar-denominated deposits (of

[4] Miller and Parkhe (1998) note that the absent relationship between FDI and foreign bank presence in developing countries might be due to the retrenchment of the level of FDI from the USA to Latin American countries after suffering the debt crises in the mid 1980s.

US banks in foreign countries) are also not subject to US reserve requirements as long as they are not payable in the US, US banks will be encouraged to establish branches in the UK. In this respect, Frankel and Morgan (1992) have noted that asymmetric deregulation of banking industries across countries can enhance the competitiveness of foreign banks and stimulate entry. Thus, locations where regulations may hinder banks' achieving maximum profits in certain activities, they search for other markets with regulations that fit the profit maximization goal. In these international markets, branches of foreign banks can circumvent the restrictions in their local markets and undertake activities prohibited in the home countries.

The size of banks and the banking sector

Countries with large banking sectors tend to have more domestic banks extending their services overseas. Among the results in the study of Grosse and Goldberg (1991), they find that foreign banks' presence in the US during 1980-1987 was positively correlated with the size of the banking systems of the countries from which foreign banks originally came. In addition, Fisher and Molyneux (1996) find that, over the period from 1980 to 1989, the size of foreign bank branches located in London was positively correlated to the size of their home countries' banking systems. They stress that this finding implies that countries with large banking systems should be the focus of the issue of reciprocity arguments from a regulatory perspective.

Furthermore, several other studies have investigated the relationship between the size of parent banks and their foreign presence overseas. This relationship is found to be positive in Tschoegel's (1983) study of 100 of the largest international banks. Similarly, Ursacki and Vertinsky (1992) find that the number of Japanese banks' branches operating in Korea and the value of their parent banks' assets in Japan exhibited a positive relationship. Confirmation of this finding comes also in the study of Focarelli and Pozzolo (2000), where a positive correlation between bank size and the degree of internationalisation is found.

Many reasons may explain why the size of banks and their foreign presence tend to be positively correlated. One reason is that when the growth of a bank's market share at home is limited, overseas expansion may provide greater opportunities. Therefore, diseconomies of scale in a local banking market can become a burden when banks' activities are not expanded (Clarke et al., 2001). Another reason may be bank market capitalization, which can be a factor stimulating cross-border expansion (Bear, 1990). High market capitalization in the home market gives a bank greater ability to develop its foreign operations, given that it has the resources to fund expansions. Banks with high market capitalization at home will be able to undertake more risky and low return loans, as well as to enter more competitive markets. An additional reason could relate to the size of multinational enterprises which may rely on large banks to provide the funds that suit their investments. Thus, large home banks, rather than small banks, are more able to fund large enterprises through their foreign entities (Focarelli and Pozzolo, 2000; Berger, Kashyap, and Scalise, 1995).

International diversification

Banks tend to spread risk by diversifying not only the supply, but the extension of their services geographically (Berger, 2000). In fact, banks based in a single country are more likely to be exposed to risks of cyclical fluctuations in the economy. Hence, when establishing foreign entities across regions or countries, adverse effects to a bank's business arising from the negative performance of (say) one of its branches in a particular geographical locality may be offset by the better performance of branches in other regions, assuming efficient diversification.

Although it may be costly and not profit-maximizing for banks to enter international markets, they may seek diversification benefits and therefore be willing to forgo some short-term profits in order to secure a platform on which to practise their banking activities in a foreign country in the hope of an improved long-term performance. For example, a reason put forward as to why banks locate in the major financial centres of New York and London is, in part, the importance these centres play in foreign exchange trading, as well as their role in the international economy. Foreign banks may seek to locate in these centres to diversify their operations while seeking access to global capital and money market businesses (Ramchander et al., 1999).

Intangible assets possession

Some banks find it worthwhile to use their intangible assets to enter the foreign banking sector (Heffernan, 1996). Intangible assets are not tradable but can be used to make profits and attract customers looking to reduce their search costs. Intangible assets, such as high-qualification expertise specializing in a particular financial service, could encourage a bank to enter a banking sector with a growing market for services that make intensive use of such expertise. Moreover, another intangible asset, reputation, can also be a reason to facilitate entry and attract customers of a target country, who are convinced of a particular bank's high standards in conducting financial services in terms of efficiency and reliability.

All and any of the factors discussed in this section may have an important influence on banks' decisions to tap into foreign banking sectors. Once banks find it possible to enter a banking sector, the form of entry through which they can practise their activities should also be considered, and this is the topic of our next section.

3.4 The form and the mode of foreign bank entry

This section discusses the forms and ways through which foreign banks may enter host countries' banking sectors. There are three main organizational forms in which foreign banks conduct their businesses in host countries: agencies, subsidiaries, and branches (Krugman and Obstfeld, 1997; Clarke et al., 2001). An agency (or a representative office) is the simplest physical form of organization in which banks can establish abroad. It is a basic banking service office established to engage in activities such as loan arrangements and fund transfers. Banks may find it worthwhile to establish an agency at the beginning of their presence as a means of examining the local market

(Clarke et al., 2001), searching investment opportunities, collecting information, and building customer-bank relationships. However, regulations placed on activities of this type of foreign bank representation do not usually permit acceptance deposits from domestic residents. Usually, agencies' funding comes from deposits of foreign residents, funds from parent banks, and funds borrowed from the interbank market. In addition, agencies are generally not allowed to make consumer loans; their funds are therefore almost always directed towards loans for commercial and industrial sectors (Clarke et al., 2001).

Since an agency is the most restricted instrument by which to conduct banking business, branches are more important in terms of permitted broader banking activities. In fact, branches stand in a mid-way position between agencies and subsidiaries, at least in terms of restrictions. Branches are recognized as the most important form of foreign bank representation in many countries, because their assets usually occupy the dominant share in total foreign bank assets. [5] However, foreign branches might be subject to the regulations of both the host and the home country. Moreover, branches may direct their activities into wholesale banking. [6]

Subsidiaries are banks located in foreign countries and are typically controlled by the home country parent bank. Subsidiaries are usually subject to the same regulations that are imposed on local banks and are not usually bound by regulations imposed on their parent banks in their home countries. Therefore, subsidiaries are the organizational form that has the nearest to equal footing with domestic bank businesses (Clarke et al., 2001) and thus it is the most flexible form of foreign bank establishment, practising a wider range of banking services. On account of this, as Miller and Parkhe (1998) note, countries that allow universal banking have a higher percentage of subsidiaries in comparison to others. Moreover, in contrast to agencies and branches, retail banking comprises the major portfolio of subsidiaries' loans in the US banking sector (Clarke et al., 2001)

[5] For example, in 1989, foreign branches' assets accounted for about 63 per cent of the total assets of foreign banks located in the US (Goldberg, 1992).
[6] Miller and Parkhe (1998) note that, in most cases, foreign branches in the US conduct wholesale banking.

Instead of setting up from scratch an institutional representation of a bank in a host country, banks may prefer to enter foreign banking industries through acquisitions or joint ventures. Acquisition of an institution could be possible through using funds to buy the stocks of an existing domestic bank (Mayer, 1997; Clarke et al., 2001). This form of entry may be preferable, especially when barriers to entry stand in the way of a bank's plans to enter a profit-promising banking sector. One of the benefits of entry through acquisition is that banks may better understand the domestic financial system and can establish relations with market participants faster than in the case of a new establishment. However, a possible disadvantage of acquisition may be the difficulty for foreign entrants to turn a failing bank into a success (Peek et al., 2000). Difficulties such as loan quality and the problem of accommodating a foreign entrant's managerial conduct may act as a constraint on improving the performance of the target bank.

Another mode of foreign bank entry may take the form of joint ventures (Kogut, 1991). This method of entry may also be considered as a way of circumventing imposing entry barriers. The benefit of joint ventures is that foreign entrants may cut the cost of information, especially when they join with domestic investors. This form of entry may also allow foreign banks to benefit from economies of scale. However, joint venture entry typically restricts the control of foreign entrants as ownership has to be shared (often equally) with a domestic bank.

In essence, therefore foreign bank entry could take many forms. These forms may depend on the size and the type of activities the parent bank intends to undertake in the host country, as well as the regulatory environment in which the foreign bank can accommodate its activities. The next section tackles the issue of foreign bank activities and how these activities are related to the domestic banking sector and economy.

3.5 The pattern of foreign bank activities

Foreign banks generally engage in both wholesale and/or retail banking services (Heffernan, 1996). Wholesale services include corporate loans, investment consultation, and Euromarkets participation. International retail services involve foreign exchange facilities, international investment facilities for personal customers, and global automated teller machines. In general, foreign banks are observed to engage much more in wholesale than in retail banking. For instance, Houpt (1980), Goldberg (1981), and Hodgkins and Goldberg (1981) find that, in the decade of the 1970s, foreign banks in the US directed a small percentage of their loans into residential mortgages and consumer loans. In a study on Japanese banks in the state of California, Zimmerman (1989) finds that the Japanese banks have mostly engaged in wholesale lending and money market activities. Among the available studies dealing with developing countries, that of Clarke et al. (2000) finds that foreign banks operating in Argentina are heavily concentrated in areas which they traditionally have come to target, such as the manufacturing sector.

Other studies that analyse the differences between foreign and domestic banks operations have also found that the former engage more in wholesale banking and allocate a greater proportion of loans to industrial and large corporations than do domestic banks. Goldberg (1981) observes that domestic banks are more oriented to retail lending, while foreign banks' lending is concentrated in wholesale business, on the basis that their activities consist of dealings mainly with multinational firms operating in the US. In a later study, Goldberg (2001) asserts that, in the US, foreign banks make 28.5 per cent of the loans for commercial and industrial sectors in the US and the total assets of these foreign banks form only 25 per cent of total banking assets. In Argentina, as Clarke et al. (2000) indicate, in the late 1990s, about 35% of foreign banks' loans went to the manufacturing sector, while domestic banks deployed less than 20% of their loans in the same sector.

There are many reasons adduced to explain why foreign banks typically choose to undertake wholesale lending and transact business with large corporations than focusing on consumer lending.

First, barriers to entry for wholesale banking are usually smaller than those for retail banking (Heffernan, 1996). For example, the lower requirements on foreign banks' capital in the US banking sector have permitted foreign banks to gain a cost of capital advantage that enables them to make large loans with prices lower than those of domestic banks (Zimmer and McCauley, 1991; Goldberg, 1992; Terrell, 1993). [7]

Second, wholesale banking requires less network facilities with the parent banks and less capital to recover in case of defaults. Moreover, the tendency towards wholesale lending is driven by the competitive endowments that foreign banks are equipped with, such as their ongoing access to funds from the parent banks.

Third, since it is observed that foreign banks tend to make wholesale loans, this may indicate that the size of these foreign banks is large. [8] This means that they are more likely to make wholesale loans to large rather than to small businesses. Keeton (1995) suggests two reasons why large banks deal with large rather than small businesses. These reasons may apply to foreign banks, given that Focarelli and Pozzolo (2000) indicate that banks that extend their operations into international industries tend to be large. One reason is that large banks are able to concentrate their loans to single borrowers with large borrowed funds. That is, large banks are less likely, in comparison to small banks, to be under the constraint of exceeding the ceiling of a share of a loan to a single borrower as a percentage in the total loans, as some regulations require. Thus, large banks are able to diversify loans among many single borrowers, while small banks tend to make large numbers of small loans (usually to small businesses and retail

[7] However, this does not necessarily mean that foreign banks' profitability will be better than that of domestic banks. DeYound and Nolle (1996) and Peek et al. (1999) find that foreign banks' performance in the US banking sector has been less than that of domestic banks.

[8] Usually, it is observable that, regardless of bank ownership, large banks allocate less of their loan portfolios to small businesses than do small banks. A great many studies undertaken on the US banking industry find that this observation holds, regardless of bank ownership (see e.g. Keeton, 1995; Levonian and Soller, 1996; Strahan and Weston, 1996).

customers). Another reason is that large banks' managers are not in a position to review each loan application, especially when the number of small loans is high. If this is the case, then banks may need to employ more staff, more equipment, and larger data base technologies in order to increase their share of small loan borrowers. In this connection, Clarke et al. (2001) find that foreign banks operating in four Latin American countries (Argentina, Chile, Colombia, and Peru) have made fewer loans to small and medium-size businesses.[9]

Although a majority of studies find that foreign banks tend to make large loans to large businesses, loans to small and medium-size businesses are not significantly absent. Foreign banks can depend on standardized contracts, which may be based on improved assessment of clients' credit worthiness and which enable them to deal more swiftly with small businesses (Clarke et al., 2001). Moreover, the growth in foreign bank lending to small and medium-size businesses may be fostered by advances in high technology and the abundance of data, which improve the credit-scoring technology and increase information on the probability of applicants' loan defaults (Mester, 1997).

By and large, foreign bank activities in the host country have mainly focused on wholesale services, which flow to large corporate businesses, usually in the trading and industrial sectors.

After discussing the issues of why foreign banks exist, what form of institution they choose to undertake their activities, and what patterns of foreign banks' activities may look like, it is worth discussing why some countries are hesitant to license more foreign banks and why, on the contrary, other nations welcome their substantial presence. The next section reviews some of the ongoing arguments as to the various attitudes of countries towards having foreign banks in their banking sectors.

[9] However, in this study, the authors indicate that, after controlling for certain factors, their estimation suggests that large foreign banks have provided more loans to small and medium-size businesses than have large domestic banks in two countries: Chile and Columbia.

3.6 Arguments concerning foreign banks' entry

The issue of having a substantial foreign bank presence is not without critiques. While liberalization and openness policies advocate foreign bank participation, the benefits to be derived from increased foreign participation are not widely agreed upon. This section outlines the arguments for and against foreign banks' presence, bearing in mind that only limited empirical evidence is available to support each argument on either side (Dages, Goldberg, and Kinney, 2000).

Arguments against foreign banks' presence

Concerns regarding foreign banks' presence can be viewed in two strands of argument: those that relate to competition and those that do not. The latter focus on infant industry and other arguments that aim to protect the domestic system. For example, host countries may believe that the openness of the financial system to foreign ownership may negatively affect the domestic financial system, especially during economic crises or at times of economic weakness, as their existence could increase the avenues for capital flight (Dages, Goldberg, and Kinney, 2000). These avenues for money flight will affect the level of credit needed to stimulate the economy. In addition, the activities of foreign banks in the economy may place an additional burden on regulators in supervising the financial system. Their existence may make supervision more complex. For example, the question of whether foreign bank activities should be only supervised or co-supervised by both the host country and the foreign bank's country of origin is a special problem, particularly when the financial system of one or both of these countries is underdeveloped.

The majority of arguments against foreign banks' presence are related to how their competition with the domestic banks will affect the latter. Generally, there are concerns that foreign bank competition affects domestic banks' profitability. The sophisticated and more competitive foreign banks can threaten the profitability of the domestic banks; in this case, the level of profits earned by domestic banks is expected to decrease, as more competitive foreign entrants are able to affect market shares of domestic banks.

Moreover, there is no guarantee that foreign banks will retain their profits in the economy as domestic banks do. Given that the shareholders of domestic banks are mostly national citizens who either re-invest or direct their dividends to consumption purposes within the economy, foreign bank shareholders are less likely do so in the host country. Hence, the increase in the stake of profits of foreign banks may, to some extent, prove a lost opportunity for the local economy to benefit from foreign banks' profits as long as these banks continue to transfer their profits abroad.

Another concern is related to the indirect impact of foreign banks' presence on credit to small businesses, since credit to small businesses provided by small banks is expected to shrink as small banks become threatened by the more competitive foreign banks. Clarke et al. (2001) observe that although foreign banks' entry might increase the supply of credit in the economy, they may, however, affect the supply of credit when some of the domestic banks will not be able to compete with foreign banks and are forced to leave the market. These authors foresee the implication of this to be that when large banks enter a foreign banking industry, or either take over or merge with small banks, the supply of credit to small businesses may decrease, especially when small banks that finance a large portion of small businesses leave the market or become merging or take-over targets. For example, after the merging and take-over activities of the 1990s, only a small portion of loans made by large banks in the US went to small business enterprises, in comparison with those made by small banks (Berger, Klapper, and Udell, 2001).

Last but not least, there are some fears that foreign banks may behave as a "cherry picker" in the local banking industry. Martinez-Peria (1999) observes that the more competitive environment brought by the entry of foreign banks may affect the loan quality of domestic banks, especially when foreign banks attract the good quality borrowers who are the customers of domestic banks. This crowding-out effect on good quality borrowers may increase the likelihood of loan losses for domestic banks and may entail a deterioration in their performance.

Arguments against foreign bank entry in some sense mirror those favouring entry that are discussed below.

Arguments favouring foreign banks' presence

Generally, Levine (1996) points out that foreign banks' presence is seemingly beneficial for the host countries' financial systems. He notes that their existence is likely to increase market competition, improve the quality of services by bringing in sophisticated banking operations, foster legal and supervisory reforms, and enhance a countries' access to foreign capital.

In essence, Glaessner and Oks (1994) argue that foreign banks' presence improves the structure of the financial system through the enhancement of auditing, transparency, and regulation. They add that their presence attracts rating agencies, auditors, and credit bureaus. Moreover, Dages, Goldberg, and Kinney (2000) note that foreign banks may act as a channel to import regulations and supervisory skills to the host county's financial system. They add that foreign banks' risk measurement skills can be beneficial for domestic banks, as an additional guide is obtained by foreign banks' experience. Furthermore, because of the increased presence of foreign banks, international coordination (represented in the Basel Committee and the Banks for International Settlements) has devised set roles and guidance for international banking and for countries hosting foreign banks, which help to resolve difficulties in supervising foreign banks' presence.[10]

The presence of foreign banks may also lead to enhance access to foreign capital. Moreover, the existence of foreign banks may raise the incentive of multinational enterprises to invest in a country, since foreign banks can help to attract more net capital inflows. Although there has been some evidence showing that banks follow their customers overseas, the reverse could also happen (Clarke et al., 2001).[11] Thus, the availability of foreign banks in a country may attract firms from the banks' country of

[10] Bank for International Settlements (http://www.bis.org/publ/bcbsc312.pdf).
[11] This argument is discussed under the following customer subsection in section 3.3.

origin since these banks can provide their corporate customers with information on investment opportunities in the host country. In sum, whichever one leads to the other, the net outcome for the host country is the enjoyment of increased capital inflows.

Increased market competition induced by the presence of foreign banks means that domestic banks are brought into a competitive environment which should motivate them to increase their efforts towards greater efficiency, better management, and higher quality of services. Domestic banks may be expected to focus more seriously on how to improve their market competitiveness through a review of their policies on customer relationships and by offering a wider range of facilities to their customers, as well as better service increasing the satisfaction of their customers. Thus, competition from foreign entry may benefit both domestic customers and domestic banks, as customers come to enjoy better banking services and better prices, and domestic banks are compelled toward greater discipline in how to be more efficient and competitive.

Foreign banks are thought to foster financial system stability (Dages, Goldberg, and Kinney, 2000; Sebastian and Hernansanz, 2000). Depositors favour the use of a wide range of national and multinational banks, allowing them to diversify their deposit portfolios across different bank ownerships. When branches of well-known international banks operate in a country, confidence in the banking system may be expected to increase, as depositors believe that these banks may provide additional resorts for their assets should the financial system face any difficulties. For instance, Dages, Goldberg, and Kinney (2000) argue that diversification of the banking system ownership led to greater stability in the credit industry during the crises and financial difficulties in Mexico.

Moreover, Crystal, Dages and Goldberg (2001) suggest that the accusation of "cherry picking", i.e. the extraction by foreign banks of good quality borrowers from domestic banks, does not always hold since foreign banks in the countries the researchers studied had relatively high loan loss provisioning compared with that of domestic banks. Crystal, Dages, and Goldberg note that high loan loss provisions suggest more aggressive action by foreign banks to deal with asset deterioration, especially during financial crises. Furthermore, they notice that foreign banks did not show any behaviour of "cutting and running" during the period of economic down-turn that hit the emerging market economies they studied. They find this behaviour to be less characteristic of foreign banks since their quality of loans is usually higher. Also, Goldberg (2001) shows that loans of foreign banks in emerging countries have been less volatile, implying that foreign banks have been stable lenders. Therefore, because of the precautions taken by foreign banks to extend their loans to safer borrowers, the argument goes that foreign banks contribute further to the stability of the financial system of the host country.

The debate over the issue of foreign banks' entry is still continuing, a situation that leaves some policy-makers with a margin of uncertainty in judging whether more benefits or more costs from foreign banks' presence will be generated in the financial system and the economy. Countries may, therefore, impose restrictions instead of total prohibition of foreign banks' entry. These restrictions could be imposed on the entering bank's organizational type, its banking activities, or may even come in the form of heavy tax impositions. In some countries, like Egypt, foreign banks have only been allowed to enter in the form of joint ventures with domestic banks (Caperio and Cull, 2000). In certain countries, foreign banks may only be allowed to take over a failing domestic bank. Other countries may impose heavy taxes on foreign banks' profits in order to deter any new foreign bank entry. These types of restrictions are imposed to reduce concerns about the potential negative effects of foreign entry. The underlying concerns generally relate to fears that foreign banks have an unfair competitive advantage over their domestic rivals.

It has been commented that restrictions limit competition and protect inefficient domestic banks (Clarke et al., 2001). A study by Barth, Caprio, and Levine (2001b) shows that tight regulations, which may include severe restrictions against entry, are usually associated with broad interest margins and high overhead costs, which may indicate high levels of inefficiency. However, restrictions against foreign banks' participation in many emerging economies have recently been relaxed because of the overall restructuring programmes which have come with liberalization policies.

Clarke et al. (2001) point out that if the main purpose of entry is to provide services for customers of the foreign bank's country of origin, then any negative impact on domestic banks should be small and domestic banks will not be hurt by competition with foreign banks. Similarly, little impact on domestic banks may be expected if the new entry is accompanied by a new range of banking services. As Clarke et al. (2001) also note, the policy-makers should evaluate the benefits and the costs of foreign banks' influence on domestic banks and financial systems and seek to discover which may dominate over the other.

However, the experience of particular countries, such as developed and developing countries that have foreign banks in their territory, may shed some light on how these foreign banks have performed, how efficient they have been, and how their performance compares with that of the domestic banks within these countries' banking sectors. This we seek to address in the next section.

3.7 Foreign banks' performance and efficiency: evidence from developed and developing countries

In addition to the issues of why foreign banks exist and what their businesses in host countries may look like, the literature also includes studies that examine how foreign banks' performance and efficiency are different from that of domestic banks. Generally, a near majority of studies on foreign banking performance and efficiency come from the US banking sector. However, the last decade has witnessed a growing focus on other developed countries, as well as developing countries, especially those that have sought to liberalize their financial systems.

In this section, we first present the findings of studies that tackle the issue of foreign banks' performance in developed and developing countries. Then the section overviews the findings of studies undertaken on the issue of foreign banks' efficiency in developed and developing countries. Furthermore, other studies that have examined the performance and efficiency of foreign banks entering banking systems through acquisitions are also relevant to this section. Finally, the section discusses the determinants affecting the performance and efficiency of cross-border banking business.

Generally, it is noticed that the performance of foreign banks in developed countries tends to be at a lower level than that of domestic banks. For example, Seth (1992) examines the performance of foreign banks (subsidiaries, branches, and agencies) in the US and compares this to the performance of US banks over the period 1980-1991. Based on return on assets and return on equity, Seth finds that the performance of foreign banks was lower than that of domestic banks, to the extent that the ratio of their performance accounted for only around 33 per cent of the total performance of domestic banks. Similarly, in a study that includes a number of developed countries, Claessens et al. (1998) find evidence that foreign banks in developed countries have lower interest margins, lower overhead expenses, and lower profitability than domestic banks.

In contrast, the literature also tends to find that foreign banks perform better than domestic banks in developing countries. For example, in a comparative study of foreign and domestic banks operating in Hungary over the period 1992-1993, Sabi (1996) finds that foreign banks were more profitable and less exposed to liquidity or credit risks than domestic banks. In addition, the study by Barajas et al. (2000) on the Colombian banking system covering the period 1991-1998 shows that, compared to domestic banks, foreign banks tended to have lower administrative costs and higher loan quality. Furthermore, Claessens et al. (2000) finds that foreign banks achieved higher profits than domestic banks in developing countries.

In comparing foreign and domestic banks' efficiency, [12] a number of studies on the US banking sector find that, on average, foreign banks tend to be less efficient than domestically owned banks. For example, Elyasiani and Mehdian (1993) use the non-

[12] The methodologies used to estimate efficiency are not the same in these studies. Some of these studies use direct measures of efficiency such as the study by DeYoung and Nolle (1996), which uses the parametric approach, and the study of Elyasiani and Mehdian (1993), which uses the non-parametric approach. However, most of the studies mentioned previously refer to efficiency as an indirect result of indicators such as change in interest margins, change in profit levels, and overhead costs. For example, in the cross-country study by Claessens et al. (1998), they show that the low bank profitability and low overhead costs associated with the presence of foreign banks can be interpreted as an improvement in the banking sector's overall efficiency.

parametric data envelopment analysis (DEA) approach to estimate the efficiency of foreign and domestic banks operating in the US in 1988. They find that foreign banks were less cost efficient than domestic banks. Nolle (1995) compares the cost efficiency of both foreign and domestic banks in the US and finds that foreign subsidiaries were less cost efficient than domestic banks in every year except for one between 1984 and 1992. Moreover, in DeYoung and Nolle's (1996) study of the profit efficiency of foreign banks operating in the US, they find that foreign banks are more profit inefficient than domestic banks. [13] Hasan and Hunter (1996) find that the Japanese multinational banks operating in the US reported, on average, a lower cost efficiency level than their domestic peers.

Little evidence is available to inform us about the efficiency of foreign banks in developing countries in comparison with their domestic peers. One exception is the study of Bhattacharayya, Lovell, and Sahay (1997) that covers seventy commercial banks over the 1986-1991 period of ongoing banking sector liberalization. The authors find that foreign banks operating in the Indian banking system are more efficient[14] than private domestic banks. [15] They also find that foreign banks exhibited temporal improvement in their performance, but no temporal improvement trend is found in the private domestic banks' performance. However, it is still difficult to draw a general conclusion about comparisons of efficiency between foreign and domestic banks in developing countries.

Some studies have looked at the performance and efficiency of foreign banks that are established in terms of acquisition. The results of these studies tend to give the same finding, that is, that foreign banks show lower levels of performance and efficiency than domestic banks in developed countries, but higher in developing countries. Houpt (1980), Goldberg (1981), and Hodgkins and Goldberg (1981) have studied how foreign banks that acquired US banks performed during the 1970s. They generally conclude that foreign banks were less profitable than their domestic peers. These studies also note that

[13] These authors, however, find that foreign bank subsidiaries are more profit efficient than domestic banks.
[14] The authors use both parametric and non-parametric approaches to calculate efficiency.
[15] They find, however, that commercial public banks have been the most efficient among all banks, but their trend of temporal performance improvement is found to be declining.

these banks did not change their pre-acquisition strategies. Similarly, referring to the 1990s, the study of Peek et al. (2000) finds that foreign banks that acquired domestic banks operating in the US experienced low performance and low efficiency. On the other hand, Crystal, Dages, and Goldberg (2001) studied a number of developing countries in Latin America and find that local banks that were acquired by foreign banks performed marginally better than those institutions that remained under the control of domestic banks. Moreover, in these countries, Crystal, Dages, and Goldberg find that foreign banks with a longer presence exhibited stronger loan growth compared to domestic owned banks. They note, however, that foreign banks that have been recently established through acquisition have the lowest loan and deposit growth rates.

The literature points out some determinants that may, in general, affect foreign banks' performance and efficiency. [16] For example, Molyneux and Seth (1998) find that three factors determined the performance of foreign banks in the US between the years 1987 and 1991. These factors are capital strength, commercial and industrial loan growth, and assets composition. They note that capital strength was shown to be the most important factor influencing foreign bank performance.

However, what makes the performance and efficiency of foreign and domestic banks differ could also be due to other factors. For example, some say that the reason why foreign banks in developed countries have lower levels of performance and are less efficient than domestic banks is perhaps because of linguistic and cultural barriers and conflicts (DeYoung and Nolle, 1996; Hasan and Hunter, 1996; Mahagan et al., 1998).

Moreover, some studies tend to show that, in comparison to domestic banks, foreign banks could be more reluctant to deal with borrowers whose qualities are unknown to the bank. Therefore, foreign banks may have lower loan losses and so register better performance and efficiency than domestic banks because foreign banks tend to adopt a more cautious loan strategy and may set higher loan loss provisions. In this connection,

[16] We combine the discussion of efficiency and performance as they are so addressed in the references. However, although the determinants may affect both of them, the effect may not show the same magnitude or indicators. For example, the number of firms may increase competition, which may lead to more cost efficiency, but diminished performance may be observed as profit margins decline.

Crystal, Dages, and Goldberg (2001) find that foreign banks operating in both Argentina and Mexico during 1994 and through to mid 1999 exhibited stronger and less volatile credit growth compared to domestic banks. Moreover, the study by Goldberg and Kinney (2000) shows that although foreign and domestic banks in Mexico and Argentina had similar lending activities during the 1990s, the growth of foreign bank lending was stronger and less volatile mainly because of better loan diversification.

In fact, the dependence on borrowed funds can also be a reason for an increase in cost inefficiency and hence the weaker performance of a bank. Using the profit efficiency model applied to US banks over the period 1985-1990, De Young and Nolle (1996) find that foreign banks were less profit-efficient than domestic banks because of their reliance on (relatively costly) purchased funds. However, they show that input inefficiency, rather than output inefficiency, is the most likely explanation of the low profitability of foreign banks.

In addition, management deficiencies are also found to be a factor affecting bank efficiency. The study by Peek et al. (2000) implies that part of what makes foreign banks in the US banking industry inefficient and poor performers is their acquisition of failing banks and the difficulty management has in initiating successful strategies to change the performance of the acquired banks. Moreover, the amount of non-performing loans of the banks acquired by foreign ownership has also been a hindrance to their successful performance.

In more focus, Berger, DeYoung, Genay, and Udell (2000) propose two main hypotheses, which indirectly test possible determinants of efficiency comparisons between foreign and domestic banks. Under the home field advantage hypothesis, domestic banks are likely to be more efficient than foreign banks partly because of organizational diseconomies[17] that parent banks of the foreign banks may suffer from. Organizational diseconomies could be operating problems, as also could problems of monitoring managerial behaviour from a distance. The authors also draw attention to

[17] Such as problems staff may experience in working in different nations and the high costs involved in persuading managers to take charge of institutions abroad.

(especially the USA) tend to be less efficient than domestic banks. When results of bank efficiency are disaggregated by country of origin, the results show that the finding that domestic banks are more efficient than foreign banks does not hold; it only holds in this study when results are not disaggregated.

To summarize, studies relating to developed countries generally find that foreign banks show lower performance and are less efficient than domestic banks, in contrast to the studies relating to developing countries. The main determinants that affect performance and efficiency are found to be bank capital, risk management, managerial skills, and cost of funds, as well as the ability of banks to overcome cross-border disadvantages such as cultural differences.

3.8 Conclusion

This chapter has reviewed the literature on the topic of foreign bank presence. This foreign bank topic is part of the theory of international banking, which consists of two components: international banking services and multinational banking. The theory of foreign bank presence is best explained by the theory of multinational enterprises, as it focuses on the determinants that drive firms to practise their business through some form of physical presence overseas. Determinants of foreign bank presence can be divided into those that are specific to the international environment (such as political, social, and geographical factors) and those that belong to banks and financial system specifics (such as following customers, regulatory asymmetries, size of banks, and banking systems). There are many forms and modes through which foreign banks may enter a banking sector, where the decision on the choice of form to be established is affected by regulations and the advantages and disadvantages that each form may have.

Foreign banks are observed to engage more in wholesale banking because this area of business tends to have less restriction compared to retail banking. Moreover, the bank's size may help it in dealing with many large single borrowers (such as corporate customers) rather than with small borrowers. Because foreign banks may affect domestic banks' profitability and may serve as an avenue for capital flight and also because of the complexity of supervising such institutes, certain countries may not favour their existence in their banking sectors. However, countries may benefit from a greater foreign bank presence as they encourage foreign direct investment, induce greater efficiency in the operation of domestic banks, and may enhance stability in the banking system.

CHAPTER 4

FINANCIAL SYSTEM EFFICIENCY: AN OVERVIEW

4.1 Introduction

Among many functions the financial system performs, there are two that are essential for any economy: one is the administration of the payments mechanism, and the other is intermediation between ultimate savers and borrowers (Mishkin, 1998). However, undertaking these functions may not be sufficient for the financial system to maintain its well being and performance. The experience of many financial systems that have experienced financial crises, such as Latin America and East Asia, suggests an essential element in the functioning of the financial system is the extent of its efficient operation. This is extensively linked to the soundness and safety of the financial system overall.

The growing interest in efficiency of the financial system emerged from the memory of the financial crises of the US Great Depression of 1929-1933, where the literature has introduced various efficiency aspects explaining the causes of financial crises. For example, Bernanke (1983) identifies the important role of the costs of intermediation (such as screening, monitoring and accounting costs) in choosing high quality borrowers. When the costs of intermediation increase (e.g. because of difficulty to identify the risk type of borrowers), the efficient allocation of credit may be affected and thereby raise the cost of credit to borrowers. Accordingly, Bernanke suggests that an increase in intermediation costs reduces the efficiency of credit allocation, because financial crises make it more difficult to identify the quality of borrowers. Therefore, the supply of funds will be squeezed as borrowers find credit more expensive. Moreover, Mishkin (1998, p. 217-219) highlights the scenario in which asymmetric information (such as adverse selection and moral hazard) contributes to the aggravation of inefficiency of allocating financial resources, especially in a period of financial distress. During financial crises, lending activities decline because it is difficult for banks to verify the risk rating of borrowers, and so asymmetric information problems would encourage the shift of loans to riskier investments.

The aim of this chapter is to outline a framework through which the efficiency of the financial system can be overviewed. Because the literature lacks a clear definition of financial system efficiency, the motivation of this chapter is then to define financial system efficiency and explain how this is important for economic growth. As will be explained in section 4.5, Tobin (1984) has identified four aspects of efficiency that are pertinent to a developed financial system. This chapter will discuss broader aspects of efficiency of the financial system, which will also incorporate the efficiency features identified by Tobin. The benefit of having a study on a wider range of efficiency aspects is to provide a clearer picture of the performance of the financial system from many of these efficiency aspects. This will also help to highlight various policy issues that are aimed at improving the overall functioning of the financial system.

This chapter contains six sections. The following section discusses why financial efficiency is important for the health of the financial system. Section 4.3 provides a general economic view of the efficiency concept and the definition of financial system efficiency. Section 4.4 addresses the different concepts of efficiency that are applied to the financial system. Section 4.5 explains the relationship between financial system efficiency and the performance of the real economy. The final section is the conclusion.

4.2 Why should we be concerned about the efficiency of the financial system?

As the core function of the financial system is to mobilize internal financial resources to finance productive investment, the efficiency of the financial system is an important determinant of its soundness. The need for financial system efficiency can be highlighted in the following:

- Over the last two decades, many financial systems have experienced severe currency and banking crises; such crises occurred in the US banking sector in 1985-92, Mexico in 1994-95, and Asia in 1997-98 (see Krugman, 1998; Mishkin, 2001). One of the major features of these crises was the lack of efficiency in these systems, as well as the lack of adequate regulations to enhance efficiency. For example, Krugman (1998) points out that among the reasons responsible for the Asian crisis was the severity of the moral hazard problem where banks had been provided with implicit guarantees, which distorted incentives towards making risky loans.[1] These types of loans, which in the case of East Asia are loans made to the real estate sector, created a boom in asset prices. When the asset market crashed, many banks faced insolvency problems because borrowers were unable to repay their loans.

- In general, international investors seeking more internationally diversified portfolios and better returns on investments may have learned from these crises that questions about how strong (and how efficient) a financial system is are really matters.[2] However, it might be hard for an investor to tell which financial system is more efficient than another, especially when it comes to developing countries. The difficulty in choosing an efficient financial system might be due to the lack of any well-known indicators that could be useful to guide international investors to answer the question of how efficient a financial system is. Chen and Khan (1997) argue that foreign incentives to invest internationally depend on the return on the foreign investment, which also depends on the level of financial development and the country's economic growth. Moreover, empirical studies by Demirgüc-Kunt and Levine (1996) and Levine (1996) on

[1] Providers of funds to Asian financial institutions (especially in Thailand) believed that their funds would be protected from risk. In addition, the owners of financial institutions concluded, through their strong political connections, that the provision of such government guarantees would be available (See Krugman, 1998).
[2] As we will analyse later financial system efficiency includes both stock market efficiency and bank efficiency.

financial development and economic growth provide a set of indicators to measure the level of development of a financial system and how it relates to economic development. From these studies, one may conclude that more developed financial systems may indicate more efficient financial systems. However, the level of development of a financial system does not mean that all efficiency aspects are mature. East Asian financial systems have been classified with relatively high levels of development, but they were inefficient in the sense that they had market imperfections that resulted in an inefficient allocation of financial resources that existed prior to the financial crises.

- The study of the efficiency of financial systems before and after crises can provide valuable information for policymakers (see Berger and Humphrey, 1997). Therefore, it is important that there should be studies evaluating financial systems from an efficiency perspective to judge the health of the financial system and its suitability to encourage capital flow. Studies that measure banks efficiency and performance, stock market efficiency and volatility, regulation and supervisory effectiveness, can help in determining the efficiency features of a financial system.

4.3 The general interpretation of efficiency and a definition of financial system efficiency

In economics, the word efficiency is always linked to the allocation of resources. Its narrow definition usually refers to resources being employed in a way that gives the maximum production of goods and services. When this is achieved, then allocation is said to be optimal, and resources waste will be absent. Generally, the concept of economic efficiency means that the economy produces goods and services that fully reflect the preferences of consumers, given that the production of these goods and services is made with minimum cost (Nicholson, 1995). In addition to this, economists may also include environmental and social aspects in the account of economic efficiency.

The concept of economic efficiency dates back to the classical school of economics (see Nicholson, 1995, p. 561-62). Adam Smith's invisible hand and laissez-faire arguments stress that when there is no government involvement in economic activities, the market mechanism will be capable of maximising individual welfare and allocating resources efficiently. This is because the market mechanism can independently coordinate between buyers and sellers interests and reach equilibrium; thus, whenever imbalances in the supply and demand of goods and services occur, the market will automatically re-adjust itself to achieve equilibrium.

The idea of the invisible hand has been the impetus for the development of welfare economics, specifically the relationship between efficient allocation and market competition. Pareto showed how economic efficiency could be reached in the context of a competitive market (see Nicholson, 1995, p.563-64). In an economy of one consumer and two commodities, *x* and *y*, efficiency occurs upon the condition when

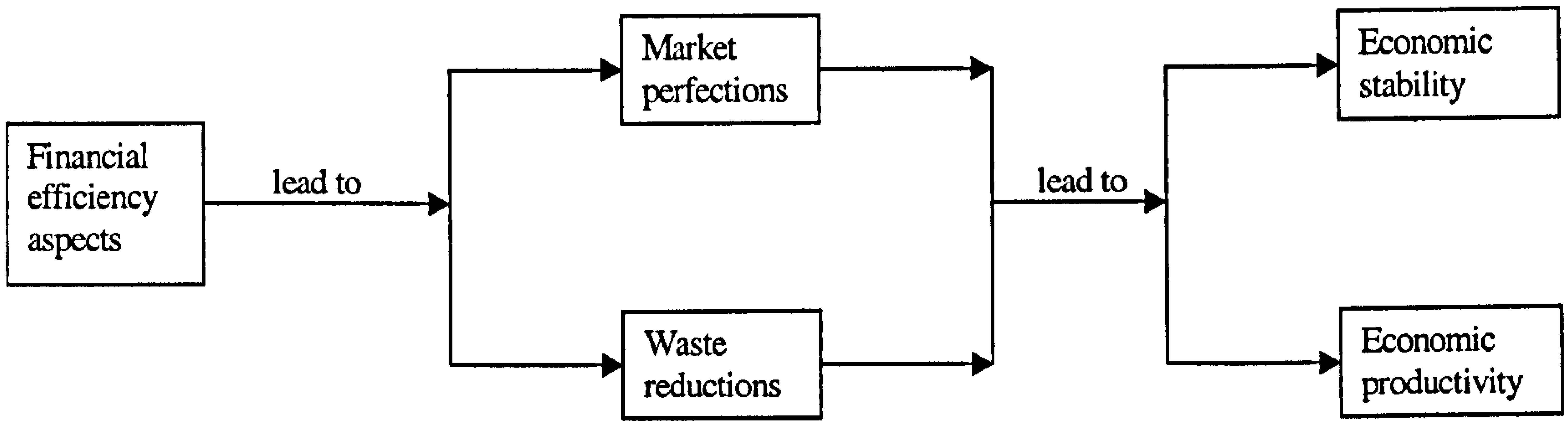
$$MRS_{y,x} = MRT_{y,x} = \frac{P_x}{P_y}$$

where $MRS_{y,x}$ is the marginal rate of substitutions between commodity *x* and *y*, given that $MRS_{y,x}$ reflects the slope of the consumer's indifference curve; $MRT_{y,x}$ is the marginal rate of transformation between both commodities, where $MRT_{y,x}$ is the slope of the production possibilities curve; and $\frac{P_x}{P_y}$ is the relative prices of *x* and *y* and is known as the slope of the budget line. Within this framework, economic efficiency occurs when *MRS*, which is the same across all consumers, is equated to *MRT*, which is the same across all commodities produced in the economy, and both *MRS* and *MRT* must be equal to $\frac{P_i}{P_j}$, the relative prices of these commodities produced in the economy. This condition is known as Pareto optimality in which the economy achieves the efficient allocation of its resources. Under this condition, it would be impossible to improve the welfare of one or more individuals without reducing the welfare of at least someone else. In other words, any reallocation of resources will make some individuals better off and others worse off.

The classical economic view of the predominance of the invisible hand, however, became increasingly questioned in the light of various developments in the early 20th century. In particular, the Great Depression created a paradigm shift that resulted in Keynes' *General Theory of Employment, Interest and Money* (1963) (see Mankiw, 1994, p. 275). The persistence of depressions has shown that the pure market economy (of the invisible hand and laissez-faire) fails to reach equilibrium and full employment of resources. Keynes stressed the significance of government involvement to revive and stimulate the economy and to help the market to increase efficient allocation of resources. The idea of Keynes has been extended to show other reasons in which market failure (failure to achieve optimal allocation) provides welfare grounds for government involvement in economic activity. For example, the government may intervene to set anti-trust laws that protect market competition.

The aim of both classical and Keynesian schools, when explaining their approaches of how to achieve an efficient allocation of resources, is to enable the economy to obtain economic stability (i.e. avoid severe fluctuations and crisis) in order to foster productivity. While the aforementioned schools of thought focused mainly on the efficiency of the real sector, they paid little attention to financial sector efficiency. It seems appropriate, however, to link between the economic efficiency of both the classical and Keynesian schools and financial system efficiency on the basic aim of achieving stability and fostering productivity. Based on this link, we can define financial system efficiency as *the various efficiency aspects that should be available to minimise market imperfections and waste in a way that enhances stability and fosters economic productivity* (see Figure 4.1). For achieving stability, when financial resources are used efficiently, then stability in the system is enhanced for the reason that market imperfections (such as price distortions) will be eliminated and the economy will reflect the fundamentals. This should help the economy avoid or minimise any adverse impacts on various sectors. For example, when a price bubble develops beyond fundamentals, then this may induce the sector to crash as the prices severely deflate. For fostering productivity, the rise of productivity in the economy may stem from the fact that efficient allocations minimise waste and avoid excess utilisation of financial resources. This means that the economy has the opportunity to shift what it saves from the resources not being wasted or excessively used to destinations that are more productive. The economy may also utilize them in future allocations.

Figure 4.1 Financial system efficiency definition

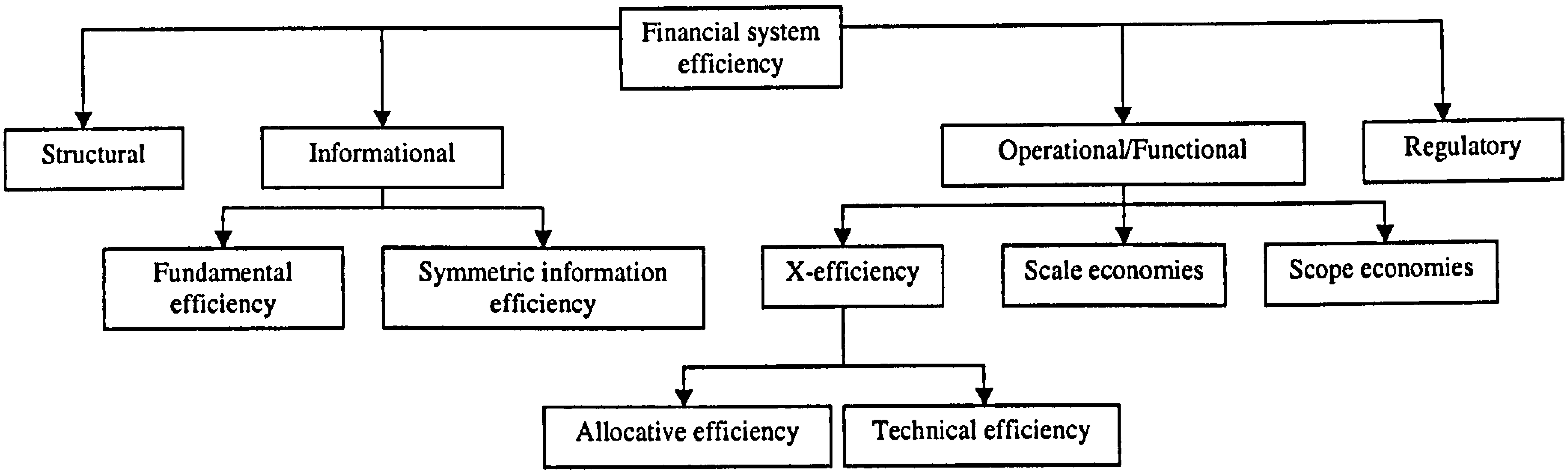


The next section presents the efficiency aspects that are used to reduce market imperfections and waste of resources.

4.4 The different aspects of efficiency as applied to the financial system

In this section, we present most of the efficiency aspects discussed in the context of the financial literature. These efficiency aspects can be considered as components of overall financial efficiency that contribute to market perfections and waste reduction, which lead to stability and productivity growth. Based on the literature, the related issues of efficiency, which apply to the financial system, are viewed in several features. James Tobin (1984) provides four aspects of financial system efficiency (see also Fry, 1988, p. 296). These are: information arbitrage efficiency, fundamental valuation efficiency, full insurance efficiency (or hedging), and functional efficiency. We argue that Tobin's efficiency aspects can be reintroduced in a framework that covers broader aspects of financial system efficiency. As in Figure 4.2, we introduce four main efficiency aspects of the financial system: structural efficiency, informational efficiency (in which Tobin's *fundamental valuations efficiency* and *informational arbitrage efficiency* concepts are discussed), operational efficiency (in which Tobin's *risk pooling* and *full insurance efficiency* concepts are discussed), and regulatory efficiency. Within these aspects, we explain, in addition to Tobin's efficiency concepts, many other concepts that may fall under each of these aspects.

Figure 4.2 Financial system efficiency aspects



In most cases, our analysis of the efficiency aspects of the financial system will have several dimensions such as financing activities by banks and primary and secondary financial markets where stocks and bonds instruments are issued and traded. Given these dimensions, we explain how efficiency aspects may affect the performance of both banking and financial markets activities.

4.4.1 Informational efficiency

Informational efficiency refers to the extent to which a financial system is able to provide information that helps in allocating financial resources to their most productive destinations. Indeed, information is one of the most important factors affecting the process of funds allocation. This is because the acquisition of information on the ability of borrowers to make acceptable earnings for 'lenders' may be the main determinant for financing activities.[3] In addition, the more information available on the quality of borrowers (i.e. their success in loans repayment and their projects' feasibility) the more funds the lenders are willing to provide for borrowers. If the lenders lack information,

[3] 'Lenders' can be financial institutions like banks, or individuals who buy bonds. The use of the word 'financier' may refer to these mentioned as lenders as well as buyers of stocks.

the risk of non-payment of the debt will increase, and risk averse lenders will be less willing to finance borrowers. In this case, informational inefficiency leads to more market imperfections, which reduces the supply of funds available for economic growth (see the analysis in section 4.5).

Informational efficiency in the financial literature has two related aspects. The first could be viewed as how parties deal with *asymmetric information* problems. The second is about the ability of the financial markets (mainly, the stock market) to reflect the financial assets prices (such as stock prices) according to *fundamentals.*

Symmetric information efficiency

Symmetric information efficiency deals with how the financial system is able to provide all relevant information for the parties engaging together in a financial deal (see Stiglitz and Weiss, 1983; Mishkin, 1998, p. 35). When the distribution of information between these parties is uneven, then this is known as an asymmetric information problem. That is, when the less informed party deals in a transaction with the more informed party, it is difficult for the less informed party to make accurate decisions.

Asymmetric information in the financial system can appear before and/or after the transaction. Pre-transaction asymmetric information problems relate to *adverse selection*; while *moral hazard* comes after the transaction (Mishkin, 1988, p. 35-36). *Adverse selection* occurs when the lack of information makes it difficult for the financier to make successful selections. In the case of banking, adverse selection exists when a bank is not able to distinguish between borrowers with low or high default probabilities. In this case, the quality of borrowers would be indistinguishable to the bank. By applying Akerlof's (1970) lemons model, the credit market will suffer from market imperfections in which the lack of information will induce lenders to raise the interest rate. The lenders tend to do that since a higher interest rate will compensate unexpected defaults. However, this will bring more low quality borrowers with high risk and drive out good quality borrowers with lower risk.

In the case of financial markets, the problem of adverse selection may appear before purchasing a firm's stocks. If the securities market fails to reflect the fundamentals in the price of stocks of an underlying firm, then information about the firm's quality will be difficult to evaluate. For example, when the firm's stock price is overvalued and does not match the firm's profitability, then investors will be reluctant to buy the stock of this firm because it is difficult for them to determine the quality of the firm. In this case, the firm might fail to raise the funds they need. On the other hand, when the overvalued firm succeeds in raising the funds it needs, then it can be said that, due to market imperfections, the stock market has failed to allocate funds to their most productive destinations.

The second sort of asymmetric information is called *moral hazard.* It appears after the parties agree to make a transaction. The hazard in the transaction exists when one of the parties engages in behaviour that is undesirable to the other party. In banking, moral hazard arises when the borrower uses the funds in activities that increase the probability of default; in financial markets, since a firm has no obligation to repay the nominal value of the stock, the incentive of firms' managers to undertake risky investments is more likely.[4]

When the funds allocation to risky destinations becomes a norm for getting high returns, instability in the economy will become more likely. If borrowers fail to repay their loans and firms' failures increase, it would be difficult for banks to meet savers withdrawals, and this could make banks insolvent. Moreover, as the likelihood of firms' failures increases, stockholders will still rush to sell shares of these firms, and the stock market may crash.

Therefore, in the absence of an efficient market, asymmetric information problems will increase market imperfections that may destabilise the financial system and the economy. In order to overcome asymmetric information problems, these informational

[4] In terms of the informational aspect, Stiglitz (1989) gives more details on advantages and disadvantages of bonds, stocks and short term finance.

efficiencies (obviously) have to be improved. The literature explains several methods that the financier might use to increase information about the quality of the funds' applicants. Among theses are screening, credit rationing, monitoring and commitment (Stiglitz, 1989; Mishkin, 1998). The first two, screening and credit rationing, are used to alleviate the adverse selection problem. The others, monitoring and commitment, are used to reduce moral hazard.

Screening is a method used by the lender and intermediaries to screen good loans from bad loans. The lenders collect information on the borrower's historical credit record and evaluate the current status of his creditworthiness as well as how successful is his future ability to repay the loan. In financial markets, the collection of information about the performance of firms is the tool used by investors to screen out the firms' quality. Therefore, in order for an investor to judge the stock price of a firm, information available on the firm's performance plays a major role in assessing its stock price. This point will be elaborated below when we talk about *fundamental efficiency.*

Credit rationing is used by banks in order to reduce the effects of adverse selection. Stiglitz and Weiss (1981) show that as uncertainty and the distribution of information widens, the lack of information on borrowers and their projects may induce the bank to increase interest rates. However, the increase in the interest rate will bring another problem to the bank. It will face riskier borrowers instead of safer borrowers since, as in Williamson (1986), high risk increases the adverse selection problem. This is because high interest rates induce the current borrowers to shift to riskier investments as the rate of return increases with the level of risk.

Instead of increasing the rate of interest, banks may use the credit rationing method, which limits the amount of loans according to the expected risk attached to a borrower. Some authors (such as Stiglitz, 1998) have voiced concern about the effect of this method since it causes the level of investment to fall. However, the existence of other sources of fund raising, such as the securities market, will help mitigate the negative impact of credit rationing on economic growth. Cho (1986) shows that the availability of the stock market can help investors facing credit rationing to raise their funds by stock issue. Nevertheless, the success of a firm to raise funds from other sources when it fails to do so from banks depends on other elements specific to the firm, such as its reputation and rating records. Also, the level of development of the capital market and the overall financial system are important factors in financing financial deficit agents to find better alternatives to raise funds. In this case, the level of investment will be much less affected (Thakor, 1996).

Monitoring is the method used by the party offering the finance in order to alleviate the problem of moral hazard. The lender oversees the behaviour of the borrower in order to ensure that borrower activities are in line with the contract. In banking, the commercial borrower is asked to provide the bank with audited accounts and other information. In financial markets, firms are enforced to publicize their audited accounts and to have investors informed of the firms' activities. This will make it easier for investors to judge on how well the firm is performing.

Commitment deals with the ways to tackle the incentive distortions that lead to moral hazard. It aims to increase the credibility of the borrowers to maintain the interest of the lenders. Many methods can be used to enhance commitments (Mishkin, 1998). For example, banks may design restrictive contracts that confine the loans to be made to only particular projects and activities. Moreover, banks may enforce borrowers to present periodical reports to monitor how the loan is spent on the project. Banks may also ask borrowers to provide collateral in order to get loans. The collateral can effectively influence the incentive of borrowers since it induces them not to use the loans in activities that increase the probability of the default; otherwise, they might lose their collateral.

The importance of informational efficiency aspect in alleviating asymmetric information problems is that they contribute to real economy efficiency by deriving social benefits. For example, Boyd and Prescott (1986) show how the screening process allows the financial system to achieve socially beneficial projects by reducing or eliminating inferior projects and diverting resources to more productive projects. Moreover, the collection of information about investors' creditworthiness creates a valuable database for intermediaries and a network of information that eases information transmission (Greenwood and Jovanovic, 1990). Therefore, the existence of private firms (such as Moody's and Standard and Poor's in the US and London based IBCA bank credit rating agencies) specializing in collecting information and evaluating the performance of firms will guide financiers who purchase such information to determine which firms are worthy of receiving funds (Mishkin, 1998).

Fundamental efficiency - Efficient stock market

Fundamental valuations efficiency is the term used by James Tobin (1984) to express how the current market prices of assets reflect the fundamentals.[5] More precisely, the market is called fundamentally efficient when it is able to set a price of a financial asset equal to the present value of the asset's future income stream. When the market is fundamentally efficient, no one will have an incentive to pay more than what the asset's future income is worth today.

Stock market efficiency can also be viewed as the stock market's ability to reflect fundamentals. Since market participants' interaction (ask and bid mechanisms) determine the price of financial assets, it is important that they have information on what the future income stream of the stock will be. If they make their decisions according to all information available in the market, then the price of the stock would be the present value of its future income stream. The more information available, the better expectation of the future income would be placed on it by investors, and the more accurate price will be set by the stock market. In this sense, the stock market is called efficient when it fully and correctly reveals information on the stock prices of the listed firms.

[5] Fundamentals refers to the analysis of evaluating the price of a stock on the basis of information on the micro-performance of the firm, such as earnings, dividends and financial statements, and on the macro-performance of the economy, such as interest rates, GNP, inflation and unemployment. The information is used to forecast the future price of the stock.

Efficiency of the stock market may be reduced when there are imbalances in the distribution of information. When there are differences in the level of information obtained, investors with more information would be able to make gains from the trade. From this perspective, Tobin's *informational arbitrage efficiency* term can be applied here. Tobin views this term as: when information is equally distributed across all market participants, the investor cannot make any profit from engaging in trade (of a financial asset). This notion can also be applied here to explain the role of information in affecting stock market activity. That is, when information is equally distributed across all stock market investors, they cannot make abnormal profits; investors can only make such profits when they have access to some information that is not known to others.

Fama (1965, 1970) has utilized the idea of fundamental efficiency to develop hypotheses (weak, semi-strong, and strong form) that assess market efficiency in terms of pricing accuracy. Moreover, Fama's efficient market hypotheses incorporate rational expectations theory to evaluate how information may be used to make abnormal profits. Accordingly, the stock market is efficient when investors cannot make use of historical information (weak form); plus the use of publicly available information (semi-strong form); and plus the use of private information (strong form) to make abnormal profits. Therefore, Fama's hypotheses assess how far the stock market reveals information so that the stock prices reflect the fundamentals.

The importance of the efficiency of the stock market in having correct price signals of its listed firms is that stock prices can affect firms' sources of finance for two reasons. One is that the firm's stock price is the cost at which funds are raised for the expansion of the firm. The higher the price, the cheaper the funds; and the lower the price, the more expensive it is for the firm to attract finance from the primary market. Another reason is that inefficient price signals may affect the net worth (the firm's capital) of the listed firm. If the firm's value is under priced, it means that the value of the firm's net worth will decline. This might affect the lenders/financiers attitude towards financing a firm that is not backed by strong capital. In addition, the undervalued firm may further suffer from the contraction of its financial sources, especially when the firm's internal sources of finance are not sufficient.

As it has been addressed before, the failure of correctly pricing the firms stocks will also increase the severity of asymmetric information (Greenwald, Stiglitz, and Weiss 1984). Lenders will find it difficult to distinguish good firms from bad firms. That is, they may lend to a firm with high stock prices when in fact the firm's stocks are overvalued; conversely, they become more reluctant to lend to a good firm when its stock prices are undervalued. When the financier feels that the stock market makes the screening of the quality of the firm more difficult, the amount of the funds raised by the firm will be less than needed and therefore stock market efficiency in channelling the funds will be weakened (Greenwald, Stiglitz, and Weiss 1984).[6] Moreover, the under-priced firm will face a moral hazard problem since managers' incentives are more likely to be geared to making riskier investments (in order to make greater profits) than in appreciating the firm's stock price.

Thus, fundamental efficiency stresses the role of stock market efficiency in setting the correct price of stocks according to information available. This will reduce price distortions caused by asymmetric information, improve market perfections, and provide stability in share prices.

4.4.2 Operational efficiency

Operational efficiency in the financial system relates to the system's ability to organise the channelling of funds with minimum cost. As we will show below, when the cost of intermediation is at minimum, this means fewer resources are utilised to channel more funds. Operational efficiency is mostly applied to financial institutions, such as banks, (although it can also relate to the operational characteristics of capital market organisations and exchanges).

---

[6] Greenwald, Stiglitz, and Weiss (1984) call the situation in which adverse selection leads to lower funds raised than needed as 'stock rationing'.

Before talking about the operational efficiency elements of financial intermediaries, it is appropriate to explain how we define the outputs of banks. As elaborated later in Chapter 4, the measurement of outputs is a controversial issue in financial studies since the production of financial institutions is characterised by its non-physical (service) nature. In banking studies, there is however two views of measuring outputs: the production and the intermediation approaches. In the production approach, banks are viewed as firms that use labour and capital to produce loans, deposits and other earning assets. In addition, this approach measures output as the number of loans and deposits accounts. The intermediation approach views banks as firms that use labour, capital and deposits to produce loans and other earning assets. The intermediation approach measures outputs in terms of their values, but not number of accounts. Therefore, the difference between both approaches lies mainly as to whether deposits should be considered among inputs or outputs; and whether banks' inputs and outputs are measured according to the number or value of accounts. Most of the banking efficiency studies adopt the intermediation approach because it is easier in terms of data availability, and because it comes at the heart of measuring the cost of intermediating deposits to the receivers of loans (Berger and Mester, 1997).

In return, the operational efficiency (which will be detailed in Chapter 5) of banks can be addressed as follows:

X-efficiency exists when banks' cost (or profit) functions do not deviate from the cost (profit) frontier that attains the minimum cost (maximum profit) of intermediation in the industry (we use cost instead of profit concept hereafter in this Chapter). In fact, X-efficiency is the most important part of operational efficiency. As Berger et al. (1993) state "[t]he one result upon which there is virtual consensus is that X-efficiency differences across banks are relatively large and dominate scale and scope efficiencies." [This is also true as in Berger and Humphery, 1991; Evanoff and Israilevich, 1991]

X-efficiency is usually decomposed into *technical* and *allocative efficiency*. In welfare economics, ***allocative efficiency*** is used to show the situation in which the prices of goods and services produced in the economy reflect their minimum cost to supply them.

Thus in perfect competition, consumers pay prices that reflect the minimum cost of production at which producers receive normal profits that are adequate to make their businesses continue supplying the products. In a market with a sole producer, the price is set above the minimum cost, where the price consumers pay deviates from being allocatively efficient. In financial studies, specifically banking, ***allocative efficiency*** denotes the ability of a bank to use inputs in optimal proportions with respect to their prices (Farrell, 1957). In banking studies, most authors, including Berger et al. (1993), find that banks inefficiencies are technical in nature rather than allocative. Therefore, many authors, such as Mester (1993) and Altunbas et al. (2000), do not decompose the X-efficiency measurements. In general, the empirical banking literature provides more attention to technical rather than allocative efficiency.

Technical efficiency relates to the avoidance of the excessive use of inputs than that which is optimal for the given level of output (Berger et al., 1993). In banking, the measurement of the optimal use of inputs, once technical efficiency is achieved, involves analysis of the cost or price of inputs. From society's point of view, society is better off if a cost inefficient bank is to improve its operational efficiency towards reducing the inefficient and unproductive usage of its inputs. There are many reasons why technical inefficiency might exist. A managerial element might have an influence on firm's operations through mistakes in choosing the optimal size of inputs. Banks may mistakenly accept to pay high rates for deposits in order to increase their deposits base that enables them to make larger profitable loans. If this is not accompanied with higher allocation of deposits in the form of profitable loans and investments, then this is an excessive input that the bank is undertaking.

Moreover, the level of competition in the banking system may influence X-efficiency. When competition increases, banks' managers may be more inclined to reduce prices to fight against potential erosion of their market share. Also, banks may have incentives to incur high costs in order to provide services that are more attractive to their customers. Banks may even channel the amounts of deposits by making risky loans or by making loans to too many low return investments. This will lead banks to face delays and probably defaults of their loans, which induce higher monitoring costs accompanied with a reduction in the amounts of interest received from loans. In this case, banks will face higher cost X-inefficiency. The higher cost X-inefficiency may lead the bank not only to be forced to increase the interest rate margin but also to set higher fees on the bank's services to its customers. This may reduce the bank's competitiveness and again expose it to potential solvency problems.

Scale economies exist when a bank operates on its decreasing long-run total average cost curve. There are many empirical studies that investigate the existence of scale and scope economies. Most of them have been undertaken on the US banking system. Generally, one might expect that large or merged banks realise greater scale economies, making them more efficient. However, empirical research has suggested that this is not the case. As an example, the survey article on US banks by Humphrey (1990) has deduced that, on average, banks operate on their constant portion of their average cost curve; where medium-sized banks, rather than large and small banks, tend to be more scale efficient. In addition, Humphrey (1992) has used different output measures of banks in which the results indicate that small banks operate with scale economies, but medium and large size banks operate on constant and decreasing scale economies respectively. In general, long-run average cost curves for banks are relatively flat; however, recent studies that look more at large banks tend to find greater evidence of scale and scope economies (see the review in Chapter 5).

Scope economies exist when it costs the same or less if one or more outputs are added (to the available output set) than if different firms produce each output separately. Scope economies may be realised when mergers or acquisitions take place between firms producing different outputs. Nevertheless, even for an individual bank producing a variety of services, the bank might be enjoying scope economies. As an example, one bank may provide loans and another bank may engage in portfolio investments. If these two banks join in one entity and produce both loans and investments, scope economies may be achieved when joint production of these two outputs are less costly than the total cost of these outputs being produced separately by these banks.

The majority of empirical studies on scope economies in banking have been undertaken on the US banking industry. The evidence to support the hypothesis that multi-product mix of banks results in lower costs than if each product is produced separately has been supported in studies such as Gilligan and Smirlock (1984) and Lawrence (1989). However, studies such as Hunter, Timme and Yang (1990) and Mester (1987) have found no strong support for the existence of scope economies in banking.

Other operational efficiency aspects

Other efficiency aspects may be classified under the operational/functional efficiency category. Among these are: small funds pooling, risk pooling and uncertainty reduction. In fact, Tobin has mentioned functional (operational) efficiency with two features, risk pooling and allocation of savings to the most productive funds' users. The latter has been explained through the feature of informational efficiency, which helps financiers choose the most productive investment opportunities. Other operational efficiency aspects discussed below are relevant to this efficiency category.

Risk pooling comes from the role of diversification and spread of assets being invested in the financial system (Tobin, 1984). Banks can spread risk across large numbers of borrowers with different risk types, different projects and different sectors of an economy. Also, financial markets allow investors to make their portfolios more efficient by choosing well diversified assets. The general idea behind risk spreading is to avoid non-systematic risk; that is, the fall in the return of an investment will be recovered by the rise of return of another.[7]

Uncertainty reduction. Uncertainty reduction has been explained by Tobin as an aspect of efficiency, which he calls 'full insurance efficiency'. Insurance efficiency implies that the financial system enables its participants to have their financial assets delivered and obtained with insurance against all future contingencies. In other words, this is called 'hedging' against uncertainty. Since the volatility of stocks, interest rates, and

[7] William Sharpe (1964) developed the capital asset pricing model that examines the systematic and non-systematic risks of holding a portfolio. Systematic risks are risks that cannot be avoided even by holding well diversified portfolios. Non-systematic risks are risks that can be eliminated within well diversified portfolios.

exchange rates impede the trade of financial assets, financial derivatives (such as forward contracts, financial futures, options, swaps, and so on) are tools that allow individuals and companies to engage in contracts that contain the delivery of a specified amount and quantity of assets on a certain date. Therefore, future financial instruments are, in general, important for financial system efficiency because they reduce the risk associated with the volatility of asset prices and provide confidence and stability in the transactions within the financial system.

Funds pooling. The law of aggregation is an important efficiency feature of the financial system, which helps maximise the level of funds intermediated in the economy. Banks are the main financial system institutions able to aggregate and pool small savings in order to make large loans (Stiglitz, 1984). Financial markets can also aggregate small funds from the new issues of reasonable face value stocks and bonds. The law of aggregation can help all society's wealth classes to participate with their funds in a way that matches their wealth capacities.

4.4.3 Structural efficiency

Efficiency is also studied from the view of market structure. Market structure usually refers to the way in which the market is organised in order to provide products for end users (Rutherford, 2000, p. 288). In the financial sector, market structure embraces market competition, the nature of products produced, and the regulatory environment. The study of market structure may also go further to include the question of whether a bank-based or a market-based financial system is more efficient in the allocation of financial resources (see Levine, 2002). For example, the US and the UK financial systems are characterised by market-based finance since the financial markets play a major role in raising funds. In contrast, Japan and most European countries are bank-based financial systems. For developing countries, where financial systems are not so advanced, it may be more preferable that their financial systems be bank-based since banks are better suited in resolving market imperfections created by information asymmetry problems, which may be more severe in developing countries.

If one considers market structure the level of competition is probably the most important aspect. In banking industries, one might consider a more competitive market as better in allocating financial resources, increasing consumer welfare and achieving market stability. Although high competition contributes to the welfare of the end users, high competition may destabilise the banking system. Stiglitz (1994) has pointed out that increased competition erodes profits and increases the insolvency threat of poorly functioning banks.

When there is a contraction in banks profitability because of high competition, banks face two main choices: one is to be more cost efficient; the other is to make riskier loans. The problem is that the second choice may tend to be dominant during periods of intense competition. Therefore, when risky loans end up defaulting, the banking system may correct itself by restructuring through takeovers and mergers by banks that are more efficient. However, this usually happens after a crash that may be harmful to the banking system and the economy. As mentioned earlier, in the mid 1980s-1992, the US savings and loan industry experienced a crisis resulting from severe competition. Moreover, among the causes of the Great Depression's financial crises was the high competition in the banking industry; therefore, one of the procedures used to restore stability to the banking industry is to limit competition and restrict entry barriers (Dziobek, 1988).

Today, many banking markets around the world appear to have an oligopolistic structure. It has been argued that having a smaller number of banks is more preferable because (Cetorelli and Peretto, 2000; Cetorelli and Gambera, 1999): (1) in terms of stability provision, an oligopolistic structure means that banks will face less threat on their profitability, which helps maintain stronger solvency. Policy makers should not be concerned about consumer welfare issues as long as domestic rates and fees charged by banks are reasonable when compared to other international banking sectors; (2) a large number of banks means that there might be banks that are poorly capitalised. These banks can be the source of inefficiency and instability to the banking system because insufficient capital may induce them to undertake risky activities

(Wachtel, 2000); finally, having a small number of banks makes it much easier for the central bank to supervise these banks.

4.4.4 Regulatory efficiency

Financial regulations include a set of rules that organise the operation of the financial sector. Its efficiency derives from how appropriate the rules in providing prudential regulations are. Regulations are said to be prudent when they offer safety and soundness to the financial system so that it is protected from financial crises. The efficiency of financial regulation comes also from effective supervision. Moreover, regulatory efficiency can also mean how these regulations enhance efficiency aspects in order to further reduce market imperfections. However, it is known that policy makers, who set regulations, are also subject to information asymmetries. Therefore, the close watch of the performance of the financial system and its regulatory effectiveness should provide feedback on how effective the regulatory framework is in achieving the goals of safety and soundness.

Appropriate surveillance, the collection of information, and the effective implementation of regulations and good supervision are important elements for enhancing the efficiency of the financial system. For example, in the case where the financial system is liberalised, when there is inadequate regulations to restrict the banking and financial activities from risks that expose the system to financial distress, it is then said that regulations are inadequate and inefficient. This is because inadequate financial supervision and regulation is one source of financial instability and crisis (Stiglitz, 1998).

Government regulations aim to improve various sorts of efficiency in the financial system. The government, through either the central bank or other regulatory body, supervises banks' activities to reduce risks, maintain solvency, and enhance/maintain the soundness of individual banks and the system overall. Central banks typically impose reserve requirements as a safety line for banks to provide enough liquidity that meets bank's daily requirement. Banks may also be asked to periodically report their financial transactions to the central bank to check that they abide by these requirements.

One important feature that aims to reinforce the stability of the financial system is that various regulations are set in order to provide guarantees to protect depositors. The deposits guarantee idea originates from the US as a way to rehabilitate and restore confidence in the financial system (e.g. after the Great Depression's).[8] Although deposit guarantees can be a source of enhancing the efficiency in the financial system, guarantees may also erode efficiency as they encourage the moral hazard problem. That is, depositors may care less to impose a discipline on banks behaviour since banks know that when they are in trouble, their depositors will not withdraw their money because their deposits are protected. Moreover, guarantees may further distort banks' incentives since they will be more inclined towards taking risky activities (Mishkin, 1998).

Stock markets, which are in most countries supervised by governments, set regulations in such a way as to enhance stock market efficiency. Regulations require listed firms to maintain high reporting and other standards. For transparency purposes, which alleviate the adverse selection problem, firms are required to make their financial reports available to the public. These types of requirements improve stock market efficiency since market participants will make their decisions according to the information being available for all market participants. Moreover, obligations on the minimum accounting standards and contract enforcements limit managerial cheating and correct incentives, which can alleviate the moral hazard problem. Therefore, market prices, which are backed by effective regulations are likely to be set on the basis of firms' performance. If this is the case, managers will always try to direct their firms towards productive activities that boost the value of their firms in the stock market; otherwise, bad performance, which will be reflected by stock market valuations.

On the other hand, some government regulations, backed by interventionist policies, have induced substantial inefficiencies. For example, during the 1950s and 1960s, many

[8] Such as that offered by the U.S. Federal Deposit Insurance Corporation (FDIC), which provides insurance to deposits up to $100,000 (Mishkin, 1997).

financial systems, especially in developing economies, suffered from heavy financial restraints (interest ceilings, high reserve requirements, and exchange rate controls) in order to finance priority sectors (such as the import substitutions industries), as well as to finance government financial requirements (such as budget deficits). These heavy financial restraints created distortions in the prices of resources being allocated in the economy. They also impeded the growth of the size of the financial system since depressed interest rates did not encourage the taking of deposits and thereby investments.

The view that financial repression in developing countries formed an impediment to economic growth is suggested in the analysis developed by McKinnon (1973) and Shaw (1973). They noted that in order to remove these distortions and promote economic growth, all forms of financial repression should be removed. However, the experience of developing countries that shifted to greater liberalisation of their financial sectors (such as in Mexico and East Asia) has provided evidence of the failure of some financial liberalisation policies (Stiglitz, 1998). This may be because the context in which financial liberalisation implemented was lacking, to some extent, a variety of regulatory efficiency and other efficiency aspects. For example, in terms of regulatory efficiency, many studies, including Fry (1997), indicate that an adequate level of supervision and regulation must be accompanied with financial liberalisation policies. Adequate supervision and regulation, must therefore, enhance efficiency resulting in stability and financial (banking) productivity.

Others, such as Stiglitz (1998), reacted to the recent Asian financial crises by advocating mild financial repression that will result in various efficiency gains. Stiglitz argues that, previously, financial restraints were a policy used by various governments to earn 'rents' that enable them to finance growth projects. [9] In mild financial repression, governments administer interest rates but let the rent be contained within the market. That is, investors and household borrowers will privately allocate the generated rents. The reason why mild financial restraints can lead to more efficiency is that it encourages higher investment since the administered interest rate is meant to be slightly

[9] Rent denotes the difference between the administered interest rate and the rate that should be set by market mechanism.

lower than the rate of return on investments. Though savings will be affected, individuals will have more incentive to seek better returns on their financial assets than deposits, given that the elasticity of savings in response to the changes in the interest rate is at least low if not close to zero. The other efficiency benefit of mild financial repression is that the low interest rate will induce safer investors to show up (compared to the situation where high interest rates attracts risk-seeking investors). Thus the likelihood of defaults will decline leading to safer and prudent financial systems (see Caprio and Summers, 1993). Therefore, according to these arguments, mild financial repression can improve the efficiency of the financial system.

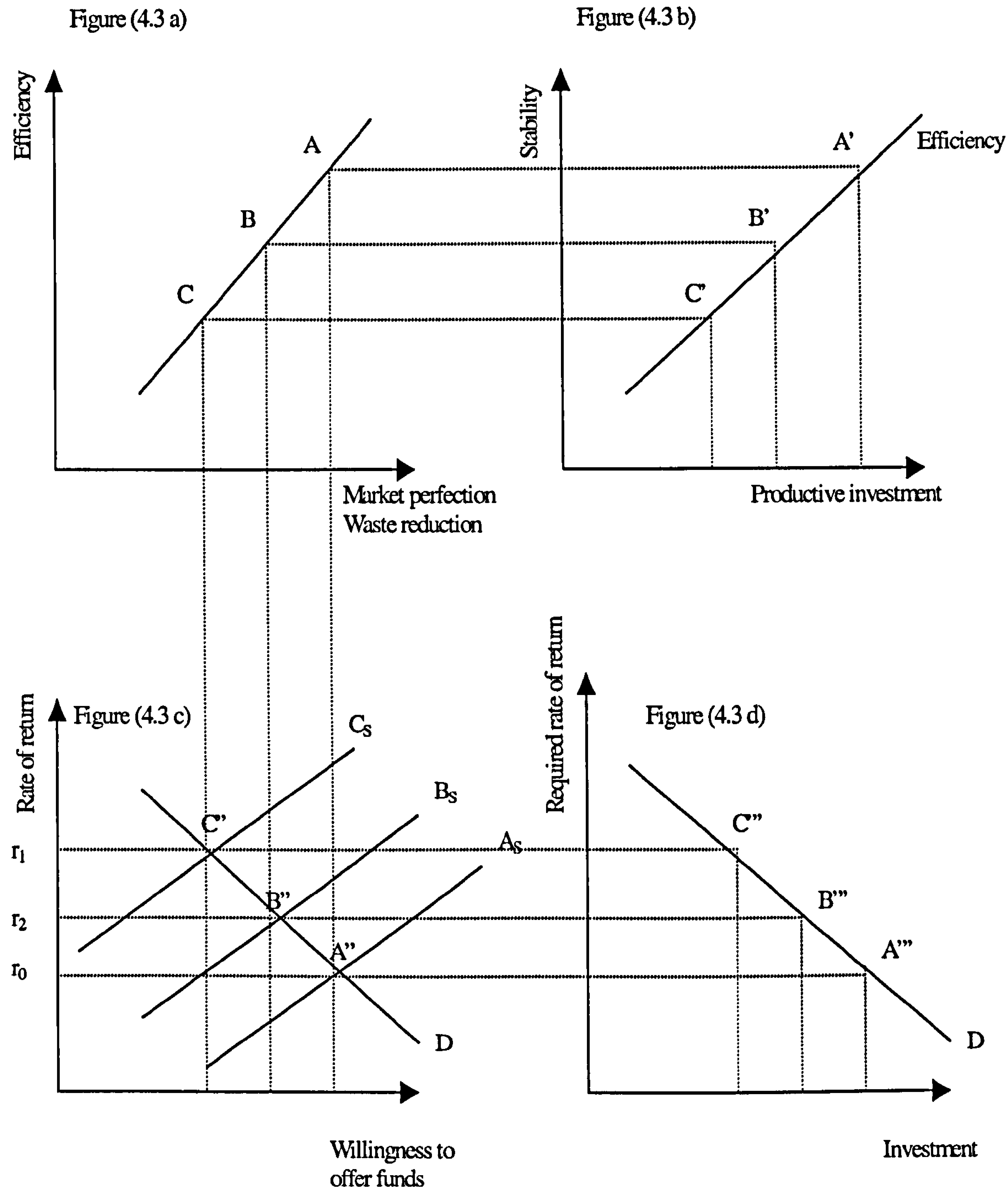
Overall, this section explains how efficiency of the financial system can have various aspects of which we classify as categories such as informational, operational, structural, and regulatory efficiency. The next section shows how the existence of the efficiency aspects in the financial system induces the latter to foster economic growth.

4.5 The relationship between financial system efficiency and the real economy

This section provides an analysis of how the level of financial efficiency affects the real economy and economic growth. [10] We assume that high (low) levels of efficiency are characterised by the existence of high (low) levels of all the aspects of efficiency explained in the above section. Figure 4.3 generally provides three levels of efficiency. The move from levels C to A (also C' to A' and so on in each diagram) means greater aspects of efficiency are available, which also indicates enhanced market perfection, less resources wasted, a greater supply of funds directed to productive investments, and (we assume) a more stable financial system. Therefore, when efficiency levels are high, economic growth is fostered. Conversely, when the level of efficiency is low, economic growth might be hurt. For simplicity, our main analysis will focus on banking and the stock market in which both of these segments are analysed from an operational and informational efficiency standpoint.

[10] We assume that this analysis is endogenous, that is, we ignore the impact of exchange rates, inflation, unemployment, and other external factors.

Figure 4.3 Financial efficiency and real economy



We start with Figure (4.3 a). The high level of efficiency enhances market perfection and waste reduction process. In banking, a higher level of cost efficiency brings prices of funds to a level more favourable to consumers. This makes the prices the consumers pay for funds closer to the costs of intermediation. In this case, if the price of funds reflects their intermediation cost, then market perfection increases because such funds' prices allow both savers and investors to make better decisions on their financing deals. Moreover, cost efficiency means banks can avoid the excess utilisation of financial resources. That is, if banks are more X-efficient, their input mix to provide the given output level will be closer to the cost frontier. Therefore, inputs will not be excessively used, which implies more resources are saved from being wasted.

Informational efficiency in banking helps in providing funds to those with better quality and successful projects, which ensures that funds are going to more productive destinations. Higher informational efficiency reduces market imperfections and makes it more possible for banks to know the quality of loan applicants. This will allow the economy to preserve resources compared to the case in which funds were going to non-productive loan applicants. Although, banks incur costs associated with improving informational efficiency, it is assumed that the amount they save from the cost of loan defaults compensate for this.

In the financial (capital) market, informational efficiency occurs when the stock market sets prices correctly. Here, investors will pay prices that reflect all available information. In this situation, the level of market perfection will increase since financiers will be able to make a better selection relating to a more accurate assessment of the quality of firm. This will help avoid the waste of financial resources because funds in this case will go to the most productive investments.

Therefore, higher efficiency (point A in Figure 4.3 a) brings about more productive investments (point A' in figure 4.3 b) because cost efficiency contributes to the provision of funds by moderate prices (interest rates) which attract more safer borrowers. Moreover, informational efficiency maintains the success of projects receiving the funds, and assures that the receivers of funds are those with successful economic activities.

Greater efficiency (point A in figure 4.a) also results in higher stability (point A' in figure 4.3 b) for many reasons. In banking, informational efficiency makes the funds' providers avoid the selection of low quality borrowers, all other things being equal. Also, the collection of information during the loan period reduces the moral hazard problem. Therefore, banks will reduce the likelihood of insolvency as the probability of default decreases. In the stock market, informational efficiency restrains asset prices to go far beyond their correct values. This, in theory, should result in a more stable stock market that also enhances investor confidence. Also, informational efficiency, which also requires information to be available to all market participants, reduces speculative activities. In this case, the current value of the price equates the present value of future income streams, which indicates that all investors will form the same expectations as to the behaviour of stock prices.

At point A" in figure (4.3 c), higher efficiency brings about a greater supply of funds since improved efficiency reduces the required rate of return on financial assets. For example, in banking, cost efficiency can narrow the net interest rate margin.[11] Informational efficiency reduces the risk premium to be set on interest rates; as more information is available on the quality of borrowers, the funds' owners are in a better position to supply more funds. The high supply of funds with a low required rate of return will induce more investment in the economy as it is shown at point A"' in figure (4.3 d).

In sum, as this Figure 4.3 shows, efficiency is an important aspect of financial system operations as it enhances the soundness of the financial system, having implications for the real economy and the goal of economic development.

4.6 Conclusions

There are many studies that tackle efficiency features of the financial system. These efficiency studies have many objectives; although, generally speaking, it is rare to find a

[11] The difference between deposits and loans interest rates.

single study that examines the efficiency of the overall financial system. In this chapter, we attempt to provide an overview of the broad literature that looks at financial system efficiency. One can see that an analysis of financial system efficiency goes beyond the analysis of market competition and government intervention. From the goals of stability and economic productivity, we show how various elements relating to banking and capital market efficiency are important. The following chapter aims to narrow the focus of this and examine the literature on banking sector efficiency.

CHAPTER 5

MEASURING BANKING SECTOR EFFICIENCY: THEORY AND EMPIRICAL EVIDENCE

5.1 Introduction

This chapter explores the theoretical background and reviews the empirical literature on the main aspects of operational efficiencies that are extensively studied in the banking literature; these aspects relate to economies of scale, economies of scope and X-efficiency. This chapter includes the following sections. Section 5.2 outlines the importance of the study of banking sector efficiency for all parties participating in the financial system. Section 5.3 identifies how the literature views a bank in the context of the theory of the firm. Section 5.4 defines banks' inputs and outputs in the production process. Section 5.5 and 5.6 illustrate the theoretical background of both economies of scale and scope. These sections also review a number of important studies examining both these aspects. Section 5.7 discusses X-efficiency and examines the different approaches and studies used to measure these types of efficiency. The section also illustrates the various functional forms that have been used in the literature to model bank efficiency. In addition, the section provides a review of recent empirical studies undertaken to measure efficiency in various banking systems. Conclusions are given in section 5.8.

5.2 The importance of operational efficiency studies

Studies on operational efficiency (scale economies, scope economies, and X-efficiency) have gained more attention by financial system policy makers and regulators, researches, managers, and owners of financial institutions.

Policy makers and regulators can benefit from a further understanding of the efficiency of banks as the performance of the banking sector can impact on certain policies implemented in the financial system. For example, bank efficiency studies are helpful in judging the extent to which deregulation aimed at improving efficiency through the removal of restrictions (e.g., interest rate restrictions and entry barriers) stimulate industry performance, create social benefit by reducing waste in resources, and increase competition that reduces the market prices of financial services (Berger and Humphrey, 1997).

Regulators consider efficiency studies to be important for market structure and performance, especially when examining if bank profitability is driven by market power or efficient operations (see Berger, 1995; Molyneux, Altunbas, and Gardener, 1996, Chapter 4). Concentrated banking sectors may make banks operating in the same industries earn high profits through setting prices of financial products and services at levels unfavourable to customers. This situation is known as the market-power hypothesis. In an alternative view, a new paradigm has incorporated the role of efficiency, which has become known as the efficient-structure hypothesis. This hypothesis suggests that more efficient banks are able to generate higher market shares and earn high profits that are mostly induced by competitive prices enabled by efficient performance rather than market power practices. Hence, testing whether the efficient-structure or market power hypothesis prevails can provide regulators with information about the appropriate reforms in the banking industry.

Studies on efficiency can also provide signals on the health of the financial industry. They can help to identify efficiency sources that could either strengthen or harm the performance of the banking industry. For example, many studies have found that strong capital levels are connected to efficient bank performance because banks that perform well are able to generate higher profits that strengthen their solvency base; on the other hand, the level of problem loans is found to be negatively related to bank efficiency (Berger and Humphrey, 1992; Hermalin and Wallace, 1994; Mester, 1996). Studies that link bank efficiency to financial soundness help to provide regulators with information about the source of inefficiency and how this may be related to banking sector risk. As such, this type of evidence can help inform public policy relating to banking sector risk.

From a research perspective, efficiency studies provide commentators on the financial system a wider range of efficiency indicators that can be compared with other measures. For instance, efficiency studies enable researchers to check how their results may be affected when implementing different modelling techniques and these measures derived from various optimisation techniques can be compared with traditional accounting indicators. Researches also aim to increase the accuracy of banks' rankings according to their efficiency measures in order to help identify best and worst practice institutions in order to help set policies that encourage the former and improve the latter (Berger and Humphrey, 1997).

Efficiency studies are important for managers since, from the point of view of business strategy, managers need to take the steps or find the reasons and the determinants for why and how they can improve their efficient performance from both the input side (by improving cost efficiency using better information technology, managerial practices, and enhancing capital) and on the output side (by improving profit efficiency through their marketing and pricing strategies).

Studies of efficiency are also important from a shareholders perspective because they appoint managers and expect them to run their financial firms efficiently. Having a wider range of best-practice benchmark indicators may help shareholders monitor their managers more effectively. It is clearly in shareholders interest that managers maintain efficient performance that ensures stable profits and soundness for the bank or banks in question.

Overall, bank efficiency studies can provide results that are of interest to financial policy makers, financial institution managers, owners, as well as academic researchers. The study of banking sector efficiency can provide useful added information as to the extent of resource employment in the banking sector, financial institutions' profitability, market power, and the safety and soundness of the financial system overall.

The next section explores the microeconomic aspects of the bank production process focusing on different approaches to defining banking inputs and outputs.

5.3 Inputs and outputs in bank production

In the context of the traditional theory of firm, banks could be viewed as financial firms that employ certain input resources and transform them into certain outputs. However, the treatment of banks in the context of the theory of firm is relatively complex, mainly because there is no consensus as to what a bank actually produces. If one considers the production process, it is important "[to] appropriately classify outputs and inputs of the financial firm by considering the criteria on which the financial firm makes economic decisions" (Sealey and Lindley, 1977, p.1251). Problems arise as it is by no means certain as to what constitutes the input or output side of bank production. For example, there is no consensus as to whether deposits should be treated as services that banks produce because deposits are items used as inputs transformed into loans.

Another difficulty associated with banks' production is related to the nature of the bank as a financial firm since its production is characterised by non-physical items, which could lead to measurement difficulty. For example, there is no consensus as to whether it is better to measure banks' output in terms of the number of accounts or the value of these accounts. A bank may appear to have a large number of accounts but when it is compared to another bank in the same sample, the value of its accounts might be less than the value of accounts of another bank with a lower number of accounts (Heffernan, 1996). In extension, this difficulty in defining bank output may lead to problems in the measurement of bank productivity. For example, is it best to use loans, deposits, or assets to measure the productivity of employees and/or branches? Moreover, even if one defines bank output, problems still remain in identifying the quality of the output.

Since balance sheet accounts are designed to give information on a bank's resources (e.g., financial capital and other liabilities) and the uses of these resources (e.g., the assets side), these accounts may also provide information on a bank's inputs (from the liabilities side) and outputs (on the assets side). As Berger and Humphrey (1990, p. 247) have stressed "[v]irtually all observers would agree that banks liabilities have some characteristics of inputs, because they provide the raw material of investable funds, and that bank assets have some characteristics of outputs as they are ultimate uses of funds that generate the bulk of the direct revenue that bank earn." Therefore, although the bank balance sheet may give a potential insight of a bank's inputs and outputs, there is however no consensus as to whether the balance sheet classification of liabilities and assets should be used in explaining the production process of a bank.

The pivotal issue in defining a bank inputs and outputs lies on one of the main items of the balance sheet, deposits -- an item that has stemmed the controversy as to whether this should be considered under inputs or the outputs classification of a bank production technology. Some studies adopt a dual approach in order to resolve how deposits should be treated. For example, Hughes and Mester (1993) and Bauer et al. (1993) have used demand deposits as outputs and time deposits as inputs, considering interest paid as a price of inputs as well as the treatment of interest paid as a part of total cost. However, other researchers have attempted to empirically test whether deposits should be classified as inputs or outputs (see for example, Hughes and Mester, 1993; Favero and Papi, 1995). The test is generally based on the idea that when the use of some inputs increases, expenditure on other inputs should decrease. The findings of these studies tends to show that deposits are negatively related to other inputs for given outputs, suggesting that deposits are better considered as inputs rather than outputs.

In the banking literature, anyhow, it appears that there are two main lines in defining inputs and outputs of banking institutions; these are: the intermediation and the production approaches (see Humphrey, 1985; Berger and Humphrey 1990). In both approaches, the treatment of deposits is clearly identified.

Basically, both the intermediation and the production approaches agree on the view that labour and physical capital items are inputs used in the banking production process. The main differences between the two approaches lies in how to view deposits and how banks' inputs and outputs should be measured. The intermediation approach treats deposits as a category of inputs since banking firm's decision-making process rely on deposits to produce earning assets such as loans (Sealy and Lindley, 1977). In contrast, the production approach sees deposits as a part of a banks' outputs on the grounds that deposits are attracted using bank resources (such as labour and capital) so as to offer customers with value-added outputs in the form of clearing, record-keeping, and security services (Bauer et al., 1993; Berger et al., 1997; Resti, 1997).

For the measurement of inputs and outputs, the intermediation approach uses the currency value of accounts and considers both operating and interest costs. In contrast, the production approach measures banks' outputs by physical quantities (such as the number of deposit accounts, loans accounts, current accounts, and so on) and considers only operating costs.

Providing that outputs in the production process are flow variables, and that most of the data taken from the financial reports are stock variables, the production approach employs data with a flow nature, while the intermediation approach uses stock data (Heffernan, 1996). For the latter approach, researchers tend to use the monetary value of inputs and outputs mainly because it is difficult to obtain detailed information on deposit and loan transactions. Accordingly, it is simpler to obtain monetary variables knowing that these proxy for the average flow of such banking variables (Resti, 1997). Studies by Benston et al. (1982b) and Berger et al. (1987) have concluded that the uses of monetary values (stocks) or physical quantities (flow) in input and output definitions show no substantial differences in efficiency results using these two different metrics.

It should be noted, however, that both the intermediation and production approaches ignore the maturity structure of loans (Heffernan, 1996), where the amount of total loans could differ when the study is multi-period in which some loans are repaid.[1] This might create some bias in the intermediation approach because of using the value of accounts especially when the changes in the amounts are large. However, the effects are less on the production approach which uses number of accounts.

Along with the intermediation and the production classification, various other approaches to defining banks' inputs and outputs have been applied in various banking studies. Among these are: the user-cost approach and the value-added approach (see Berger and Humphrey, 1990).

The user-cost approach emphasizes how a category in the bank balance sheet adds to the net contribution of total revenue. Under this approach, a category in the banks' assets is considered an output if its returns exceed the opportunity cost of funds. If not then this category is considered an input. Likewise, a category of the banks liability is considered an output if its costs are less than the opportunity cost of the funds; otherwise it is an input.

The value-added approach claims that a category, whether it is in the liabilities or assets sides of the balance sheet, should be considered as a bank's output if the category generates an important value added to a bank. On the other hand, a bank's activities from which the bank creates low added value are treated as unimportant outputs, intermediate outputs, or inputs. For example, balance sheet items such as loans and deposits are expected to be treated as a banks' output since they add a significant amount to the majority of banks' value-added; however, purchased funds are considered as inputs, and government securities are classified as 'unimportant' banks' outputs because of their low value added.

[1] Moreover, the change in the interest rate may change the deposits and loans significantly.

In practice, the intermediation approach is the most widely used in the bank efficiency literature. Many studies adopt this approach for various reasons.[2] Firstly, it conforms with the microeconomic theory of intermediation since this approach emphasizes that funds deposited are intermediated to lenders with minimum costs (Berger et al., 1987; Ferrier and Lovell, 1990). Secondly, Kaparakis et al. (1994) adds that it is better to use the intermediation approach when large banks are to be included in the sample; this is because they fund a large share of their assets from non-deposit sources. Thirdly, data on the number of accounts is difficult to obtain, as this information is usually proprietary in nature. Typically, the production approach has been used to study the efficiencies of branches of financial institutions because branches deals with customer documents and process them for the financial institution as a whole, and the manager of these branches have little influence over banks funding and investment decisions (Berger and Humphrey, 1997).

In conclusion, the main objective of this section is to explore the main microeconomic aspects of the bank production process, with a focus on the different approaches to the definition of banking inputs and outputs. The definition and the measurement of banks' inputs and outputs are still a controversial issue in the banking literature. Moreover, none of the above mentioned approaches is perfect because they do not fully capture the overall role of financial institutions, which consists of the provision of transaction services, document processing services, and the intermediation of funds from financial surplus agents to deficit agents (Berger and Humphrey, 1997). The main approach adopted in the literature, however, is the intermediation approach (see Chapter 6).

The following three sections discuss the theoretical background, measurements methods, as well as the results of the empirical studies on economies of scale, economies of scope, and X-efficiency in banking.

[2] For example, Berger and Mester, 1997; Altunbas et al., 2000; Allen and Rai, 1996.

5.4 Economies of scale in banking

A firm enjoys economies of scale when the production of one more units of an output leads to a decline in unit production costs. Usually, economists view economies of scale as the average total cost in relation to the quantity produced.

Before explaining the relationship between costs and outputs, it is important to clarify the relationship between costs and inputs (Binger and Hoffman, 1988). We know that the long run total cost as a function of inputs is always homogeneous of degree 1; that is, long run total costs and inputs increase in the same proportion as $w(\alpha L)+r(\alpha K)=\alpha(wL+rK)$, where α is a positive constant greater than 1, L and K are inputs, w and r are prices of input L and input K respectively. This equation says that when inputs are multiplied by α, costs are multiplied by the same constant.

In addition, costs and outputs can be explained in the context of long-run rather than short-run relationships, given that in the long-run costs allow for simultaneous changes in all inputs of production whereas in the short-run at least one input may be fixed.

We can explain scale economies by breaking down total costs into fixed and variable costs. Variable costs are direct costs that change with the level of production. Fixed costs are overheads or unavoidable cost (at least in the short-run) and are independent of output.

We move now to explain the relationship between costs and outputs. Usually, this relationship may take one of three aspects: increasing returns, constant returns, or decreasing returns to scale. These are illustrated as follows.

5.4.1 Increasing returns and economies to scale

Increasing returns to scale implies that when multiplying all inputs by $\alpha > 1$, output multiplies by more than α:

$$x(\alpha K, \alpha L) = \alpha^{k} x(K, L) > \alpha x(K, L), \; for \; \alpha, k > 1.$$

In other words, knowing that long-run total costs and inputs change in the same proportion, if inputs are multiplied by a positive constant, α, costs are multiplied by the same constant, but output is multiplied by more than that constant. This means that costs increase by a slower rate than the rate of increase of the output. As Figure 5.1 shows, the total cost function is concave and is increasing in a diminishing rate with respect to an increase in the output.

Figure 5.1 Long-run cost functions for an increasing returns to scale

Third Party material excluded from digitised copy. Please refer to original text to see this material.

Source: Adapted from Binger and Hoffman (1988, p. 260)

Thus the marginal cost, or $\frac{\Delta LRTC}{\Delta x}$, must be declining. Moreover, the long-run average cost, $\frac{LRTC}{x}$, lie above the long-run marginal cost since the spread of the fixed costs, or the costs concerning each additional unit of output, is attributed with higher fixed cost than variable costs.

5.4.2 Constant returns to scale

Constant returns to scale implies that changes in total costs are proportional to changes in output with a relationship that takes fixed proportion. Thus, as Figure 5.2 shows, the long-run total cost is a linear function of output, and the long-run average and marginal costs are constant, or independent of output where the curve of the long run marginal cost takes a constant pattern at b since $\frac{\Delta LRTC}{\Delta x} = b$, and the long run average cost curve, $\frac{LRTC}{x}$, is also fixed at b.

Figure 5.2 Long-run cost functions for a constant returns to scale

Third Party material excluded from digitised copy. Please refer to original text to see this material.

Source: Adapted from Binger and Hoffman (1988, p. 259)

As output increases, the decrease in the per unit cost attributed to the spread of fixed cost is offset by an increase in the variable cost, resulting in a constant proportionate change in total cost with respect to a change in output levels, shown at b.

5.4.3 Decreasing returns to scale

When a production function exhibits decreasing returns to scale, this can be indicated as

$$x(\alpha K,\alpha L)=\alpha^{k}x(K,L)<\alpha x(K,L),\ for\ \alpha>1,k<1.$$

Given that costs are proportional to outputs, if inputs are multiplied by a positive constant, costs are multiplied by the same constant, and output is multiplied by less than that constant. This means that costs increase at a higher proportion than the increase in outputs (see Figure 5.3). This leads the long-run average cost, which is $\frac{LRTC}{x}$, and the long-run marginal cost, which is $\frac{\Delta LRTC}{\Delta x}$, to be upward sloping, meaning that the spread of the fixed cost is exhausted and the increase in total cost is largely attributed to an increase in the variable costs.

Figure 5.3 Long-run cost functions for a decreasing returns to scale

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Source: Adapted from Binger and Hoffman (1988, p. 201)

In general, scale economies can be measured as

$$SE = \frac{\textit{Long-run average cost}}{\textit{Long-run marginal cost}} = \frac{f(x)}{x(\Delta LRTC / \Delta x)}.$$

This relationship between costs and inputs is simply the elasticity of cost with respect to output, where SE>1, SE=1, and SE<1, imply increasing, constant, and decreasing returns to scale respectively (Molyneux, Altunbas, and Gardener, 1996, p. 139).

In theory, average and marginal costs are usually thought to be represented by U-shaped curves in relation with output (see Figure 5.4). This can be explained by putting together the analysis of cost and output with references to fixed and variable costs. The fall in total costs in the stage of increasing returns to scale is driven by a decline in both fixed and variable cost since, for the initial level of outputs, the share of fixed cost is larger than that of variable cost per unit produced. In essence, "[a]verage costs decline initially as fixed cost are spread over additional units of output. Average costs eventually rise as production runs up against capacity constraint" (Besanko, Dranove, and Shanley, 2000, p. 72). Thus, as fixed costs are spread over an increasing volume of outputs, the share of fixed costs decline much faster than the variable cost and the latter is responsible for pulling the average cost curve upward, since the contribution of variable cost to the total cost per additional unit of output produced outweighs the share of average fixed cost. Therefore, the fall in average and marginal costs, shown in the decreasing portion of these cost curves, is largely due to fixed cost, however, an increase in the costs curves is largely due to an the increase in variable costs.

Figure 5.4 Long-run cost functions with a U-shaped long-run average cost function

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Source: Adapted from Binger and Hoffman (1988, p. 262)

The analysis above assumes that the bank is a single product producer, and other factors, such as technical progress, are held constant. However, in the case of a multi-product bank, the construction of the cost function to analyze scale economies is not straightforward because there is no longer one unit of measurement (e.g., one output). One way to measure the effects on costs for a multi-product firm is to estimate product-specific economies of scale of q_i; that is, we see what impact on costs an increase in one output has holding the other outputs constant (see Baumol, Panzar, and Willing, 1982, Chapters 3 and 4; Panzar and Willing 1977). In this case the multi-product cost function is similar to the single product cost function except that it measures scale effects of specific outputs assuming that all levels of other outputs are fixed. Therefore, if the ratio of change in the total cost with respect to the change in the specific output, holding other outputs constant, is less than one then scale economies exist; if this ratio is equal to one, constant returns to scale prevails; if the ratio is greater than one, decreasing returns to scale exist since the average cost curve would be rising.

There are several reasons why bank business may be characterised by scale economies. Forestieri (1993) identifies a number of possible factors that may bring about scale economies in banking. 1. Administrative procedures associated with monitoring and screening of borrowers are likely to fall when a firm applies such techniques to a larger number of customers, and this should lead to a declining average cost for loan granting [Arrow, 1965; Williamson, 1975; Berger et al., 1987; Shaffer, 1991; Humphrey, 1991]. 2. Larger banks may exploit their size by employing specialized labour (technical and managerial labour) that adopts more efficient organisational forms (Clark, 1988). 3. As bank size increases, the use of IT (information technology) helps better utilisation of resources because of imperfect divisibility of investments and the facilitation of more flexible production process (Landi, 1990). 4. Some inputs may have excess capacity so that an increase in output only accounts for the exploitation of this capacity, given that the increase in output does not require an increase in all inputs over the entire production period (Bell and Murphy, 1968).[3]

Estimates of the degree of scale economies have important implications for firm expansions as well as for policy implications (Binger and Hoffman, 1988). For banks, it is important to discover the relationship between scale and cost so that the bank, with the scale and output information, knows whether such an expansion leads to an increase or decrease in costs. From a policy perspective, a firm's return to scale has important

[3] On the other hand, there are several factors that may lead to scale diseconomies. Most of these could be related to size of the firm. Large firms pay more labour cost than small firm. Moreover, large firms find it more difficult to monitor and evaluate employees' activities and tie it with their performance. In this case the contribution of the employee to the firm profit is not straightforward.

implication for market structure and entry policies. For example, it is generally expected by economists that when an industry exhibits substantial increasing returns to scale then the industry could become monopolistic in structure. Moreover, a market with a certain group of firms that enjoy increasing returns to scale could be viewed as a concentrated market. This is because economies of scale allow firms to offer more competitive prices and thus capture a larger share in the market. This has implications for merger and competition policy if one knows that economies of scale are important over time. For instance, it helps inform policymakers to identify at what levels of concentration further M & A activity maybe prohibited? It also informs policy concerning how new entry be encouraged in the sector, for instance, if optimal size is very large this may limit entry of new firms.

The following subsections shed light on empirical studies that examine economies of scale. Most of these studies are undertaken on the US banking sector, with a more limited European literature.

5.4.4 Studies on scale economies [4]

In this subsection, we focus on selected studies that have provided important contributions to the empirical banking economies of scale literature.

5.4.4.1 Scale economies studies on US financial industry

One of the earliest studies on economies of scale in banking was undertaken by Alhadeff (1954), where he uses earning assets as outputs (specifically loans plus investments) to investigate cost differences between branches and unit banks in the state of California. The author uses interest bearing deposits, labour cost, and miscellaneous costs as inputs. Over the study period 1938-1950, his general finding reveals the existence of economies of scale with small and large banks exhibiting increasing returns

[4] For a more comprehensive summary of the literature see appendix 1.

to scale and constant returns to scale for medium sized banks. The results also show that branch banks had higher average costs than unit banks.

Horvitz (1963) followes Alhadeff in using the same framework with data covering the period 1940-1960. Horvitz's findings replicate Alhadeff results in that Horvitz's findings suggest that small banks ($5 million deposits) do not achieve any substantial cost reductions until they grew to over $500 million in deposits size. Moreover, the average cost of branch banks was found to be generally higher, on average, than unit banks over the study period.

The use of earning assts as a measure of bank output, however, received criticism.[5] For example, it has been argued that this definition of output does not differentiate between important types of banks' credits since it treats all loans types as one output. Moreover, the consideration of only two outputs has received criticism since this discarded the role of other important variables, something that may result in an overestimation of costs relating to the specific output. Given their limitations, later studies incorporate total assets as a measure of banks outputs, for example, studies like Schweiger and McGee (1961) and Gramley (1962) use multiple regression techniques to account and control for factors that affect costs other than bank size. In essence, they employed total deposits, growth in assets, and fee income as independent variables in their cost regressions. Both studies find that the scale coefficient (the coefficient of the bank size independent variable in relation to cost dependent variable) is negative with a magnitude indicating substantial economies of scale. These studies generally find that large banks have greater cost advantage than smaller banks. [6]

Overall, the studies undertaken prior to 1965 tend to show that scale economies exist in US banking. Moreover, these studies find that large US banks benefit from scale economies, although there was less agreement on evidence of scale economies for small and medium sized banks. The following covers the findings of empirical studies undertaken on scale economies in US banking from the mid-1960s to the late 1970s.

---

[5] See Greenbaum, 1967.
[6] See Benston, 1972.

Benston (1965a) uses a Cobb-Douglas cost function to investigate scale economies in US banking over the period 1959-1961. The banks investigated in his sample were relatively small banks ranging from \$3.4 to \$55 million in total assets size. Bank outputs were defined as demand deposits, time deposits, mortgage loans, instalment loans, business loans and securities. The finding of this study suggest that scale economies is present but at a low level for all banking services. Moreover, the study indicates that bank branches also exhibit scale economies for these banks with up to three branches, where as costs increased for banks with larger branch size.[7]

Bell and Murphy (1968) follow Benston's (1965a) study in using the same functional form, the Cobb-Douglas cost function, but use a larger sample of 238 banks with data from the Functional Cost Analysis programme of New York, Philadelphia and Boston. Bell and Murphy found evidence of economies of scale for most bank services including demand deposits and business and mortgage loans.

Longbrake and Haslem (1975) use the Functional Cost Analysis (FCA) data for the year 1968 on 989 banks. They find that unit banks have the lowest cost. However, as the number of accounts and average deposit size increase, unit banks are shown to have higher costs. Moreover, they find that the number of offices operated by branch banks have little effect on the average cost per dollar of demand deposits. They also find that average costs decline for all banks except unit banks that are not affiliated with holding companies.

In general, studies undertaken between the mid-1960s up until the late 1970s tend to find that scale economies are evident for small and medium-sized banks, while large bank exhibit diseconomies of scale. Moreover, these studies' general findings tend to show that branch banks operate at higher average cost than unit banks.

[7] Greenbaum (1967) criticises the use of total asset approach on the ground that total assets aggregation in measuring output ignores the indivudual weight of important categories in the total assets. He, instead, used weighted output index. Greenbaum findings indicated that average cost decline for small banks and increase for large banks, suggesting a U-shaped average cost. Moreover, the findings asserted that branch bank operating cost were higher than unit bank costs.

Other issues motivated later studies on scale economies. Most importantly, in realisation of the shortcomings of the Cobb-Douglas cost function approach, later studies tend to use the more flexible translog cost function approach. (This issue is discussed in section 5.7).

Benston et al. (1982a) and Berger et al. (1987) assert that the previous literature does not distinguish between scale economies at the level of the bank branch office and at the level of the banking firm. They note that holding the number of branches fixed in the cost or production equation does not provide the possibility that both the number and size of braches may expand as production increases. This, they argue, lead the early studies (based on scale economies at the branch level) to misleading results. Benston et al. (1982a) and Berger et al. (1987) therefore control for branch size and expansion effects and they generally find that scale economies occur at branch level and scale diseconomies at the banking firm level.[8]

Kolari and Zardkoohi (1987) used US Federal Reserve Functional Cost Analysis (FCA) data for the period 1979-1983. The authors estimate three different models representing various aspects of bank production. The first defines bank output as the Dollar value of demand and time deposits, the second uses an output as the Dollar value of loans and securities, and the third model specifies output as the Dollar value of loans and total deposits. Generally, the main findings suggest that cost curves for all US banks are U-shaped. These findings also indicate that unit banks have relatively flat cost curves (constant returns to scale), while branch banks exhibit U-shaped cost curves but these tend to be more upward sloping – suggesting scale economies at relatively low levels of output.

Humphrey (1987) examined scale economies by investigating cost dispersion among similar banks sizes. He pointed out that the source of cost difference across banks size could be explained by scale economies across different banks sizes and cost variations

[8] Using deterministic translog cost function to measure scale economies, expansion-path scale economies and expansion path sub-additivity, Berger et al. (1987) found that scale economies are shown slightly at the branch level, but large scale diseconomies at the banking firm level.

across similar sized banks. In a sample of 13,959 US banks observed over 1980, 1982, and 1984, Humphrey divides bank data into 13 size classes and looks at the average cost of these banks. The author find that the variation in average costs between banks that have the highest cost in comparison with those having the lowest cost is two to four times larger than the observed differences in the average cost across bank size classes. Moreover, the result on cost economies does not show strong evidence of competitive advantage for large banks over small banks.

Studies that use the translog functional form to model US bank costs, and mostly undertaken in the 1980s, suggest that the estimated cost function is characterised by a U-shaped average cost curve. Although these studies do not consistently show the optimal size for a US banking firm, they suggest that scale economies exist at relatively low bank size levels, somewhere between $25 and $200 million in deposits. As with the earlier Cobb-Douglas studies, while scale economies are found at low levels of bank output they seem to disappear when banks become larger.

Most of the studies on bank cost functions during the 1980s and 1990s focus on identifying the bank size where economies of scale are realised. For example, Humphrey's (1990) survey suggests that very large banks do not tend to exhibit economies of scale. Berger et al. (1993b) refers to various other studies (including Berger et al., 1987; Ferrier and Lovell, 1990; Berger and Humphrey, 1991; Bauer et al., 1993) that focus on estimating the minimum level of the U-shaped average cost curve. Taken together, these studies tend to show that banks with assets between $75 million and $300 million tend to have the minimum average cost. Berger et al. (1993b) notes that for larger banks with assets over $1 billion [summarising the findings of Hunter and Timme (1986, 1991), Noulas et al. (1990), and Hunter et al. (1990)] tend to find that that minimum efficient scale is achieved for banks with assets between $2 billion and $10 billion.

While differences in methodological approaches may be one reason for the differences in results, the above evidence indicates little evidence of substantial economies for large banks. Other evidence from the US banking sector that use alternative nonparametric approaches (see section ***5.6.1.1***) to model bank costs find that increasing returns to scale are evident for banks at least up to $500 million in assets size, and constant returns to scale thereafter (McAllister and McManus, 1993; Mitchell and Onvural, 1996).

5.4.4.2 Empirical Evidence on Economies of Scale in European Banking

Similar to the earlier studies, most of the European cost economies studies prior to the mid-1980s used the Cobb-Douglas and CES functional forms to model bank costs. From the mid-1980s the literature uses more flexible functional forms, such as the translog to model bank's production process. Overall, while there is greater evidence of scale economies in European banking compared to the US, there remains little evidence to support the view that scale economies are prevalent for large banks.

Fanjul and Maravall (1985), for instance, study 83 commercial banks and 54 savings banks and use the Cobb-Douglas functional form to estimate scale economies in the Spanish banking market in 1979. The authors find significant economies of scale when focusing on accounts per branch. Also, the findings suggest the existence of economies of scale with respect to deposits per account. When estimated for the number of branches the findings report constant returns to scale, however. Rodriguez, Alvarez and Gomez (1993) also examine scale economies for 64 Spanish savings banks in 1990. Using a hybrid translog function, the results revealed scale economies for medium-sized saving banks, but scale diseconomies were reported for larger institutions.

Gathon and Grosjean (1991) studied Belgian banks and find decreasing returns to scale for the four largest banks, but decreasing returns to scale for all other bank sizes in this financial system. However, Pallage (1991) find scale diseconomies for large Belgian banks.

Levy-Garboua (1997) examines the French banking markets using a sample containing 94 banks for 1974. With a methodology combining both the production and intermediation approaches, their results indicate the evidence of increasing returns to scale. Dietsch (1993) examines the cost structure in the French banking markets in 1987. Using a sample of 343 banks Dietsch find that banks enjoyed scale economies across all output ranges.

On studying the banking cost structure in Italy, the general findings suggest that Italian banks generally exhibit scale economies (see for example, Cossutta et al., 1988; Baldini and Landi, 1990; and Conigliani et al., 1991). Cossutta et al. findings suggest evidence of scale economies at the plant level. At the firm level, increasing returns to scale were reported for large banks, given that small banks show constant returns to scale. The studies by Baldini and Landi (1988) and Conigliganin et al. (1991) suggest the existence of scale economies at the plant level but only for smaller banks at the firm level.

Vennet (1993) uses the translog functional form and studies 2600 credit institutions operating in the EU banking industry for the year 1991. The author found scale economies were realised for bank assets sizes in the range between $3 and $10 billion. In the cross-country studies on scale economies undertaken by Altunbas and Molyneux (1996) on four European countries (France, Germany, Italy, and Spain) the authors find that scale economies were evident across a wide range of bank sizes. Their finding also indicate strong evidence of economies of scale across all output sizes for French, German, and Spanish banking systems, except for the Italian banks, which tended to exhibit constant returns to scale.

Cavello and Rossi (2001) examine the cost features of 442 European banks over the period 1992-1997, and find evidence that scale economies existed in the main banking systems although they were more pronounced for small-sized banks. On studying 15 European countries over the period 1989-1997, Altunbas, Gardener, Molyneux, and Moore (2001) find that economies of scale were extensive across the smallest banks and banks that range between ECU 1 billion and ECU 5 billion size.

In sum, the findings from the European studies reveal greater evidence of economies of scale in banking than the US, however, there is no consensus as to the level of output at which these economies are exhausted.

5.5 Economies of scope

Economies of scope exist when a firm achieves cost savings by increasing the variety of products and services the firm produces. Unlike economies of scale, which is related to declining average costs for additional unit output produced, economies of scope are related to a decline in the total cost when outputs are produced together in a single firm relative to producing them separately in different firms; or it is cheaper for a firm to produce varieties of outputs by one branch rather than producing them by different branches (See Baumel et al., 1988; Sinkey, 1992; Binger and Hoffman, 1988).

This can be shown in a formal way as follows. Suppose that there are two branches of a single financial firm, A and B, where branch A produces product X and branch B produces product Y; the cost functions for these products in each branch is given by $TC^A(Q_x,0)$ and $TC^B(0,Q_y)$ respectively. If the financial firm finds that both products X and Y should be produced by only one branch, say A, then the total cost of producing both products by branch A becomes $TC^A(Q_x,Q_y)$, and economies of scope could then be achieved when:

$$TC^A(Q_x,Q_y) < [TC^A(Q_x,0) + TC^B(0,Q_y)]$$

This mathematical expression says that the total cost of the joint production of both products X and Y, $TC(Q_x,Q_y)$, produced by a single branch, branch A, is less than the sum of total cost of each product produced separately in both branches A and B.

The extent to which scope economies can exist for a firm (or a branch) can be measured by:

$$S = \frac{[TC^A(Q_x,0) + TC^B(0,Q_y)] - TC^A(Q_x,Q_y)}{[TC^A(Q_x,0) + TC^B(0,Q_y)]}$$

A negative value for S would mean diseconomies of scope because it is more expensive to combine the production of X and Y in one branch, A. It is only when S is positive that economies of scope exist, and the closer the value of S to one the more important it is to limit the production of the two outputs to a single branch A.

Several reasons are explained for why joint products by an entity may be less costly than producing the same products separately by different entities. Berger et al. (1987) point out several reasons such as: fixed costs can be spread over a wider range of outputs and levels, information economies since information on one type of output lead to a reuse of the same information on the other type of output that share similar characteristics, and risk reductions obtained through more diversified outputs. Mester (1994) stresses that the most important source of economies of scope is the share of a great deal of inputs usage in the production of several outputs. As fixed inputs are heavily used in the production of both outputs, the firm witnesses a decline in average costs as fixed costs can be spread across multiple outputs. From this perception, one can find a connection between economies of scale and scope, where the spreading of fixed cost over a wider range of output volumes lead to greater cost savings per unit of outputs.

The advantage of joint production is that it may lead to economies of scope on both the input and the consumption side (Molyneux, Altunbas, and Gardener, 1996). From the input side, the firm can use the same types of inputs that are used in the production of the parts (or services) that the joint outputs have in common. If this is achieved then firms can realise internal economies of scope. On the other hand, from the consumption side, customers benefit from the provision of both joint products through more competitive prices. Moreover, customers that simultaneously consume joint products provided by one firm will realise economies of scope as these customers save much of the transaction costs embodied in information and transportation expenses. In this case the firm is creating external economies of scope that benefit their customers.

After this presentation of the theoretical background of economies of scope, we provide various empirical studies examining the existence of scope economy in various banking systems.

5.5.1 Studies on economies of scope in banking [9]

The results of cost economies studies on the US banking industry generally suggest a weak presence of economies of scope. Gilligan and Smirlock (1984) use a sample of 2700 US banks with balance sheet data covering the period 1973-1978 to examine scope economies. They use two outputs defined in terms of either liabilities (demand and time deposits), or assets (securities and loans outstanding). Their results indicate that the costs of producing one output depend on the level of other outputs implying the existence of economies of scope. Mester (1987a) reviews a number of studies that investigate economies of scale and scope in US banks between the period 1983 and 1986. Mester infers that the surveyed studies tend to find no evidence of the existence of economies of scope in US banking. Lawrence and Shay (1986) examine economies of scope in US banking over the period 1978-1982. Using a generalised functional form and three outputs (deposits, investments, and loans), the authors find that cost complementarities are present in the joint production of these outputs, thus suggesting the presence of the economies of scope. On the other hand, Hunter, Timme, and Yang (1990) examine economies of scope in a sample of 311 large US banks at the end 1986. The results indicate non presence of a sub-additive cost function which also indicates the absence of cost complementarities. Their conclusion stresses that there is no evidence found for the presence of economies of scope in large US banks.[10]

[9] For a more comprehensive summary of the literature see appendix 1.

[10] Pulley and Humphrey (1993) showed that cost complementaries are less evident for both deposit and loan products, while the spread of the fixed cost over both these products are shared in the order of 4 to 5 per cent.

As in the case of the US literature, little evidence is available on scope economies in European banking. Altunbas and Molyneux (1996) tested for the presence of economies of scope on the basis that joint production of loans and securities is less costly than their production in separate banks. The study covers a number of European countries, and the results they find are almost mixed. In France, the authors found that medium-sized banks show economies of scope. In Spain, banks with less than $1 billion in assets are found to enjoy substantial economies of scope, while German banks of the same assets size experience scope diseconomies. Moreover, irrespective of the size class, Italian banks realise diseconomies of scope. Hardwick (1989) tests for economies of scope on a sample of 79 UK building societies in the year 1985. For two financial products, mortgage accounts and outstanding share and deposit accounts, Hardwick measures economies of scope by taking the derivatives of each product's marginal cost with respect to changes in the output of other product. The author does not find evidence of economies or diseconomies of scope for societies with assets sizes greater than £1.5 billion. Drake (1992) also examines for economies of scope in the UK building societies sector in the year 1988. The results show that building societies with assets ranging from £500 million to £5 billion exhibited diseconomies of scope. In this study, the author also tested for specific product economies of scope. Drake found that unsecured consumer lending and secured commercial lending exhibited scope economies, while the results suggested that mortgage lending realised diseconomies of scope. In Vennet's (1993) study on cost economies of credit institutions operating in the EU, large banks were found to enjoy economies of scope. Moreover, Lang and Welzel (1996) examined cost economies for German cooperative banks. Using the standard translog cost function, the authors found that economies of scope were prevalent for the largest banks.

Overall, there is rather limited evidence on scope economies in banking. This relates to the difficulty in estimating scope economies, as estimates tend to be sensitive to different output and input specifications (see Molyneux, Altunbas, and Gardner, 1996). In addition, scale economies seem to be more prevalent than scope economies, given evidence from the empirical literature. The next section examines efficiencies unrelated to size (scale) and product mix (scope) and these are known as X-efficiency.

5.6 The economics of X-Efficiency

In the previous sections, we have shown that scale economies are realised by producing outputs at levels where a bank operates on the decreasing average cost. Scope economies are achieved when a bank jointly produces outputs that result in cost savings compared to the cost of separate production of these outputs. In fact, until the late 1980s, the focus of the literature was extensively directed towards the study of scale and scope economies. These studies are mostly concerned with the issue of inefficiency due to non-exploitation in the utilized output mix or scale of production.[11] In this sense, Fukuyama (1993) and Drake and Simper (1999) assert that, in these studies, it is implicitly or sometimes explicitly assumed that banks are efficient, that is, their input mix is at the cost frontier, or their output mix is on the production frontier.

X-(in)efficiency is part of operational efficiency and is a term introduced by Leibenstein (1966). Leibenstein's view on X-inefficiency is based on the description of a firm that produces at less than the optimum level. In general, the banking literature considers X-inefficiency as having two components: allocative inefficiency and technical inefficiency. Allocative inefficiency reflects the failure to choose an optimal input mix in reaction to relative input prices. Technical inefficiency exists when employing an excessive level of inputs for certain output production (Berger and Humphrey, 1997; Kaparakis et al., 1994; Allen and Rai, 1996).

Figure 5.5 presents graphically how both technical and allocative efficiency can be viewed (see Coelli et al., 1998). The figure shows the isoquant curve (QQ`) which presents here the combination of two factors (inputs) of production (X_1 and X_2) that can be used to produce a given quantity of a product Q with a given state of technology. The slope of the isoquant reflects the substitutability among the factor inputs quantities in the production process. The isoquant is convex to the origin because substitution of one production factor to another is not a perfect substitution, and the marginal rate of

[11] See Humphrey (1990) and McAllister and McManus (1993)

substitution declines as we move down from left to right (use less X_1 and more X_2). The line (w_1w_2) is the isocost line that shows the combinations of the two factor inputs (X_1 and X_2), which can be bought by the prevailing prices w_1 and w_2 respectively. The production function $Q = f(X_1, X_2)$ of the firm is characterized by the constant returns to scale. Point D, where the isoquant is tangential to the isocost line, shows the least cost combination of inputs for producing the given level of output Q.

Figure 5.5 Technical and allocative efficiencies

Third Party material excluded from digitised copy. Please refer to original text to see this material.

Source: Adapted from Coelli et al. (1998, p. 135)

The measure of the technical inefficiency could be viewed as the ratio of 0B to 0A, where this portion captures the distance at which the firm can reduce its input usage in order the reach the optimum usage of inputs at the isoquant (Cummins and Weiss, 1998; Farrell, 1957). However, although the reduction of this excess use of input will locate the firm's inputs' mix on the isoquant at B, this does not ensure the production of Q level at the minimum cost, which give rise to the idea of allocative inefficiency measured by the ratio 0C/0B. The firm needs to move along on the isoquant to point D where both allocative and technical efficiency is achieved at the tangency of both isoquant and isocost curves.

The banking efficiency literature considers that the term X-inefficiency incorporates both allocative and technical inefficiency. Failure to achieve both technical and allocative efficiencies lead to X-inefficiency, which is defined as the deviation of bank's cost (or profit) function form the best practice cost (or profit) function. This best practice cost function is the frontier towards which the firm cost function should move in order to become more X-efficient. The next subsection sheds more light on how to examine X-efficiencies in banking.

Estimating X-efficiency in banking

The efficient frontier, or the benchmark, is estimated depending on the objective that a financial firm wishes to pursue. One can estimate cost or profit X-efficiency depending on whether one wishes to estimate X-efficiency on the input side (cost X-efficiency) or input and output side (profit X-efficiency).

If cost minimisation is the banks' objectives, then cost efficiency shows how close the estimated cost function of a financial firm is to the estimated best practice cost function. If profit maximisation is the main objective, profit X-efficiency estimates how close a banks profit function is to the maximum or the best performing bank's profit function in the industry. Berger and Mester (1997) suggest two concepts of profit efficiency: standard profit and alternative profit efficiency. The two profit efficiency concepts measure how close a bank is to achieving the maximum possible profit given particular levels of input and output prices.

Measures of profit X-efficiency are believed to be superior to cost X-efficiency as profit X-efficiency measures allow us to take into account inefficiency from both input and output sides (Berger, Hancock, Humphrey, 1993; Berger and Mester, 1997). In this sense, banks can be cost efficient from the view of their input side, but they might be inefficient with respect to outputs. For instance, a bank may minimise costs (thus making it cost efficient) but if it does not maximise revenues it will not be profit efficient. Alternatively, banks can be cost inefficient yet profit efficient if they have high costs that result in greater revenues (Berger and Mester, 1997).

5.6.1 Studies and approaches to measuring X-efficiency

Studies on efficiency measurement date back to the late 1950s, specifically to the work of Farrell (1957). He measured inefficiency by calculating the deviation of the actual behaviour from the optimum. By following Farrell, nearly all approaches to efficiency measurement concentrate on his idea of which there must be a frontier representing an optimum capacity, and the deviation from the frontier is considered as inefficiency. However, the estimation approach of the optimum benchmark or the frontier and the measurement of the distance of which the estimated observations are placed away from this theoretical estimated frontier is the area where many empirical studies have differed. There are two main statistical approaches used to measure the efficient frontier: non-parametric and parametric approaches.

The following subsections discuss the nature of these two main approaches and the way the efficient (or best-practice) frontier is calculated. The benefits and shortcomings of these approaches are also addressed.

5.6.1.1 Non-parametric approach

Non-parametric (or linear-programming) approaches (see Aigner and Chu, 1968; Afriat, 1972; Richmond, 1974; Berger and Humphrey, 1997) specify no functional form to estimate the best-practice frontier. It assigns the best practice banks on the frontier and other banks are considered less efficient relative to the ones defining the frontier. In fact, most non-parametric approaches do not allow for any random disturbances, so no stochastic term is included in the model.[12] Deviations of the data from the frontier are the inefficiency residuals that are strictly one-sided and negative for the production frontier model and positive for the cost frontier. This is because the data cannot lie above the estimated maximum production function or fall below the minimum cost function. The major disadvantage of the deterministic frontier approach is that because it does not take into account random noise, the inefficiency term may be overestimated since the latter may include random noise.

Non-parametric studies manly use Data Envelopment Analysis (DEA), which is a linear programming technique utilized to construct the frontier and measure efficiency. This technique as constructed by Farrell (1957) has been subject to many extensions (Charnes et al., 1978; Banker et al. 1984). DEA approximates the efficient frontier through the envelope of hyper-planes in the input space. It uses a linear programming algorithm method to measure how far a given observed input vector is from the frontier; the inefficiency of the firm is computed as the ratio of the firm's input costs relative to the least input cost of the best practice firm which lies on the efficient frontier (Evanoff and Israilevich, 1991).

Some researches such as Elyasiani and Mehdian (1990b) envisage DEA, and therefore the deterministic frontier, as having the advantage of no standard specification of what functional form must be used. In essence, however, there are a number of drawbacks concerning this approach. First, since DEA is a non-parametric approach in which the frontier estimates are deterministic, it does not allow for errors or any stochastic variables to enter the model; therefore, any deviation from the estimated frontier is considered an inefficiency (Elysiani and Mehdian, 1990a). The problem here is that the calculated efficiency might contain information of data shocks or measurement errors which may resulte in misestimation of inefficiency. Second, DEA does not estimate the model parameters, and there is no test that makes the researcher sure of how accurate the estimation is (Fakuyama, 1993; Mester, 1996); therefore, because DEA does not produce standard errors, inferences are not available (Greene, 1993). Third, inefficiency

[12] It also assumes that the frontier is fixed for all observations in the sample.

represents only an upper bound of the DEA estimates, a matter that makes comparison between banks unreliable (Schmidt, 1986).

Other non-parametric approaches to efficiency measurement

The Free Disposal Hull (FDH) approach as introduced by Deprins et al. (1984) develops the DEA technique. This approach has been gaining increased acceptance, as it is seen as an alternative non-parametric approach competing with the DEA technique to measure inefficiency (DeBerger, Ferrier, and Kerstnes, 1995). FDH differs from DEA in that it does not take into consideration the convexity assumption, which is a property related to the production possibility set. In referring to (Tulkens, 1993), Berger and Humphrey (1997, p. 177) state that "...the points on lines connecting the DEA vertices are not included in the frontier. Instead, the FDH production possibilities set is composed only of the DEA vertices and the free disposal hull points interior to these vertices. Because the FDH frontier is either congruent with or interior to the DEA frontier, FDH will typically generate larger estimates of average efficiency than DEA...". Similar to DEA, however, the principal shortcoming of the FDH is that it ignores random error. However, FDH considers the variation of efficiency over time and makes no assumption as to the type of distribution of the inefficiency component, thus the measured distance between the estimated observation and the frontier is wholly considered as inefficiency (Berger and Humphrey, 1997).

5.6.1.2 Parametric approaches to measuring inefficiency

The parametric approach assumes an explicit functional form to estimate the frontier of either cost or profit functions (See Berger and Humphrey, 1997). The parametric method is stochastic, in that it allows random disturbance along with an inefficiency residual to be accounted for when estimating the efficient frontier. There are various parametric techniques that have been used to estimate bank efficiency, the most common of which is known as the stochastic frontier approach.

Stochastic frontier approach

The stochastic frontier model was developed by Aigner et al. (1977) and, later, by Jondrow et al. (1982). Realising the disadvantages of the deterministic frontier approach, especially the non-consideration of random noise, this induced a significant development in the efficiency measurement literature; that is, the estimation of a frontier comprising both inefficiency and stochastic (or random noise) terms (Aigner et al., 1977; Meeusen and Van den Broeck, 1977). The reason why one includes a stochastic term is to account for random noise that can either increase or decrease the frontier due to luck or other measurement error factors (Berger and Humphrey, 1991). In the case of the stochastic frontier one assumes that the frontier shifts from one observation to another. Here, the inefficiency term implies that, in the case of cost studies, inefficiency raises costs above the minimum estimated cost function (the cost frontier). Inefficiency also decreases profit below the profit frontier if one is studying profit efficiency.

In the stochastic frontier approach, strong distributional assumptions are necessarily needed to decompose the residual into inefficiency and noise components. The distributional assumption for the stochastic term component is typically characterised by a two-sided normal distribution; while, the inefficiency term is always assumed to be a one-sided distribution representing the shortfall of output from the production frontier or the increase of the cost beyond the cost frontier. Aigner, Lovell and Schmidt (1997) provide two ways of estimating the inefficiency, assuming the distribution of the inefficiency term takes a half-normal distribution in one estimation and an exponential distribution in another. Meeusen and Broeck (1977) consider inefficiency to take only the exponential distribution. Cebenoyan et al. (1993) and Berger and DeYoung (1997) use the truncated normal distribution, while the gamma distribution is considered by Richmond (1974), Stevenson (1980) and Greene (1990).

One difficulty related to the stochastic frontier approach is that there is no consensus as to the type of distribution one should choose to arrive at the inefficiency measure. Greene (1990) suggest that the distributional assumptions do not have much impact on the efficiency estimates. Also, in Aigner et al.'s (1977) study, little difference in inefficiency scores is found when different assumptions (the half-normal and the exponential distributions) are used to derive inefficiency measures. Moreover, in their study on the German banking, Altunbas and Molyneux (1994) compared inefficiency results derived using the half-normal, truncated normal, exponential and gamma distributions; they find that efficiency estimates are insensitive to these different distributions. Overall, the literature suggests that different distributional assumptions tend to yield similar inefficiency scores.

Other econometric approaches

Another alternative econometric approach to deriving X-efficiency is known as the distribution–free approach (DEA). This assumes that the inefficiency term is stable and does not change over time; whereas other coefficients and variables are allowed to vary, leaving the random error component to average out over time (see Berger, 1993; Berger and Humphrey, 1992a). Therefore, unlike stochastic frontier approach, the DEA places no specific type of distribution on the inefficiency term. This approach usually requires a panel data set so that the cancellation of the error term finds enough time to retain a zero value.

In their estimation of a global cost function, Allen and Rai (1996) applied both the stochastic frontier and the DEA and find that the latter approach overestimates the magnitude of inefficiency relative to its correspondent of the stochastic frontier approach. They explain that this overestimation by the DEA is due to the disappearance of the random error term over time, leaving only the inefficiency measure. This might be considered as a shortcoming in the DEA, especially, when the technique transfers the net of the error term, not cancelled out over time (i.e. due to structural change), to the inefficiency term (Allen and Rai, 1996). Therefore, some authors truncate the extreme scores of inefficiency to get rid of values that are associated with non-cancelled random noise (see Berger and Mester, 1997).

In the third econometric approach related to the error term components, Berger and Humphrey (1991 and 1992b) propose the thick frontier approach. This approach is also used in Bauer et al. (1993) and Mahajan et al.'s (1996) studies. The thick frontier approach estimates the cost function of banks in the lowest average cost quartile, which is the thick frontier, and compares it with the cost function of banks in the highest average cost quartile. It then decomposes the deviations into random noise and an inefficiency residual. In order to distinguish between both error terms, the thick frontier approach assumes that the random noise is embodied within the lowest and the highest average cost quartile, which appears as the deviations from the predicted costs of each quartile. In the result, the differences between the lowest and the highest average cost quartiles are measured as the inefficiency component. This approach avoids making any assumptions on how the error components are distributed. A critique to this approach is raised in Kaparakis et al's. (1994) study that stresses the inconsistency between the ordering of the firms in the sample according to the lowest average total cost per Dollar of assets, and according to the nature of multiple-input/multiple-output estimated cost functions. That is these authors' argument is that, since input prices are not the same across firms, this ordering will be subject to bias. Moreover, the approach does not allow for efficiency estimates for each bank, it only estimates the overall efficiency for the sample in question.

In defence of their approach, Berger and Humphrey (1992) have pointed out that the assumptions used in the thick frontier approach seems to be no worse than those for other techniques; for instance the stochastic approach sets arbitrary distribution assumptions on the inefficiency term (for example, a half-normal distribution), and the DEA method assumes zero random error. In addition, even if the deviation from the predicted estimates represent inefficiency rather than the random term as maintained, the thick frontier approach still, they argue, produces a valid comparison of average inefficiency of high and low cost quartile banks.

After identifying the main analytical approaches used in the literature to estimate X-efficiency, the following subsection outlines the main empirical findings.

5.6.1.3 Comparisons between parametric and non-parametric empirical results

The bulk of banking studies that use the econometric approach (stochastic frontier, thick frontier, and distribution-free approaches) -although they use different assumptions regarding the error term components- typically find similar X-inefficiency results. That is, for different studies using various data sets, mean levels of banking sector inefficiency lie within the range of 5 to 30 per cent. In contrast, DEA studies report a wider divergence of banking X-inefficiency, averaging from less than 5 per cent to more than 50 per cent (Berger et al., 1993).

In their review of the financial sector X-efficiency literature, Berger and Humphrey (1997) surveyed some 130 studies. This survey examines various studies covering different financial institution (such as banks, bank branches, saving and loans, and other financial service firms), using different approaches to measuring efficiency (parametric and non-parametric approaches), and also covering different countries and regions. In their survey, Berger and Humphrey find that the mean inefficiency across all studies included in the survey was 27 per cent, with standard deviation of 13 per cent. This means that, on average, financial firms could produce 27 per cent more outputs, given current inputs if they operated as efficiently as the most efficient firms. In a comparison between the parametric and non-parametric approaches, the mean inefficiency measures of these approaches are relatively close to one another, 39 per cent and 19 per cent, for the non-parametric and parametric studies respectively. However, the standard deviation of the non-parametric studies, at 17 per cent, is higher than that of the parametric approach, at 6 per cent. This indicates that the mean efficiency found by parametric studies are more likely to be closer to each other compared to those found in non-parametric studies.

With regard to the rankings of inefficiency estimates, there are only a few studies that calculate the range between the efficiency estimates for both parametric and non-parametric approaches. In Berger and Humphrey's (1997) survey, only two studies report a Spearman rank correlation coefficient (R) between the DEA and SFA estimates: the study by Ferrier and Lovell (1990) found R=0.02, and the study by Eisenbeis et al.

(1996) that found R ranging between 0.44 and 0.59. These findings suggest a weak ordinal association between the two approaches.

Berger et al. (1993) stress that there is no rule or standard guide to researchers in choosing the most reliable approach that fully describes the nature of the banking data. Also, the level of inefficiency measured by using different approaches may make more confusion especially when there is a large dispersion in the reported scores. Thus, in most cases, studies using both econometric and linear programming approaches end up with mixed results. For example, Ferrier and Lovell (1990) use both approaches on the same set of data and concludes that both approaches yield large differences in inefficiency scores. Therefore, comparisons between inefficiency scores using different techniques appear to yield conflicting results due to differences in assumptions made, data, and the nature of the industry studies.

In addition, the literature does not provide guidance as to what level of inefficiency is acceptable or harmful. Again, this might be due to the lack of general agreement on measurement techniques. Overall, most techniques have, at least, the advantage of providing an indication of whether or not inefficiency is present in the industry. Moreover, although it is not clear as to what level of inefficiency is regarded as 'good' or 'bad' from a policy perspective, one can only just assume that higher levels of inefficiency are worse than lower levels.

However, because the interpretation of bank efficiency levels has important implication for both owners, managers, regulatory and policy-decisions makers, Bauer et al. (1998) propose a set of consistency tests. These tests are stated as follows. 1. Efficiency scores obtained using different approaches should yield comparable statistical means, standard deviations, and distributional properties. 2. Different approaches should approximately rank the efficiency of the financial institutions in the same order. 3. Relating to point 2, different approaches should generate similar estimates on best and worst practice financial institutions so that it eases the identification of successful and unsuccessful financial firms. 4. In order to identify the effect of the implementation of regulatory policies, different approaches should generate consistent results of efficiency overtime.

5. Efficiency scores should be consistent with market competitive condition since banks of old establishment and matured ones are expected to be more competitive than newer banks. 6. Efficiency results should match the financial ratios that are used to evaluate the profitability and the performance of these firms.

Bauer et al. (1998) suggested that consistency conditions 1 through 3 should be used to check how the efficiency scores from different approaches could arrive at a degree where they are mutually consistent and could provide useful insights for policy questions. Moreover, consistency conditions 4 through 6 can evaluate the extent to which efficiency scores from different approaches yield credible and reasonable measures. Overall, these consistency conditions tests, if passed, should increase the confidence in efficiency scores using different approaches and advance judgment on the efficiency features of the financial industry under study.

The above provides an overview of the main findings of the parametric and non-parametric studies. As we intend to use the parametric approach to estimate X-efficiency in GCC banking, the following section outlines the main types of functional forms used in the parametric banking studies.

5.6.1.4 Estimating efficient frontiers and the choice of functional forms

Parametric approaches have been widely used in the bank efficiency literature and they all rely on the choice of an appropriate cost or (more recently) profit function. Early studies extensively used the Cobb-Douglas functional form,[13] although since the mid-1970s until the late-1980s more studies have adopted the translog functional form to model bank costs.[14] More recently the Fourier Flexible functional form has been the preferred choice for estimating cost and profit efficiencies in banking.

[13] The other functional forms used are Constant Elasticity of Substitution or CES [Arrow et al. (1961)], and Leontief [Diewert (1971)], Box-Cox transformations of the translog model [Clark (1984)], hybrid translog function [Mester (1992) and Molyneux et al. (1996)], and Fuss normalised quadratic variable profit function [Berger, Hancock, and Humphrey (1993)].

[14] Although there exist some studies that used the translog specification (for example Turati, 2001; Maudos et al. 2002), the majority of researchers nowadays tend to adopt the of Fourier flexible form.

The introduction of the Cobb-Douglas cost function in banking efficiency studies was initiated by Benston (1965a) where he used this approach to estimate economies of scale in US banking. The use of the Cobb-Douglas function serves as a first order approximation of an arbitrary or unknown functional form. This function is assumed to be linearly homogenous and strictly quasi-concave where the elasticity of substitution of inputs must be equal to one.

For most studies published in the late 1960s until the late 1970s, the use of the Cobb-Douglas functional form was the main approach and these studies estimated the relevant cost functions in order to obtain economies of scale estimates (for example, Bell and Murphy, 1968; Schweitzer, 1972; Murphy, 1972; Kalish and Gilbert, 1973; Mullineaux, 1975 and 1978).

In later studies, Benston, Hanweck, and Humphrey (1982b) and Berger and Humphrey (1994) claimed that the Cobb-Douglas functional form suffered from certain shortcomings; namely, the function was not the most appropriate model to estimate a cost function that exhibited a U-shaped average cost curve since the Cobb-Douglas specification only allowed for one aspect of the estimation of increasing, decreasing, or constant average cost for all banks. Moreover, the Cobb-Douglas functional form does not allow for estimates of economies of scope (Berger and Humphrey, 1994). In addition, the use of Cobb-Douglas may yield biased results as it cannot appropriately account for major differences in banks sizes across samples.

Benston, Hanweck, and Humphrey (1982) suggested a more flexible cost function, the translog functional form, to estimate bank costs. The authors claimed that the translog function overcomes the shortcomings of the Cobb-Douglas functional form in the sense that the translog form is more able to account for U-shaped average costs, and to estimate the cost function across firms of different sizes in an industry. Moreover, the translog model allows homogeneity of the degree one by simply imposing restrictions on the translog model parameters (McAllister and McManus, 1993). In practice, a great deal of research in the banking efficiency literature has used the translog functional form to estimate the cost characteristics of banking firms (see for example, Kwan and Eisenbeis,1996; Altunbas et al., 2000; Berger and DeYoung, 2000).

However, some studies have cast doubts on the result of efficiency, scale, and scope economies obtained using the translog model. For example, White (1980) hinted that least square estimates of the second series expansion of the translog function do not accurately match the expansion points on the path of the Taylor series expansion, a reason that generates biased estimates. In support of White's (1982) view, McAllister and McManus (1993) argue that the use of translog form may give poor approximations of the industry's estimated cost function since it may not account for differences in bank size, and the lack of global approximation lead some banks' size cost functions to be forced to lie on other locations on the estimated cost curve. For example, if a sample of different sizes has a large number of small banks that may operate on the decreasing portion of the average cost curve, and a small number of large banks that operate at constant economies to scale, the translog function may misestimate this by showing large banks operating with diseconomies of scale. Moreover, McAllister and McManus (1993) add that the translog approximation may not compute economies of scope accurately because translog approximations may behave poorly away from the average mean product mix as large banks tend to have very different product mixes from the average.

In showing the shortcomings of the translog function, McAllister and McManus (1993) tested four model specifications: translog, kernel, spline, and Fourier functional form.[15] As shown in Figure 5.6, these authors find that (within the global approximation) all specifications behaved well except the translog model where its cost function started to

[15] As defined by McAllister and McManus (1993, p. 395) "The kernel regression technique [Hardle (1990) in a recent survey] builds a global estimate of the cost function by forming weighted averages over localized regions. Although this technique comes closest to the goal of 'letting the data speak for themselves', it has the disadvantage of requiring very large samples to obtain accurate results, especially in applications in which there are more than a few explanatory variables." " The linear spline estimation technique [Porier (1976)] approximates the unknown cost function by a piecewise linear function. The grid of knot points becomes finer as the sample size increases, enabling the spline to approximate any continuous function. Experimentation suggested that a very simple spline-augmented translog function was an adequate approximation to the unknown cost function for purposes of the present comparison, with three knot points for each of the output variables."

suffer from bias caused by large outputs resulting in incorrect estimates of average costs for large sized banks.

Figure 5.6 Translog, Kernel, Spline, and Fourier cost functions' estimates

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Source: Adapted form McAllister and McManus (1993, p. 396)

In addition, non-parametric approaches such as Kernel regression as well as the Fourier Flexible form (which are to be considered semi-parametric approaches) overcome this shortcoming of the translog function because they allow more flexibility and freedom for the shaping of the cost function given different bank sizes.

The principal technique that has been used in the recent banking efficiency studies to overcome the disadvantages of the translog function is the Fourier Flexible form. In contrast to the translog model, which has limitations in estimating the global approximation of the true function when compared to sample subsets sample (see McAllister and McManus, 1993), the Fourier Flexible form can provide more accurate approximations to the true function over the whole range of data (Gallant, 1982). As the Fourier Flexible form adds trigonometric terms (which are mutually orthogonal over the [0,2π] interval) to the translog specification, '[the] linear combination of sine and cosine functions called a Fourier series can represent exactly any well-behaved multivariate function…' (Mitchell and Onvural,1996, p. 140).[16] Thus, the use of trigonometric terms will narrow the edge between the approximated function and the true path of data (Gallant, 1982 and 1984; Mitchell and Onvural, 1996; Berger and Mester, 1997).[17]

Moreover, the Fourier Flexible form is considered suitable to use when data are used to draw a relationship between variables that have an unknown functional form that estimate the relationship. Thus a researcher lacking knowledge of the true form of a cost function may avoid functional misspecification by positing a Fourier series.

However, it should be noted that although it has been argued that the translog has deficiencies regarding the estimation of global approximations, a study by Altunbas and Chakravarty (2001) indicates that, in general, while the Fourier Flexible form is better in terms of goodness of fit, its forecasting ability is worse. This may imply that the use of translog form could be justified by its predictive abilities.

Overall, the development of various functional forms to estimate cost and profit efficiency in banking is still ongoing. The greater use of the Fourier Flexible functional form complemented with translog estimates provides another consistency test for efficiency measure.

[16] This is possible because the sine and cosine functions are mutually orthogonal and functions space-spanning; hence, representing an arbitrary function by a Fourier series is analogous to representing n-vectors.

[17] Berger and Mester (1997) assure that Fourier-flexible has improved the fit of the data in every application they undertook.

As the above section (section 5.61) presents a review on the main approaches to estimate efficiency, it remains to present a review on recent studies that measure efficiency in banking systems.

5.6.2 Empirical studies on X-efficiency in banking sectors [18]

In this section we present recent empirical studies undertaken to examine X-efficiency in banking systems. We start with the empirical evidence on US banking efficiency studies, European studies, and then concluded by discussing recent studies that examine banking sector efficiency in other systems.

5.6.2.1 Studies on X-efficiency in US banking system

A number of selected recent studies in US banking are mentioned here. Other studies are briefly reported in appendix 2. In general, the findings indicate the existence of cost inefficiency in the US banking system at less than 30 per cent; and typically in the 5 to 25 per cent range. Profit inefficiencies are found to be much larger in the 40 to 50 per cent range.

For instance, Kaparakis, Miller, and Noulas (1994) use data covering 5548 US banks with assets over $50 million in the year 1986. By employing the stochastic cost frontier and using the translog function, the authors find overall mean inefficiencies of 10%. Their findings also suggest that banks generally become less efficient with increasing size. Moreover, by moving to more competitive environment, the authors find that banking may become costly and more inefficient. The study by Mester (1993) investigate efficiency in mutual and stock Saving and Loan (S & L) using 1991 data on US S & Ls. By employing the stochastic cost frontier approach, the results suggest that deregulation of interest rate and increased competition may, to some extent, contribute to shifting of large number of costly and inefficient S & L institutions from mutual to stock ownership, a finding suggesting that more competition needed to wipe out the

[18] For a more comprehensive summary on X-efficiency studies see appendix 2.

remaining inefficient firms. Moreover, the author found that capital-assts ratios are positively correlated - but uninsured deposits are negatively correlated with efficiency.

Berger, Hancock, and Humphrey (1993) study US commercial banks with data including three panels of 384 to 599 banks each covering the period 1984-1989. There results, which are obtained using the DFA approach, show that the mean profit X-efficiency ranged from 52 to 66 per cent, suggesting that larger banks are found to be more X-efficient than smaller banks. Moreover, the authors find that technical inefficiencies dominate allocative inefficiency, suggesting that banks are not in particular poor in choosing input and output plans, but they are poor in running and carrying out these plans. Moreover, the findings suggest that most of the profit inefficiencies stemmed from revenue deficiencies, rather than excessive costs.

Elyasiani and Mehdian (1995) aim at identifying how cost efficiency in US small and large banks differs in order to explore the relationship between size and productive efficiency and to examine how changes in regulations affect efficiency during the 1980s. The sample contains 150 US banks studied for the period between 1979 and 1986. Generally, using the DEA approach, their findings show that the mean efficiency estimated ranged between 95 and 97 per cent. Their findings also showed that in pre-deregulation era, small banks were more efficient than large banks, and in post-deregulation both small and large banks were almost equally efficient.

Kwan and Eisenbeis (1995) study 254 US banks holding companies observed on semi-annual for the years 1986 to 1991. Using the approach of SFA, the authors find that the mean inefficiency is declining over the study period. Their results suggest also that small-sized banks reported less efficiency (81 per cent) than their larger counterpart (92 per cent).

Mester (1997) studies 214 banks in the Third Federal Reserve District over the years 1991-1992. Using the SFA and accounting for risk and quality factors in banking outputs, the author finds that although the studied banks are operating at cost-efficient output levels and product mixes, there appear to be a significant level of X-inefficiency at the banks considered.

Berger and Mester (1997) examine 6000 US banks over a six-year period 1990-1995. They employed three efficiency concepts - cost, standard profit, and alternative profit efficiencies. Using DFA to estimate these efficiency types, and using the preferred model including risk and quality variables, the efficiency scores are found to be 86, 54, and 46 percent respectively. Moreover, the authors find that the profit inefficiencies are not positively correlated with accounting profits.

Rogers (1998) used the parametric approach to estimate the efficiency of 10,000 commercial banks over the period 1991-1995. The author estimate stochastic translog cost, revenue, and profit frontiers where each included net non-interest income as a measure of non-traditional output. Under all three frontier specifications, the restricted model which omitted these activities is rejected in favor of the unrestricted model. Overall, mean cost efficiency is found to range between 71 and 76 per cent for the unrestricted model, compared to 65 and 66 per cent for the restricted one. The mean revenue efficiency ranged between 41 and 44 per cent and 50 and 51 per cent for the restricted and unrestricted models respectively. The mean profit efficiency ranged between 60 and 71 per cent and 65 and 68 per cent for the restricted and unrestricted models respectively.

The recent efficiency studies undertaken in the US banking sector suggest the existence of inefficiency on both input and output sides, with mixed results on whether small or large banks are more efficient. The studies from the US have also initiated the importance of considering risk and quality factors in modelling bank efficiency. It is agreed that it is important to consider risk and quality factors in modelling bank efficiency as it helps researchers reach better conclusions on bank's efficiency status given that these elements are viewed as important elements influencing banking performance. The next subsection presents more evidence from European banking X-efficiency studies.

5.6.2.2 Studies on X-efficiency in European banking systems

Important evidence on cost and profit efficiencies has been obtained on individual and cross-country European banking studies. The following gives a number of recent X-efficiency studies on the European banking market.

Berg et al. (1993) use the DEA approach to study banking sector efficiency in Finland, Norway and Sweden in the year 1990. The authors find that the largest banks in Sweden are among the most efficient units in the whole sample, whereas only one large Finnish bank and no large Norwegian banks score efficiency above 90 per cent. In later study, Berg et al. (1995) examine the efficiency of banks in Denmark, Finland, Sweden, and Norway. The authors also find that large banks in Sweden as well as in Denmark are the most efficient units in the pooled sample.

Altunbas et al. (1994a) study the efficiency of German banks with a data set containing 196 banks covering the year 1988. Using the stochastic frontier approach, the authors find that the mean cost inefficiency in German banking is around 24%.

Pastor et al. (1995) estimate efficiency of banking sectors in eight European countries using non-parametric approach. Their findings show that the most efficient banks in the sample are those of France (95%), Spain (82%), and Belgium (80%); while the least efficient banks come from Germany (65%), Austria (60%), and UK (53%).

European Commission (1997) study a sample obtained from the IBCA Bankscope database on 10 European Union countries covering the years 1987 (with 295 banks) to 1994 (with 1451 banks). Using the SFA approach, the study find that the average efficiency ranges between 71 and 77 per cent over the five years and these decreased over the last four years of the study period.

Altunbas et al. (1999) use the Fourier Flexible functional form to estimate the characteristics of banking costs in European countries over the period 1988-1995. The authors find the mean cost X-inefficiency at the level of 25 per cent. The authors also examine whether large banks are more X-efficient than small banks. They do not find any evidence supporting this claim for their sample of European banks.

Dietsch and Weill (1998) study the efficiency of 11 European Union countries using data on 661 commercial banks, mutual, and saving banks covering the period of 1992-1996. Generally, their estimates on the efficiency and productivity over the study period suggest an increase in cost and profit inefficiency levels.

Casu and Girardone (1998) studied a sample consisting of 32 Italian banking groups and 78 bank parent companies and subsidiaries in 1995. Using both SFA and DEA approaches, the authors find that the mean efficiency using SFA is 92 per cent for banking groups, and 94 per cent for bank's parents and companies; for the DEA estimates, the mean efficiency is found to be 88 efficiency for banking group and 90 per cent for bank parent and companies.

Turati (2001) estimates the cost efficiency in European banking markets from 1992 to 1999. The author specifies three different translog cost functions. All the three models consider 3 inputs (labour, physical capital, and deposits) and 2 outputs (loans and other earning assets). Correlation between scores obtained with different specifications of the cost function is very high. The author also finds no major differences in mean efficiency among European countries. Mean efficiency across countries shows a decline from 1992 to 1998 and an increase from 1998 to 1999. Turati suggests that the low correlations between cost efficiency scores and profitability may indicate the presence of market power in the banking industry.

Maudos et al. (2002) examine both cost and profit efficiencies using a sample of banks for ten countries of the European Union, using IBCA information for the period 1993–1996. Using panel data frontier approaches, the authors find high levels of efficiency in costs and lower levels in profits. This result suggests the importance of inefficiencies on the revenue side of banking activity. Also, their results show low but positive correlation between the rankings of cost and profit efficiency. In their analysis, the authors use four groups of variables: size, specialisation, other characteristics specific to each bank, and characteristics of the markets in which they operate. The results show that medium-sized banks enjoy the highest levels of efficiency in both costs and profits; the type of banking specialisation is not significant in explaining differences of efficiency between banks; and the banks with a higher loans/assets ratio are more efficient. Overall, the authors conclude that there is a notably wide range of variation in efficiency levels in the banking systems of the European Union, with variation in terms of profit efficiency being greater than in terms of cost efficiency.

Overall, recent studies on European banking systems tend to find that cost inefficiency levels are around 25 per cent or lower, and in the majority of cases this tends to be decreasing over time, probably due to greater competitiveness within the European integrated market. The more limited evidence on profit inefficiency suggests that this is typically higher, around 40-50 per cent. However, the literature provides little consensus to the size of banks that appears to be the most efficient. The next subsection provides additional evidence on efficiency studies undertaken in other banking markets.

5.6.2.3 Studies on X-efficiency in other banking systems

The banking literature also includes a growing number of efficiency studies on banking systems of non-US and Europe markets. Various recent studies are as follows.

Altunbas et al. (1995b) examine efficiency in the Turkish banking industry over the period 1991-1993. Using the stochastic frontier approach, the authors find that cost inefficiency is relatively high: 46, 32, and 49 per cent for the years 1991, 1992, and 1993. The authors' results suggest also that there is no major difference in the level of efficiency found between public and private banks.

Allen and Rai (1996) use data on 194 banks in 24 countries for the period 1988-1992. The authors' results, which are based on using both SFA and DFA, suggest that large banks operating in countries that prohibit the functional integration of commercial and investment banking have the largest level of input inefficiency with a mean level of 27.5 per cent. Moreover, in 15 countries, input inefficiencies are larger than output inefficiencies.

Bhattacharyya, Lovell, and Sahay (1997) use data on 70 Indian commercial banks over the period 1986-1991. Using the DEA approach, the authors find efficiency to be at the level of 80 per cent for the sample. Publicly owned banks report higher efficiency levels (87%) than privately owned banks (75%) and foreign owned counterparts.

Taylor, Thompson, Thrall, and Dharmapala (1997) study 13 Mexican commercial banks over the period 1989-1991. Using the DEA approach, the study finds the mean efficiency to range between 69 and 75 per cent, with a decreasing trend over the three years under study.

Kraft and Tirtiroglu (1998) obtain data from the audited final accounts on 43 Croatian commercial banks in 1994 and 1995. The authors estimate X-efficiency and scale-efficiencies for both old and new state and private banks. The degree of cost X-efficiency in their estimates range from 54.7 to 87.9 per cent. Amongst the 43 banks, 27 banks have efficiency levels above 80 per cent. New banks are shown to be more X-inefficient and more scale-inefficient than either old privatised banks or old state banks. However, new private banks are highly profitable.

Worthingron (1998) studies 22 Australian building societies in the period 1992-1995. Using the stochastic cost frontier function, the results indicate that building societies' were 20 per cent cost inefficient. Moreover, the results also show the branch and agency networks, asset size, and non-core commercial activities contribute in determining inefficiency; wile capital adequacy restrictions do not have significant influence on estimated efficiencies. Moreover, the study indicates that cost efficiency has been improving during the period studied.

Srivastava (1999) estimates the efficiency of 85 Indian commercial and public banks over the period 1994-1995. The findings suggest that the mean cost efficiencies of private and public banks are 98.18 and 98.11 per cent respectively. Moreover, the mean efficiency of recent entrants (mostly foreign banks as well as some private banks) reports higher level than current banks. In terms of bank size, the highest cost efficiency is, generally, reported for middle-sized banks, followed by small and large sized banks.

Altunbas, Liu, Molyneux, and Seth (2000) use data on Japanese banks (nearly 130 banks in 1993 and 1994, and 121 in 1995). By implementing the SFA and the Fourier Flexible functional form with risk and quality factors taking into the consideration, the study finds that inefficiency estimates derived from the two models are similar and range between 5 and 7 per cent. They also find that the level of financial capital has the biggest influence on the scale efficiency estimates. X-inefficiency estimates, in contrast, appear less sensitive to risk and quality factors.

Intarachote (2000) use a sample on 15 Thai banks, 14 foreign banks, and other finance and specialized institutions. The author's results, which uses the DEA approach, report that inefficiency is found to range from 26 to 48 per cent for national banks, 33 to 50 per cent for the foreign banks, and 6 to 14 per cent for the finance and specialized institutions.

Isik and Hassan (2002) estimate cost and profit efficiencies for Turkish banks over the 1988–1996 period. Over these years, they find that the overall cost and profit efficiencies for the Turkish banks are 72 and 83 per cent respectively. The results also indicate that the production efficiencies of the industry consistently have declined over time. Moreover, their analysis suggests that the relationship between bank size and efficiency is strongly negative. In general, they also find that foreign banks operating in Turkey seem to be significantly more efficient than their domestic peers. In addition, private banks are found to be more efficient than public banks in terms of all types of efficiency.

Al-Jarrah (2002) studies 82 banks operating in Jordan, Egypt, Saudi Arabia, and Bahrain over the period 1992-2000. He uses stochastic frontier and Fourier Flexible forms to estimate cost and profit efficiency (standard and alternative profit efficiency) in these Arab banking markets. Cost efficiencies are found to be 95 per cent, standard profit 66 percent, and alternative profit efficiencies 58 per cent. The author finds Islamic banks operating in these countries to be the most cost and profit efficient banks. Geographically, Bahrain banks are the most efficient, while Jordanian banks are found to be the least efficient.

Overall, the results on efficiency in the above mentioned studies tend to be similar to that of the US and European literature. The increasing number of studies on banking sector efficiency in systems other than the US and Europe indicates the growing interest in examining banking sector efficiency especially because of greater banking sector deregulation.

5.7 Conclusions

This chapter reviews the literature on economies of scale, economies of scope, and X-efficiency in banking. The chapter points out the importance of efficiency studies and illustrates how bank inputs and outputs are defined. In addition, the chapter provides a theoretical overview on economies of scale and scope and X-efficiency, how these are measured, and the empirical studies undertaken to examine these efficiency aspects.

Overall, the literature finds that cost inefficiency tends to vary around 5 to 25 per cent range, with the more recent studies finding lower inefficiencies, typically around 5 to 15 per cent. There have been fewer studies that examine profit inefficiency in banking and these tend to find much higher levels of inefficiency. In essence, it seem that there is much greater variation in banks ability to maximise profits compared with measuring costs. The scale economies literature generally finds that these economies are small, around 5 per cent and these can occur at various levels of bank outputs. The limited evidence on scope economies suggests that large banks may be able to exploit scope economies more than their smaller competitors, although one has to be cautious about scope estimates given the problem associated with their estimation.

On the basis of this chapter, the following chapter presents the parametric frontier methodology to be used in this thesis to examine the efficiency of GCC banking markets.

CHAPTER 6

METHODOLOGY AND DATA

6.1 Introduction

This chapter describes the methodology used to carry out the efficiency analysis, and the following chapter presents the results based on the implementation of the methodology outlined in the present chapter.

In this chapter the analysis will mainly explain the econometric approach, which will principally be used to estimate banking X-inefficiency using various inefficiency concepts: cost, standard profit, and alternative profit inefficiency. Moreover, the efficiency analysis will be extended to cover scale economies and scale efficiency. In addition, we also outline the approach to be taken to analyse the determinants of these efficiencies.

Section 6.2 discusses efficiency concepts and their functions. In section 6.3, we describe the methodology of inefficiency calculation. Section 6.4 discusses the functional forms used in the empirical estimation. Section 6.5 analyses the variables and the data used in the efficiency estimation. In section 6.6, we illustrate the method employed to measure both scale economies and scale inefficiency, which are features mainly related to the cost function. To explore the possible variables determining inefficiency in the GCC banking industry, section 6.7 introduces the logistic regression model, an approach used to evaluate inefficiency determinants. The last section contains the conclusions.

6.2 The X-efficiency concepts

In a first step to measure efficiency in this thesis, it is vital to identify the sort of efficiency upon which a banking industry is assessed. Here, our focus is to measure X-inefficiency, where X-inefficiency refers to the deviation from the frontier that gives the maximum attainable outcome, given the employed resources. In following Berger and Mester (1997), we estimate X-inefficiency in the GCC banking industry on the basis of three inefficiency concepts: cost inefficiency, standard profit inefficiency, and alternative profit inefficiency. Cost efficiency is the widely used measure of bank efficiency from the input side (for example, Altunbas et al., 2000; Lang and Welzel, 1996; Kwan and Eisenbeis, 1995; Berger and Mester, 1997) and profit efficiency measures focus on both input and output sides (incorporating both costs and revenues).

It is important to consider both cost and profit inefficiencies in evaluating efficiency in GCC banking markets. Measurement of GCC cost efficiency is important because it tells us how well GCC banks have been doing on the input side of the financial production process, in which banks employ the available resources to produce a given level of outputs. Measurement of profit efficiency is also important because it is believed that firms may not only err on the input side by choosing non-optimal input mix, but also err on the output side by producing output mixes that make them deviate from the optimal obtainable profit in the industry. Moreover, profit efficiency is '... based on [the] more accepted economic goal of profit maximization, which requires that the same amount of managerial attention be paid to raising a marginal dollar of revenue as to reducing a marginal dollar of cost' (Berger and Mester, 1997, p. 900). Therefore, it is important to examine both cost and profit inefficiencies as they provide a collective analysis of X-efficiency that helps explore more factors that may enhance or diminish banking efficiency from both the input and output sides of the production process.[1]

[1] For example, ceilings on deposit and loan prices could affect both cost and profit functions of the banking industry.

Cost inefficiency

Under the same market conditions and for the same output bundle produced, the cost inefficiency concept views inefficiency as the distance at which the estimated cost function of a financial firm is located away from the least cost function that belongs to the best practice firm in an underlying industry. Thus if the measured cost inefficiency for a banking industry is 15 per cent, this means that banks should use their inputs as efficiently as possible in order to gain a reduction of 15 per cent in their costs in order to make their cost functions reach the minimum cost function of the best practice bank.

Cost inefficiency is derived from the cost function.[2] Basically, the cost function describes a relationship between a cost variable and a set of explanatory variables plus the random and inefficiency factors. The cost function can be written in a natural logarithm form as

$$\ln TC = f(Q,P,Z) + \ln u_c + \ln v_c \tag{6.1}$$

where ln TC is the total cost variable, f stands for some functional form, Q is the vector of outputs, P is the vector of prices of input variables, Z is the set of other likely important exogenous variables, ln u_c is the inefficiency factor that reflects X-inefficiency and raises cost above the industry's optimal cost, and ln v_c is the random error incorporated to capture luck and measurement error, which may temporarily increase or decrease a bank's costs.

Standard profit inefficiency

Standard profit inefficiency focuses on how a bank's profits are compared to the profits of the best practice firm operating in a market where banks use the same inputs, produce the same output bundles, and face the same (market) conditions. In fact, standard profit

[2] The formal calculation of the inefficiency is illustrated in the next section.

inefficiency shows the percentage by which a bank needs to increase profits so that it moves to the profits of the best practice bank. Thus, if a standard profit efficiency average score is 60 per cent, this implies that bank *i* is losing 40 per cent of its profits, probably because of its excessive use of inputs and other deficiencies in generating revenues.[3]

Calculation of standard profit inefficiency is derived from some specified profit function that can be written in a basic form with logs as

$$\ln(\pi+\theta) = f(G,P,Z) + \ln u_{\pi} + \ln v_{\pi} \tag{6.2}$$

where π is the profit variable, θ is a constant added to the firm's profits so that its natural log is positive, f stands for some functional form, G is the vector of prices of output variables, P is the vector of prices of input variables, Z is the set of other likely important exogenous variables, $\ln u_{\pi}$ is the inefficiency factor reflecting X-inefficiency that decreases bank profits at a level under that of the industry's optimal or best practice firm's profits, and $\ln v_{\pi}$ is the random error incorporated to capture luck and measurement error, that may temporarily increase or decrease a bank's profits.

Note that the standard profit function regresses profits on the same set of variables that appear in the cost function, except that it takes output prices as given rather than output levels. This also makes it necessary to calculate the standard profit inefficiency on the basis of how banks choose output levels for the given output prices, a matter that allows for standard profits to capture inefficiency stemming from the non-optimal choice of outputs when responding to these prices.

[3] In addition, a score of 100% denotes the arrival of the profit efficiency at the profit frontier, and a negative score indicates that a bank is throwing away more than 100% of its potential profits (Berger and Mester, 1997).

Alternative profit inefficiency

Alternative profit inefficiency (as developed by Berger and Mester, 1997) reflects how far away a firm's profit function is from the maximum profit function earned by the best practice firm, given the same inputs used and outputs produced within the same prevailing market conditions. Generally, alternative profit efficiency is identical to standard profit efficiency, except that the concept of alternative profit efficiency is introduced to account for the effects of output prices on profit efficiency. That is, because output quantities are held constant in the alternative profit function, the level of inefficiency in the alternative profit model differs in response to the prices of output, which are set free to vary.

The calculation of alternative profit inefficiency is based on the profit function written in the log form as

$$\ln(\pi + \theta) = f(Q, P, Z) + \ln u_{\pi} + \ln v_{\pi} \tag{6.3}$$

where the arguments in Eq. 6.3 are the same as for the standard profit function (of Eq. 6.2), except that the output quantities, Q, replaces prices of outputs, G.

The usefulness of the alternative profit inefficiency concept stems from several factors. Alternative profit inefficiency can be an appropriate approach to account for differences in the unmeasured output quality across banks. For example, banks may incur some costs to improve financial products and services' quality, which in turn may be reflected in higher prices or higher revenues. Because alternative profit inefficiency is already embodied with cost inefficiency, and because alternative profit captures the effects of output prices on inefficiency, expenditure on quality improvements may be offset by higher revenues from higher output prices set, yielding more sensible measurement of inefficiency. In other words, just looking at cost inefficiency means that such factors as increases in expenditure aimed at improving service quality may result in additives to cost inefficiency. However, if this increase in spending results in higher revenues then profit efficiency will be improved.

Moreover, alternative profit inefficiency alleviates the problem of scale bias and avoids the problem of output price inaccuracy, which are problems related to the standard profit method. The problem of scale bias usually emerges from differences in bank sizes and outputs levels because the standard profit method does not control output levels. With alternative profit inefficiency measures, this problem is less severe because comparisons are made between a bank's ability to generate profits for a given level of outputs.

With regard to output price information, proxy measures are usually used for the output prices. Since it is often difficult to obtain prices for the outputs under study, the standard profit inefficiency measures may have an inherent price inaccuracy problem that affects the reliability of the inefficiency estimates. For the same reason, taking output levels instead of output prices allows the alternative profit efficiency measures to avoid this problem of price inaccuracy.

The alternative profit function could be a more appropriate measure of inefficiency when banks have market power that enables them to set higher prices for given output levels. On the other hand, in a more competitive market, the standard profit function seems also plausible since banks tend to be price takers, regardless of the output level they produce. In both cases, it is advisable to estimate both the standard and alternative profit functions together as they provide insights into the level of profit inefficiency given the prevailing condition of market competitiveness.

In fact, the closeness of the two profit inefficiency measures could indicate the extent of market power practised by banks. For example, if both standard and alternative profits are applied in an industry with high market power, alternative profit will capture this market condition and set the level of inefficiency much higher than that measured by standard profit. By contrast, if the industry witnesses less market power, alternative profits will report inefficiency to be about the same as that derived from standard profits estimation.

Moreover, the closeness of the profit efficiency measures derived from both profit concepts may be an indication that the shortcomings of scale bias and output price inaccuracy related to standard profit efficiency do not considerably affect the measurement of profit inefficiencies.

It should, however, be noted that profit inefficiency is expected to be greater than cost inefficiency since profit inefficiency accounts for inefficiencies on both the input and output sides of financial production. Moreover, alternative profit inefficiency is expected to be greater than standard profit inefficiency because the former captures a wider source of inefficiencies such as those related to output qualities and market power.

Having explained the efficiency concepts to be used in the empirical part of this thesis, the following outlines the methodology used to estimate these efficiency concepts.

6.3 Methodological framework used to measure X-efficiency concepts

In this section we discuss the methodology used to estimate inefficiency in the GCC banking industry. As noted in Chapter 5, there are two main approaches to estimate inefficiency: parametric and non-parametric approaches (see also the review by Berger and Humphrey, 1997). Each of these approaches also includes various modelling approaches: for example, the stochastic frontier, distribution-free, and thick frontier approaches are parametric techniques used to derive efficiencies; data envelopment analysis and the disposal hull approach are the main non-parametric approaches.

Our choice in estimating GCC banking sector efficiency is the parametric approach. In our opinion, use of the parametric approach adds more statistical sense to the efficiency estimation because the stochastic nature (or randomness), representing deviation from the true population path, is always present when a random sample is tested to obtain a general inference about a population.

Regarding the choice of techniques among the parametric approaches, we use the stochastic frontier technique as it has the advantage of considering the distribution on both error term composites. As mentioned in Chapter 5, non-consideration of distributional assumptions may lead to an inexact separation of the inefficiency and the random error terms, which may in turn produce an overestimation of inefficiency, especially when the random error term is not cancelled out over time. This problem is present also, to some extent, in the distribution-free technique (Allen and Rai, 1996). Moreover, the thick frontier technique may encounter bias when ordering banks to construct the quartiles according to input prices. Because these prices are not the same across banks, inefficiency measures might be overestimated as well (Kaparakis et al., 1994).

Therefore, to estimate X-inefficiency in the GCC banking industry, we use the stochastic frontier technique, the methodology of which we discuss below in more detail.

A stochastic frontier, as typically explained for the cost function (e.g. stochastic cost frontier) can be constructed to estimate a theoretical least cost function for the industry, which will be attributed as the efficient cost function that belongs to the best practice firm.[4] Accordingly, the estimated best practice firm is said to employ the minimum amount of inputs to produce the given level of outputs.

In a formal way, the single equation stochastic cost function can be given in a logarithmic form for N firms as

[4] As this section's main purpose is to describe the methodology used to calculate the efficiency measure for some frontier function, the functional form for the frontier function specification will be discussed in the following section.

$$\ln TC_i = f(Q_i, P_i) + \ln \varepsilon_i, \qquad i = 1, \ldots, N, \tag{6.4}$$

where $\ln TC_i$ is the observed total cost of bank i, Q_i is the vector of its output levels, and P_i is the vector of input prices the bank i pays. The cost function $\ln TC_i = f(Q_i, P_i)$ gives an indirect representation of the feasible technology; it relates the firm's cost to output levels and input prices, and shows the minimum cost of producing the output vector Q, given the price vector P (Varian, 1992). So, the minimum predicted cost for the industry is explained by $f(Q_i, P_i)$, which is the cost frontier portion in Eq. (6.4) and is considered to be the industry's benchmark of the most efficient firm. The deviation of banks' costs from the cost frontier is explained by the error term ε_i, which consists in a logarithmic form of

$$\varepsilon_i = v_i + u_i, \tag{6.5}$$

where v_i is the statistical noise that represents random fluctuations due to measurement error and luck factors and u_i is the inefficiency term which is supposed to result from mistakes in the choices of input mix that are specific to the firm's practice.

It should be noted that the inefficiency factor, u, is the X-inefficiency measure which represents both technical inefficiency, which occurs when employing excessive inputs beyond the level needed to produce the given output level, Q_i; and allocative inefficiency, which occurs when failing to react optimally to relative prices of inputs, P_i.[5]

[5] The studies of Berger and Humphrey (1997), Altunbas et al., (2000), as well as others have considered the inefficiency term as reflecting both technical and allocative inefficiencies without disentangling them from each other.

On the basis of the cost function specified, we would like to estimate u_i for each GCC bank in the sample to help us compare performance across time and banks.[6]

In order to obtain the measurement of inefficiency estimates, u_i, it is essential to determine how both error term components, v_i and u_i, are assumed to be distributed. The benefit of setting certain distributional assumptions on both error term components lies in the fact that it enables the methodology used to identify these components and disentangle them from each other.

Following Aigner, Lovell, and Schmidt (1977), we assume the distribution of the error term v_i has an identical two-sided normal distribution representing statistical noise which is believed to be independently distributed with zero mean and σ_v^2 variance, that is, $v_i \sim IIN(0, \sigma_v^2)$. The rationale behind this type of distribution is to allow for a pure randomness of the v component upon which this component can either take positive or negative values according to the nature of luck and factors out of management control that affect bank performance.

On the other hand, for the inefficiency part, u_i, we adopt two possible distributions in which the inefficiency term u_i may take: the half-normal distribution and the exponential distribution.

As for the half-normal distribution, u_i is assumed to be a non-negative or one-sided error term representing inefficiency and assumed to be distributed independently of the v_i term.

Formally,[7]

[6] The functional form from which inefficiency, u, will be derived is the Fourier flexible form explained in the next section.
[7] see LIMPDEP version 7.0 user's manual, 1998, p. 753.

$$f(u)=\left(\frac{2}{\pi}\right)^2 \exp\left[-\tfrac{1}{2}(u/\sigma_u)^2\right], \tag{6.6}$$

$$E[u]=\left(\frac{\sigma_u\phi(0)}{\Phi(0)}\right)=\left(\frac{2}{\pi}\right)^2\sigma_u, \tag{6.7}$$

$$Var[u]=\left[1-\frac{2}{\pi}\right]\sigma_u^2, \tag{6.8}$$

where $f(.)$ is the distribution function, $E[.]$ is the mean, $Var[.]$ is the variance, $\sigma=(\sigma_u^2+\sigma_v^2)^{1/2}$, and ϕ and Φ are the standard normal distribution and the standard normal density functions respectively.

The rationale behind using the half normal distribution lies in the perception that the deviation from the frontier should take one side off the cost frontier, and that the cost frontier would have no mean if there should exist observations that fall anywhere under the cost frontier.

The exponential distribution of u_i has the same feature as the half-normal distribution; that is, it is positive one-sided and is independently distributed. However, we use the exponential distribution because this type of distribution is restricted to random variables that can be seen to take only positive values, such as u in our case (Newbold, Carlson, and Thorne, 2003, p. 204). The exponential distribution function, its mean, and variance are shown respectively as

$$f(u)=\theta\exp(-\theta u), \tag{6.9}$$

$$E[u]=\frac{1}{\theta}, \tag{6.10}$$

and

$$Var[u]=\frac{1}{\theta^2}, \tag{6.11}$$

where $f(.)$ is the distributions function, $E[.]$ is the mean, $Var[.]$ is the variance, and θ is the mean rate at which events occur.

As the choice of distribution is arbitrary, it is beneficial to measure inefficiency according to more than one type of distribution since the closeness of alternative efficiency estimates adds more robustness and confidence to the estimated results of inefficiency levels.

It should be noted that the approach of Aigner, Lovell, and Schmidt (1977) does not, however, estimate the u term directly. Accordingly, Jondrow et al. (1982) developed Aigner et al.'s model by providing an explicit formula, which shows that the ratio of variability, σ, for both v and u can be used to calculate the firm's relative inefficiency. This ratio is utilized for the error term portion of the estimated cost function in a way that calculates the inefficiency term given the estimate of the whole error term for each firm in each observation. That is, the level of inefficiency for each bank is calculated by the mean of the conditional distribution of u_i given ε_i. The mean of this conditional distribution for the half-normal model can be shown as

$$E(u \mid \varepsilon) = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{1-\Phi(\varepsilon_i\lambda/\sigma)} + \left(\frac{\varepsilon_i\lambda}{\sigma}\right)\right], \tag{6.12}$$

while for the exponential model

$$E(u \mid \varepsilon) = (\varepsilon - \theta\sigma_v^2) + \sigma_v\left[\frac{\phi[(\varepsilon-\theta\sigma_v^2)/\sigma_v]}{\Phi[(\varepsilon-\theta\sigma_v^2)/\sigma_v]}\right] \tag{6.13}$$

Greene (1993) claims that the mean of the conditional distribution $E(u \mid \varepsilon)$ is unbiased. Nevertheless, this mean is an inconsistent estimator of u_i because, regardless of the number of observations, the variance of the estimator remains non-zero.

After defining the distributional assumptions and the way inefficiency is calculated, we need to estimate the cost function (6.4) in order to obtain the parameters that yield the frontier as well as the estimates of inefficiency explained above.

To estimate the cost function model (6.4), we use the maximum likelihood estimation technique. In fact, this technique is widely implemented in efficiency parametric studies and is preferred over the ordinary least square method. Greene (1993) argues that the maximum likelihood technique is very useful in treating the distributional models of the random noise and the inefficiency components. The log-likelihood function can be written as

$$\ln L = \frac{N}{2}\ln\frac{2}{\pi} - N\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\varepsilon_i^2 + \sum_{i=1}^{N}\ln\left[\Phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)\right], \tag{6.14}$$

where N is the number of banks, $\varepsilon_i = u_i + v_i$, $\sigma = \left(\sigma_u^2 + \sigma_v^2\right)^{1/2}$, $\lambda = \sigma_u / \sigma_v$, and ϕ and Φ are the standard normal distribution and the standard normal density functions respectively. The maximum likelihood estimation operates in a way that finds the minimum of the log likelihood function in order to obtain the estimates of the cost function (6.4).

However, the cost function given in Eq. (6.4) is not our functional form from which to derive efficiency; it is only used here to simplify our explanation of the methodology used to derive efficiency estimates. The next section will discuss how to specify our functional form used to estimate the cost and profit frontiers from which cost and profit inefficiency measures are derived.

6.4 Functional form specification

As just stated, this section is devoted to showing how our stochastic cost and profit functional forms are constructed. Although most studies use the translog functional form to estimate inefficiency, this form is not applied here because of certain limitations. Instead, we use the Fourier Flexible model to specify the cost and profit functions and to obtain inefficiency measures. To arrive at this functional form some steps will also be explained.

A large number of banking studies have used the translog function expressed in a stochastic framework to estimate the cost frontier function (see, for example, Kwan and Eisenbeis, 1996; Altunbas et al., 2000). The translog model is a flexible functional form and is expanded by a second-order Taylor series (see Greene, 2000, p. 217). The flexibility of the translog model is demonstrated in its usefulness for approximating the second-order effect of an unknown functional form (Berndt and Christensen, 1973). This flexibility serves as an advantage for banking efficiency studies because it is difficult to identify exactly the functional form that fits the banking cost and production technology (Kaparakis et al., 1994). Moreover, the translog model allows homogeneity of degree one by simply imposing restrictions on the translog model parameter (McAllister and McManus, 1993).

However, since the translog form is said to be less global because of the bias that makes some observations follow the pattern of other dominant observations, the more recent semi-parametric functional form known as the Fourier Flexible form has been suggested as a preferred approach as it corrects for the translog model's ill fit on the true path of data (Gallant, 1981, 1982; Mitchell and Onvural, 1993). In essence, the Fourier Flexible functional form adds more global approximation and flexibility to the translog form by adding the trigonometric terms to the translog specification. This means that the frontier to be estimated will provide a greater flexibility 'by allowing for many inflection points and by including essentially orthogonal trigonometric terms that help the frontier fit the data wherever it is most needed' (Berger and Humphrey, 1997, p. 179).

On account of these advantages, the Fourier Flexible specification has recently become the more acceptable and increasingly applied parametric functional form in measuring banking inefficiency. Before we set the specification of the Fourier functional form, it should be noted that because the Fourier Flexible form is a translog form extended with trigonometric terms, it is appropriate to note certain features related to the translog form that also apply to the Fourier form as well.[8]

One thing to note regarding the translog function is that as the number of the inputs (also variables) increases, multicollinearity will likely be severe (Greene, 1980). Berndt and Christensen (1973) show how the use of factor demand equations may overcome this problem.[9] Moreover, some studies using the translog function drop the most likely interactive terms causing multicollinearity (see for examples Lang and Welzel, 1996). Doing this might not totally remove multicollinearity problems and its continuing presence may induce an increase in standard errors, which may yield a number of non-significant coefficients.

Second, we should note that in a number of studies factor (input) share equations are used along with translog models (see e.g. Noulas et al., 1990). However, in our estimation, we exclude factor share equations from our model as they embody Shephard's Lemma or Hotelling's Lemma restrictions, which make unfavourable assumptions regarding the allocative efficiency (see Berger and Mester, 1997). Moreover, since inefficiency decomposition (into allocative and technical inefficiency) requires restrictive distributional assumptions, we prefer to keep inefficiency estimation non-decomposed and assume that the whole inefficiency residual component, as noted before, is the X-inefficiency measure (see Kaparakis et al., 1994).

As our Fourier Flexible functional specification consists of the standard translog specification plus the trigonometric terms, as well as the terms of X-inefficiency and the random error, in constructing our Fourier functional model, we first show the core functions of our model along with the residuals, which include both inefficiency and the random error terms. Then we write the function in a translog form, which includes its interactive terms. We then add the trigonometric terms in order to reach the stochastic Fourier Flexible form.

---

[8] However, Altunbas and Chakravarty (2001) note that although the Fourier Flexible form has a better fit than the translog, the former, they find, provides weaker predictive power.
[9] Econometricians generally suggest that one way of reducing the multicollinearity problem is to increase the number of observations.

To start building our Flexible functional form we recall the cost and profit functions explained in section (6.2). These functions are rewritten as

$\ln TC = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln Q_i + \sum_{j=1}^{n} \beta_j \ln P_j + \varepsilon_i$ is the cost function,

$\ln(\pi + \theta) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln G_i + \sum_{j=1}^{n} \beta_j \ln P + \varepsilon_i$ is the standard profit function, and

$\ln(\pi + \theta) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln Q_i + \sum_{j=1}^{n} \beta_j \ln P_j + \varepsilon_i$ is the alternative profit function,

where TC is the cost variable, π is the profit variable, θ is a constant added to the firm's profits so that its natural log is positive, Q is the vector of outputs, P is the vector of prices of input variables, G is the vector of prices of output variables, and ε_i is the stochastic error term where $\varepsilon_i = u_i + v_i$.

The basic functions given above are developed in a multi-product translog specification. To save repetition, we typically continue showing the construction of our model using the cost function. The translog cost function is written as

$$\begin{aligned}\ln TC = {} & \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln Q_i + \sum_{j=1}^{n} \beta_j \ln P_j \\ & + \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln P_i \ln P_j\right] \\ & + \sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \ln Q_i \ln P_j + \varepsilon_i\end{aligned} \tag{6.14}$$

In order to reach our Fourier Flexible form, we transform output variables into the Fourier first and second order trigonometric terms, and, because input prices are attributed with little variations, they are left to be separately described in the translog portion.

As a result of this transformation, which adds the trigonometric terms to the translog form, the model becomes the Fourier Flexible form shown as

$$
\begin{aligned}
\ln TC = {} & \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln Q_i + \sum_{j=1}^{n} \beta_j \ln P_j \\
& + \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln P_i \ln P_j \right] \\
& + \sum_{i=1}^{n}\sum_{j=1}^{n} \rho_{ij} \ln Q_i \ln P_j + \sum_{i=1}^{n} [a_i \cos(z_i) + b_i \sin(z_i)] \\
& + \sum_{i=1}^{n}\sum_{j=1}^{n} [a_{ij} \cos(z_i + z_j) + b_{ij} \sin(z_i + z_j)] + \varepsilon_i
\end{aligned}
\tag{6.15}
$$

where z_i is the adjusted value of the natural log of the output Q_i so that z_i span the interval $[0.1 * 2\pi, 0.9 * 2\pi]$.[10]

Eq. (6.15) is the standard model used to estimate the cost function and derive efficiency estimates using the Fourier Flexible form. At this point, it should be noted that recent studies have added additional sets of variables in their standard Fourier form, mainly financial capital, asset quality, and time trend variables. These variables are included to account for risk, loan quality, and technical progress respectively when measuring inefficiency.[11]

A financial capital variable has recently been included in a number of cost and profit efficiency studies. An adequate level of financial capital may indicate the ability of banks to absorb losses and work as a cushion against any insolvency risks, resulting in more efficient performance. Moreover, in order to lessen cost inefficiencies, financial capital could be an alternative source to finance a bank's portfolio instead of relying on

[10] The ends of the $[0,2\pi]$ interval are cut off by 10% so that the z_n span $[0.1 * 2\pi, 0.9 * 2\pi]$ to reduce the approximation problems near endpoints (Gallant, 1981). The formula for z_n is $\{0.2\pi - \mu * a + \mu *$ variable$\}$ where $[a,b]$ is the range of the variable being transformed, and $\mu \equiv (0.9 * 2\pi - 0.1 * 2\pi)/(b-a)$ (see Berger and Mester, 1997).

[11] In addition, environmental variables such as fixed assets and off-balance sheet variables have also been included in these studies (see e.g. Berger and Mester, 1997; Altunbas et al., 2000).

debt finances, which incur interest payments.[12] Inclusion of financial capital in cost and profit efficiency estimation can also take into account a bank's typical risk preferences (Berger and Mester, 1997). For example, banks' managements that obtain capital beyond their profit maximization schemes may be classified as risk averse banks. However, on the other hand, these banks may have more incentives to engage in riskier activities incurring volatile profits, which may result in inefficiency when negative profits dominate the outcomes of their operations.

Recent studies (such as those of Altunbas et al., 2000; Berger and Mester, 1997; Mester, 1996) have shown the importance of considering asset quality in the efficiency measurement. Higher loan problems (proxied by non-performing loans or loan provisions) may mean that there is an amount of loans extended to low-quality borrowers that face repayment difficulty. Moreover, high loan problems can cast doubts on the screening and monitoring methods of a bank. For these reasons, the loan problems factor is expected to be a possible reason for distancing a bank from the efficient frontier.

A time trend variable has also been incorporated in various recent studies (such as those of Altunbas et al., 2000; and Lang and Welzel, 1996) to account for disembodied technical change. As the method of production changes over time, the time trend captures the factors of technological change, improvements in skills through learning by doing and training, as well as organizational and regulatory changes that may affect the efficient use of input resources (Altunbas, 2000; Baltagi and Griffin, 1988).

Technical progress causes the bank's total cost to shift inward over time with respect to the production of a given output, Q, holding input prices and regulatory conditions unchanged. This could be measured by taking the partial derivative of the estimated specified model, say total cost equation, with respect to the included time variable (*T*), which could be shown as follows: $\textit{Technical change} = \frac{\partial \ln TC}{\partial T} = t_1 + t_{11}T$, where t_1 and

[12] Banks treat paid interest on debt as cost, but paid dividends on capital are not considered as costs (Berger and Mester, 1997).

t_{11} are parameters capturing the pure effect of technical progress on the declining cost, holding the output proportions unchanged.

By considering the above-mentioned variables, we arrive at our preferred model, which can be written as[13]

$$\begin{aligned}\ln TC &= \alpha_0 + \sum_{i=1}^{2}\alpha_i \ln Q_i + \sum_{i=1}^{2}\beta_i \ln P_i \\ &+ \kappa_1 \ln E + \nu_1 \ln PROV + \tau_1 T \\ &+ \frac{1}{2}\left[\sum_{i=1}^{2}\sum_{j=1}^{2}\delta_{ij}\ln Q_i \ln Q_j + \sum_{i=1}^{2}\sum_{j=1}^{2}\gamma_{ij}\ln P_i \ln P_j + t_{11}T^2\right] \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}\rho_{ij}\ln Q_i \ln P_j + \sum_{i=1}^{2}[a_i\cos(z_i) + b_i\sin(z_i)] \\ &+ \sum_{i=1}^{2}\sum_{j=1}^{2}[a_{ij}\cos(z_i + z_j) + b_{ij}\sin(z_i + z_j)] + \varepsilon_i,\end{aligned} \tag{6.16}$$

where E is equity capital, $PROV$ is total loan provisions, and T is time trend. Since both risk and asset quality have been the variables under focus to measure the health of the banking system, we estimate Eq. (6.16) both including and excluding risk and quality factors in order to see how far these factors have an effect on the inefficiency estimates for our sample of Gulf banks. We call the model that excludes risk and quality factors the traditional model.

Note that, in Eq. (6.16), when estimating the profit functions, TC is replaced by profits *(PROF)* on the left-hand side for both the alternative and the standard profit functions. Moreover, the right-hand side of Eq. (6.16) is identical for both cost and alternative profit functions. However, for the standard profit function, we only replace the output quantities with output prices.[14]

[13] This preferred model is chosen from the feedback of our estimation experiment noted in the next chapter. In fact, the availability of data on the variables, how well the model behaves in the estimation process, and the validity of the model to pass the structural tests determined our model choice.
[14] For instance, the standard profit function is shown as

Eq. (6.16) may be characterized by increasing, constant, or decreasing returns to scale, which means that because the degree of returns to scale is not known, the model might be non-homogeneous. Thus, homogeneity restrictions are imposed on the translog portion of Eq. (6.16) to ensure that the cost function (as well as the profit functions) is linearly homogeneous in input prices. The homogeneity restrictions are shown as

$$\sum_{i=1}^{2} \beta_i = 1,$$

$$\sum_{i=1}^{2} \gamma_{ij} = 0 \quad \text{for all } j,$$

and $$\sum_{i=1}^{2} \rho_{ij} = 0 \quad \text{for all } j.$$

Moreover, Young's theorem requires symmetry of the second order parameters of the translog cost function, that is:

$$\delta_{ij} = \delta_{ji} \quad \text{for all } i,j,$$

and $$\gamma_{ij} = \gamma_{ji} \quad \text{fore all } i,j.$$

When solving for linear homogeneity restrictions, both the cost and the profit models are normalized by the price of labour (P_2) (see e.g. Greene, 1993; Berger and Mester, 1997; Altunbas et at., 2000). This can ensure that, on the efficient frontier, when input

---

$$\begin{aligned}
\ln PROF = {} & \alpha_0 + \sum_{i=1}^{2} \alpha_i \ln G_i + \sum_{i=1}^{2} \beta_i \ln P_i \\
& + \kappa_1 \ln E + v_1 \ln PROV + \tau_1 T \\
& + \frac{1}{2}\left[\sum_{i=1}^{2}\sum_{j=1}^{2} \delta_{ij} \ln G_i \ln G_j + \sum_{i=1}^{2}\sum_{j=1}^{2} \gamma_{ij} \ln P_i \ln P_j + t_{11} T^2 \right] \\
& + \sum_{i=1}^{2}\sum_{j=1}^{2} \rho_{ij} \ln G_i \ln P_j + \sum_{i=1}^{2} [a_i \cos(z_i) + b_i \sin(z_i)] \\
& + \sum_{i=1}^{2}\sum_{j=1}^{2} [a_{ij} \cos(z_i + z_j) + b_{ij} \sin(z_i + z_j)] + \varepsilon_i,
\end{aligned}$$

where $\ln PROF$ is ln (π+θ) given that π is the profit variable, θ is a constant added to the firm's profit so that the natural log of profits is positive, and G is the output price variable.

prices double, costs will exactly double by the same proportion as well, which would leave the input quantities unaffected.

Moreover, for the alternative profit function, homogeneity restrictions will serve to keep the relationship between input prices and profits in an equivalent fashion, although they need not to be imposed on the alternative profit function (Berger and Mester, 1997; Berger and DeYoung, 2000).

In this section we have explained how we specified our functional form to estimate inefficiency. We have shown that the Fourier functional form (Eq. 6.16) is the preferred model used to estimate cost, standard, and alternative profit functions. We obtain the parameters of these functions, as well as their inefficiency estimates, using Maximum Likelihood Estimation (MLE) regression. The next section explains the variables and the data used in the estimation of our Fourier Flexible model given in Eq. (6.16).

6.5 Data and variables definition

Having specified our models in the previous section, this part of the chapter details various aspects of the data, outlines the models' variables, and discusses the use of the panel data technique in our analysis of GCC bank efficiency.

The Data

The study contains a balanced time series cross-sectional dataset, which consists of 93 GCC banks covering the six-year period from 1995 to 2000. The source of our data is mainly the London-based IBCA bank credit rating agency's database (Bankscope, Jan., 2002). Data on foreign banks operating in the UAE are taken from Financial Position of Commercial Banks in the UAE (1995-2000), published by the Emirates Banks' Association. Further, data on foreign banks operating in Qatar have been acquired from the annual financial statements of these banks.

The majority of data in our sample relates to commercial banks, with the exception of seven specialized banks, that are included to enhance the total number of observations in order to reduce the impact of multicollenearity among variables.[15 & 16]

Table 6.1 shows the percentage of the total bank assets for each country included in the sample relative to the total assets of the banking industry in each country in the year 2000. The table indicates that the sample constitutes at least 89 per cent of the total banking industry's assets in Qatar, the UAE, Saudi Arabia, and Kuwait.

The percentage of assets of Bahraini banks included in the sample is about half of the total bank assets of Bahrain's banking industry, (as the rest belongs to the offshore banking units and other financial institutions, for which data are unavailable). Moreover, the sample contains 64 per cent of the total Omani bank assets as data on the remaining banks are not available.

Table 6.1 Total assets of banks in each country in the sample relative to the total assets of the banking industry by country, 2000 - ('000 US dollars)

Sources: Bankscope (Jan., 2002), financial reports of banks in the UAE and Qatar, and the annual reports published by the central banks in each country.

[15] The list of GCC banks included in the sample is given in appendix 3.
[16] According to the bank classification adopted in Qatar and by the UAE central bank authorities, Islamic banks are considered as commercial banks.

Figure 6.1 shows the share of the bank assets of each country included in the sample relative to the total bank assets of the whole sample in the year 2000. With only 9 Saudi banks, the figure indicates that the Saudi banks occupy the largest share of total assets of banks included in the sample. UAE banks occupy the second largest share in the sample, given that the number of UAE banks in the sample is 43, the highest among all GCC countries included in the sample.

Figure 6.1 Total assets of individual GCC country banks in the sample as a share in the total banking industry's assets for the underlying GCC country – Year 2000

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Sources: Bankscope (Jan., 2002) and financial reports published by banks in the UAE and Qatar.

Table 6.2 gives a break-down of the sample according to the size of banks in terms of total assets. Small banks are banks with assets of less than $300 million; banks whose assets fall in the range between $300 million and $1 billion are classified as medium-sized banks; and if a bank's assets are greater than $1 billion, we define these as large banks. The reason for classifying bank sizes according to these ranges is that it allows a greater number of banks in each country to be allocated within their correspondent size class. According to this size classification, Table 6.2 shows that large banks take the major stake in the total size of the GCC banks, about 89 per cent compared to 8 and 3 per cent for medium and small banks respectively. The number of banks for each size class is also shown in Table 6.2.

Table 6.2 Total GCC bank assets according to size class

Third Party material excluded from digitised copy. Please refer to original text to see this material.

Sources: Bankscope (Jan., 2002) and financial reports of UAE and Qatar banks.
Note: Total no. of banks is 93.

Another classification according to bank size is attempted in order to see if the general conclusion regarding the inefficiency estimates would change.[17] These size classes are given in Table 6.3.

Table 6.3 Total GCC bank assets according to size class

Third Party material excluded from digitised copy. Please refer to original text to see this material.

Sources: Bankscope (Jan., 2002) and financial reports of UAE and Qatar banks.
Total no. of banks is 93.

[17] This arrangement would allocate foreign banks in the range of small to medium-size classes.

With regard to foreign banks' data, the study has access only to information relating to banks in the UAE and Qatar. Owing to the strict privacy policies, access to such information relating to foreign banks in Oman and Bahrain has not been possible. Regulations governing both the Kuwaiti and Saudi banking systems forbade the operation of foreign banks during the study period. Consequently, the study will consider foreign banks operating in both Qatari and UAE banking industries as representative of how well foreign banks are performing in the GCC banking system.

Figure 6.2 shows that the total assets of the foreign commercial banks is about a quarter of the total assets of GCC banks included in the sample.

Figure 6.2 Percentage of total foreign and national banks' assets in the GCC banking sample – Year 2000

Third Party material excluded from digitised copy.
Please refer to original text to see this material.

Sources: Bankscope (Jan., 2002) and financial reports published by banks in the UAE and Qatar.

The figure indicates that the ratio of foreign banks' assets in the total GCC banks included in the sample is 7 per cent, which is nearly equivalent to both Qatar and Oman banks assets shares in the sample.[18] The number of foreign banks in the sample is 30, equivalent to 32 per cent of the total 93 banks included in the sample.

[18] Our sample indicates that the ratio of foreign banks' assets in the total assets of UAE and Qatari banks in the sample is 14 and 26 respectively.

The variables

Table 6.4 defines the variables used in the specification of cost and profit functions of Eq. (6.16).

Table 6.4 Descriptive statistics of the outputs, inputs, and control variables used in the Eq. (6.16)

Variable	Description	Mean	St. Dev.	Min.	Max.
Dependent variables					
TC	Total cost includes interest expenses and operating costs ('000 US dollars)	155,717.4	262,769.4	1,736.9	1,728,938.2
PROF	Profits include revenues from loans and other earning assets less total cost ('000 US dollars)	62,480.4	81,198.4	6,472.7	522,098.6
Prices of inputs					
P1	Price of deposits	0.0469	0.0100	0.0218	0.0728
P2	Price of labour	0.0182	0.0078	0.0047	0.0530
Output quantities					
Q1	Total loans ('000 US dollars)	1256245	2090826	2807	1728938
Q2	Other earning assets ('000 US dollars)	1397850	2488464	13456	14409000
Prices of Outputs					
G1	Price of loans	0.1413	0.0394	0.0509	0.2000
G2	Price of other earning assets	0.0402	0.0333	0.0018	0.2959
Control variables					
E	Total equity ('000 US dollars)	307,501.8	451,200.0	0.3	2,297,223.0
PROV	Total provisions ('000 US dollars)	12,688.1	42,650.2	0.3	882,099.1
T	Time trend			5.0	10.0

Sources: Bankscope (Jan., 2002), financial reports of banks in the UAE and Qatar, and annual reports published by the central banks of the GCC countries.

These variables are given along with their descriptive statistics including sample means and standard deviations. Both cost and alternative profit functions specify two outputs, two inputs, two input prices, and two output prices variables used in the standard profit functions, as well as risk, asset quality, and technical progress variables.

The specifications of outputs and inputs are viewed from the assets and liabilities sides respectively, which conforms with the intermediation approach to modelling banking production (Sealey and Lindley, 1977).[19] The output variables are total loans, denoted by Q_1; and other earning assets, denoted by Q_2, which reflects investments or securities categories.

Two prices of inputs are considered: prices of borrowed funds, denoted by $P1$; and prices of labour, denoted by $P2$. These are calculated as follows. $P1$ is obtained by the division of interest paid by the borrowed funds, where borrowed funds are the total of all interest bearing deposits. $P2$ is a proxy of labour price computed as the ratio of staff costs to total assets.[20 & 21]

The dependent variable of the cost function, denoted by TC, is obtained from the sum of interest expenses and the staff costs, where both of these comprise the vast majority of the banking total cost. Variable profits, denoted by $PROF$, are calculated as the revenues from loans and other earning assets less total cost.

To control for bank risk, we use financial capital, denoted by E. The variable $PROV$ is the loan loss provisions taken as a proxy for loan (or assets) quality.[22] The model also includes time trend, denoted T, which accounts for technical progress.

[19] Approaches to defining bank inputs and outputs are discussed in Chapter 5.
[20] Price of labour is usually computed by the division of staff cost by the number of staff. However, owing to the non-availability of data on staff numbers we follow Altunbas et al. (2000) to calculate the price of labour as a ratio of staff cost to the total assets.
[21] The majority of studies also include the price of fixed assets. However, for many banks considered in this research (especially foreign banks in the UAE) there are no data on fixed assets expenses (for example, depreciations) to calculate the price of fixed assets. We are therefore forced to confine the number of inputs to borrowed funds and labour.
[22] Among categories of loan loss provisions, loan loss level and the non-performing loan data, only the loan loss provisions category is available for the entire sample.

Why panel data technique?

It will be appropriate to consider the nature of the data under study and to employ the estimation technique that, accordingly, considers the data nature. Our data set can be classified under a type known by econometricians as panel data. Panel data refers to 'the pooling of observations on a cross-section of households, countries, firms, etc. over several time periods' (Baltagi, 2001, p. 1). Thus, our panel data combines both time series (6 years) and cross-sections (93 banks) together. Generally, the analysis of panel data simply tends to have more of the cross-section than time-series characteristics since $N > T$; that is, the number of the observed firms in panel data is greater than the number of the observed time (see Greene, 2000).

In estimating our model we use panel data approaches. The main benefits using such approaches are (see Batagi, 2001; Hsiao, 1985 and 1986; Solon, 1989):

- Panel data approaches can help control for heterogeneity across the data sample. Differences in size, ownership type, and so on can be more accurately controlled for using panel data approach.

- Because panel data have more N relative to fewer T, the domination of the cross-section over the time series gives much variability and more informative data. This helps overcome the multicollinearity problem, which usually plagues time-series data. In this case, panel data estimators are more statistically efficient.

- Panel data better study technical changes and technical efficiency over time, given that technical changes and technical efficiency are specified by an appropriate parametric model.

- Panel data techniques are better able to treat dynamic changes, adjustments, and intertemporal changes that occur from one point to another within the period studied.

According to the panel data literature, the treatment of the firm specific characteristics is modelled by either fixed or random effects approaches (see Greene, 2000, Chapter 14; Gujarati, 2003, Chapter 16). In the fixed effects approach, the regression model is allowed to differ among banks in order to capture some special characteristics of each bank. Thus, the differences across banks can be captured by differences in the constant term. In the random effects approach, the firm specific characteristics are captured to reflect an intercept which is assumed to be a random disturbance drawn from a much larger population with a constant mean value. The deviation from this constant mean is the individual intercept.

In our stochastic frontier framework, the random effects approach is more appropriate than the fixed effects approach since the fixed effects approach may induce a large loss in the degrees of freedom when the number of units is large. (However, in the following chapter we also test if the bank's individual effects are present so as to compare fixed and random effects approaches).

The rest of this chapter discusses the methods used to estimate other features that accompany the analysis of the GCC banking industry's inefficiency. These are: scale economies, scale inefficiency, as well as the determinants of inefficiency levels. (Note, we do not consider scope economies due to the limitations associated with estimating theses economies).

6.6 Scale economies and scale inefficiency

The empirical work of this thesis is also interested in measuring both scale economies and scale inefficiency in the GCC banking industry. It is important to know whether GCC banks are exploiting their scale capabilities to the extent that the cost of a unit output produced reaches the minimum possible level. Moreover, it is also important to know how the measure of X-inefficiency is compared to scale inefficiency, a task that helps in examining if the cost savings are more usually achieved through altering the size of banks or through managerial improvements. In this section, we first introduce how economies of scale are measured using our Eq. (6.16), followed by scale inefficiency.

Economies of scale

Economies of scale explains by how much a proportional change in outputs level would lead to a change in total cost. In other words, economies of scale express the total cost elasticity with respect to output, which can be obtained by differentiating the cost function with respect to output. Thus, for the two outputs in our banking sample, economies of scale solved for Eq. (6.16) are given as

$$\begin{aligned} Scale\ economies = \sum_{i=1}^{2}\frac{\partial \ln TC}{\partial \ln Q_i} = & \sum_{i=1}^{2}\alpha_i + \sum_{i=1}^{2}\sum_{j=1}^{2}\delta_{ij}\ln Q_j + \sum_{i=1}^{2}\sum_{j=1}^{2}\rho_{ij}\ln P_i \\ & + \mu_i\sum_{i=1}^{2}\left[-a_i\sin(Z_i) + b_i\cos(Z_i)\right] \\ & + 2\mu_i\sum_{i=1}^{2}\sum_{j=1}^{2}\left[-a_{ij}\sin(Z_i+Z_j) + b_{ij}\cos(Z_i+Z_j)\right]. \end{aligned} \tag{6.17}$$

If $\sum_{i=1}^{n}\frac{\partial \ln TC}{\partial \ln Q_i} = 1$, this shows that a proportional change in outputs yields the same proportional change in total cost. This is known as constant returns to scale or constant economies of scale. When the measurement $\sum_{i=1}^{n}\frac{\partial \ln TC}{\partial \ln Q_i}<1$, this means that a proportional change in outputs leads to a change in the total cost with a proportional change less than that of output. In this case the relationship between output and total cost is said to exhibit increasing returns to scale, implying economies of scale. If $\sum_{i=1}^{n}\frac{\partial \ln TC}{\partial \ln Q_i}>1$, this means that a proportional change in outputs leads to a more than proportional change in total cost. This relationship is known as decreasing returns to scale, which implies diseconomies of scale.

Scale inefficiency

Evanoff and Israilevich (1995) distinguish between scale economies and scale inefficiency. While the former measures the change in total cost with respect to the change in the output level, scale inefficiency measures how much a bank needs to change its output levels so that it moves to the size of the scale-efficient bank, which has the minimum efficient scale. In other words, as Evanoff and Israilevich (1995, p. 1037) put it, '[s]cale inefficiency, *I*, can be measured as the aggregate cost of *F* inefficient firms ... relative to the cost of a single efficient firm where *F*= size of the efficient relative to the inefficient one.'

Obviously, these authors direct the attention to the fact that scale economies and X-inefficiency cannot be directly compared. Thus, scale economies need to be transformed into a scale inefficiency measure. By adopting the approach suggested by Evanoff and Israilevich (1995), scale inefficiency can be measured as

$$\textit{Scale inefficiency} = e^{(.5/c)(1-\varepsilon)^2} - 1,$$

where ε is the first derivative of the cost function, that is the scale elasticity with respect to output, and c is the second derivative of the cost function with respect to output. The interpretation of scale inefficiency measures is the percentage that cost could decline if banks were to move to minimum efficient scale.

We use our Eq. (6.16) to obtain the arguments that we will incorporate in the scale inefficiency equation. Once we obtain the measures of scale inefficiency, we will compare these with the X-inefficiency measure to judge which is more important for GCC banks.

6.7 Inefficiency determinants & logistic regression

After measuring cost and profit inefficiency levels in the GCC banking sector, one may need to go a step further and investigate the sources or the possible determinants of inefficiency in the industry. In order to do this, we need to employ the most likely influential variables and the appropriate econometric technique. Table 6.5 presents the descriptive statistics of the variables that are examined as possible inefficiency determinants. Most of these variables have been used in studies such as those of Mester (1996), Altunbas et al. (2000), and Girardone et al. (2000).

Table 6.5 Descriptive statistics of the variables used in the logistic regression model

Variable	Description	Mean	St. Dev.	Min	Max
CN	Cost inefficiency half-normal	0.0839	0.0469	0.0144	0.4467
SN	Standard profit inefficiency half-normal	0.3312	0.1849	0.0361	.9897
E	Total equity ('000 US dollars)	307,501	451,200	0.3000	2,297,223
ROA	Rate of return on assets	0.0480	0.0404	0.0064	0.2841
PROV	The ratio of provisions to total loans	0.0170	0.0361	0.0000	0.3735
FOREIGN	Foreign banks – dummy variable	-	-	0	1
LTA	Ratio of loans to total assets	0.4966	0.1926	0.0105	0.9059
FIX	Fixed assets ('000 US dollars)	39,767	103,297	1	899,873
TA	Total assets ('000 US dollars)	2,832,011	4,704,465	31,616	26,699,785
TBGDP	Total bank assets as a ratio to GDP	1.7739	1.8609	0.1935	7.3626

Sources: Bankscope (Jan., 2002), financial reports of banks in the UAE and Qatar, and annual reports published by the central banks authorities in the GCC countries.

The inefficiency variables (CN and SN) are the measured cost and profit inefficiencies derived from the traditional Fourier Flexible cost and profit functions that exclude risk and asset quality variables. We use inefficiency estimates derived from the traditional rather than the preferred model because we want to avoid double consideration of the risk and quality factors.

Basically, the authors of various studies (e.g. Mester, 1996; Altunbas et al., 2000) believe that factors of risk and quality are important variables determining inefficiency levels. Accordingly, our inefficiency determinant model mainly includes *E*(=financial capital) and *PROV*(=loan loss provisions); these variables are used again as proxies for risk and loan quality respectively.

Here, it is expected that the sign of *E* is negative, indicating that the more inefficient banks have more risk that may be attributed to inadequate capital maintained in their operations. In other words, efficient banks have lower risk and are more able to generate profits that help in accumulating more retained earnings added to the financial capital (this assumes that dividends are unchanged).

In relation to bank capital, risk, and bank returns, we also include the variable ROA(=rate of return on assets), which is used as a proxy for performance. ROA is expected to be inversely related to cost and profit inefficiency on the grounds that the more inefficient firms are believed to employ their inputs in non-productive outputs that earn low returns.

With regard to the loan quality variable, the sign of the loan quality (*PROV*) is expected to be positive, showing that the more inefficient firms have higher provisions, indicating that they face loan problems and, thus, regulations force them to increase their loan provisions in accordance with deteriorating loan quality.

Moreover, because we wish to consider whether foreign banks are more efficient than their domestic competitors in the GCC, we include the dummy variable FOREIGN(=foreign banks), which consists of a value of one if the bank is foreign and zero otherwise. With regard to the GCC banking data of our sample, the foreign bank dummy variable and inefficiency variables are expected to be positively related since foreign banks operate under restrictions relating to bank size and branching limits, as well as tax impositions that may add to their costs.

Other independent variables are also considered in order to capture additional characteristics of bank and industry specifics. These are: L/TA(=net loans/total assets), FIX(=fixed assets), (TA=total assets), and TBGDP(=total banking assets/GDP). Variables L/TA, FIX, TA, and TBGDP respectively, control for balance sheet mix, bank size, and market size factors that may be influential in influencing banking sector inefficiency.

Overall, in order to investigate the determinants of GCC bank inefficiency we estimate the following model

$$INEFF = f\ (E, ROA, PROV,\ FOREIGN,\ LTA, FIX, TA,\ TBGDP) \qquad (6.18)$$

As mentioned, this model will be estimated using the logistic functional form. The general form of the logistic model is written as

$$\hat{E}(u_i \setminus \varepsilon_i) = \frac{\exp(X_i'\gamma)}{1+\exp(X_i'\gamma)} + \xi_i \qquad (6.19)$$

where X_i is a vector of independent variables for the ith firm, γ is the parameter vector, and ξ_i is a normally distributed error term.

Since the inefficiency variables are the dependent variables with values falling between zero and one, the logistic functional form is preferred here (compared with ordinary least squares methods) because the former is generally used to estimate models where the dependent variables are bounded between zero and one.

Following Mester (1996), the interpretation of the logistic function results only tells us about correlation relationships and do not tell us anything about causality. Nevertheless, the logistic regression is also preferred over the simple correlation method because it is possible to take other variables into consideration when estimating inefficiency determinants.

6.8 Conclusions

This chapter explains the econometric approach which will principally be used to estimate banking X-inefficiencies in GCC banking markets. The chapter describes the functional forms, the data, and the variables used in the efficiency estimation. The chapter also illustrates the methods used to measure both scale economies and scale inefficiency, which are features mainly related to the cost characteristics of banks. To explore the possible variables determining inefficiency in the GCC banking industry, the chapter outlines a logistic regression model approach that will be used to evaluate these efficiency determinants. The next chapter presents the results based on the implementation of the methodology outlined in the present chapter.

CHAPTER 7

EFFICIENCY IN GCC BANKING: EMPIRICAL EVIDENCE

7.1 Introduction

This chapter empirically analyses the efficiency of the GCC banking sector over the period 1995-2000.[1] It presents and compares the results using three efficiency concepts: cost efficiency, standard profit efficiency, and alternative profit efficiency.[2]

Figure 7.1 shows the framework of the empirical work and notes how the inefficiency concepts will be estimated under the assumptions of half-normal and exponential distributions, as set out in the inefficiency component of the residual term. Within this framework, two main specifications are examined: the traditional and the preferred specification that controls for risk and quality factors. In addition, the inefficiency analysis is extended to evaluate both scale economies and scale inefficiencies. In order to investigate the determinants of inefficiency in GCC banking systems, we use a logistic model approach to see if banking sector inefficiency is related to various bank-specific and market-specific features.

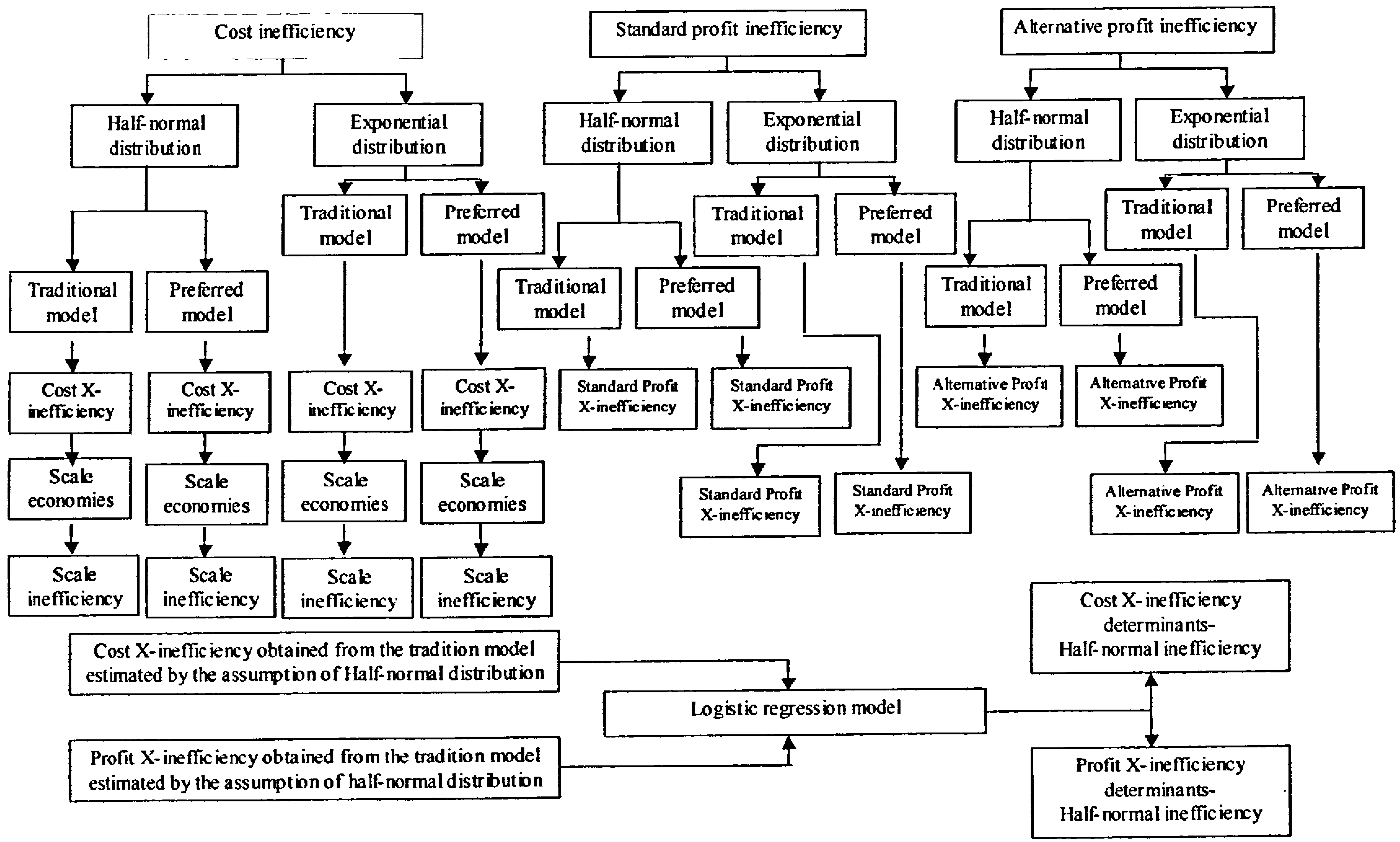
The findings of the empirical analysis show that the level of inefficiency in the GCC banking industry ranges between 8 and 10 per cent for the cost inefficiencies, and between 30 and 32 per cent for the profit inefficiencies. Both half-normal and exponential distributions tend to provide similar inefficiency measures for the profit functions. Because the cost inefficiency measures, using both distributional assumptions, tend to differ, we use the distribution-free approach, that yields

[1] The empirical work in this chapter is carried out using *LIMDEP* econometric software version 7.0.
[2] Hereafter, efficiency/inefficiency are referred to using the term X-efficiency/X-inefficiency.

inefficiency measures similar to those obtained using the half-normal model. The findings show also that the risk and quality factors provide information influencing bank inefficiency levels when we use either the cost or profit function models. When risk and quality factors are considered, the mean inefficiency measures show a slight decrease. Foreign banks are found to be less cost efficient, but more profit efficient than national banks. This suggests that foreign banks focus more on revenue generating than do their domestic counterparts. As foreign banks tend to have a different business mix (high end retail clients, large corporate banking services, and so), it is perhaps not surprising that they are found to be less cost efficient but more profit efficient. Moreover, foreign banks have less scale economies and are less scale efficient. In the logistic regression, cost and profit inefficiency is found to be negatively related with the risk variable. It also provides evidence that inefficiency is positively related to loan quality variables, suggesting that banks with enhanced financial capital and high loan quality are more efficient.

The chapter is divided into seven sections. Section 7.2 discusses the model specifications as well as the results of the model's parameter estimation. Section 7.3 examines the model in terms of the relevant structural tests. Section 7.4 discusses the inefficiency results. Scale economies and scale inefficiency results are outlined in section 7.5. In addition, the relationship between bank ownership type, bank size, and inefficiency differences across countries are also covered in these sections. Section 7.6 examines the results of the logistic regression model used to determine the potential correlates of the inefficiency measures. The last section provides a summary and conclusions.

Figure 7.1 The framework of the empirical analysis of GCC banking inefficiency



Traditional Model = Fourier Flexible form with two outputs, two inputs, and time trend
Preferred Model = Traditional model with addition of risk and asset quality variables. This model is given in Eq. (7.1)

7.2 Model specification

Following the methodology of Berger and Mester (1997), we evaluate three inefficiency concepts: cost inefficiency, profit inefficiency, and alternative profit inefficiency. Each of these efficiency measures are estimated using the corresponding functional forms shown below.

In essence, this section undertakes two main tasks: first, it outlines the functional forms used to estimate and derive cost, standard profit, and alternative profit inefficiency measures. Then, the section presents and analyses the results of parameter estimates using the different model specifications.

The results are based on two specifications: the preferred model and the traditional model. For both the preferred and the traditional models we use the Fourier Flexible form; however, the preferred model differs from the traditional specification by including the equity and loan provisions variables, which are considered in this research as proxies for risk and loan quality factors respectively.

Recalling Eq. (6.16) from the previous chapter, the preferred model, typically shown below for the cost function, is written as[3]

[3] Eq. (7.1) is estimated using a panel data random effects model (see Chapter 6). In fact, after undertaking many estimation experiments, this model with the variables it contains is the model that behaved well considering all efficiency concept functions, distributional assumptions, and model specifications, as well as the structural tests shown in the next section. Obviously, owing to convergence problems, variables such as the interactive terms of risk and quality factors are not included in the model. In fact, some success in convergence happened by partial inclusion of the interactive terms of the quality variable; nevertheless, this did not last across all model specifications and distributional assumptions. In addition, the fixed assets variable caused convergence problems as well. Fixed assets are, however, included in the logistic model specification for examining the determinants of the inefficiency in the GCC banking system. In order to keep the comparison consistent and on the same footing across specifications and distributional assumptions, we prefer to undertake our estimation using Eq. (7.1).

$$
\begin{aligned}
\ln TC = {} & \alpha_0 + \sum_{i=1}^{2} \alpha_i \ln Q_i + \sum_{i=1}^{2} \beta_i \ln P_i \\
& + \kappa_1 \ln E + \nu_1 \ln PROV + \tau_1 T \\
& + \frac{1}{2}\left[\sum_{i=1}^{2}\sum_{j=1}^{2} \delta_{ij} \ln Q_i \ln Q_j + \sum_{i=1}^{2}\sum_{j=1}^{2} \gamma_{ij} \ln P_i \ln P_j + t_{11} T^2 \right] \\
& + \sum_{i=1}^{2}\sum_{j=1}^{2} \rho_{ij} \ln Q_i \ln P_j + \sum_{i=1}^{2} [a_i \cos(z_i) + b_i \sin(z_i)] \\
& + \sum_{i=1}^{2}\sum_{j=1}^{2} [a_{ij} \cos(z_i + z_j) + b_{ij} \sin(z_i + z_j)] + \varepsilon_i,
\end{aligned}
\tag{7.1}
$$

where

TC	=	total cost (financial and operating costs);
$PROF$	=	profits (including revenues from loans and other earning assets less total cost) used in case of profit functions as a dependent variable;
Q_1	=	total loans;
Q_2	=	total other earning assets (including items such as securities and non-loan earning financial assets);
P_1	=	price of direct inputs (deposits);
P_2	=	price of indirect input (mainly labour cost);
E	=	financial capital;[4]
$PROV$	=	loan loss provisions;
G_1	=	price of loans in case of the standard profit function;

[4] We may equivalently use equity and financial capital to refer to the same thing, knowing that equity is in fact a component of financial capital, which includes retained profits and reserves.

G_2	=	price of other earning assets in case of the standard profit function;
T	=	time trend used as a proxy for technical progress;
ε_i	=	stochastic error term, where $\varepsilon = u + \upsilon$;
z_i	=	adjusted values of the log output ln Q_i so they span the interval $[0,2\pi]$;
$\alpha, \beta, \kappa, v, \tau, \rho\ \delta, \gamma,$ a, and b	=	parameters to be estimated;
Ln (variable)	=	the natural logarithm of a variable;
i	=	1, 2; and
j	=	1, 2.

The alternative profit function equation is identical to the cost function shown in Eq. (7.1), except that the left-hand side variable, *TC*, must be changed to the *PROF* variable. With regard of the standard profit function, the *PROF* variable replaces *TC* in Eq. (7.1) as well. In addition, on the right-hand side of the standard profit function, the price of loans (G_1) and the price of other earning assets (G_2) are included in the place of loan quantity variable (Q_1) and the other earning assets quantity (Q_2) respectively.

We first estimate the cost, standard profit, and alternative profit functions using the preferred specification [Eq. (7.1)]. Then we estimate the traditional model, which has the same arguments as the preferred model except that both risk and quality variables are excluded. Setting up these two model specifications may provide us with information on what difference these important factors could make when estimating cost and profit levels, cost and profit inefficiency levels, and scale economies and scale inefficiencies.

Moreover, for a robustness check, both the traditional and the preferred model are estimated separately with half-normal and exponential distribution models, as we assume that the inefficiency term component, u, of the error term, ε, follows either one of these types of distributions.

For both the preferred and the traditional specifications, the standard restrictions of homogeneity and duality are to be imposed on Eq. (7.1) as follows:[5]

Homogeneity restrictions

$$\sum_{i=1}^{2} \beta_i = 1,$$

$$\sum_{i=1}^{2} \gamma_{ij} = 0 \qquad \text{for all } j,$$

$$\text{and} \quad \sum_{i=1}^{2} \rho_{ij} = 0 \qquad \text{for all } j.$$

Duality, or symmetry restrictions

$$\delta_{ij} = \delta_{ji} \qquad \text{for all } i,j,$$

$$\gamma_{ij} = \gamma_{ji} \qquad \text{for all } i,j.$$

As these restrictions are imposed on the translog portion of Eq. (7.1), the linear homogeneity restrictions result in a normalization of the variables TC and P_1 by the price of labour, P_2. Homogeneity restrictions allow proportion change of input and output to be equivalent. For example, doubling inputs leads to doubling of outputs.

[5] The restrictions are imposed in the same way on the cost, standard, and alternative profit functions. Moreover, in following Berger and Mester (1997), we impose no homogeneity restrictions on the output prices variables of the standard profit function.

Moreover, the duality restrictions help avoid the repetitions of variables that are cross-products.[6]

The estimates in Tables (7.1a to 7.1f) show the maximum likelihood parameter estimation (MLE) of the Fourier Flexible cost, standard profit, and alternative profit functions, which are estimated for both the traditional and preferred model specifications. In Tables (7.1a to 7.1f), we see that the estimates of the model parameters are quite similar across model specifications (traditional and preferred model, as well as models excluding only risk variable [equity] or loan quality variable [provisions] from the preferred model).

Moreover, model parameters are also fairly similar for the half-normal and exponential models for each of the efficiency concepts, and this similarity has also been found in a number of earlier studies (see e.g. Aigner et al., 1977). This indicates that the choice between half-normal and exponential distributions has little impact on our model parameters, which may also result in the similarity between the inefficiency scores, as we will discuss later in this chapter.[7]

The results in Tables (7.1a to 7.1f) show that the functions' estimated coefficients mostly have consistent signs. To be specific, the input prices (P_1 and P_2) have positive effects on costs, implying that higher input prices lead to greater costs [see Tables (7.1a and 7.1b)]. Moreover, in Tables (7.1e and 7.1f), the positive relationship between the prices of inputs (P_1) and alternative profits may be explained by the fact that when the price of deposits increases, loan prices also increase, resulting in higher profits. Because output quantities are set as given in the alternative profit function and prices of output are left to move freely, changes in output prices induced by input price movements may bring the latter and profits into close relationship.

---

[6] Fourier terms as well as any other terms, such as risk, quality, and time trend term, are not normalized because they are not multiplicative with input prices.

[7] This is only true for the two profit efficiency measures, standard and alternative profits. However, as section 7.4 elaborates, efficiency measures of the half-normal and exponential cost functions are not as close as each other's efficiency measures. Thus, the approach of the distribution-free is used to decide which efficiency measure to rely on.

Table 7.1 (a to f) Maximum likelihood parameter estimation for cost, standard, and alternative profit functions

Table 7.1 a

Maximum likelihood parameter estimation for the cost function using the half-normal model

Variable	Parameter	Traditional model			Preferred model Eq. (7.1)			No Equity			No Provisions		
		Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value
Constant	α_0	-5.6534	10.7210	-0.5270	-4.8803	15.6330	-0.3120	-5.7154	10.5360	-0.5420	-4.7678	15.1940	-0.3140
lnQ1	α_1	-0.8054	1.3327	-0.6040	-0.5292	2.3481	-0.2250	-0.7993	1.3166	-0.6070	-0.5465	2.2376	-0.2440
lnQ2	α_2	2.7250	1.2191	2.2350	2.4391	1.0588	2.3040	2.7296	1.2320	2.2160	2.4317	1.0610	2.2920
lnP1	β_1	0.5019	0.0428	11.7340	0.4970	0.0434	11.4650	0.5017	0.0435	11.5440	0.4989	0.0428	11.6620
lnE	κ	-	-	-	-0.0225	0.0025	-8.8460	-	-	-	-0.0228	0.0025	-9.0970
lnPROV	ν	-	-	-	0.0020	0.0028	0.7270	0.0019	0.0027	0.7110	-	-	-
T	τ	-0.0421	0.0443	-0.9500	-0.0458	0.0454	-1.0090	-0.0427	0.0442	-0.9680	-0.0454	0.0449	-1.0110
TT	τ^2	0.0055	0.0058	0.9510	0.0060	0.0060	1.0070	0.0056	0.0058	0.9660	0.0060	0.0059	1.0110
lnQ1 lnQ1	δ_{11}	0.2953	0.1067	2.7680	0.2824	0.1883	1.5000	0.2944	0.1055	2.7920	0.2847	0.1789	1.5910
lnQ1 lnQ2	δ_{12}	-0.1555	0.0130	-11.9200	-0.1704	0.0130	-13.1340	-0.1553	0.0133	-11.7000	-0.1706	0.0126	-13.5170
lnQ2 lnQ2	δ_{22}	-0.0456	0.0915	-0.4990	-0.0056	0.0791	-0.0710	-0.0460	0.0924	-0.4970	-0.0051	0.0792	-0.0640
lnP1 lnP1	γ_{11}	0.2491	0.0527	4.7240	0.2433	0.0502	4.8500	0.2513	0.0534	4.7040	0.2421	0.0490	4.9430
lnP1 lnP2	γ_{12}	-0.2115	0.0340	-6.2220	-0.2010	0.0343	-5.8670	-0.2123	0.0345	-6.1600	-0.2001	0.0339	-5.9030
lnP1 lnQ1	ρ_{11}	0.1004	0.0301	3.3320	0.0965	0.0291	3.3110	0.0990	0.0302	3.2740	0.0981	0.0294	3.3370
lnP1 lnQ2	ρ_{12}	-0.1043	0.0311	-3.3510	-0.0972	0.0303	-3.2090	-0.1024	0.0307	-3.3340	-0.0992	0.0310	-3.2030
lnP2 lnQ1	ρ_{21}	0.0154	0.0219	0.7050	0.0036	0.0205	0.1750	0.0163	0.0227	0.7180	0.0034	0.0204	0.1650
cos(z1)	a_1	-0.6970	0.5260	-1.3250	-0.5632	0.8861	-0.6360	-0.6931	0.5204	-1.3320	-0.5739	0.8453	-0.6790
sin(z1)	b_1	0.0153	0.1420	0.1080	-0.0244	0.2566	-0.0950	0.0130	0.1412	0.0920	-0.0178	0.2346	-0.0760
cos(z2)	a_2	0.7306	0.3138	2.3280	0.6695	0.2817	2.3770	0.7317	0.3159	2.3170	0.6661	0.2811	2.3700
sin(z2)	b_2	0.0744	0.0828	0.8980	-0.0358	0.0633	-0.5660	0.0757	0.0863	0.8770	-0.0304	0.0640	-0.4750
cos(z1+z1)	a_{11}	-0.1552	0.0681	-2.2790	-0.1374	0.1131	-1.2150	-0.1544	0.0678	-2.2790	-0.1393	0.1067	-1.3060
sin(z1+z1)	b_{11}	0.0668	0.0452	1.4790	0.0227	0.0739	0.3080	0.0656	0.0450	1.4580	0.0260	0.0685	0.3800
cos(z1+z2)	a_{12}	0.0059	0.0273	0.2180	0.0522	0.0342	1.5250	0.0057	0.0270	0.2120	0.0493	0.0331	1.4880
sin(z1+z2)	b_{12}	-0.0627	0.0346	-1.8120	-0.0269	0.0348	-0.7740	-0.0626	0.0347	-1.8030	-0.0269	0.0349	-0.7710
cos(z2+z2)	a_{22}	0.1444	0.0472	3.0600	0.0970	0.0453	2.1390	0.1453	0.0463	3.1380	0.0968	0.0458	2.1140
sin(z2+z2)	b_{22}	0.0414	0.0270	1.5330	0.0055	0.0234	0.2330	0.0422	0.0273	1.5460	0.0064	0.0237	0.2700
	λ	1.5721	0.3254	4.8320	1.5728	0.2800	5.6170	1.5601	0.3198	4.8780	1.5697	0.2860	5.4890
	σ	0.0061	0.0003	19.6390	0.0056	0.0003	21.7900	0.0060	0.0003	19.7230	0.0056	0.0003	22.2840
	θ	-	-	-	-	-	-	-	-	-	-	-	-
	σ_v	-	-	-	-	-	-	-	-	-	-	-	-
lnP2	β_2	0.4981			0.5030			0.4983			0.5011		
lnP2 lnP2	γ_{22}	0.2115			0.2010			0.2123			0.2001		
lnP2 lnQ2	ρ_{22}	-0.0154			-0.0036			-0.0163			-0.0034		
Variance components	$\sigma^2(v)$	0.00606			0.0056			0.00603			0.00565		
	$\sigma^2(u)$	0.00953			0.0089			0.00940			0.00886		
Log Likelihood function		527.3			542.4			527.9			541.2		
Function converged at iteration		11			20			12			19		
R^2		0.996			0.996			0.996			0.996		

Traditional Model = Fourier Flexible form with two outputs, two inputs, and time trend

Preferred Model = Traditional model with the addition of risk and asset quality variables. This model is given in Eq. (7.1)

Table 7.1 b

Maximum likelihood parameter estimation for the cost function using the exponential model

Variable	Parameter	Traditional model			Preferred model Eq (7.1)			No Equity			No Provisions		
		Coefficient	Std Error	T-value	Coefficient	Std Error	T-value	Coefficient	Std. Error	T-value	Coefficient	Std. Error	T-value
Constant	α_0	-5.6237	17.1930	-0.3270	-4.8586	53.6430	-0.0910	-5.6865	18.0310	-0.3150	-4.7452	52.6630	-0.0900
lnQ1	α_1	-0.8015	2.5010	-0.3200	-0.5482	6.2148	-0.0880	-0.7934	2.7731	-0.2860	-0.5632	6.2024	-0.0910
lnQ2	α_2	2.7054	1.5721	1.7210	2.4012	7.5186	0.3190	2.7082	1.5516	1.7450	2.3968	7.4175	0.3230
lnP1	β_1	0.4864	0.0511	9.5140	0.4618	0.1987	2.3240	0.4860	0.0509	9.5480	0.4627	0.1985	2.3320
lnE	κ	-	-	-	-0.0436	0.2582	-0.1690	-	-	-	-0.0394	0.0091	-4.3340
lnPROV	ν	-	-	-	0.0009	0.0101	0.0850	0.0005	0.0035	0.1400	-	-	-
T	τ	-0.0388	0.0559	-0.6950	0.0092	0.0334	0.2740	-0.0387	0.0566	-0.6840	-0.0430	0.2496	-0.1720
TT	τ^2	0.0064	0.0074	0.8760	-0.0393	0.0092	-4.2600	0.0064	0.0075	0.8580	0.0091	0.0323	0.2820
lnQ1 lnQ1	δ_{11}	0.2918	0.1993	1.4640	0.2819	0.5141	0.5480	0.2907	0.2208	1.3170	0.2831	0.5142	0.5510
lnQ1 lnQ2	δ_{12}	-0.1511	0.0148	-10.2170	-0.1600	0.0507	-3.1560	-0.1508	0.0145	-10.3640	-0.1597	0.0509	-3.1410
lnQ2 lnQ2	δ_{22}	-0.0461	0.1208	-0.3820	-0.0139	0.5871	-0.0240	-0.0464	0.1188	-0.3910	-0.0140	0.5783	-0.0240
lnP1 lnP1	γ_{11}	0.2417	0.0672	3.5980	0.2548	0.2797	0.9110	0.2422	0.0677	3.5790	0.2543	0.2790	0.9110
lnP1 lnP2	γ_{12}	-0.2052	0.0377	-5.4490	-0.2102	0.1807	-1.1630	-0.2048	0.0376	-5.4490	-0.2103	0.1775	-1.1840
lnP1 lnQ1	ρ_{11}	0.1147	0.0364	3.1540	0.1115	0.1309	0.8520	0.1153	0.0383	3.0110	0.1120	0.1289	0.8690
lnP1 lnQ2	ρ_{12}	-0.1106	0.0382	-2.8910	-0.1026	0.1343	-0.7640	-0.1108	0.0390	-2.8430	-0.1032	0.1318	-0.7830
lnP2 lnQ1	ρ_{21}	0.0105	0.0242	0.4320	0.0117	0.0904	0.1300	0.0094	0.0255	0.3680	0.0121	0.0902	0.1350
cos(z1)	a_1	-0.7000	0.9582	-0.7310	-0.6102	2.3965	-0.2550	-0.6977	1.0608	-0.6580	-0.6161	2.3978	-0.2570
sin(z1)	b_1	-0.0063	0.2757	-0.0230	0.0033	0.8169	0.0040	-0.0047	0.3051	-0.0150	0.0030	0.8107	0.0040
cos(z2)	a_2	0.6572	0.3932	1.6720	0.5808	1.8585	0.3130	0.6573	0.3903	1.6840	0.5798	1.8280	0.3170
sin(z2)	b_2	0.1052	0.1138	0.9240	0.0561	0.5664	0.0990	0.1047	0.1127	0.9290	0.0567	0.5481	0.1040
cos(z1+z1)	a_{11}	-0.1416	0.1211	-1.1690	-0.1194	0.3229	-0.3700	-0.1419	0.1324	-1.0720	-0.1205	0.3241	-0.3720
sin(z1+z1)	b_{11}	0.0584	0.0842	0.6930	0.0599	0.2294	0.2610	0.0591	0.0918	0.6440	0.0604	0.2299	0.2630
cos(z1+z2)	a_{12}	-0.0205	0.0327	-0.6260	-0.0186	0.1137	-0.1640	-0.0209	0.0321	-0.6490	-0.0195	0.1146	-0.1700
sin(z1+z2)	b_{12}	-0.0832	0.0406	-2.0500	-0.0923	0.1770	-0.5220	-0.0827	0.0415	-1.9920	-0.0928	0.1779	-0.5210
cos(z2+z2)	a_{22}	0.1243	0.0563	2.2060	0.1162	0.2566	0.4530	0.1243	0.0550	2.2600	0.1164	0.2542	0.4580
sin(z2+z2)	b_{22}	0.0602	0.0361	1.6690	0.0608	0.1363	0.4460	0.0592	0.0354	1.6720	0.0612	0.1356	0.4510
	λ	-	-	-	-	-	-	-	-	-	-	-	-
	σ	-	-	-	-	-	-	-	-	-	-	-	-
	θ	16.4384	8.8535	1.8570	18.0046	35.5920	0.5060	16.4953	9.1904	1.7950	17.8491	35.2380	0.5070
	σ_v	0.0070	0.0004	15.6420	0.0317	0.0102	3.1180	0.0068	0.0005	15.0210	0.0319	0.0102	3.1230
lnP2	β_2	0.5136			0.5382			0.5140			0.5373		
lnP2 lnP2	γ_{22}	0.2052			0.2102			0.2048			0.2103		
lnP2 lnQ2	ρ_{22}	-0.0105			-0.0117			-0.0094			-0.0121		
Variance components	$\sigma^2(v)$	0.00005			0.00100			0.00005			0.00102		
	$\sigma^2(u)$	0.00370			0.00308			0.00368			0.00314		
Log Likelihood function		455.2			198.7			456.1			197.8		
Function converged at iteration		5			11			5			10		
R^2		0.996			0.996			0.996			0.996		

Table 7.1 b

Maximum likelihood parameter estimation for the cost function using the exponential model

Variable	Parameter	Traditional model			Preferred model Eq (7.1)			No Equity			No Provisions		
		Coefficient	Std Error	*T*-value	Coefficient	Std. Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value
Constant	α_0	-5.6237	17.1930	-0.3270	-4.8586	53.6430	-0.0910	-5.6865	18.0310	-0.3150	-4.7452	52.6630	-0.0900
lnQ1	α_1	-0.8015	2.5010	-0.3200	-0.5482	6.2148	-0.0880	-0.7934	2.7731	-0.2860	-0.5632	6.2024	-0.0910
lnQ2	α_2	2.7054	1.5721	1.7210	2.4012	7.5186	0.3190	2.7082	1.5516	1.7450	2.3968	7.4175	0.3230
lnP1	β_1	0.4864	0.0511	9.5140	0.4618	0.1987	2.3240	0.4860	0.0509	9.5480	0.4627	0.1985	2.3320
lnE	κ	-	-	-	-0.0436	0.2582	-0.1690	-	-	-	-0.0394	0.0091	-4.3340
lnPROV	ν	-	-	-	0.0009	0.0101	0.0850	0.0005	0.0035	0.1400	-	-	-
T	τ	-0.0388	0.0559	-0.6950	0.0092	0.0334	0.2740	-0.0387	0.0566	-0.6840	-0.0430	0.2496	-0.1720
TT	τ^2	0.0064	0.0074	0.8760	-0.0393	0.0092	-4.2600	0.0064	0.0075	0.8580	0.0091	0.0323	0.2820
lnQ1 lnQ1	δ_{11}	0.2918	0.1993	1.4640	0.2819	0.5141	0.5480	0.2907	0.2208	1.3170	0.2831	0.5142	0.5510
lnQ1 lnQ2	δ_{12}	-0.1511	0.0148	-10.2170	-0.1600	0.0507	-3.1560	-0.1508	0.0145	-10.3640	-0.1597	0.0509	-3.1410
lnQ2 lnQ2	δ_{22}	-0.0461	0.1208	-0.3820	-0.0139	0.5871	-0.0240	-0.0464	0.1188	-0.3910	-0.0140	0.5783	-0.0240
lnP1 lnP1	γ_{11}	0.2417	0.0672	3.5980	0.2548	0.2797	0.9110	0.2422	0.0677	3.5790	0.2543	0.2790	0.9110
lnP1 lnP2	γ_{12}	-0.2052	0.0377	-5.4490	-0.2102	0.1807	-1.1630	-0.2048	0.0376	-5.4490	-0.2103	0.1775	-1.1840
lnP1 lnQ1	ρ_{11}	0.1147	0.0364	3.1540	0.1115	0.1309	0.8520	0.1153	0.0383	3.0110	0.1120	0.1289	0.8690
lnP1 lnQ2	ρ_{12}	-0.1106	0.0382	-2.8910	-0.1026	0.1343	-0.7640	-0.1108	0.0390	-2.8430	-0.1032	0.1318	-0.7830
lnP2 lnQ1	ρ_{21}	0.0105	0.0242	0.4320	0.0117	0.0904	0.1300	0.0094	0.0255	0.3680	0.0121	0.0902	0.1350
cos(z1)	a_1	-0.7000	0.9582	-0.7310	-0.6102	2.3965	-0.2550	-0.6977	1.0608	-0.6580	-0.6161	2.3978	-0.2570
sin(z1)	b_1	-0.0063	0.2757	-0.0230	0.0033	0.8169	0.0040	-0.0047	0.3051	-0.0150	0.0030	0.8107	0.0040
cos(z2)	a_2	0.6572	0.3932	1.6720	0.5808	1.8585	0.3130	0.6573	0.3903	1.6840	0.5798	1.8280	0.3170
sin(z2)	b_2	0.1052	0.1138	0.9240	0.0561	0.5664	0.0990	0.1047	0.1127	0.9290	0.0567	0.5481	0.1040
cos(z1+z1)	a_{11}	-0.1416	0.1211	-1.1690	-0.1194	0.3229	-0.3700	-0.1419	0.1324	-1.0720	-0.1205	0.3241	-0.3720
sin(z1+z1)	b_{11}	0.0584	0.0842	0.6930	0.0599	0.2294	0.2610	0.0591	0.0918	0.6440	0.0604	0.2299	0.2630
cos(z1+z2)	a_{12}	-0.0205	0.0327	-0.6260	-0.0186	0.1137	-0.1640	-0.0209	0.0321	-0.6490	-0.0195	0.1146	-0.1700
sin(z1+z2)	b_{12}	-0.0832	0.0406	-2.0500	-0.0923	0.1770	-0.5220	-0.0827	0.0415	-1.9920	-0.0928	0.1779	-0.5210
cos(z2+z2)	a_{22}	0.1243	0.0563	2.2060	0.1162	0.2566	0.4530	0.1243	0.0550	2.2600	0.1164	0.2542	0.4580
sin(z2+z2)	b_{22}	0.0602	0.0361	1.6690	0.0608	0.1363	0.4460	0.0592	0.0354	1.6720	0.0612	0.1356	0.4510
	λ	-	-	-	-	-	-	-	-	-	-	-	-
	σ	-	-	-	-	-	-	-	-	-	-	-	-
	θ	16.4384	8.8535	1.8570	18.0046	35.5920	0.5060	16.4953	9.1904	1.7950	17.8491	35.2380	0.5070
	σ_v	0.0070	0.0004	15.6420	0.0317	0.0102	3.1180	0.0068	0.0005	15.0210	0.0319	0.0102	3.1230
lnP2	β_2	0.5136			0.5382			0.5140			0.5373		
lnP2 lnP2	γ_{22}	0.2052			0.2102			0.2048			0.2103		
lnP2 lnQ2	ρ_{22}	-0.0105			-0.0117			-0.0094			-0.0121		
Variance components	$\sigma^2(v)$	0.00005			0.00100			0.00005			0.00102		
	$\sigma^2(u)$	0.00370			0.00308			0.00368			0.00314		
Log Likelihood function		455.2			198.7			456.1			197.8		
Function converged at iteration		5			11			5			10		
R^2		0.996			0.996			0.996			0.996		

Table 7.1 c

Maximum likelihood parameter estimation for the standard profit function using the half-normal model

		Traditional model			Preferred model Eq (7 1)			No Equity			No Provisions		
Variable	Parameter	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std. Error	*T*-value
Constant	α_0	10 6471	0 8133	13 0910	9.4045	0 8779	10 7120	10.3903	0 8179	12.7040	9.7141	0 8797	11 0420
lnG1	α_1	-1 9503	0.6740	-2.8940	-2.1203	0 6741	-3 1460	-2 0625	0 6749	-3 0560	-1 9997	0 6812	-2.9350
lnG2	α_2	0 1650	0 3629	0.4550	0.0711	0.3802	0 1870	0.1249	0 3662	0 3410	0 1215	0.3757	0.3230
lnP1	β_1	1 3330	0 0903	14.7610	1.3433	0 0965	13 9220	1.3263	0 0888	14 9400	1 3511	0 0993	13.6070
lnE	κ	-	-	-	0.0601	0 0064	9 4450	-	-	-	0 0585	0.0059	9.9780
lnPROV	ν	-	-	-	0.0117	0.0044	2 6860	0.0108	0 0038	2.7960	-	-	-
T	τ	0 1243	0.0662	1.8770	0.1202	0.0646	1 8600	0.1173	0 0644	1.8210	0 1276	0 0659	1 9350
TT	τ^2	-0 0158	0.0089	-1.7800	-0.0157	0.0086	-1 8250	-0 0150	0 0086	-1.7440	-0 0164	0 0088	-1.8700
lnG1 lnG1	δ_{11}	-0 7008	0 3586	-1.9550	-0.7884	0 3827	-2 0600	-0.7193	0 3805	-1.8910	-0 7702	0.3771	-2.0430
lnG1 lnG2	δ_{12}	0 0102	0.0403	0 2530	0.0073	0 0421	0 1740	0 0064	0 0407	0.1570	0 0134	0.0414	0.3230
lnG2 lnG2	δ_{22}	0 0971	0 0366	2 6510	0.0849	0.0370	2 2980	0 0935	0 0367	2 5450	0 0891	0 0364	2.4470
lnP1 lnP1	γ_{11}	0 4361	0.1215	3.5900	0.4413	0.1249	3 5340	0 4492	0 1249	3 5970	0 4285	0 1230	3.4850
lnP1 lnP2	γ_{12}	0 1597	0 0859	1 8590	0.1714	0 0829	2 0680	0.1615	0 0832	1 9410	0 1685	0.0854	1.9730
lnP1 lnG1	ρ_{11}	-0 2994	0.1375	-2 1780	-0.2887	0.1435	-2 0120	-0.3139	0.1414	-2.2200	-0 2736	0.1415	-1.9330
lnP1 lnG2	ρ_{12}	-0 0139	0.0705	-0 1970	-0.0267	0.0724	-0 3690	-0 0201	0 0704	-0.2850	-0 0202	0.0720	-0 2810
lnP2 lnG1	ρ_{21}	0 0983	0.0576	1 7080	0.0946	0 0591	1 6020	0 0961	0 0577	1 6660	0 0964	0.0586	1 6450
cos(z1)	a_1	0 6841	0 0682	10.0280	0.6418	0.0783	8 1940	0 6707	0 0755	8 8870	0 6560	0 0688	9.5370
sin(z1)	b_1	-0 6949	0.0675	-10.3010	-0.6397	0.0755	-8 4780	-0 6815	0 0685	-9.9550	-0 6553	0 0726	-9 0220
cos(z2)	a_2	-0 0275	0 0477	-0.5760	-0.0237	0.0552	-0 4290	-0 0186	0 0486	-0.3830	-0.0323	0 0537	-0 6010
sin(z2)	b_2	-0 8446	0 0641	-13.1750	-0.7735	0 0654	-11 8280	-0.8414	0 0647	-13 0080	-0.7761	0.0640	-12 1180
cos(z1+z1)	a_{11}	0 0566	0 0483	1.1710	0.0530	0 0506	1 0490	0.0557	0 0495	1.1240	0 0530	0.0485	1.0930
sin(z1+z1)	b_{11}	-0 1611	0.0360	-4 4740	-0.1513	0 0390	-3 8760	-0.1607	0 0375	-4.2810	-0 1525	0.0366	-4.1640
cos(z1+z2)	a_{12}	0 0117	0.0585	0.2000	0.0007	0 0621	0 0110	0 0072	0 0600	0.1200	0 0051	0.0598	0 0850
sin(z1+z2)	b_{12}	-0 2043	0 0674	-3 0310	-0.2071	0.0741	-2 7940	-0.2117	0 0690	-3.0670	-0 1998	0 0713	-2 8040
cos(z2+z2)	a_{22}	0 0314	0 0392	0 8010	0 0455	0 0399	1 1410	0 0429	0 0395	1 0850	0 0341	0 0397	0 8580
sin(z2+z2)	b_{22}	-0 1402	0 0375	-3 7440	-0 1201	0 0421	-2 8530	-0 1357	0 0388	-3 4950	-0 1251	0 0401	-3 1200
	λ	7 2100	2 2215	3 2460	5.9792	1.6861	3 5460	7 0114	2 1344	3 2850	6 2850	1 8185	3.4560
	σ	0 0223	0 0011	20 0310	0 0213	0 0011	19 3230	0 0220	0 0011	19 7520	0 0216	0 0011	19 4130
	θ	-	-	-	-	-	-	-	-	-	-	-	-
	σ_v	-	-	-	-	-	-	-	-	-	-	-	-
lnP2	β_2	-0 3330			-0 3433			-0 3263			-0 3511		
lnP2 lnP2	γ_{22}	-0 1597			-0.1714			-0 1615			-0 1685		
lnP2 lnQ2	ρ_{22}	-0 0983			-0 0946			-0 0961			-0 0964		
Variance components	$\sigma^2(v)$	0 02227			0 02129			0 02196			0 02165		
	$\sigma^2(u)$	0 16058			0 12732			0 15398			0 13604		
Log Likelihood function		116 6			136.6			121 9			130 2		
Function converged at iteration		29			29			30			29		
R^2		0 959			0 963			0 960			0 962		

Table 7.1 e

Maximum likelihood parameter estimation for the alternative profit function using the half-normal model

Variable	Parameter	Traditional model			Preferred model Eq (7 1)			No Equity			No Provisions		
		Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value
Constant	α_0	12.6909	37.4410	0.3390	16.5441	34.8460	0.4750	12.7897	38.0130	0.3360	16.3033	35.9250	0.4540
lnQ1	α_1	1.5433	3.7596	0.4100	1.3510	3.9883	0.3390	1.6818	4.5546	0.3690	1.2154	3.4139	0.3560
lnQ2	α_2	-1.9204	4.3169	-0.4450	-2.4241	3.5370	-0.6850	-2.0715	3.7442	-0.5530	-2.2559	4.1893	-0.5390
lnP1	β_1	1.0928	0.0896	12.1970	1.1120	0.0896	12.4100	1.0891	0.0892	12.2060	1.1151	0.0901	12.3740
lnE	κ	-	-	-	0.0656	0.0066	9.8890	-	-	-	0.0635	0.0071	8.9560
lnPROV	ν	-	-	-	0.0165	0.0038	4.3460	0.0158	0.0038	4.1330	-	-	-
T	τ	0.0997	0.0702	1.4190	0.0970	0.0692	1.4010	0.0923	0.0710	1.3000	0.1044	0.0688	1.5170
TT	τ^2	-0.0133	0.0094	-1.4100	-0.0134	0.0093	-1.4500	-0.0126	0.0095	-1.3280	-0.0141	0.0092	-1.5280
lnQ1 lnQ1	δ_{11}	-0.0810	0.3060	-0.2650	-0.0928	0.3205	-0.2900	-0.0931	0.3681	-0.2530	-0.0783	0.2776	-0.2820
lnQ1 lnQ2	δ_{12}	-0.0506	0.0313	-1.6190	-0.0273	0.0316	-0.8640	-0.0512	0.0325	-1.5730	-0.0285	0.0312	-0.9140
lnQ2 lnQ2	δ_{22}	0.2181	0.3352	0.6510	0.2313	0.2707	0.8550	0.2319	0.2872	0.8080	0.2178	0.3240	0.6720
lnP1 lnP1	γ_{11}	-0.2630	0.1665	-1.5800	-0.2534	0.1681	-1.5070	-0.2579	0.1685	-1.5300	-0.2555	0.1654	-1.5450
lnP1 lnP2	γ_{12}	0.1822	0.0746	2.4410	0.1861	0.0702	2.6500	0.1859	0.0707	2.6300	0.1827	0.0747	2.4470
lnP1 lnQ1	ρ_{11}	-0.0157	0.0692	-0.2270	-0.0364	0.0704	-0.5170	-0.0329	0.0700	-0.4700	-0.0169	0.0694	-0.2440
lnP1 lnQ2	ρ_{12}	-0.0711	0.0527	-1.3500	-0.0501	0.0550	-0.9110	-0.0514	0.0541	-0.9490	-0.0714	0.0533	-1.3380
lnP2 lnQ1	ρ_{21}	-0.0689	0.0427	-1.6110	-0.0530	0.0379	-1.3980	-0.0596	0.0376	-1.5840	-0.0623	0.0441	-1.4130
cos(z1)	a_1	1.0803	1.4067	0.7680	1.0522	1.4950	0.7040	1.1486	1.6888	0.6800	0.9759	1.2945	0.7540
sin(z1)	b_1	-0.2264	0.4882	-0.4640	-0.2525	0.4988	-0.5060	-0.1994	0.5814	-0.3430	-0.2772	0.4343	-0.6380
cos(z2)	a_2	-0.5098	1.0619	-0.4800	-0.6445	0.8681	-0.7420	-0.5554	0.9188	-0.6040	-0.5958	1.0187	-0.5850
sin(z2)	b_2	-0.0549	0.3099	-0.1770	-0.0741	0.2746	-0.2700	-0.0686	0.2795	-0.2450	-0.0583	0.3178	-0.1830
cos(z1+z1)	a_{11}	0.0891	0.1588	0.5610	0.0782	0.1781	0.4390	0.1018	0.1899	0.5360	0.0651	0.1549	0.4210
sin(z1+z1)	b_{11}	-0.0888	0.1106	-0.8030	-0.0805	0.1149	-0.7000	-0.0950	0.1295	-0.7330	-0.0742	0.1021	-0.7260
cos(z1+z2)	a_{12}	-0.0348	0.0679	-0.5130	-0.0259	0.0673	-0.3850	-0.0342	0.0669	-0.5110	-0.0294	0.0660	-0.4450
sin(z1+z2)	b_{12}	-0.0432	0.0913	-0.4730	-0.0993	0.0969	-1.0250	-0.0503	0.0987	-0.5100	-0.0914	0.0899	-1.0170
cos(z2+z2)	a_{22}	-0.0303	0.1303	-0.2330	-0.0279	0.1102	-0.2530	-0.0248	0.1150	-0.2150	-0.0325	0.1257	-0.2580
sin(z2+z2)	b_{22}	0.0346	0.0795	0.4350	0.0545	0.0674	0.8080	0.0369	0.0708	0.5210	0.0518	0.0805	0.6430
	λ	7.0035	2.0261	3.4570	5.7688	1.8094	3.1880	6.3914	2.0186	3.1660	6.3306	1.8600	3.4040
	σ	0.0276	0.0015	18.9990	0.0261	0.0014	18.5730	0.0270	0.0014	18.7780	0.0267	0.0014	18.6840
	θ	-	-	-	-	-	-	-	-	-	-	-	-
	σ_v	-	-	-	-	-	-	-	-	-	-	-	-
lnP2	β_2	-0.0928			-0.1120			-0.0891			-0.1151		
lnP2 lnP2	γ_{22}	-0.1822			-0.1861			-0.1859			-0.1827		
lnP2 lnQ2	ρ_{22}	0.0689			0.0530			0.0596			0.0623		
Variance components	$\sigma^2(v)$	0.02756			0.0261			0.0270			0.0267		
	$\sigma^2(u)$	0.19301			0.1505			0.1727			0.1690		
Log Likelihood function		58.4			81.6			67.7			71.1		
Function converged at iteration		46			43			42			39		
R^2		0.948			0.956			0.951			0.953		

Table 7.1 f

Maximum likelihood parameter estimation for the alternative profit function using the exponential model

		Traditional model			Preferred model Eq (7 1)			No Equity			No Provisions		
Variable	Parameter	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value	Coefficient	Std Error	*T*-value
Constant	α_0	12.7064	37.4480	0.3390	16.5232	34.8970	0.4730	12.6374	37.5340	0.3370	16.3382	35.7840	0.4570
lnQ1	α_1	1.5419	3.7626	0.4100	1.3493	3.9995	0.3370	1.6468	4.4460	0.3700	1.1927	3.3758	0.3530
lnQ2	α_2	-1.9216	4.3167	-0.4450	-2.4193	3.5387	-0.6840	-2.0148	3.7367	-0.5390	-2.2407	4.1852	-0.5350
lnP1	β_1	1.0928	0.0896	12.1980	1.1120	0.0896	12.4070	1.0897	0.0893	12.1990	1.1152	0.0901	12.3720
lnE	κ	-	-	-	0.0656	0.0066	9.8710	-	-	-	0.0635	0.0071	8.9550
lnPROV	ν	-	-	-	0.0165	0.0038	4.3370	0.0158	0.0038	4.1330	-	-	-
T	τ	0.0997	0.0702	1.4200	0.0969	0.0692	1.4000	0.0923	0.0710	1.3000	0.1043	0.0688	1.5160
TT	τ^2	-0.0133	0.0094	-1.4110	-0.0134	0.0093	-1.4490	-0.0126	0.0095	-1.3280	-0.0141	0.0092	-1.5270
lnQ1 lnQ1	δ_{11}	-0.0809	0.3062	-0.2640	-0.0926	0.3213	-0.2880	-0.0904	0.3597	-0.2510	-0.0766	0.2748	-0.2790
lnQ1 lnQ2	δ_{12}	-0.0506	0.0313	-1.6180	-0.0273	0.0316	-0.8630	-0.0511	0.0324	-1.5750	-0.0284	0.0312	-0.9120
lnQ2 lnQ2	δ_{22}	0.2181	0.3352	0.6510	0.2309	0.2708	0.8530	0.2275	0.2864	0.7940	0.2166	0.3236	0.6690
lnP1 lnP1	γ_{11}	-0.2628	0.1665	-1.5790	-0.2533	0.1681	-1.5070	-0.2579	0.1685	-1.5300	-0.2561	0.1654	-1.5480
lnP1 lnP2	γ_{12}	0.1822	0.0746	2.4410	0.1861	0.0702	2.6500	0.1863	0.0707	2.6340	0.1827	0.0747	2.4460
lnP1 lnQ1	ρ_{11}	-0.0156	0.0692	-0.2260	-0.0364	0.0705	-0.5160	-0.0330	0.0700	-0.4720	-0.0170	0.0694	-0.2440
lnP1 lnQ2	ρ_{12}	-0.0712	0.0527	-1.3510	-0.0501	0.0550	-0.9110	-0.0512	0.0541	-0.9460	-0.0715	0.0533	-1.3400
lnP2 lnQ1	ρ_{21}	-0.0689	0.0427	-1.6120	-0.0530	0.0379	-1.3970	-0.0596	0.0376	-1.5860	-0.0624	0.0441	-1.4150
cos(z1)	a_1	1.0798	1.4077	0.7670	1.0515	1.4987	0.7020	1.1354	1.6530	0.6870	0.9672	1.2826	0.7540
sin(z1)	b_1	-0.2262	0.4886	-0.4630	-0.2521	0.5001	-0.5040	-0.1994	0.5655	-0.3530	-0.2767	0.4285	-0.6460
cos(z2)	a_2	-0.5100	1.0618	-0.4800	-0.6432	0.8685	-0.7410	-0.5406	0.9165	-0.5900	-0.5916	1.0176	-0.5810
sin(z2)	b_2	-0.0548	0.3099	-0.1770	-0.0743	0.2747	-0.2700	-0.0688	0.2790	-0.2470	-0.0577	0.3176	-0.1820
cos(z1+z1)	a_{11}	0.0889	0.1589	0.5600	0.0780	0.1785	0.4370	0.1002	0.1865	0.5370	0.0640	0.1538	0.4160
sin(z1+z1)	b_{11}	-0.0887	0.1107	-0.8020	-0.0804	0.1152	-0.6980	-0.0950	0.1262	-0.7530	-0.0740	0.1010	-0.7330
cos(z1+z2)	a_{12}	-0.0348	0.0679	-0.5120	-0.0258	0.0673	-0.3830	-0.0340	0.0668	-0.5100	-0.0293	0.0661	-0.4440
sin(z1+z2)	b_{12}	-0.0432	0.0913	-0.4730	-0.0994	0.0969	-1.0250	-0.0503	0.0986	-0.5100	-0.0914	0.0900	-1.0160
cos(z2+z2)	a_{22}	-0.0304	0.1303	-0.2330	-0.0276	0.1102	-0.2510	-0.0226	0.1148	-0.1970	-0.0318	0.1256	-0.2530
sin(z2+z2)	b_{22}	0.0346	0.0795	0.4350	0.0545	0.0674	0.8080	0.0368	0.0707	0.5210	0.0519	0.0805	0.6450
	λ	-	-	-	-	-	-	-	-	-	-	-	-
	σ	-	-	-	-	-	-	-	-	-	-	-	-
	θ	7.0070	2.0275	3.4560	5.7554	1.8033	3.1920	6.3765	2.0097	3.1730	6.3397	1.8641	3.4010
	σ_v	0.0276	0.0015	19.0010	0.0261	0.0014	18.5640	0.0270	0.0014	18.7910	0.0267	0.0014	18.6860
lnP2	β_2	-0.0928			-0.1120			-0.0897			-0.1152		
lnP2 lnP2	γ_{22}	-0.1822			-0.1861			-0.1863			-0.1827		
lnP2 lnQ2	ρ_{22}	0.0689			0.0530			0.0596			0.0624		
Variance components	$\sigma^2(v)$	0.00076			0.0007			0.0007			0.0007		
	$\sigma^2(u)$	0.02037			0.0302			0.0246			0.0249		
Log Likelihood function		58.4			81.6			67.7			71.1		
Function converged at iteration		50			54			44			50		
R^2		0.948			0.956			0.951			0.953		

In addition, the negative coefficients on the price of loans (G_l) in relation to standard profits clearly indicate that an increase in the price of loans would decrease the level of profits [see Tables (7.1c and 7.1d)]. At first glance, this result might look odd since profits may be expected to increase as prices rise. However, because the standard profit function takes the price of output as given and leaves the quantity of output to move freely, this means that at higher prices banks face a lower demand for output; hence, at this given higher price of output, banks' profits may decrease. Thus, this negative relationship between loan prices and profits could indicate that the quantity of loans demanded (rather than the price of loans) is more influential in driving GCC banking profits. An alternative explanation suggests that an increase in loan prices may result in a reduction in the quantity of loans demanded, reducing profits by a greater proportion than would be added by any loan price increases. The main finding is that an increase in prices results in lower levels of standard profits.

The results also show that the risk variable (*E*) has consistent effects on both cost and profit functions. That is, since the estimated coefficient of equity is negative in the cost model, this might inform us that low levels of financial capital could contribute to increasing costs because of reliance on borrowed funds, while high levels of capital indicate the opposite.[8] Moreover, the positive coefficients of the (*E*) variable in the profit models indicate that as bank financial capital increases, banks secure greater profits as risk exposure lessens. Besides, the positive relationship between financial capital and profits may derive from the fact that profits add to financial capital in the form of retained profits, given that profits are not allocated as dividends. Thus, the stronger the financial capital base, the greater is banks' access to sources of internal finance, and therefore their opportunity of generating profits.

The coefficients reported in the tables also reveal some other interesting relationships. For example, Tables (7.1a and 7.1b) also show that the loan quality proxy (*PROV*) has the expected relationship with costs (since bad loans increase the cost burden of banks),

[8] High capital also means less risk exposure, which may place a low burden on the cost function compared to the case when capital is low and risk is high.

and Tables (7.1c to 7.1f) indicate that *PROV* is positively related to profits. This is probably because more profitable banks have the ability to make greater provisions. (However, one could also argue that one may expect an inverse relationship as banks tend to be more profitable when provisions fall.)

It may also be noted from Tables (7.1a to 7.1f) that the cost, standard profit, and alternative profit functions fit the data reasonably well. The adjusted R^2 reported over the six years for all model specifications ranges from 94.8 to 99.6 per cent. This means that the explanatory variables explain most of the variation in the dependent variables.

Tables (7.1a to 7.1f) also present both inefficiency and random error variances, denoted as $\sigma^2(u)$ and $\sigma^2(\upsilon)$ respectively. Among all inefficiency concepts, the lowest inefficiency variance as a ratio in the total error term variance amounted to around 62 per cent (for the cost function estimated using the half-normal distribution). For the standard profit and alternative profit functions estimated using the half-normal distribution, the inefficiency term variance ratio accounted for around 86 per cent and 88 per cent of total variance respectively. On the other hand, compared to the half-normal estimation, the exponential model estimated for the three efficiency concepts reports higher inefficiency variances: around 75 to 98 per cent for the cost function, and around 98 per cent for both the standard and alternative profit functions. These results suggest that, for both the half normal and exponential estimations across different efficiency concepts and specifications, the majority of the total variances of the stochastic error term ε is accounted for by the variances in the inefficiency component u, rather than the variances in the random error υ. This suggests that the deviations from the best practice bank's cost and profit functions have much more to do with managerial factors (X-inefficiency) than with luck and other factors that are incorporated in the random error term.

In sum, this section discusses the parameter estimation and finds that most coefficients have consistent signs. Moreover, the models show relatively strong explanatory power. Also, the inefficiency term is found to be the major factor explaining the distance of banks from the best practice cost and profit frontiers. The following section discusses various tests undertaken on the estimated models.

7.3 Structural tests

In this section we verify the validity of our preferred model shown in Eq. (7.1) in order to check how reliable this model is in estimating the inefficiency of the GCC banking industry. Thus, the section discusses the results of several statistical tests aimed at evaluating the robustness of the cost and profit models that have been estimated. This includes tests on data poolability, heteroskedasticity, random effects, test of the translog versus Fourier specification, as well as other tests that check the validity of the existence of the risk, loan quality, and technical progress variables in the preferred model.

Undertaking the estimation of the model using pooled time series cross-section data usually requires a test to check if it is permissible to pool both dimensions of the data, an issue that arises when one is using panel data (Baltagi, 2001). The checking of the data poolability is mainly in order to detect whether or not the parameters of the model are the same (or stable) across time and bank observations, especially when data are pooled. This can be tested using the poolability test, which is an application of a generalized Chow's (1960) test. The residual sum squared of the restricted model, which is obtained from the OLS pooled model estimated for Eq. (7.1), and the total value of the unrestricted residual sum of squares, which is obtained from individual OLS regressions of 93 banks across each year of the study period, are calculated to carry out Chow's poolability test. As shown in Tables (7.2a to 7.2c), the test which is undertaken for the cost, standard profit, and alternative profit functions yields observed F-statistics of 1.05, 0.96, and 0.63 respectively, which are distributed as $F(120, 414)$. Under the null hypothesis: $H_0 : \beta_t = \beta$ for $t = 1, \ldots, T$, the test does not reject poolability at the 1 per cent level of significance. Therefore, our poolability test suggests that pooling our data in order to estimate Eq. (7.1) is valid, which also implies that the estimated model parameters are stable over time and bank observation.

As our data sample has a 'panel' dimension with a large cross-section (93 banks estimated over 6 years), the inclusion of banks of different sizes in the sample may give rise to concern of heteroskedasticity in the error term. We apply the Goldfeld-Quandt test (1965) to check whether or not the heteroskedasticity problem is present in the model. If not, then the test indicates that disturbance variances are homoskedastic, or, in other words, constant across observations. For the cost, standard profit, and alternative profit functions, Tables (7.2a to 7.2c) show that because the calculated test values are less than the critical value, the Goldfeld-Quandt test does not reject the null hypothesis of homoskedasticity at 1 per cent level of significance.

Table 7.2 (a to c) Structural tests for model (7.1)

Table (7.2a) Structural tests of the cost function Eq. (7.1)

Test Performed		Test Statistics	Degrees of Freedom		Critical value		H0 Hypothesis	Decision
Poolability	Chow's test	1.0592	((T-1)K)= ((N-K)T))=	120 414	F(.01,120,414)=	1.383	H0.Bt=B data is poolable or betas are stable	Not-rejected
Heteroskedasticity	Goldfeld-Quant test	0 8299	n1= n2=	253 253	F(a,N1-K,N2-K)=(.01,279-26,279-26)=	1.341	H0 Disturbances of the variances are constant	Not-rejected
Heteroskedasticity	LM test	18.391	k=	1	X2(0 01,1)=	6 635	H0 Disturbances of the variances are the constant	Rejected
Random effects	LM test	358 09	k=	1	X2(0 01,1)	6 635	H0 No individual effects	Rejected
Translog Form	F-test Chi-squared test	3 397 33 993	K=	10	F(.01,10,536) X2(10)=	2 354 23 209	H0 Translogn form is valid or Fourier terms =0	Rejected Rejected
DROP PROV,E,T,TT	F-test	13.327	k=	4	F(01,4,532)=	3.355	H0 estimated model is better when these are dropped	Rejected
DROP PROV,E	F-test	25 059	k=	2	F(.01,2,532)=	4 645	H0 estimated model is better when these are dropped	Rejected
DROP PROV,T,TT	F-test	1 4102	k=	3	F(01,3,532)=	3 819	H0 estimated model is better when these are dropped	Not-rejected
DROP E,T,TT	F-test	17.729	k=	3	F(.01,3,532)=	3 819	H0 estimated model is better when these are dropped	Rejected
DROP RPOV	F-test	0 1815	k=	1	F(01,1,532)=	6 683	H0 estimated model is better when these are dropped	Not-rejected
DROP E	F-test	50 024	k=	1	F(.01,1,532)=	6 683	H0 estimated model is better when these are dropped	Rejected
DROP T,TT	F-test	1.9974	k=	2	F(.01,2,532)=	4 645	H0 estimated model is better when these are dropped	Not-rejected

Table (7.2b) Structural tests of the standard profit function Eq. (7.1)

Test Performed		Test Statistics	Degrees of Freedom		Critical value		H0 Hypothesis	Decision
Poolability	Chow's test	0 9676	((T-1)K)= ((N-K)T))=	120 414	F(.01,120,414)=	1.383	H0 Bt=B data is poolable or betas are stable	Not-rejected
Heteroskedasticity	Goldfeld-Quant test	0 7705	n1= n2=	253 253	F(a,N1-K,N2-K)=(01,279-26,279-26)=	1 341	H0 Disturbances of the variances are the constant	Not-rejected
Heteroskedasticity	LM test	6.4003	k=	1	X2(0.01,1)=	6 635	H0 Disturbances of the variances are the constant	Not-rejected
Random effects	LM test	400 64	k=	1	X2(0 01,1)	6 635	H0 No individual effects	Rejected
Translog Form	F-test Chi-squared test	661.00 6610.0	K=	10	F(.01,10,536) X2(10)=	2 354 23.209	H0 Translogn form is valid or Fourier terms =0	Rejected Rejected
DROP PROV,E,T,TT	F-test	19 025	k=	4	F(01,4,532)=	3.355	H0 estimated model is better when these are dropped	Rejected
DROP PROV,E	F-test	35 174	k=	2	F(01,2,532)=	4 645	H0 estimated model is better when these are dropped	Rejected
DROP PROV,T,TT	F-test	6.3979	k=	3	F(01,3,532)=	3 819	H0 estimated model is better when these are dropped	Not-rejected
DROP E,T,TT	F-test	20 255	k=	3	F(.01,3,532)=	3 819	H0 estimated model is better when these are dropped	Rejected
DROP RPOV	F-test	14 096	k=	1	F(01,1,532)=	6.683	H0 estimated model is better when these are dropped	Not-rejected
DROP E	F-test	56.348	k=	1	F(.01,1,532)=	6 683	H0 estimated model is better when these are dropped	Rejected
DROP T,TT	F-test	2 0251	k=	2	F(01,2,532)=	4 645	H0 estimated model is better when these are dropped	Not-rejected

Table (7.2c) Structural tests of the alternative profit function Eq. (7.1)

Test Performed		Test Statistics	Degrees of Freedom		Critical value		H0 Hypothesis	Decision
Poolability	Chow's test	0 6303	((T-1)K)= ((N-K)T))=	120 414	F(01,120,414)=	1 383	H0 Bt=B data is poolable or betas are stable	Not-rejected
Heteroskedasticity	Goldfeld-Quant test	0 9854	n1= n2=	253 253	F(a,N1-K,N2-K)=(01,279-26,279-26)=	1 341	H0 Disturbances of the variances are the constant	Not-rejected
Heteroskedasticity	LM test	2 8519	k=	1	X2(0 01,1)=	6 635	H0 Disturbances of the variances are the constant	Not-rejected
Random effects	LM test	375 520	k=	1	X2(1)	6 635	H0 No individual effects	Rejected
Translog Form	F-test Chi-squared test	6.9483 69 4850	K=	10	F(01,10,532) X2(10)=	2 354 23 209	H0 Translogn form is valid or Fourier terms =0	Rejected Rejected
DROP PROV,E,T,TT	F-test	25 5201	k=	4	F(01,4,532)=	3.355	H0 estimated model is better when these are dropped	Rejected
DROP PROV,E	F-test	50 1498	k=	2	F(.01,2,532)=	4 645	H0 estimated model is better when these are dropped	Rejected
DROP PROV,T,TT	F-test	12 7859	k=	3	F(01,3,532)=	3 819	H0 estimated model is better when these are dropped	Not-rejected
DROP E,T,TT	F-test	20 5604	k=	3	F(01,3,532)=	3.819	H0.estimated model is better when these are dropped	Rejected
DROP RPOV	F-test	36 8296	k=	1	F(.01,1,532)=	6 683	H0 estimated model is better when these are dropped	Not-rejected
DROP E	F-test	60 6521	k=	1	F(.01,1,532)=	6.683	H0 estimated model is better when these are dropped	Rejected
DROP T,TT	F-test	0.55111	k=	2	F(.01,2,532)=	4 645	H0 estimated model is better when these are dropped	Not-rejected

As it is widely recommended to conduct more than one test for checking heteroskedasticity, the LM test for dependent-variable heteroskedasticity is another useful test to carry out here.[9 & 10] At the 1 per cent level of significance, Tables (7.2a to 7.2c) show that the LM test does not reject the null hypothesis of homoskedasticity in both standard and alternative profit functions; however, for the cost function the test is rejected. Assuming homoskedasticity in the disturbance term when heteroskedasticity is present would still produce consistent but not inefficient estimates (Baltagi, 2001).

Overall, the heteroskedasticity tests of these models tend to indicate that the estimation may be viewed as free from heteroskedasticity since both the Goldfeld-Quandt test and the LM test, if taken together, suggest that the error term is apparently not positively correlated with any of the explanatory variables. This implies that the various model specifications do not have serious heteroskedasticity problems.

Given that we are estimating models using panel data, it is also important to investigate whether fixed or random effects estimation must be undertaken. A number of studies that estimate translog and Fourier Flexible models suggest that it is not appropriate to work under the framework of the panel fixed effects model since this induces a substantial loss in the degree of freedom, especially when the number of cross-sections is large (see e.g. Lang and Walzel, 1996; Altunbas et al., 2000). However, before dismissing the fixed effects model, it is important to undertake the random effects test since it can, at least, provide information as to whether individual effects are present or not. The test undertaken here to check the existence of random effects is the Lagrange Multiplier test, devised by Breusch and Pagan (1980). Tables (7.2a to 7.2c) show that the LM test rejects the hypothesis of no individual effects at the 1 per cent level of significance, for the cost, standard profit, and alternative profit functions. In this case, the LM test suggests that there is considerable heterogeneity across banks and that the random effects model is the method to be used to control for the effects of the

[9] White's (1980) test of heteroskedasticity is not appropriate for our model since this test causes a loss in the degree of freedom if applied.
[10] See Thomas, 1997, Chapter 10.

differences across bank observations in our sample. Based on the LM test, we conclude that the random effects model is the appropriate panel estimation approach.

With regard to the choice of the functional form, the Fourier Flexible form is tested against the translog model. Using the F-test, Tables (7.2a to 7.2c) show that the hypothesis that the translog model is valid was rejected at the 1 per cent significance level for the cost, standard profit, and alternative profit functions. The results show the superiority of the Fourier Flexible form over the translog model since the presence of the Fourier trigonometric terms in the model is compelling.

Additional tests are also undertaken to check if the exclusion of the risk (*E*) and asset quality (*PROV*) variables, as well as the technical progress variables (*T* and *TT*), has no statistical significant effects on the model specification shown in Eq. (7.1). The F-test evaluated at the 1 per cent level of significance rejects the null hypothesis that these variables have a zero effect on the dependent variables in each efficiency concept function. In other words, the existence of these variables in the model are important for our inefficiency analysis.

Generally, as the structural tests imply, this section concludes that Eq. (7.1) for the cost, standard profit, and alternative profit functions (that have the Fourier Flexible functional form and incorporate banks' asset quality, risk, and time trend variables) are econometrically valid for our efficiency analysis. The inefficiency measures derived from estimating the aforementioned model are discussed in the following section.

7.4 Inefficiency estimates

This section discusses the results of our inefficiency estimates in three subsections. The first subsection tackles the issue of how our inefficiency estimates vary according to different distributional assumptions and model specifications. The second is devoted to an interpretation of the results for the cost, standard profit, and alternative profit measures taken separately. The final subsection examines inefficiency results across different Gulf banking industries and according to bank size and ownership type.

7.4.1 Inefficiency results - Different distributional assumptions and model specifications

The aim of this subsection is to examine how mean inefficiencies differ according to the various distributional assumptions and model specifications used. The need to examine variations of inefficiency across distributions derives from the fact that the greater the similarity of mean inefficiencies is, the more reliable the results are. Moreover, examination of how the mean inefficiency differs across model specifications gives us information on how the exclusion of the risk and quality variables from the preferred specification [Eq. (7.1)] would affect mean inefficiency estimations in the GCC banking industry.

As previously noted, the purpose in using both the half-normal and exponential distribution models is to give more strength and robustness to the results of the efficiency analysis. In section 7.2, we have seen that robustness existed in the parameter estimates, where both half-normal and exponential models yielded similar coefficients for each specification. Generally, this robustness has extended to the mean inefficiency scores as well. In fact, and in accordance with the literature, the mean inefficiency of the half-normal and the exponential distributions does not differ much (see e.g. Greene, 1990; Aigner et al., 1977; Altunbas and Molyneux, 1994). To be specific, we generally notice from Table 7.3 that both the half-normal and exponential models yield very similar mean inefficiency scores for the standard and alternative profit specifications estimated for both distributional assumptions.

Surprisingly, however, the mean cost inefficiency of the half-normal and the exponential models appears to differ since the exponential model reports inefficiency measures around 10 per cent higher than the half-normal model. This difference could be attributed to the calculation of the error term variance components since Tables (7.1a to 7.1f) show that the exponential model reports a contribution of the inefficiency term in the total error term higher than its correspondent of the half-normal model (as pointed out in section 7.2). This being so, one may need to decide which cost inefficiency measures derived for both distribution models reflect the factual cost inefficiency in the GCC banking industry. For this purpose, we need another robustness check in order to avoid the mistake of adopting results that do not reflect the nearest factual inefficiency estimate. Thus, we undertake the distributional-free approach, a method that avoids making any assumptions on the way the error term components are distributed. As mentioned in Chapter 5, the distribution-free approach has, however, the drawback of possible overestimation of inefficiency, specifically, when the assumption that the random errors cancel out over time is violated.[11] To avoid this, truncation of the inefficiency results is used here in order to remove extreme scores that could exist from non-cancelled random errors.

[11] Fortunately, the use of panel data may help in alleviating this drawback since one of the virtues of the panel data technique is that it averages the noise out of the expression for u_i, leaving the error term to account for the inefficiency term plus a term that its average tend to be zero (Greene, 1993).

Table 7.3 The mean inefficiency estimates

Half-normal distribution model	Traditional model	Preferred model Eq. (7.1)	No Equity	No Provisions
Cost function	0.0839	0.0807	0.0837	0.0806
	(0.0469)	(0 0456)	(0.0468)	(0.0457)
Standard profit function	0.3312	0.2904	0.3235	0.3014
	(0.1849)	(0.1694)	(0.1813)	(0.1739)
Alternative profit function	0.3594	0.3138	0.3383	0.3345
	(0.2101)	(0.1865)	(0.1995)	(0.1974)

Exponential distribution model	Traditional model	Preferred model Eq. (7.1)	No Equity	No Provisions
Cost function	0.1804	0.2194	0.1798	0.2202
	(0.0272)	(0.0086)	(0.0274)	(0.0087)
Standard profit function	0.3068	0.3013	0.3056	0.3034
	(0.0909)	(0.0993)	(0.0920)	(0.0979)
Alternative profit function	0.3225	0.3180	0.3218	0.3207
	(0.1008)	(0.1082)	(0.1044)	(0.1055)

Distribution free approach undertaken for the cost function	Traditional model	Preferred model Eq. (7.1)	No Equity	No Provisions
Truncation 5%	0.1262	0.1229	0.1264	0.1222
Truncation 10%	0.0927	0.0854	0.0936	0.0861

Standard deviations are shown in the brackets

We tried 5 per cent and 10 per cent truncations; that is, each bank in the top 5 per cent (or 10 per cent) and in the bottom 5 per cent (or 10 per cent) of the distribution of the error term is assigned with a value of a bank that is just at the highest and the lowest 5 per cent (or 10 per cent) of the error term distribution.

The cost inefficiency measure is then calculated by the given equation[12]

$$INEFF_i = 1 - \exp(\min(\ln u) - \ln u_i) \tag{7.2}$$

where min(ln u) is the minimum residual among all estimated residuals of banks in the sample.

Indeed, a 10 per cent truncation produces about the same mean cost inefficiency as that of the half-normal model, and a 5 per cent truncation yields inefficiency measures 3 percentage points higher than the half-normal mean cost inefficiency (see Table 7.3). Therefore, as is shown in Table 7.3, the results of the distribution-free approach are strongly consistent with the cost inefficiency results estimated using the half-normal model.

As far as model specification is concerned, we notice from Table 7.3 that when estimating over each specification (traditional, preferred, preferred with no equity, and preferred with no provisions specifications), the inefficiency scores as well as the dispersions around inefficiency means tend to be similar.[13] However, the elimination of equity and provisions variables from the preferred model resulted in a slight difference in the inefficiency means. For example, for the traditional model, Table 7.3 shows that the elimination of the *E* and *PROV* variables from the preferred model slightly increased the mean inefficiency for all efficiency concepts estimated by the half-normal and

[12] This equation is drawn from Allen and Rai (1996), DeYoung (1997), and Berger and Mester (1997).

[13] For example, both the half-normal and the exponential models produced a mean inefficiency around the average of 30 per cent for the standard profit function estimated by both the traditional and preferred specifications, and around the average of 32 per cent for alternative profit inefficiency for the underlying specifications.

exponential models, save for cost function estimates derived by the exponential model. Similarly, the individual elimination of E and $PROV$ variables from the preferred model also resulted in a slight increase in the inefficiency levels across all efficiency concepts and distributional assumptions, except for the cost function estimates of the exponential model. This elimination process suggests that the control for risk and quality factors in the inefficiency models removes any over-estimation of inefficiency scores when these two factors are not taken into account. Moreover, the mean inefficiency results show that the exclusion of the E variable from the preferred specification results in higher inefficiency levels than does the exclusion of the $PROV$ variable from the same specification.

7.4.2 *Interpretation of inefficiencies – Cost, standard profit, and alternative profit models*

The mean cost efficiency from the preferred model is about 91 per cent (from the half-normal and distribution-free estimates). In other words, about 9 per cent of costs are wasted on average relative to a best-practice bank. The economic interpretation of the cost inefficiency level is that, given their particular output level and mix, on average, banks need to reduce their production costs by roughly 9 per cent in order to use their inputs as efficiently as possible. Overall, the levels of the mean cost inefficiency are consistent with inefficiency levels found by parametric studies on European, Japanese, and US banking markets. For example, Altunbas et al. (2000), Berger and Mester, (1997), Faur et al. (1993), Ferrier and Lovell, (1990), and Berger (1993) found the average cost inefficiency of commercial banks to range from anywhere between 5 per cent (Altunbas et al., 2002 on European banks) and 40 per cent (Berger, 1993 on US banks). There is a general consensus, however, that cost inefficiency typically ranges between 5 per cent and 15 per cent (see Berger and Humphrey, 1997).

On the profit side, the mean inefficiencies derived from the standard and alternative profit functions are close to each other, given that the mean inefficiency of the alternative profit function is about 3 per cent higher than standard profit mean inefficiency scores. The interpretation of the inefficiencies on the profit side is not so different from the cost side. In both standard and alternative profit functions, the inefficiency results indicate that nearly one third of the profits that could be earned by the best practice bank are lost to inefficiency. The profit level of inefficiency is found to conform to the findings of a number of previous studies that found profit inefficiency to fall in the same range; for example, Lozano (1997) found that the average profit inefficiency of the Spanish depository institutions was 28 per cent. In contrast, profit inefficiency is found to be higher in the US banking sector. So, for instance, Berger and Mester (1997) report profit inefficiency ranging between 46 and 54 per cent. In general, profit inefficiency in US banking is found to be, on average, around 36 per cent (see the review by Berger and DeYoung, 1997).

In addition, the standard deviations of mean inefficiencies suggest a tighter dispersion around the mean in the case of the exponential model compared to dispersions of mean inefficiency scores estimated by the half-normal model. That is, in the half-normal estimation, the measure of dispersion is higher than exponential estimation by .02-.04, .07-.08, and .08-.10 per cent respectively for the cost, standard profit, and alternative profit estimates. For the efficiency concepts, standard deviations report a tighter dispersion in the case of cost inefficiency than profit inefficiencies.

It is worth mentioning that, in accordance with previous profit efficiency studies (for example, those of Berger and Mester, 1997; and Al-Jarrah, 2002), the results in Table 7.3 show that profit inefficiencies are higher than cost inefficiencies. This finding is consistent across distributional assumptions and various model specifications. As mentioned in Chapter 6, the cost inefficiency calculates wastes of resources only on the input side. Profit inefficiency accounts for inefficiencies on both the input and output sides. This generally results in higher inefficiency estimates on the profit side. Furthermore, when banks face higher operating costs that may be reflected in bank product prices, the profit function can also capture this source of inefficiency. In addition, profits are more variable than costs and can be affected more dramatically on account of economic downturns, unforeseen losses, and so on. Given the greater variation in profitability, it is therefore less surprising that inefficiency tends to be much larger compared to cost inefficiency.

In accordance with the literature, it is also observed from Table 7.3 that mean alternative profit inefficiencies are higher than standard profit inefficiencies. The standard profit function takes prices of outputs as given and leaves the output quantities to change freely. In contrast, the alternative profit function allows output prices to move freely and takes output levels as given. This implies that the alternative profit function may report inefficiency levels higher than standard profit inefficiencies because of market power conditions, service quality, and other endogenous or exogenous sources that may affect output prices and profitability. For markets with high levels of concentration, such as in the GCC banking industry, the standard profit function is less able to take into account the ability of banks to exercise market power without much change in output levels, whereas the alternative profit estimates are believed to capture this phenomenon. Moreover, when banks tend to offer services of low quality with low prices relative to the best practice bank, the alternative profit function can capture this source of inefficiency. Given these reasons, alternative profit inefficiency estimates are often likely to be higher than standard profit inefficiency estimates.

In the examination of how mean inefficiency levels change over time, Figure 7.2 shows the pattern of mean inefficiencies for all banks in our sample using different distributional assumptions with the three different efficiency concepts. Generally, the figure shows a similar pattern in inefficiency levels over time with no discernable increase or decrease, although they tend to show a slight decrease in profit inefficiency over 1995-1997 and 1999-2000. The year 1998 witnessed a rise in loan loss problems (mostly due to the effect of a sharp oil prices decrease in 1998, see Chapter 2) resulting in a noticeable increase in profit inefficiency. Overall, however, both cost and profit efficiency seem to be relatively stable over time, indicating that market conditions, such as the competitive environment and regulatory changes, did not much affect industry's cost and profit functions during the second half of the 1990s.

In order to further check whether different approaches yield similar results, we undertake a rank-order correlation analysis. This was among the consistency condition tests suggested by Bauer et al. (1998). The coefficients of the rank-order correlation can also be interpreted as the extent to which the estimates of some variables are placed in the same order. Therefore, a high coefficient of the rank-order correlation suggests that the order of banks in terms of the estimates of the underlying variables is most likely placed in the same way.

We undertake the rank-order correlation test for inefficiency results across specifications, distribution, and efficiency concepts.[14] In Table 7.4 rank-order is measured to check how the order of inefficiency estimated by other used model specifications is similar to the preferred model. Table 7.4 shows that the inefficiency measures estimated by different specifications for each efficiency concept report at least 97 per cent correlation with the mean inefficiency of the preferred model, except for the cost inefficiency estimated by the exponential distribution, which reports a rank-order correlation of no less than 85 per cent.

[14] Rank-order correlation is also known as Spearman's rank correlation and is calculated using the following formula

$$r_s = 1 - 6\left[\frac{\sum d_i^2}{n(n^2-1)}\right],$$

where d_i is the difference in the ranks assigned to two different variables of the ith observation, and n is the number of observations (See Gujarati, 1995, p. 372).

Figure 7.2 Inefficiency in the GCC banking industry over the study period according to the preferred model

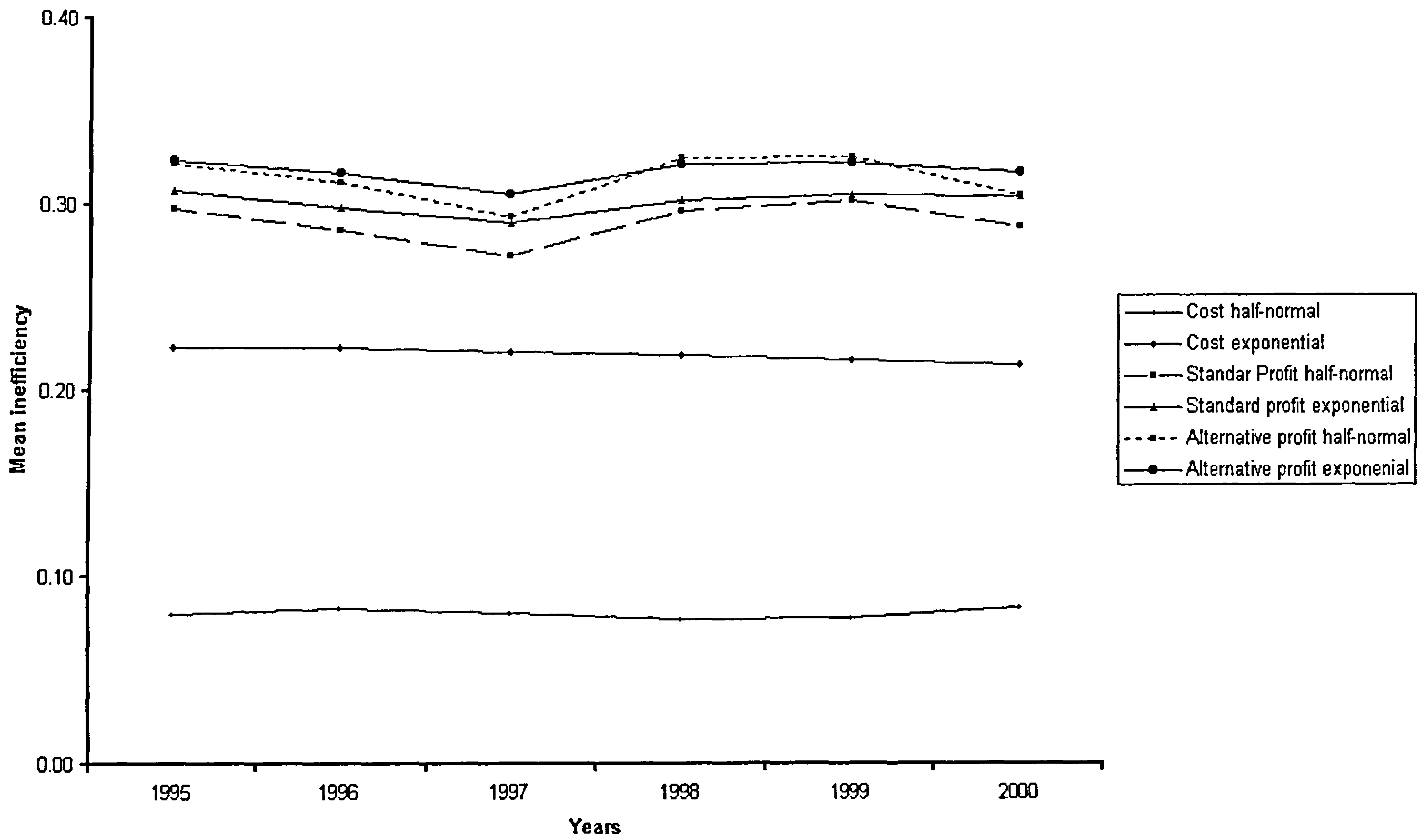


Table 7.4 Rank-order correlations of preferred model (7.1) and traditional, no equity, and no provision specifications

Half-normal model

	Cost ineff.	Std. profit ineff.	Alt. profit Ineff.
	Preferred Model (7.1)	Preferred Model (7.1)	Preferred Model (7.1)
Traditional model	0.9663	0.9687	0.9712
No Equity	0.9685	0.9814	0.9889
No Provisions	0.9966	0.9876	0.9812

Exponential model

	Cost ineff.	Std. profit ineff.	Alt. profit Ineff.
	Preferred Model (7.1)	Preferred Model (7.1)	Preferred Model (7.1)
Traditional model	0.8585	0.9694	0.9712
No Equity	0.8602	0.9820	0.9888
No Provisions	0.9995	0.9880	0.9812

Table 7.5 shows rank correlation coefficient, using different distributional assumptions. The table indicates that both half-normal and exponential distributions yield high rank correlation coefficients, of about 99 per cent, for the profit inefficiency estimates. Moreover, the cost inefficiency estimates show that both half-normal and exponential models also yield quite high rank-order coefficients, of 70 per cent.

Table 7.5 Rank correlation across distribution

		Half-normal dist.		
		Cost ineff.	Std. profit ineff.	Alt. profit Ineff.
Exponential dist.	Cost ineff.	0.76		
	Std. profit ineff.		0.99	
	Alt. profit ineff.			0.99

We also measured rank-order correlation between cost inefficiency estimates of half-normal and distribution-free approaches. The rank-order coefficient reported a low value of 21 per cent. When the rank correlation is measured for values of half-normal cost inefficiency estimates with its correspondent of the truncated observations in the distribution-free, the rank-order correlation coefficient reports a high value of 0.99 per cent.

Rank-order correlations have also been undertaken to check how inefficiency estimates are similarly ranked across inefficiency concepts. The results (shown in Tables 7.6 and 7.7) indicate that the estimates of standard and alternative profit inefficiencies are very similarly ranked, at about 80 per cent. However, profit inefficiencies and cost inefficiency correlation show a very low-rank coefficient, ranging between 1 and 15 per cent. This suggests that firms that are cost-efficient are not necessarily the same as those that are profit-efficient.

Table 7.6 Rank correlation across inefficiency concepts

		Half normal dist.	
		Cost ineff.	Std. profit ineff.
Half-normal dist.	Std. profit ineff.	0.15	
	Alt. profit ineff.	0.09	0.80

Table 7.7 Rank correlation across inefficiency concepts

		Exponential dist.	
		Cost ineff.	Std. profit ineff.
Exponential dist.	Std. profit ineff.	0.05	
	Alt. profit ineff.	0.01	0.80

In sum, rank-order coefficients show a close association between cost inefficiency estimates across different specifications and also for profit inefficiency estimates. The association is also close for inefficiency estimates using different distributional assumptions. Both profit concepts (standard and alternative profit inefficiency) reported high-rank correlation; nevertheless, the profit and cost inefficiency rank correlation is low.

Rank-order correlation is also calculated to show how consistent our inefficiency measures are compared with standard financial ratios such as return on assets and costs (see Bauer et al., 1998). One may expect high efficiency to coincide with high profitability and low costs as banks approach the best practice bank profit and cost characteristics. Table 7.8 shows measures of rank-order correlation estimated for inefficiency measures and the rate of return on assets (ROA). The table also shows the rank correlation between inefficiency measures and the ratio of total cost to total assets (TCTA).

Table 7.8 Rank correlation of inefficiency and both ROA and TCTA*

	Cost ineff.	Std. profit ineff.	Alt. profit ineff.
ROA	0.01	-0.44	-0.53
TCTA	0.29	0.24	0.23

* Preferred model under the half-normal distribution.

Table 7.8 shows that the relationship between inefficiency and profitability is almost consistent. That is, profit inefficiency and profitability are negatively related, which means that less inefficient banks are more profitable. Although cost inefficiency and profitability are positive, this does not mean that more cost inefficient banks are more profitable since the rank coefficient implies that the relationship is almost absent at a coefficient of 1 per cent.

In the same way, Table 7.8 shows that the relationship between costs (TCTA) and inefficiency measures is consistent because the positive rank correlation, which ranges between 20 and 30 per cent, suggests that the more cost- and profit-inefficient banks incur higher costs.

7.4.3 Inefficiency estimates across GCC countries, bank size, and ownership type

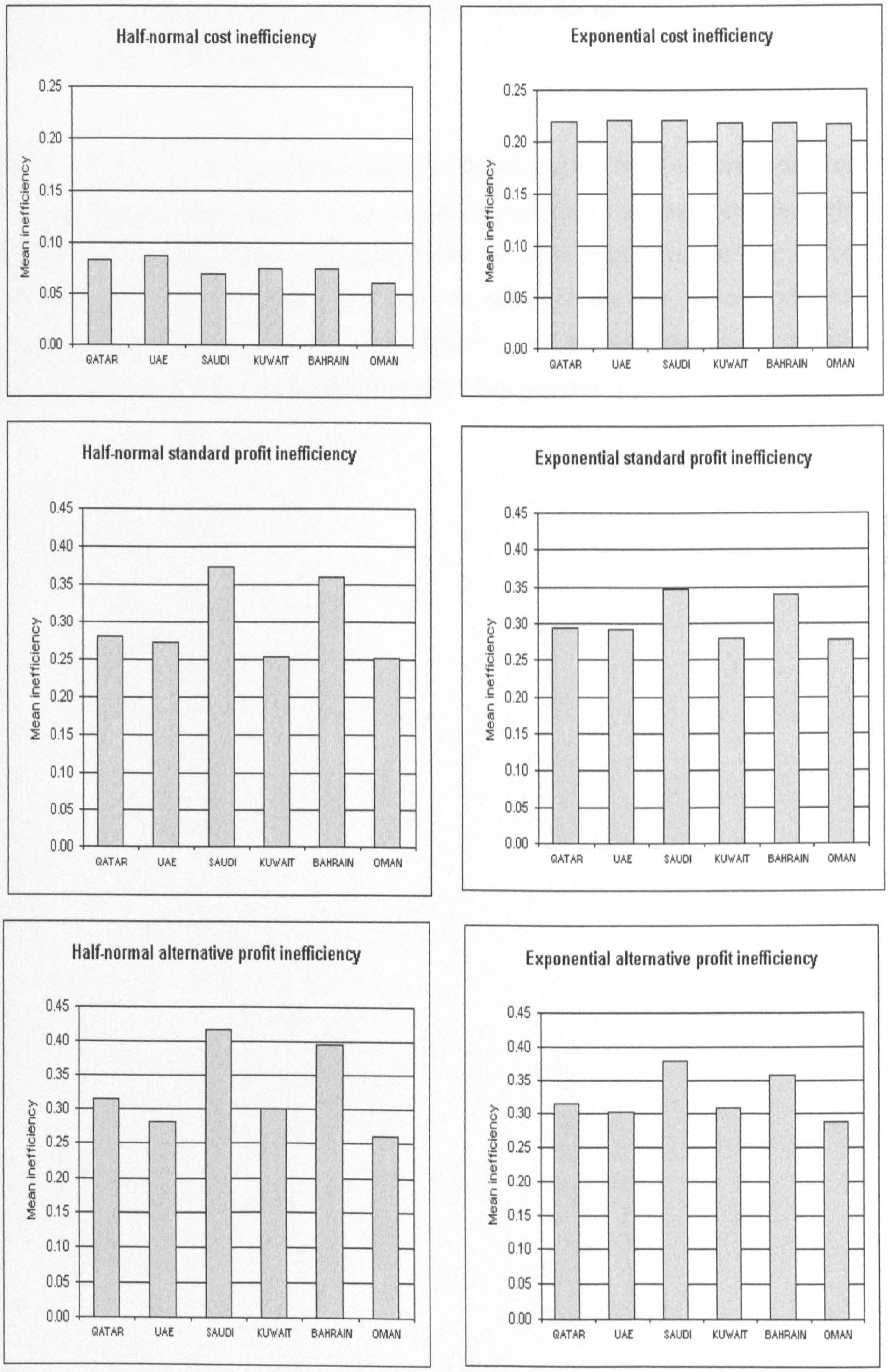
Regarding the efficiency comparisons across GCC countries, Figure 7.3 shows the mean inefficiencies for each country. These inefficiency measures are presented for the preferred Eq. (7.1) cost and profit inefficiencies estimated under both half-normal and exponential distributions.

In general, cost inefficiency estimates across GCC countries are more or less similar to each other. This is true for both half-normal and exponential model estimations. Nevertheless, because the use of the distribution-free approach provides closer inefficiency estimates to the half-normal model, we tend to accept the half-normal estimates as those that are more likely to reflect the 'actual' level of inefficiency in GCC banking markets.

Figure 7.3 indicates that Omani banks appear to be the least cost inefficient (i.e. the most efficient), scoring a level of 7.1 per cent cost inefficiency. The next least cost inefficient banks are Saudi banks, with cost inefficiency levels of 7.9 per cent. Bahraini and Kuwaiti banks occupy the middle ground of GCC cost inefficiency with levels of 7.5 per cent. Qatari and UAE banks have been the most cost inefficient with cost inefficiency levels of 8.3 and 8.8 per cent respectively.

On the profit side, standard and alternative profit inefficiencies across GCC countries tend to vary. In general, Figure 7.3 shows that banks from Saudi Arabia and Bahrain are the most profit inefficient, with a profit inefficiency difference of at least 7 percentage points higher than for other GCC countries' banks.

Figure 7.3 Cost and profit inefficiencies across GCC countries - Preferred model Eq. 7.1



Omani banks remain the least profit inefficient, while the rest of the GCC countries' banks fall in the middle positions.

Figures (7.4 and 7.5) show comparisons of inefficiency across bank ownership and bank assets size respectively. Figure 7.4 shows that foreign banks are more cost inefficient than national banks; in contrast, foreign banks are more profit efficient than national banks. Figure 7.5 presents efficiency scores according to the size of bank; generally, large banks are less cost inefficient than medium-sized and small banks. On the other hand, small banks are less profit inefficient than medium and large banks.

Figure 7.4 Foreign and national bank inefficiencies

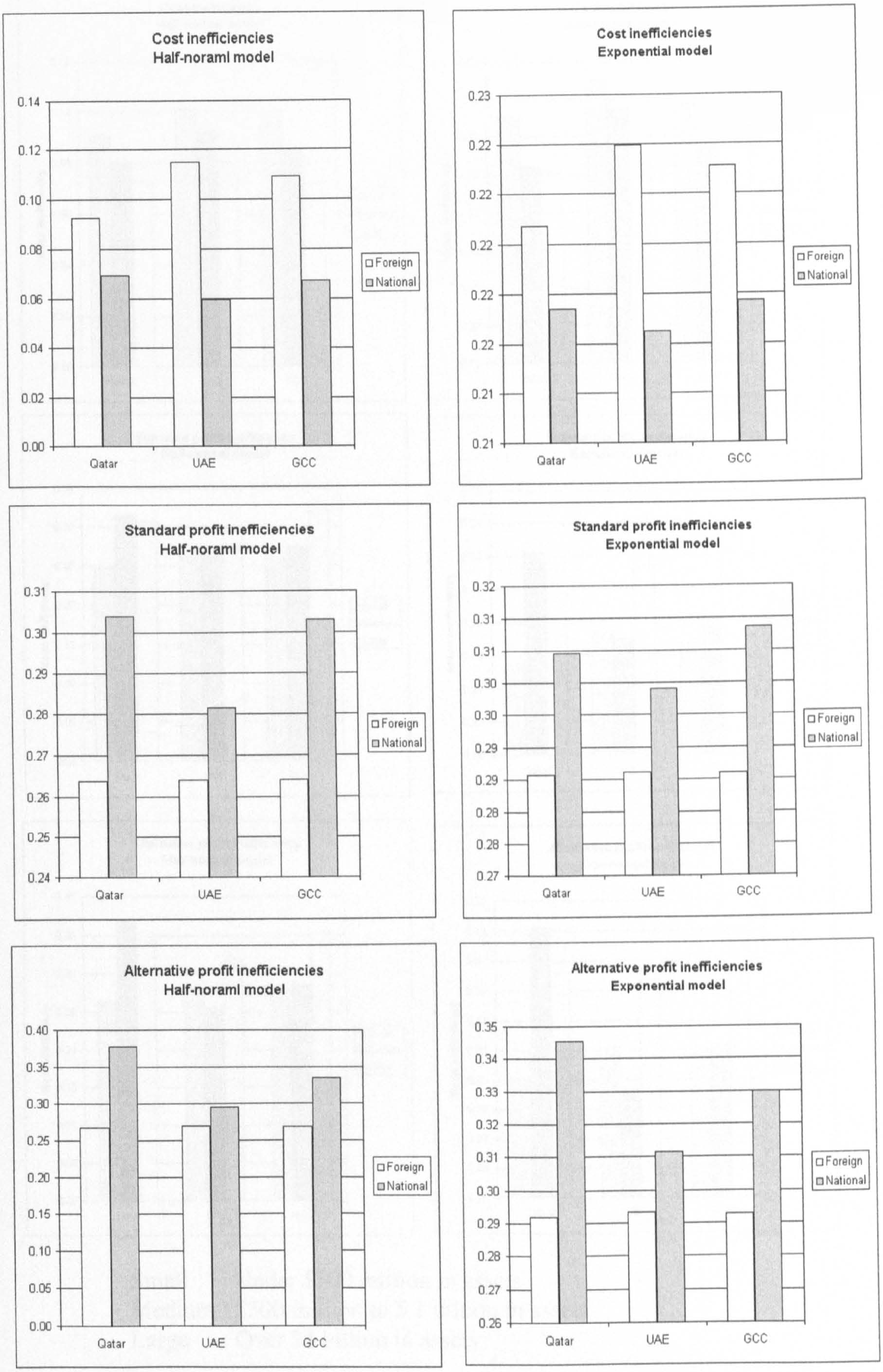
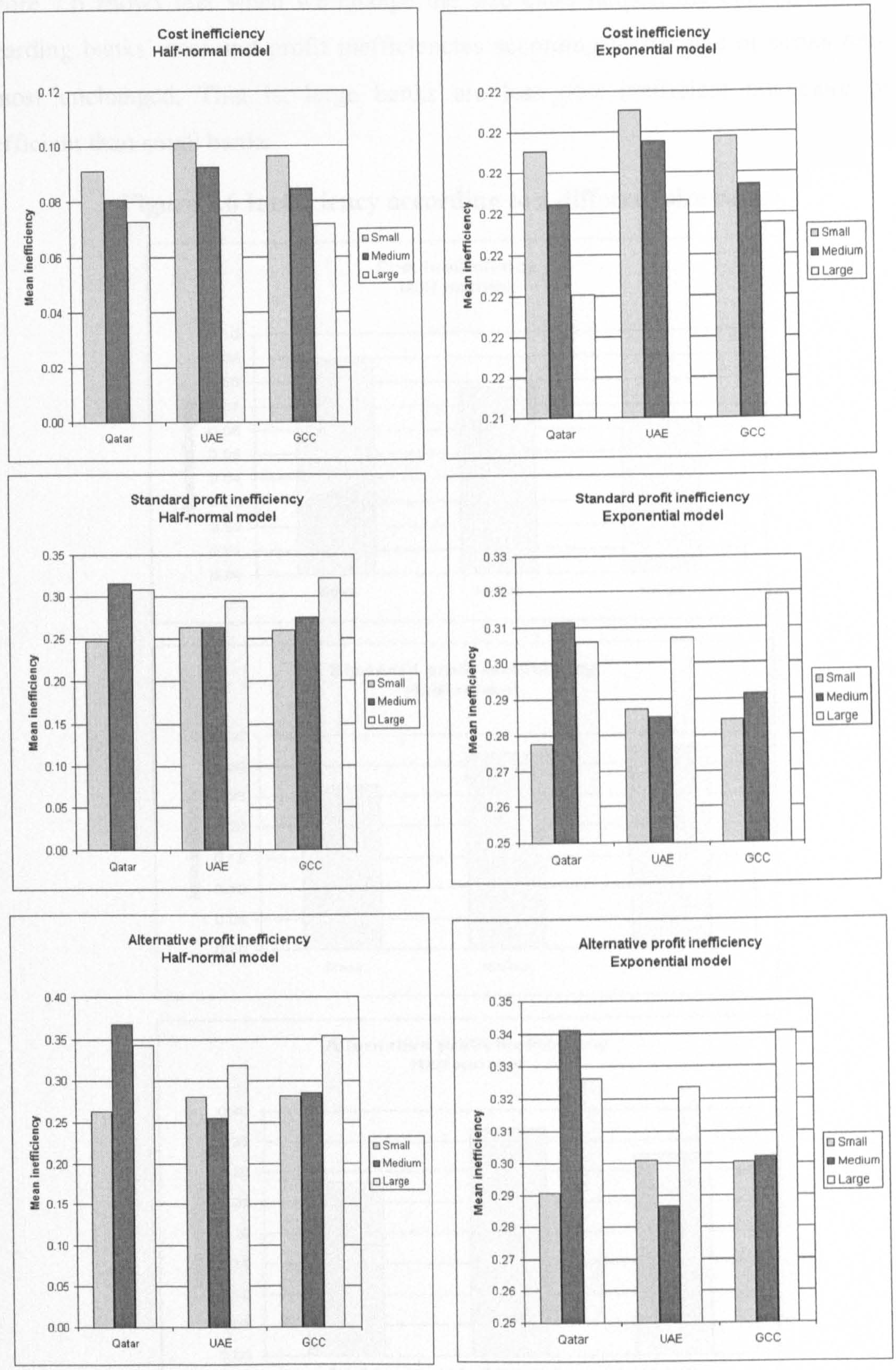


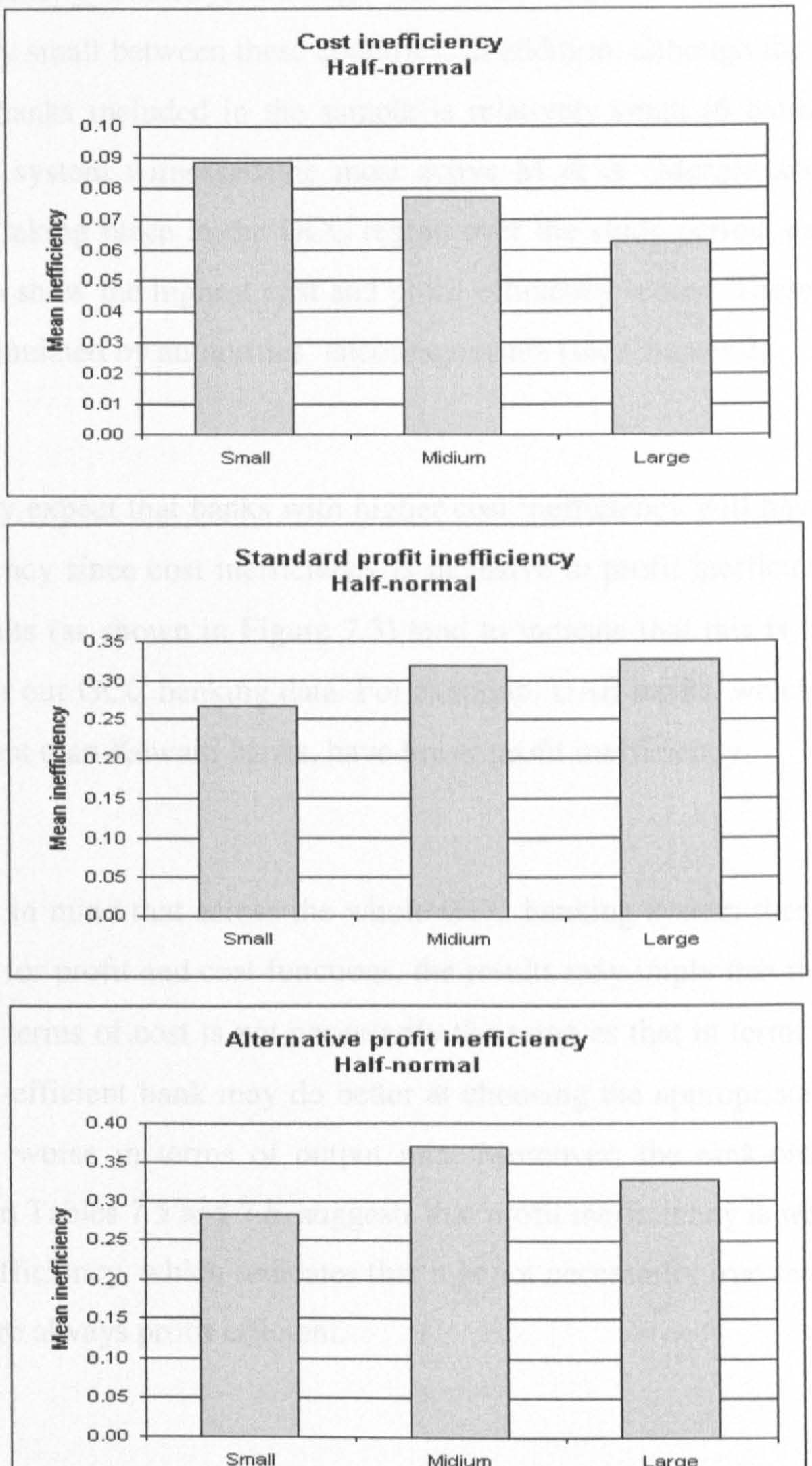
Figure 7.5 Cost and profit inefficiencies across bank size



Small : Under $300 million in assets
Medium : $300 million to $ 1 billion in assets
Large : Over $1 billion in assets

Figure 7.6 shows that when we change the size class ranges, the conclusions drawn regarding banks' cost and profit inefficiencies according to the size of banks remains almost unchanged. That is, large banks are less cost inefficient but more profit inefficient than small banks.

Figure 7.6 Inefficiency according to a different size range



Small : Under $1 billion in assets size
Medium : $1 billion to $ 5 billion in assets size
Large : Over $5 billion in assets size

It would be appropriate to discuss the following points regarding these findings:

- It may not be surprising that Omani banks are the least cost inefficient in the GCC banking industry, although the differences in inefficiency scores are relatively small between these countries. In addition, although the number of the Omani banks included in the sample is relatively small (6 banks), the Omani banking system witnessed the most active M & A (Merger and Acquisition) activity taking place in the GCC region over the study period, enabling Omani banks to show the highest cost and profit efficiency scores. These mergers have been stimulated by authorities' encouragements (see Chapter 2).

- One may expect that banks with higher cost inefficiency will have higher profit inefficiency since cost inefficiency is inclusive in profit inefficiency. However, our results (as shown in Figure 7.3) tend to indicate that this is not the case, at least, for our GCC banking data. For example, UAE banks, which are more cost inefficient than Kuwaiti banks, have lower profit inefficiency.

- Bearing in mind that across the whole GCC banking system there is a common frontier for profit and cost functions, the results may imply that the best-practice bank in terms of cost is not necessarily the same as that in terms of profit since the cost-efficient bank may do better at choosing the appropriate input mix but may do worse in terms of output mix. Moreover, the rank-order correlation, shown in Tables 7.5 and 7.6, suggests that profit inefficiency is weakly related to cost inefficiency, which indicates that it is not necessarily true that cost-efficient banks are always profit efficient.

- The question why one country's banks are more cost or profit efficient than another can be related to the size of banks in a country. For instance, with reference to Figures (7.5 and 7.6), countries that have relatively small banks, such as the UAE and Qatar, tend to show higher cost inefficiency but lower profit inefficiency. On the other hand, banking industries that are dominated by larger banks, such as those in Saudi Arabia, Bahrain, and Kuwait, tend to show lower cost inefficiency but higher profit inefficiency. In fact, large banks may have lower cost inefficiencies because their per unit cost decreases as the scale increases. However, scale effects may induce profit inefficiency because large banks may face more difficulty in generating revenues efficiently. Berger and Mester (1997, p. 936) state that '[t]he cost and profit efficiency results together seem to imply that as banks grow larger, they are equally able to control costs, but it becomes harder to create revenues efficiently.' Moreover, this finding is consistent with the conventional fact that small banks typically have higher profitability ratios than larger banks. Having said this, however, the scale effects that induce profit inefficiency are unlikely to be large.

- This scale effect could also explain differences in the inefficiency of foreign and national banks. For instance, the majority of foreign banks operating in the GCC countries are classified in terms of size as small to medium-sized banks. Therefore, as is shown in Figure 7.4, foreign banks are found to be less cost efficient but more profit efficient than national banks.

This section has discussed the results of the cost and profit inefficiency estimates for GCC banks. The following sections provide further analysis of the issue of scale by examining scale economies and scale inefficiency in Gulf banking.

7.5 Scale economies and Scale inefficiency

This section estimates economies of scale and scale inefficiency in the GCC banking markets. For our economies of scale estimates, we only compute economies of scale from the cost function, and not from the profit functions (see Berger and Mester, 1997).[15] Moreover, because scale inefficiency estimates depend on economies of scale, we also provide results on scale efficiency derived from the cost function.

7.5.1 Economies of scale

Scale economies measure how a unit change in output affects total costs.[16] The economies of scale results shown in Tables (7.9 to 7.13) are calculated for both the traditional and preferred specifications estimated using both half-normal and exponential distribution models.

The economies of scale results derived from the traditional model specification show that the exponential model provides similar economies of scale estimates to those of the half-normal model (Tables 7.9 to 7.13).

With reference to the cross-country scale economies comparisons, the results in Table 7.9 show that both the half-normal and exponential models assign Saudi and Kuwaiti banks as realising scale economies over the period under study. Moreover, Bahraini banks experience constant returns to scale. However, UAE, Omani, and Qatari banks exhibit scale diseconomies.

[15] A number of studies derive scale economies from the profit function (See for example, Al-Jarrah, 2002).

[16] Scale economies are calculated using the following equation:

$$\begin{aligned} \textit{Scale economies} &= \sum_{i=1}^{2} \frac{\partial \ln TC}{\partial \ln Q_i} = \sum_{i=1}^{2} \alpha_i + \sum_{i=1}^{2}\sum_{j=1}^{2} \delta_{ij} \ln Q_j + \sum_{i=1}^{2}\sum_{j=1}^{2} \rho_{ij} \ln P_i \\ &+ \mu_i \sum_{i=1}^{2} \left[-a_i \sin(Z_i) + b_i \cos(Z_i)\right] \\ &+ 2\mu_i \sum_{i=1}^{2}\sum_{j=1}^{2} \left[-a_{ij} \sin(Z_i + Z_j) + b_{ij} \cos(Z_i + Z_j)\right]. \end{aligned}$$

If *scale economies* >1, <1, or $=1$, then there are diseconomies, economies of scale, or constant returns to scale respectively.

Table 7.9 Scale economies in the GCC banking industry - by country

	Half-normal		Exponential	
	Traditional model	Preferred Eq. (7.1)	Traditional model	Preferred Eq. (7.1)
GCC	1.167	1.108	1.188	1.177
QATAR	1.288	1.222	1.303	1.281
UAE	1.228	1.166	1.246	1.229
SAUDI ARABIA	0.956	0.903	0.986	0.995
KUWAIT	0.924	0.886	0.954	0.970
BAHRAIN	1.072	1.027	1.096	1.097
OMAN	1.342	1.256	1.361	1.329

When risk and quality factors are taken into account, the preferred model again shows that Saudi and Kuwaiti banks exhibit scale economies at slightly higher levels. Bahraini banks are also close to unity, indicating constant returns to scale. UAE, Omani, and Qatari banks have not much been influenced by the introduction of risk and quality factors since these countries continue to exhibit scale diseconomies. In sum, closeness to unity of the scale estimates of banks in Saudi Arabia, Kuwait, as well as Bahrain may lead us to deduce that these countries' banks tend to show the range between economies and constant returns to scale, unlike banks in the UAE, Oman, and Qatar, that apparently show diseconomies of scale. Overall, on average, the sample shows that the GCC banking industry has been exhibiting scale diseconomies driven mostly by banks that belong to the GCC countries' exhibiting scale diseconomies (namely those in the UAE, Oman, and Qatar).

If we look at the size dimension of the GCC banks, the results of both half-normal and exponential distribution shown in Table 7.10 strongly indicate that scale elasticities (economies) of large GCC banks are much closer to unity than those of small and medium banks. Moreover, small banks show more scale diseconomies than medium-size banks. These results tend to be the same when deriving economies of scale estimates from the model that includes risk and quality factors.

Table 7.10 Scale economies in the GCC banking industry - by assets size

Size	Assets range	Half-normal		Exponential	
		Traditional	Preferred Eq. (7.1)	Traditional	Preferred Eq. (7.1)
Small	Under $300 million	1.301	1.246	1.312	1.291
Medium	$300 million to $1 billion	1.235	1.174	1.252	1.235
Large	Over $1 billion	1.048	0.989	1.076	1.076

Moreover, Table 7.11 shows that when we change the size class ranges, the aforementioned conclusion about scale economies analysis according to size remains the same. That is, small banks exhibit scale diseconomies, medium banks almost always exhibit constant returns to scale, and large banks enjoy scale economies.[17]

[17] Because of similarities in results, we only show the scale economies estimates derived form the preferred Eq. (7.1) for these size ranges.

Table 7.11 Scale economies according to bank size

		Half-normal
Size	Assets range	Preferred
Small	Under $1 billion	1.221
Medium	$1 billion to $5 billion	1.079
Large	Over $5 billion	0.934

Scale economies have also been calculated for the foreign banks operating in the GCC banking system (specifically, foreign banks operating in the UAE and Qatar). In comparison to the GCC national banks, as Table 7.12 shows, foreign banks have on average been operating with higher scale diseconomies than national banks over the six-years study period. These results are indicated across the two distributional models used here and with and without the risk and quality variables.

Table 7.12 Scale economies of foreign and national banks

		Half-normal		Exponential	
		Traditional Model	Preferred Model Eq. (7.1)	Traditional Model	Preferred Model Eq. (7.1)
GCC	Foreign	1.263	1.204	1.277	1.258
	National	1.122	1.063	1.145	1.139
QATAR	Foreign	1.278	1.224	1.290	1.272
	National	1.300	1.220	1.321	1.294
UAE	Foreign	1.257	1.196	1.272	1.253
	National	1.198	1.134	1.218	1.203
Qatar & UAE	Foreign	1.263	1.204	1.277	1.258
	National	1.221	1.153	1.241	1.224

Moreover, if we extend the analysis to comparison between foreign and national banks in the UAE and Qatar, the results do not differ so much. In Qatar, foreign banks have less scale diseconomies than national banks. However, an exception is noted when risk and quality is controlled for in the model in which foreign banks turned out to have more scale diseconomies than national banks when the half-normal model is considered. In the UAE case, foreign banks exhibit greater scale diseconomies than their national peers, and this result appears to hold in both the half-normal and exponential models and with and without the consideration of the risk and quality variables. In addition, aggregated results from both the Qatari and UAE banking industries show that foreign banks operating in both of these countries experience greater scale diseconomies than national banks.

Table 7.13 shows that scale diseconomies decline as the assets sizes of both national and foreign banks increase. The table shows that scale diseconomies of foreign banks have been higher than for national banks across the three different size categories of banks, except for medium asset size banks. Overall, foreign banks show evidence of smaller diseconomies of scale than do national banks in the medium asset size range. This result suggests that medium-size foreign banks are the preferred magnitude from a scale economies standpoint.

Table 7.13 Scale economies of foreign and national banks according to the size of bank

		Half-normal		Exponential	
		Traditional Model	Preferred Model Eq. (7.1)	Traditional Model	Preferred Model Eq. (7.1)
Small	Foreign	1.323	1.265	1.333	1.309
	National	1.137	1.109	1.149	1.149
Medium	Foreign	1.216	1.157	1.233	1.218
	National	1.243	1.181	1.260	1.243
Large	Foreign	1.131	1.069	1.153	1.144
	National	1.039	0.981	1.068	1.069

In addition, it also seems that only large national banks realize scale economies – a finding that confirms our earlier observation relating to size and cost economies.

7.5.2 Scale inefficiency

The reason for computing scale inefficiency is that we cannot compare between the estimates of both scale economy and X-inefficiency since they measure different aspects of a bank's cost characteristics. That is, scale economies is a measure of scale elasticities that expresses a percentage change in the total cost with respect to a percentage change in output, and X-inefficiency expresses the percentage of the cost function that bank i needs to alter so that it can reach the cost function of the industry's best-practice bank.

In order to find a common ground for comparing the estimates of these two concepts, scale economies must be transformed into scale inefficiency, which expresses the percentage the cost function ought to change if a bank needs to move to the minimum efficient scale.

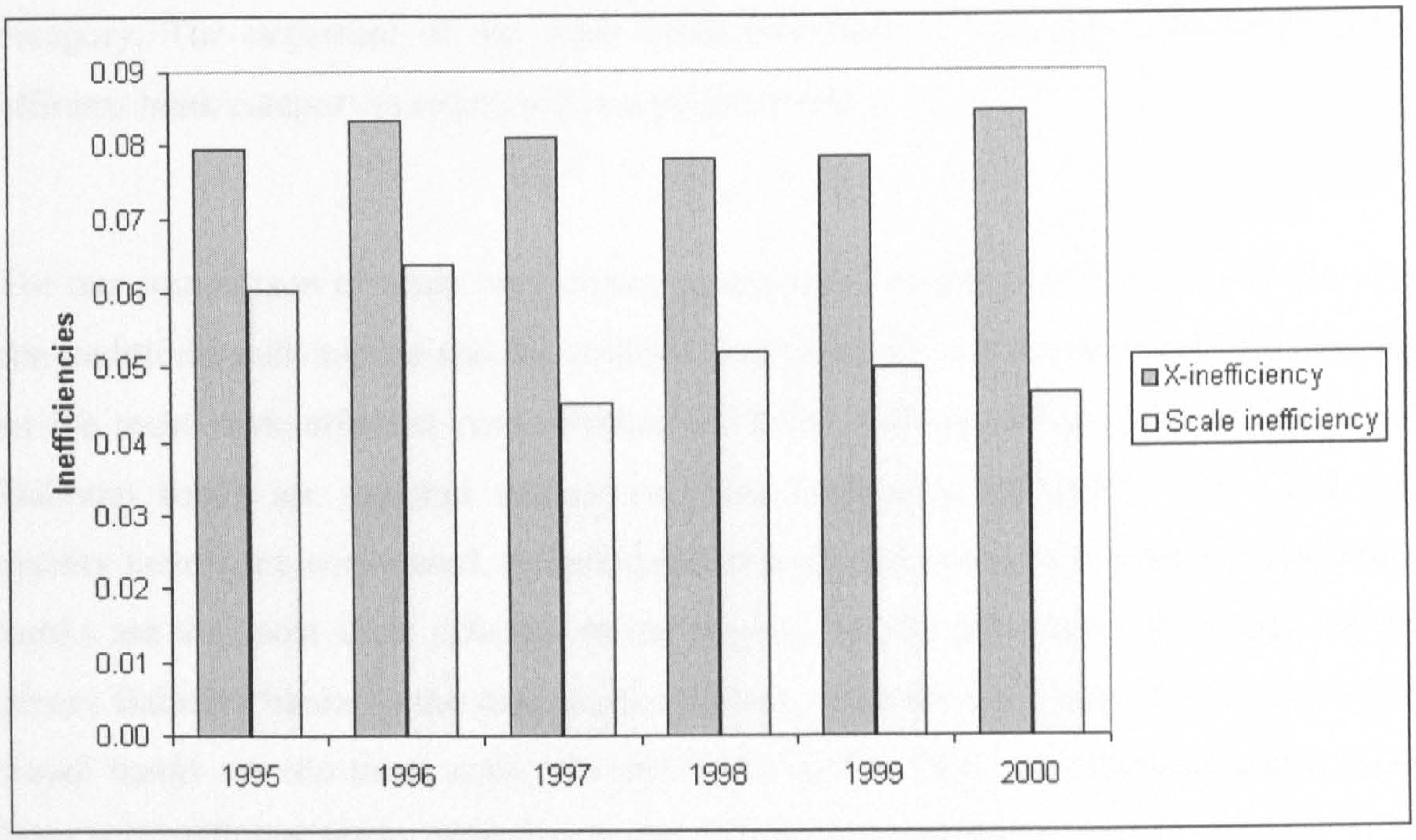
Scale inefficiency is calculated here using the approach of Evanoff and Israilevich (1995).[18] Figure 7.7 shows comparisons between X-inefficiency and scale inefficiency derived from the preferred cost model estimated under half-normal distribution assumptions. X-inefficiencies are consistently larger than scale inefficiencies during the study period.[19]

---

[18] Scale inefficiencies is calculated as $\textit{Scale inefficiency} = e^{(.5/c)(1-\varepsilon)^2} - 1$, where ε is the first derivate of the cost function, that is the scale elasticity with respect to output; and c is the second derivative of the cost function with respect to output. After taking the first and second derivates of Eq. (7.1) in terms of output quantities, we omit the trigonometric terms in the calculations of scale inefficiency to avoid getting negative scale inefficiency values that may diminish the economic logic.

[19] This finding is also true in the case of exponential distribution.

Figure 7.7 comparisons between scale inefficiency and X-inefficiency



Many other studies find that cost X-inefficiency dominates both scale and scope efficiencies. For example, Berger and Humphrey (1991) find that X-inefficiencies dominate scale and scope inefficiencies in commercial banking.[20] In contrast, Altunbas et al. (2000) find that scale inefficiencies dominate X-inefficiencies in Japanese banking.

The results contained in Tables (7.14 to 7.18) show that the half-normal and exponential models provide similar measures of scale inefficiencies. The slight difference between the models falls within the range 0.00 to 0.005. However, with consideration of the risk and quality factors, the difference increases and stands in the range of 0.05 to 0.1, higher for the exponential model.

In the scale inefficiency measure, as the value of scale inefficiency of a bank approaches closer to zero, the bank is considered to approach closer to the most scale-efficient bank. In other words, a bank that has the lowest scale inefficiency score, that is

[20] Berger (1993), Berger and Humphrey (1991), and Evanoff and Israilevich (1991) reach the same conclusions.

the closest to the value zero, is the closest to the efficient scale in an underlying category. The departure of the scale inefficient bank in category *j* from the scale efficient bank category is expressed as a positive value.

For our comparison of scale inefficiency across GCC countries, Table 7.14 shows that the traditional half-normal and the traditional exponential cost models rank Saudi banks as the most scale-efficient banks within the GCC banking sample. The Kuwaiti and Bahraini banks are together the second most scale-efficient banks. When risk and quality factors are considered, the preferred exponential model still indicates that Saudi banks are the most scale efficient in the region, but the preferred half-normal model places Bahraini banks as the most scale efficient. Overall, the results tend to show that Saudi banks are the most scale-efficient banks in the GCC, and Bahrain and Kuwait have scale efficient banks operating in their banking industries.

Table 7.14 Scale inefficiency in the GCC banking industry - by country

	Half-normal		Exponential	
	Traditional model	Preferred Eq. (7.1)	Traditional model	Preferred Eq. (7.1)
QATAR	0.180	0.093	0.206	0.159
UAE	0.110	0.051	0.131	0.102
SAUDI ARABIA	0.004	0.017	0.000	0.000
KUWAIT	0.012	0.024	0.004	0.002
BAHRAIN	0.011	0.001	0.019	0.018
OMAN	0.264	0.126	0.304	0.224

UAE banks appear to take the middle position regarding scale inefficiency within the GCC countries. However, applying both the traditional and preferred specifications estimated using both distributional models, the results indicate that Qatari and Omani banks are among the least scale-efficient banks in the region.

The results also show that large banks have scale advantages over medium and small banks. When looking at scale inefficiency in terms of the size of banks, the results in Table 7.15 indicate that large banks are the most scale efficient in the GCC banking industry. The results also show that medium-size banks are more scale efficient than their smaller counterparts. These results are consistent over both distributional models and with both traditional and preferred specifications.

Table 7.15 Scale inefficiency in the GCC banking industry - by size

		Half-normal		Exponential	
Size	Assets range	Traditional	Preferred Eq. (7.1)	Traditional	Preferred Eq. (7.1)
Small	Under $300 million	0.036	0.025	0.038	0.033
Medium	$300 million to $1 billion	0.023	0.013	0.026	0.023
Large	Over $1 billion	0.001	0.000	0.003	0.003

However, when we change the size range, as shown in Table 7.16, medium banks become the most scale-efficient, although large banks still come very close to the scale-efficient banks. Small banks consistently exhibit their lowest scale efficiency relative to the other bank size classes.

Table 7.16 Scale inefficiency according to bank size

		Half-normal
Size	Assets range	Preferred
Small	Under \$1 billion	0.0171
Medium	\$1 billion to \$5 billion	0.0009
Large	Over \$5 billion	0.0054

With regard to scale-inefficiency comparisons between foreign and national banks, Table 7.17 shows that GCC national banks are much more scale efficient than foreign banks. This finding is consistent over the period under study and across both distributional models and specifications.

Table 7.17 Scale inefficiency of foreign and national bank

		Half-normal		Exponential	
		Traditional Model	Preferred Model Eq. (7.1)	Traditional Model	Preferred Model Eq. (7.1)
GCC	Foreign	0.148	0.078	0.169	0.132
	National	0.030	0.007	0.044	0.037
QATAR	Foreign	0.168	0.095	0.187	0.148
	National	0.198	0.091	0.233	0.175
UAE	Foreign	0.141	0.072	0.162	0.127
	National	0.082	0.033	0.102	0.080
Average: QATAR & UAE	Foreign	0.148	0.078	0.169	0.132
	National	0.103	0.043	0.126	0.098

Moreover, national banks in the UAE are found to be more scale efficient than foreign banks. However, as the table shows, in the case of Qatar, foreign banks are more scale efficient than their national peers. Only in the half-normal distribution, and with consideration of risk and quality factors, foreign banks appear to be less scale efficient than national banks in Qatar. Overall, the results in this table tend to show, however, that foreign banks are less scale efficient than national banks operating in the GCC banking industry.

Table 7.18 shows that large banks (including both national and foreign banks) tend to be more scale efficient than banks of all other sizes. In fact, small foreign banks are shown to be the most scale inefficient banks across all GCC bank sizes. However, medium-size foreign banks operating in the GCC appear to be more scale efficient than national banks of the same size.

Table 7.18 Scale inefficiency of foreign and national banks according to the size of bank

		Half-normal		Exponential	
		Traditional Model	Preferred Model Eq. (7.1)	Traditional Model	Preferred Model Eq. (7.1)
Small	Foreign	0.232	0.135	0.253	0.195
	National	0.038	0.022	0.046	0.042
Medium	Foreign	0.098	0.046	0.117	0.093
	National	0.126	0.061	0.148	0.116
Large	Foreign	0.035	0.009	0.049	0.040
	National	0.003	0.001	0.009	0.009

Combining the findings on scale economies and scale inefficiencies, the results indicate that, overall, large and medium banks (including large foreign banks) better exploit their resources in terms of realizing economies of scale and scale efficiency, while small banks have substantial diseconomies and scale inefficiencies. This suggests that small banks should work to increase their size in order to reach the optimal scale, possibly through further M & A activities. The findings also indicate that countries that have, on average, large banks tend to exhibit greater scale economies and less scale inefficiencies in their banking industry.

The finding that large banks (mostly national banks) realize greater scale economies and are more scale efficient than small banks in the GCC banking industries could be explained in a number of ways. One explanation is related to the issue of the bank's age. Older banks have increased their size over time and become more scale exploitive. This explanation applies also to foreign banks since long-established foreign banks have increased their size over time in response to economic development and this has lead to a better exploitation of resources and per unit cost reduction. For example, foreign banks that were established in the 1950s, 1960s, and 1970s were probably better placed to profit from financing the early stages of the GCC growth period during the 1970s.[21] However, while a share of large foreign banks' domestic financing has, to a certain extent, been crowded out by national banks established mainly in the 1970s and 1980s, large foreign banks still have a sizable share of the foreign banks' sector (about 70 per cent of total foreign bank assets, which represent nearly 24 per cent of the total GCC commercial banking market).

Large banks are also seen to have greater geographical coverage in the GCC than small and medium-size banks. Large foreign banks have opened branches, particularly in Qatar and the UAE. They may, perhaps, have been able to realize economies from this type of expansion.

Another issue relating to the exploitation of economies and scale efficiencies is that if a country is over-banked with commercial banks, scale economies and scale efficiencies might be affected. This being so, banks that operate in relatively small GCC countries

[21] The start of sizable influxes of oil revenue in the seventies enabled the establishment of national banks with both public and private, or purely private, ownership.

that appear over-banked,[22] such as Qatar and the UAE, may have less opportunity to exploit scale and other cost efficiencies. Countries that restrict new bank licensing, like Saudi Arabia and Kuwait, potentially provide greater scale and cost efficiency advantages for locally operating banks.

Large banks may realize greater scale economies and scale efficiencies because they have greater access to better information technology. This relates to such things as more sophisticated ATM networks, better credit scoring systems, and improved internal and external monitoring and screening systems. Taken together, big banks in the GCC may have technology, managerial, and other advantages over smaller banks, resulting ultimately in improved cost performance.

Since most of the large banks operating in the GCC banking sector are national banks, it is important to note that many of these have been established and promoted by government regulation and ownership. For example, Saudi Arabia and Kuwait have strict regulations governing the establishment of any new commercial banks and similarly, in Qatar and the UAE, there are limits associated with foreign bank branches. This limitation has created the opportunity for existing banks, especially national banks, to expand their services to absorb increases in the demand for credit and other banking services.

Apart from the regulations that inhibit foreign bank presence or/and expansion in these markets, foreign banks may also suffer from other limitations. For instance, GCC governments mainly favour national banks to fulfil the majority, if not all, of their government financing needs. This can have the adverse effect of distorting the price mechanism since the choice of the government may not be based on market disciplines. Government practice, in this case, could also affect foreign interests in investing in the country, discourage potential foreign banking, and distort national bank competitiveness.

[22] Some concerns have been raised regarding countries with banking systems crowded with a great many banks licensed to operate, as this could be at the expense of profitability and might lead to severe competition, compelling some banks to engage in riskier activities such as extending loans to low-quality customers.

Foreign banking business is also characterized mainly by relatively narrow banking activities such as money transfer and the facilitation of moderate commercial trade between the host country and the country of origin, while larger foreign banks undertake a broad range of commercial and corporate banking activities. Many of the smaller foreign banks have located in the GCC to remit transfers mainly for expatriate workers. However, because deposits of expatriate workers (in most cases) tend to be relatively small, and because there are restrictions on foreign banks' operations, many foreign banks have remained small in size. This may explain these banks' inability to grow in order to realize greater scale economies and scale efficiency.

Nonetheless, small banks (mostly foreign banks) can find ways to continue their business alongside large banks. '[S]mall banks are better at relationship banking than large banks due to superior information and greater discretion in applying information' (Chen, Mason, and Higgins, 2001). Moreover, loan officers at large banks tend to be more strict in following bank rules and criteria than their counterparts in small banks (Nakamura, 1994; Udell, 1989). This also suggests why smaller banks (or foreign banks) are able to survive under restrictions and less efficient performance.

Finally, this section also briefly reports on the findings of the contribution of technical progress in cost reduction of GCC banks in the period of study. Basically, the use of technology is expected to have an important impact on the banks' costs function. The results on technical progress are shown in Table 7.19. These estimates are measured as the elasticity of total cost with respect to time, or the change in the total cost with respect to the change in time *t*.

Table 7.19 Technical progress in the GCC banking industry

	Half-normal	Exponential
Traditional model	-0.0008	0.0095
Preferred model Eq. (7.1)	-0.0006	-0.2856

The results in Table 7.19 shows that, during the study period, technical progress estimated using the half-normal model contributed to annual cost reduction for GCC banks on average of around 0.08% and 0.06%. These results appear low compared with those of other studies. For instance, Altunbas et al. (2001) finds that in the previous literature on technical change this has ranged between a 3 and 5 per cent annual contribution to cost reduction in US and European banks. Generally, it seems from the results that technical progress has not contributed substantially to cost reduction in the GCC banking industry during the 1990s since technical change estimates are low.

This section has discussed the estimates of economies of scale and scale inefficiency, as well as technical progress in the GCC banking markets. In the investigation of the determinants of inefficiency in the GCC banking industry, the next section explores the correlation between inefficiency and the variables concerning banks' and industry's characteristics using a logistic regression model.

7.6 Logistic regression

The final part of the empirical analysis examines the determinants of banking sector inefficiency in GCC banking systems over the study period 1995-2000. For this purpose, we use the logistic regression model, in which we regress inefficiency variables (cost inefficiency and profit inefficiency measures) on a variety of bank and market-specific variables that we believe are most likely to influence inefficiency levels.

As noted earlier, logistic regressions' estimated coefficients indicate relationships in terms of correlation rather than the power and size of impact or the causality relationship. The logistic regression model is preferred over the linear regression approach since the former is more appropriate to model the relationship between variables for which a dependent variable is bounded between zero and one, the range in which inefficiency scores fall.

In order to avoid double consideration of risk and asset quality variables when examining inefficiency determinants, the logistic model is estimated using inefficiency measures derived from the frontier estimation of the traditional model that does not incorporate equity and provisions. In addition, the estimates of inefficiency used here are for the traditional cost and profit functions estimated using the half-normal distribution.

As for the logistic parameters estimates, the results in Tables (7.20 and 7.21) show that the correlation between the cost and profit inefficiency measures regressed on the same set of the independent variables almost conform to the expectations we mentioned in Chapter 6.[23]

[23] Although R^2 has a low value, it is not an appropriate measure of closeness of fit in the context of logistic regressions (see Thomas, 1997).

Starting with the relationship between inefficiency and financial capital, in both cost and profit inefficiency determinants, the coefficient E is negative and is significantly different form zero. This indicates that banks with low inefficiency levels tend to hold higher levels of capital. Note that in our previous analysis in section (7.4.2), we found that if we remove the capital variable from our preferred model, this results in a slight increase in the level of cost and profit inefficiency. This means that when financial capital is introduced in the model, it controls and takes into consideration the fact that banks with strengthened capital have a better cushion against risk and this seems to make them become more efficient. However, one must caution that this does not necessarily mean that efficient banks should always have higher capital and thus have lower risk (Mester, 1996). This is because higher levels of financial capital level may distort managers' incentives in a way that makes them keener to take riskier activities (moral hazard). Generally, in this analysis, the results suggest that more efficient GCC banks generate higher earnings, which are translated into higher levels of capital.

Table 7.20 The logistic regression parameter estimation

Dependent variable: Cost inefficiency half-normal (CN)

Variable	Coefficient	Std. error	*T*-value
Constant	8.13E-02	7.71E-03	10.548
EIQUITY	-4.12E-08	1.33E-08	-3.108
ROA	-0.2483704	4.46E-02	-5.575
PROV	8.23E-02	4.27E-02	1.928
FOREIGN	2.59E-02	4.18E-03	6.203
LTA	-5.05E-02	8.63E-03	-5.857
FIX	2.42E-08	2.57E-08	0.943
TA	2.14E-09	1.35E-09	1.59
TBGDP	-3.45E-04	8.38E-04	-0.412
CN[-1][24]	0.4309832	3.61E-02	11.942
Durbin-Watson Statistic =	1.91243	Rho =	0.04379
Adjusted R-squared =	0.46058		
Observations =	558		

[24] The lagged dependent variable is used to remove auto-correlation.

Table 7.21 The logistic regression parameter estimation

Dependent variable Standard profit inefficiency half-normal (SN)

Variable	Coefficient	Std. error	*T*-value
Constant	0.2588479	3.13E-02	8.274
E	-2.66E-07	5.65E-08	-4.705
ROA	-1.006451	0.18616	-5.406
PROV	1.11E-02	0.18015	0.062
FOREIGN	8.52E-03	1.62E-02	0.527
LTA	-7.09E-02	3.53E-02	-2.005
FIX	-9.05E-08	1.08E-07	-0.838
TA	2.33E-08	5.74E-09	4.069
TBGDP	1.99E-03	3.53E-03	0.564
SN[-1][25]	0.5102705	3.46E-02	14.757
Durbin-Watson Statistic =	1.95567	Rho =	0.02217
Adjusted R-squared =	0.38476		
Observations =	558		

[25] The lagged dependent variable is used to remove auto-correlation.

The results also show that accounting profits (denoted as ROA) is negative and is significantly different form zero as well. The ROA coefficient in both cost and profit inefficiency regressions confirms that more efficient banks may be expected to achieve, on average, better accounting profits performance than less efficient banks. Therefore, this may underline the perception that more efficient banks can consolidate their capital through better profits performance, enabling them to accumulate higher capital, in turn making them less risky firms.

With respect to loan quality, both the cost and profit inefficiency dependent variables are positively correlated with the level of provisions (PROV); the PROV variable is significant at the 10 per cent level in the cost inefficiency regression but insignificant in the profit inefficiency regression. This positive correlation suggests that inefficient banks are forced by regulation to increase the level of provisions when their loans are facing defaults problems. In other words, a high level of provisions indicates loan quality deterioration and, as a result, inefficiency generally increases in response to the higher level of problem loans. This may also suggests that efficient banks with lower levels of loan provisions are better at evaluating credit risk (see Mester, 1996; Berger and DeYoung, 1997; Altunbas et al., 2000).

Turning to the issue of ownership, the binary variable FOREIGN shows a positive and statistically significant relationship with cost inefficiency but a statistically insignificant relationship with profit inefficiency. Taking at least the relationship between cost inefficiency and the variable FOREIGN, we infer that the existence of foreign banks has contributed to the inefficiency level in the GCC banking industry during the study period. This result is consistent with our previous finding reported in sections (7.4 and 7.5), that foreign banks operating in GCC countries tend to be less cost efficient than their national peers. In fact, regulatory restrictions on foreign bank business, such as restrictions on bank size, taxes, and bank branching, could also be the main factors inducing foreign banks to contribute to inefficiency in the GCC banking industry.

As for the rest of the control variables, the negative correlation between the loan to assets ratio (LTA) and the inefficiency levels indicate that banks with higher proportions of lending business in their balance sheets are more efficient. This result contrasts with previous studies' findings (for example, Altunbas et al., 2000, found a positive correlation between inefficiency and the loan to assets ratio in the case of Japanese banks). This result, however, may indicate that the GCC countries' larger banks have emphasized lending business during the second half of the 1990s in order to respond to market demand.

Moreover, total assets (TA), which approximates the size of a bank, shows a clearer relationship between bank profit inefficiency and bank size (than bank cost inefficiency and bank size). As we previously noted, large banks usually experience higher profit inefficiency than small banks, here, the statistically significant and positive relationship between TA and profit inefficiency indicate that as banks increase in size, their profit inefficiency increases. Nevertheless, this relationship is not evident in case of cost inefficiency since the TA coefficient is not significant, although its sign is positive.

Taken together, the main results from our logistic regression are that the strengthening of financial capital is a central element explaining bank efficiency in the GCC region. On the other hand, the erosion in loan quality reduces banking sector efficiency. Overall, the policy implication is that regulations in the region need to focus on building a safe and sound banking system with adequate and prudential rules, and this should ultimately feed into improved banking sector efficiency levels.

7.7 Summary and conclusions

In this chapter we have discussed the results of the empirical analysis undertaken concerning banking sector efficiency in the GCC banking industry over the period 1995-2000. We estimate three efficiency measures: cost efficiency, standard profit, and alternative profit efficiency. The efficiency measures are obtained by applying the stochastic Fourier Flexible model, using both the half-normal and exponential distributions.

Two model specifications are used: the traditional model (specified as the Fourier Flexible including two outputs, two input prices, and a time trend) and the preferred model (the same as the former with the addition of variables controlling for asset quality and risk) given in Eq. (7.1).

In estimating the models, the parameter estimates are generally found to conform to expectations. Moreover, the inefficiency factor is found to make a major contribution in the stochastic term, suggesting that managerial errors in choosing optimal input and output mixes are responsible for deviations from the best-practice firm's cost and profit frontiers.

The findings also show that the risk and quality factors provide information influencing bank inefficiency levels when we use either the cost or profit function models. When risk and quality factors are considered, the mean inefficiency measures show a slight decrease.

Although the parameter estimates are similar, the exponential and half-normal efficiency estimates derived from the standard and alternative profit estimates are comparable, but this is less so for the cost efficiency estimates. In order to arrive at consistent results, we compare efficiency estimates derived using the distributional-free approach with our cost inefficiency measures. As they tend to be almost identical, this suggests that the half-normal cost inefficiency measures are more likely to concur with 'actual' cost inefficiencies in the system.

The mean cost efficiency from the preferred model is about 91 per cent (from the half-normal and distribution-free estimates). Both the standard and alternative profit inefficiency results indicate that nearly one third of the profits that could be earned by the best-practice bank are lost to inefficiency. Overall, the levels of the mean cost and profit inefficiency are consistent with inefficiency levels found in previous parametric studies on European and US banking.

Foreign banks are found to be less cost efficient, but more profit efficient than national banks. This suggests that foreign banks focus more on revenue generating than do their domestic counterparts. As foreign banks tend to have a different business mix (high end retail clients, large corporate banking services, and so on), it is perhaps not surprising that they are found to be less cost efficient but more profit efficient. Moreover, large banks are less cost inefficient than medium and small banks, and small banks are less profit inefficient than medium and large banks.

The rank-order coefficients show a close association between inefficiency estimates across different model specifications. The association is also close for inefficiency estimates using different distributional assumptions. Rank correlation is also undertaken to investigate the relationship between estimated inefficiency results and financial ratios. The relationship between inefficiency and profitability is almost consistent. That is, profit inefficiency and profitability are negatively related. In the same way, the relationship between costs and inefficiency measures is consistent because the positive rank correlation suggests that the more cost- and profit-inefficient banks incur higher costs.

The sample shows that the GCC banking industry has been exhibiting scale diseconomies driven mostly by banks from the UAE, Oman, and Qatar. Scale economies are prevalent for larger GCC banks and diseconomies prevail for small and medium-sized banks. Moreover, small banks show more scale diseconomies than do medium-size banks. Scale economies have also been calculated for the foreign banks operating in the GCC countries. The results show that foreign banks have been operating with, on average, higher scale diseconomies than national banks over the six-years study period.

In comparing between cost X-inefficiency and scale inefficiency, X-inefficiencies are consistently larger than scale inefficiencies during the study period. This result is found to be consistent with findings of other studies such as those of Berger (1993), Berger and Humphery (1991), and Evanoff and Israilevich (1991). This result also suggests that banks need to improve their managerial practices as a priority in order to increase the efficiency of their performance.

When looking at scale inefficiency in terms of the size of banks, the results indicate that large banks are the most scale efficient in the GCC banking industry. The results also show that medium-size banks are more scale efficient than their smaller counterparts. With regard to scale inefficiency comparisons between foreign and national banks, GCC national banks are found to be much more scale efficient than foreign banks.

Lastly, in the logistic regression, cost and profit inefficiencies are found to be negatively related with risk. There is also evidence that inefficiency is positively related to loan quality variables, suggesting that banks with enhanced financial capital and high loan quality are more efficient.

CHAPTER 8

CONCLUSIONS

8.1 Introduction and summary

This thesis examines the efficiency of GCC banking systems between 1995-2000. Given the ongoing deregulation process, it is important to have an indication of the efficiency features of GCC banks in order to evaluate the influence of financial reforms that aim to improve the soundness and enhance competitiveness of the GCC financial systems overall.

Over the last thirty years, GCC countries have made significant progress in laying down the foundations of modern economies and financial systems. Driven largely by oil export revenues, the GCC region has experienced among the highest rates of economic growth and per capita income in the world. Their economies have been characterised by low domestic income diversification, low inflation rates, stable exchange rate policies, high dependence on foreign labour, and a major role played by governments in the economic growth process. Certain policies have been undertaken aimed at restructuring these economies. Such reforms have aimed at economic diversification, the privatisation of public enterprises, the encouragement of greater participation of endogenous labour, and the relaxation and reformation of investment rules aimed at encouraging foreign direct investment.

With regard to financial system development, GCC countries have made significant progress in building the infrastructure of their financial systems resulting in high rates of financial deepening that fulfil the growing financial needs of the real sector of the economy. Before the existence of domestic banks, the early presence of foreign banks helped in shaping the banking and financial systems of the GCC countries, and induced them to set out or establish a more modern financial sector architecture. Foreign banks in most GCC countries continue to offer financial services that aim to mobilize oil revenues and private sector funds to more productive destinations, and these nowadays compete together with a significant number of domestic operators.

The GCC financial systems are characterised by relatively high levels of financial deepening, capitalisation, and deposit bases, and have experienced increasing levels of profitability in the second half of the 1990s. GCC financial systems have also adopted reform policies aimed at unleashing competitive forces and improving the regulatory structure of the respective markets. GCC countries have also made significant progress by increasing the openness within the banking systems, coordinating between monetary policies of intra-GCC countries, and by adopting appropriate legislation that will enable them to obtain economic and monetary union and a single GCC currency by 2010. All these measures should help enhance the competitive environment in which banks operate. As a result, there is a greater pressure in banks to improve their cost of operation and to adopt the use of more sophisticated technologies that increase productivity with the ultimate goal of improving overall efficiency of the financial system.

The empirical analysis undertaken in this thesis seeks to examine the cost and profit efficiency features of GCC banks, both domestic and foreign, in order to help inform the debate concerning financial sector reform in these countries.

8.2 The main findings on GCC banks efficiency

This thesis examines the cost and profit efficiency features of 93 banks between 1995 and 2000. The main findings of our study are as follows:

- The results show that the mean cost efficiency of GCC banks is about 91 per cent. The economic interpretation of the cost inefficiency level is that, given their particular output level and mix, on average, banks need to reduce their production costs by roughly 9 per cent in order to use their inputs as efficiently as possible. In addition, the profit efficiency results derived from both standard and alternative profit functions indicate that nearly one third of the profits that could be earned by the best practice bank are lost to inefficiency.

- There are no major differences in banks efficiency levels among GCC countries. Moreover, the mean efficiency across countries shows almost a stable trend over the study period 1995-2000. Moreover, our analysis suggests that the low and negative correlations between cost efficiency scores and profitability may indicate the presence of market power in the banking industry. The results also show that profit inefficiency and accounting profitability (ROA) are negatively related.

- Comparisons of inefficiency levels across bank ownership type and assets size reveal various findings. Domestic banks are more cost efficient than foreign banks; in contrast, foreign banks are more profit efficient than domestic banks. In terms of bank size, large banks are found to be more cost efficient than medium-sized and small banks. On the other hand, small banks are more profit efficient than medium and large banks. In fact, scale effects may induce profit inefficiency because large banks may face more difficulty in generating revenues efficiently compared to small banks. Moreover, this finding is also consistent with the conventional fact that small banks typically have higher profitability ratios than larger banks. Therefore, countries that have relatively small banks, such as the UAE and Qatar, tend to show higher cost inefficiency but lower profit inefficiency. On the other hand, banking industries that are dominated by larger banks, such as those in Saudi Arabia, Bahrain, and Kuwait, tend to show lower cost inefficiency but higher profit inefficiency. This scale effect could also explain differences in the inefficiency of foreign and domestic banks. For instance, the majority of foreign banks operating in the GCC countries are classified in terms of size as small to medium banks. Therefore, foreign banks are found to be less cost efficient but more profit efficient than domestic banks.

- The results also indicate that foreign banks have on average been operating with higher scale diseconomies than domestic banks over the six-year study period. Moreover, scale diseconomies decline as the assets sizes of both domestic and foreign banks increase. With regard to scale-inefficiency comparisons between foreign and domestic banks, the results also show that GCC domestic banks are much more scale efficient than foreign banks.

- Combining the findings on scale economies and scale inefficiencies, the results indicate that, overall, large and medium banks (including large foreign banks) better exploit their resources in terms of realizing economies of scale and scale efficiency, while small banks have substantial diseconomies and scale inefficiencies. This suggests that small banks should work on increasing their size in order to reach the optimal scale, possibly through further M & A activities. The findings also indicate that countries that have, on average, large banks tend to exhibit greater scale economies and less scale inefficiencies in their banking industry.

- Comparing cost X-inefficiency and scale inefficiency, X-inefficiencies are consistently larger than scale inefficiencies during the study period. This result seems to suggest that banks need to focus on improving their managerial practices as a priority in order to increase the efficiency of their performance. Gaines from improving X-efficiency (on both the cost and profit sides) will yield greater performance improvements than if banks just look to get larger per se.

- In the logistic regression, cost and profit inefficiency are found to be negatively related to risk variable. There is also evidence that inefficiency is positively related to loan quality variables, confirming that banks with enhanced financial capital and high loan quality are more efficient.

8.3 Recommendations

In the light of the findings of this study, a number of recommendations can be provided.

GCC banks need to reduce the inefficient utilization of resources in order to improve their efficiency status and increase their productivity. Our results suggest that while attention needs to be paid to bank costs and revenues, greater attention needs to be paid to the latter (given the high levels of profit inefficiency in the system). To ensure more effective control of costs, banks need to adopt and update such things as information systems and transaction processing systems. In order to improve revenues greater strategic attention needs to be paid to product innovation, pricing, and improvement in staff skills (including selling skills). Greater attention needs to be paid to improving customer satisfaction in order to boost revenues and therefore profits.

Probably restrictions (such as licensing limit and branching limits, as well as taxes) imposed on foreign bank activities may have to be reduced so that they can expand in order to exploit scale economies. It is suggested that increasing competitive pressures may eventually force GCC governments to remove all restrictions imposed on foreign banks so they can compete on an equal footing with their domestic peers. The gradual removal of restriction on foreign banks operation could act as an impetus for greater domestic bank competition and as precursor for greater financial system openness.

Moreover, GCC governments need not differentiate between foreign and domestic banks in seeking finances to fulfil their financial needs from domestic banks rather than foreign banks. Instead, these governments should follow market rules to encourage financial service quality which leads all banks to compete in providing better service quality for all users of financial service products in the region.

Finally, greater consolidation in the industry could be encouraged between GCC banks. (While consolidation may increase profit inefficiency, these inefficiencies are unlikely to be much bigger than other sized banks). In essence, larger GCC banks will be in a position to realise greater scale economies. Moreover, larger bank size and levels of banking sector competition will help allay policy-makers fears concerning greater financial system openness.

8.4 Contributions and limitations

The main contribution of this thesis is that it is the first to examine the efficiency of domestic and foreign banks across GCC banking sectors. While recent studies (such as Al-Jarrah, 2002; and Al-Shammari, forthcoming), examine various aspects of Arab and GCC banks, neither tackle the issue of foreign banks efficiency. GCC banking systems, to some extent, share similar policies, regulations, and economic characteristics. They also aim to reach more integrated financial systems as they approach the economic and monetary union arrangements. Understanding the role and behaviour of foreign banks is, therefore, important as it helps us to evaluate their possible role in the future integration process. It also informs us as to the possible direction of policy concerning banking sector reconfiguration in the future.

Last but not least, the thesis is not without its limitations and these have already been mentioned in Chapters 6 and 7. These are as follows.

The sample set of data used to estimate the GCC banking efficiency includes data on foreign banks operating in Qatar and the UAE. Although foreign banks are prohibited from banking in Saudi Arabia and Kuwait, there are foreign banks licensed to operate in Oman and Bahrain; however, financial information on all these foreign banks is not available. Unavailability of data, therefore, may bias our results.

The stochastic frontier approach (using either the translog and/or Fourier Flexible forms) can be sensitive to number of variables included in the models that used to estimate cost and profit efficiencies. As a result, risk and asset quality interactive terms are not included in the model. Moreover, a wider range of variables (such as off-balance sheet items and fixed assets) that may capture more accurately the production features of GCC banks are not included in our analysis.

Although this thesis discusses economies of scope (as a feature of operational efficiency), we do not estimates these for our sample of GCC banks. This is because of data limitations. Further efforts could be made to obtain appropriate information in order to investigate such scope economies.

Finally, this thesis uses parametric rather than non-parametric approach to estimate bank efficiency, and the bank production process is defined using the intermediation rather than the production approach. It may be appropriate for non-parametric estimates of bank efficiency to be undertaken for comparison purposes. In addition, we could also consider using production approaches to model efficiency and also the influence of technical change in GCC banking systems.

BIBLIOGRAPHY

Afriat, S. (1972) "Efficiency Estimation of Production Functions", *International Economic Review*, 13, 568-598.

Aigner, D. and S. Chu (1968) "On Estimating the Industry Production Function", *American Economic Review*, 58, 826-839.

Aigner, D., C. Lovell, and P. Schmidt (1977) "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, 6 (May), 21-37.

Akerlof, G. (1970) "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84(3), 488-500.

Alhadeff, D.A. (1954) "*Monopoly and Competition in Banking*", Berkeley, University of California Press.

Aliber, R., (1984) "International Banking: A Survey", *Journal of Money, Credit and Banking*, 16, 661-712.

Al-jarrah, I. M. (2002) "*Efficiency in Arabian Banking*", PhD, University of Wales, Bangor.

Allen, F. and A. Santomero (1998) "The Theory of Financial Intermediation", *Journal of Banking and Finance*, 21, 1461-1485.

Allen, L. and A. Rai (1996) "Operational Efficiency in Banking: An International Comparison", *Journal of Banking and Finance*, 20, 655 - 672.

Al-Sahlawi, K. (1997) "*The Role of Industrial Development Banks in Financing and Promoting Technological Change: The Case of the Saudi Industrial Development fund*", PhD, University of Wales, Bangor.

Al-Shammary, S. (2003) "*Structure-Conduct-Performance and the Efficiency of the GCC Banking Markets*", PhD (forthcoming), University of Wales-Bangor.

Al-Sharrah, R. (1999) "*The Specialised Banks in the Arab Gulf Countries*", Al-Falah Publishing, Kuwait. (In Arabic)

Al-Suhaimi, J. (2001) "Consolidation, Competition, Foreign Presence and Systemic Stability in the Saudi Banking Industry", *BIS Papers*, 4 August 2001.

Altunbas, Y. and S. Chakravarty (2001) "Frontier Cost Functions and Bank Efficiency", *Economics Letters*, 72, 233-240.

Altunbas, Y. and P. Molyneux (1996) "Economies of Scale and Scope in European Banking", *Applied Financial Economics*, 6, 367-375.

Altunbas, Y. and P. Molyneux (1994b) "Sensitivity of Stochastic Frontier Estimation to Distributional Assumptions: the Case of the German Banks", *Institute of European Finance Publication/Bangor*, Discussion paper, Un-published.

Altunbas, Y., E. Gardener, and P. Molyneux (1996) "Cost Economies and Efficiency in EU Banking Systems", *Journal of Economics and Business*, 48, 217-230.

Altunbas, Y., E. Gardener, P. Molyneux, and B. Moore (1998) "Efficiency in European Banking", 'University of Wales, Bangor, Un-published', (WP 2/98).

Altunbas Y., E. Gardener, P. Molyneux., and B. Moore (2001) "Efficiency in European Banking", *European Economic Review*, 45(10) 1931-1955

Altunbas, Y., L. Evans, and P. Molyneux (1994) "Universal Banks, Ownership and Efficiency: A Stochastic Frontier Analysis of the German Banking Market", University of Durham Economics Department, Un-published working paper.

Altunbas, Y., L. Evans, and P. Molyneux (1997) "Bank Ownership and Efficiency", 'Institute of European Finance, University of Wales, Bangor, Unpublished', (WP 18/97).

Altunbas, Y., L. Evans, and P. Molyneux (1997) "Bank Ownership and Efficiency", Institute of European Finance, University of Wales, Bangor, Unpublished, (WP 18/97).

Altunbas, Y., L. Evans, and P. Molyneux (2001) "Bank Ownership and Efficiency", *Journal of Money, Credit and Banking*, 33(4), 926-54.

Altunbas, Y., M. Liu, P. Molyneux and R. Seth (2000) "Efficiency and Risk in Japanese Banking", *Journal of Banking & Finance*, 24, 1605-1628.

Altunbas, Y., O. Fazylov, and P. Molyneux (2002) "Evidence on the Bank Lending Channel in Europe", *Journal of Banking & Finance*, 26, 2093–2110

Altunbas, Y., P. Molyneux, and N. Murphy (1995b) "Privatisation, Efficiency and Public Ownership in Turkey - An Analysis of the Banking Industry 1991-1993", *Institute of European Finance Publication/Bangor.*

Aly, H., R. Grabowski, C. Pasurka, and N. Rangan (1990) "Technical, Scale, and Allocative Efficiencies in U.S. Banking: An Empirical Investigation", *The Review of Economics and Statistics*, 72, pp. 211-218.

Anthony, J. (1975) "*Arab States of the Lower Gulf: People, Politics, Petroleum*", Middle East Institute, Washington.

Arrow, K., H. Chenery, B. Minhas, and R. Solow. (1961) "Capital Labor Substitution and Economic Efficiency", *The Review of Economics and Statistics*, 43, No. 3, August 1961, 225-250.

Athanassopoulos, A.D. (1997) "Service Quality and Operating Efficiency Synergies for Management Control in the Provision of Financial Services: Evidence from Greek Bank Branches", *European Journal of Operational Research*, 98, 300 - 313.

Baer, H. (1990) "Foreign Competition in U.S. Banking Markets", *Economic Perspectives*, Federal Reserve Banks of Chicago, 14, 22-29.

Bahrain Monetary Agency (1994) "*Bahrain, An International Financial Centre*", Bahrain.

Bahrain Monetary Agency (2001) "*Annual Report*", Kingdom Bahrain.

Bahrain Monetary Agency (2002) "*Bahrain Dinar Digest*", Vol. 1 (January), Bahrain.

Bain, J. (1956) "*Barriers to New Competition*", (Harvard University Press, Cambridge, Massachusetts).

Baldini, D. and A. Landi (1990) "Economie di scala e complementarieta' di costo nell", industria bancaria italiana', L'Industria, 1, 25-45.

Baltagi, B. (2001) "*Econometric Analysis of Panel Data*", Second edition, John Wiley & Sons, New York.

Baltagi, B. and J.M. Griffin (1988) "A General Index of Technical Change", *Journal of Political Economy*, Vol. 96, No. 1, 20-41.

Bank for International Settlements (2001) "Development and Restructuring of the Saudi Banking System", BIS Policy Papers, Sep. 1999.

Bank for International Settlements (http://www.bis.org/publ/bcbsc312.pdf).

Banker, R., A. Charnes, and W. Cooper (1984) "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", Management Science, 30(9), 1078-1092.

Barajas, A., R. Steiner, and N. Salazar (2000) "Foreign Investment in Colombia's Financial Sector", in Claessens, S. and Marion Jansen, (eds.) TheInternationalization of Financial Services: Issues and Lessons for Developing Countries, Boston, Mass.: Kluwer Academic Press.

Barnes, P. and C. Dodds (1983) "The Structure and Performance of the UK Building Society Industry 1970-78", *Journal of Business, Finance and Accounting*, 10, 37-56.

Barth, J. R., G. Caprio, and R. Levine (2001b) "Bank Regulation and Supervision: What Works Best?", World Bank, mimeo.

Battese, G. and T. Coelli (1992) "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", The Journal of I Productivity Analysis, 3, 153-169.

Battese, G. and T. Coelli (1993) "A Stochastic Frontier Production Function Incorporating A Model For Technical Inefficiency Effects", Centre for Efficiency and Productivity Analysis (CEP A), Department of Econometrics, University of New England, Armidale, Australia, (WP 69/93).

Battese, G. and T. Coelli (1995) "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", *Empirical Economies*, 20, 325-332.

Bauer, P., A. Berger, G. Ferrier, and D. Humphrey (1997) "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods", *Finance and Economics Discussion Series*, Federal Reserve Board, USA, 50, pp. 85-144.

Bauer, P., A. Berger, G. Ferrier, and D. Humphrey (1998) "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods", *Journal of Economics and Business* Volume, 50(2), 85-114.

Bauer, P.W., Berger A.N., and D.B. Humphrey (1993) "Efficiency and Productivity Growth in U.S. Banking", in Fried H.O., Lovell C.A.K. and S.S. Schmidt (Eds.) *'The Measurement of Productive Efficiency: Techniques and Applications'* Oxford University Press, UK.

Baumol, W., J. Panzar, and R. Willig (1988) "*Contestable Markets and the Theory of Industry Structure*", Harcourt Brace Jovanovich, New York, (1988) Revised Edition.

Baumol, W.J., Panzar, J.C., and R.D. Willig (1982) "Contestable Markets and the Theory of Industry Structure", Revised Edition, New York, Harcourt Brace Jovanovich.

Bell, F. and N. Murphy (1968), "Economies of Scale and Division of Labour in Commercial Banking", *Southern Economic Journal*, October, 131-139.

Benston G.J. (1965a) "Economies of Scale and Marginal Costs in Banking Cperations", *National Banking Review*, 2(4), June, 507 - 549.

Benston, G. (1965b) "Branch Banking and Economies of Scale", *Journal of Finance*, May, 312-331.

Benston, G., A. Berger, G. Hanweck and D. Humphrey (1983) "Economies of Scale and Scope", Proceedings of the Conference on Bank Structure and Competition, Chicago, Federal Reserve Bank of Chicago.

Benston, G., G. Hanweck, and D. Humphrey (1982a) "Scale Economies in Banking: a Restructuring and Reassessment", *Journal of Money, Credit and Banking*, 14, 435-456.

Benston, G., G. Hanweck, and D. Humphrey, (1982b) "Operating Costs in Commercial Banking", Federal Reserve Bank of Atlanta, *Economic Review*, pp. 6-21.

Berg, S., F. Forsund and N. Bukh (1995) "Banking Efficiency in the Nordic Countries: a few-country Malmquist index analysis", mimeo.

Berg, S., F. Forsund, and E. Jansen, (1992) "Malmquist Indices of Productivity Growth during the Deregulation of Norwegian Banking, 1980-89", Scand. J. of Economics, 94, Supplement, 211-228.

Berg, So, F. Forsund, L. Hjalmarsson and M. Suominen (1993) "Banking Efficiency in the Nordic Countries", *Journal of Banking and Finance*, 17 (2-3), April, 371-388.

Berger A.N. (1993) 'Distribution-Free estimates of efficiency in the US banking industry and tests of the standard distributional assumptions', *Journal of Productivity Analysis*, No. 4, 261-292, September.

Berger A.N., Hanweck G.A. and D.B. Humphrey (1987) "Competitive Viability in Banking. Scale, Scope and Product Mix Economies", *Journal of Monetary Economics*, 20, 501-520.

Berger, A. (1995) "The Profit-structure Relationship in Banking-tests of Market Power and Efficient-structure Hypothesis", *Journal of Money, Credit and Banking*, 27(2), 404-431.

Berger, A. (1995) "The Profit-Structure Relationship in Banking - Tests of Market Power and Efficient Structure Hypotheses", *Journal of Money Credit and Banking*, 27, 404-31

Berger, A. and D. Humphrey (1990) "Measurement and Efficiency Issues in Commercial Banking", *Finance and Economics Discussion Series*, 151, Federal Reserve Board, USA.

Berger, A. and D. Humphrey (1991) "The Dominance of Inefficiencies Over Scale and Product Mix Economies in Banking", *Journal of Monetary Economics*, 28(1), 117- 148.

Berger, A. and D. Humphrey (1992) "Measurement and Efficiency Issues in Commercial Banking", in Zvi Griliches (Eds.) 'Output Measurement in the Service Sectors' National Bureau of Economic Research Studies in Income and Wealth, 56, University of Chicago Press, Chicago.

Berger, A. and D. Humphrey (1997) "Efficiency of Financial Institutions: International Survey and Directions for Future Research", *European Journal of Operational Research*, 98, 175-212.

Berger, A. and L. Mester (1997) "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?", *Journal of Banking & Finance*, 21, 895- 947.

Berger, A. and L. Mester (1997) "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions", *Journal of Banking and Finance*, 21, 895-947.

Berger, A. and R. DeYoung (1998) "Problem Loans and Cost Efficiency in Commercial Banks", Federal Reserve Board, Washington, D.C., (WP 8/1997).

Berger, A. and R. DeYoung (2000) "The effects of geographic expansion on bank efficiency", Federal Reserve Bank of Chicago, *Working Paper Series*, No. WP-00-14.

Berger, A., G. Hanweck, and D. Humphrey (1987) "Competitive Viability in banking: scale, scope and product mix economies", *Journal of Monetary Economics*, 20, 501-520.

Berger, A., G. Hanweck, and D. Humphrey (1993a) "Bank Efficiency Derived from the Profit Function", *Journal of Banking & Finance*, 17, 317-347.

Berger, A., Hanweck G., and D. Humphrey (1987) "Competitive Viability in Banking. Scale, Scope and Product Mix Economies", *Journal of Monetary Economics*, 20, 501-520.

Berger, A., J. Leusner, and J. Mingo (1997b) "The Efficiency of Bank Branches", *Journal of Monetary Economics*, 40.

Berger, A., Kashyap, A., and Scalise, J. (1995) "The Transformation of the U.S. Banking Industry: What a Long Strange Trip It's Been", Brookings Papers on Economic Activity, 2: 55-218.

Berger, A., L. Klapper, and G. F. Udell (2000) "The Ability of Banks to Lend to Informationally Opaque Small Businesses", Mimeo. 35

Berger, A., R. DeYoung, and G. Udell (2000) "Efficiency Barriers to the Consolidation of the European Financial Services Industry", *European Financial Management*, 6.

Berger, A., R. DeYoung, H. Genay, and G. Udell (2000) "Globalization of Financial Institutions: Evidence from Cross-Border Banking Performance", *Brookings Papers on Economic Activity*, 2: 23-158.

Berger, A., W. Hunter, and S. Timme (1993b) "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present, and Future", *Journal of Banking & Finance*, 17, 221-249.

Berger, A.N., and D. Humphrey (1994) "Bank Scale Economies, Mergers, Concentration, and Efficiency: The U.S. Experience" Financial Institution Centre, The Wharton School, University of Pennsylvania.

Berger, A.N., and D. Humphrey (1997) "Efficiency of Financial Institutions: International Survey and Directions for Future Research", *European Journal of Operational Research*, 98, 175-212.

Berger, A.N., Hancock, D., and D.B. Humphrey (1993) "Bank Efficiency Derived from the Profit function", *Journal of Banking and Finance*, 17, 317-347.

Bernanke, B. (1981) "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression"", *American Economic Review*, 73(3), 257-276.

Berndt, E. and L. Christensen (1973) "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing, 1929-1968", *Journal of Econometrics*, 1, 81-114.

Besanko, D., D. Dranove, and M. Shanley (2000) "*Economics of Strategy*", John Wiley & Sons, US.

Bhattacharya, A., C. Lovell, and P. Shay (1997) "The Impact of Liberalisation on the Productive Efficiency of Indian Banks", *European Journal of Operational Research*, 98, 332-345.

Binger, B. and E. Hoffman (1988) "Microeconomics with Calculas", Harper Collins Publishers, US.

Bourke, P. (1989) "Concentration and Other Determinants of Bank Profitability in Europe, North America and Australia", *Journal of Banking & Finance*, 13, 65-79.

Boyd, J. and E. Prescott (1986) "Financial Intermediation-Coalitions", *Journal of Economic Theory*, 38(2), 211-232.

Breusch, T. and A. Pagan (1980) "The LM Test and Its applications to Model Specification in Econometrics", *Review of Economic Studies*, 47, 239-254.

Caprio, G. and R. Cull (2000) "Bank Privatization and Regulation for Egypt", Egyptian Center for Economic Studies.

Caprio, G., and Summers, L. (1993) "Finance and Its Reform: Beyond Laissea-Faire", World Bank.

Casson, M. (1989) "The Economic Theory of Multinational Banking: An Internationalisation Approach", *University of Reading Working Paper*, No. 133, Reading.

Casu, B. and C. Girardone (1998) "A Comparative Study of the Cost Efficiency of Italian Bank Conglomerates", Institute of European Finance, *University of Wales, Bangor, Un-published*, (WP 3/98).

Casu, B. and P. Molyneux (1999) " A Comparative Study of Efficiency in European Banking", 'European Association of University Teachers of Banking and Finance, Wolpertinger Seminar, Lisbon-Portugal'.

Cavello, L. and S. Rossi (2001) "Scale and Scope Economies in the European Banking Systems", Journal of Multinational Financial Management, 11, 515–531.

Central Bank of Kuwait (2001) "*Annual Report*", Kuwait.

Central Bank of Oman (2001) "*Annual Report*", Oman.

Central Bank of Qatar (2001) "*Annual Report*", Qatar.

Cetorelli N. and M. Gambera. (1999) "Banking Market Structure, Financial Dependence and Growth: International Evidence from Industry Data", Wharton School Center for Financial Institutions, University of Pennsylvania, 1999/10, http://fic.wharton.upenn.edu/fic/papers/00/0019.pdf

Cetorelli, N. and Peretto, P. (2000) "Oligopoly Banking and Capital Accumulation", Federal Reserve Bank of Chicago,< http://www.chicagofed.org/publications/workingpapers/papers/wp2000_12.pdf>

Charnes, A., W. Cooper, and E. Rhodes (1978) "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2, 429-444.

Chen, Y., J. Mason, and E. Higgins (2001) "Economies of Scale in the Banking Industry: The Effects of Loan Specialization", Department of Business Administration and Education, Emporia State University, Emporia, KS.

Chen, Z., and Khan, M. (1997) "Patterns of Capital Flows to Emerging Markets: A Theoretical Perspective", INF, WP/97/13, < http://www.imf.org/external/pubs/ft/wp/wp9713.pdf>.

Cho, K., S. Krishnan, and D. Nigh (1987) "The State of Foreign Banking Presence in the United States", *International Journal of Bank Marketing*, 5(2): 59-75.

Chow, G.C. (1960) "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", *Econometrica*, 28, 591-605.

Claessens, S., A. Demirgüç-Kunt, and H. Huizinga, (1998) "How Does Foreign Entry Affect the Domestic Banking Market?", *Journal of Banking and Finance*, Vol. 25, 891-911.

Clark, J. (1984) "Estimation of Economies of Scale in Banking Using a Generalised Functional Form", *Journal of Money, Credit, and Banking*, 16(1), 53-68.

Clark, J. (1988) "Economies of Scale and Scope At Depository Financial Institutions: A Review of the Literature", Federal Reserve Bank of Kansas City, *Economic Review,* 73, 16-33.

Clark, J. (1996) "Economic Cost, Scale Efficiency, and Competitive Viability in Banking", *Journal of Money, Credit, and Banking* 28 (3).

Clarke, G., R. Cull, M. Martinez, and S. Sanchez (2001) "Bank Lending to Small Businesses in Latin America: Does Bank Origin Matter?", *Mimeo*, World Bank, Washington DC.

Coelli, T., D. Prasada, and G. Battese (1998) "*An Introduction to Efficiency and Productivity Analysis*", Kluwer Academic Publishers.

Conigliani, C., R. De Bonis, G. Motta, G. Parigi (1991) "Economie di scala e di diversificazione nel sistema bancario", Banca d'Italia, Temi di discussione 150.

Cooper, J. (1980) "Economics of Scale in the UK Building Society Industry", *Investment Analysis*, 55, 31-36.

Cossutta, D., M.L, Di Battista, C. Giannini and G. Urga (1988) "Processo produttivo e struttura dei costi nell'industria bancaria italiana", in Banca e Mercato a cura di F. Cesarini. M. Grillo, M. Monti, M. Onado (Bologna: eds. Il Mulino).

Crystal, J. (1990) "*Oil and Politics in the Gulf: Rulers and Merchants in Kuwait and Qatar*", Cambridge University Press, Cambridge.

Crystal, J., B. Dages, and L.Goldberg. (2001) "Does Foreign Ownership Contribute to Sounder Banks in Emerging Markets? The Latin American Experience", Federal Reserve Bank of New York, *Staff Reports*, No. 137, New York.

Cummins, J. and M. Weiss (1998) "Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency Methods", *The Wharton School*, University of Pennsylvania.

Dages, G., L. Goldberg, and D. Kinney (2000) Federal Reserve Bank of New York Economic Policy Review, September.

DeBorger, B., G. Ferrier, and K. Kerstens (1995) "The Choice of a Technical Efficiency Measure on the Free Disposal Hull reference Technology: A Comparison Using Banking Data", *Working paper,* University of Arkansas, USA.

Demirgüc-Kunt, A. and R. Levine (1996) "Stock Market Development and Financial Intermediaries: Stylised Facts", The World Bank Economic Review, 10(2), 291-321.

Demirgüc-Kunt, A., and R. Levine (2001). Bank-based and market-based financial systems: Cross-country comparisons. in "Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets, and Development", (Asli Demirguc-Kunt and Ross Levine, Eds.), pp. 81-140. MIT Press, Cambridge, MA.

Demsetz, H. (1973) "Industry Structure, Market Rivalry and Public Policy", *Journal of Law and Economics*, 16, 1-9.

Deprins, D., L. Simar and H. Tulkens (1984), "Measuring Labour-Efficiency in Post Offices", in M. Marchand, P. Pestieau and H. Tulkens (Eds.), *The Performance of Public Enterprises: Concepts and Measurements*, North-Holland, Amsterdam.

DeYoung, R. (1997) "Comment on Operational Efficiency in Banking: An International Perspective", *Journal of Banking & Finance*, 21, 1325-1329.

DeYoung, R. (1998) "Management Quality and X-Inefficiency in National Banks", *Journal of Financial Services Research*, 13, 5-22.

DeYoung, R. and D. E. Nolle (1996) "Foreign-Owned Banks In The United States: Earning Market Share Of Buying It?", *Journal of Money, Credit, and Banking*, 28(4), 622-36.

DeYoung, R. and D. Nolle (1996) "Foreign-owned banks in the US: Earning market share or buying it?", *Journal of Money, Credit and Banking*, 28(4), 622-636.

Dietsch, M. (1988) "Economies d'echelle et economies d'envergure dans les banques de depots francaises", 'mimeo, Istitut d'Etudes Politiques de Strasbourq'.

Dietsch, M. (1993) "Economies of Scale and Scope in French Commercial Banking Industry", *Journal of Productivity Analysis*, 1.

Dietsch, M. and L. Weill (1998) "Banking Efficiency and European Integration: Productivity, Cost and Profit Approaches", Paper presented aNhe 21st Colloquium of the SUERF, Frankfurt, 15-17.

Diewert, W. (1973) "Functional Forms for Profit and Transformation Functions", *Journal of Economic Theory*, 7, 284-316.

Drake, L. (1992) "Economies of Scale and Scope in UK Building Societies: An Application of the Translog Multiproduct Cost Function", *Applied Financial Economics*, 2, 211-219.

Drake, L. and R. Simper (1999) "X-efficiency and Scale Economies in Policing: A Comparative Study Using the Distribution Free Approach and DEA", Loughborough University, Economic Paper (7/97).

Dziobek, C. (1998) "Market-Based Policy Instruments for Systemaic Bank Restructuring", IMF working Paper WP/98/113.

Economist Intelligence Unit (1991) "Country Profile: Bahrain, Qatar, 1991-92", London.

Economist Intelligence Unit (1991) "*Country Profile: Saudi Arabia*, 1991-92", London.

Economist Intelligence Unit (1991) "Country Profile: United Arab Emirates, 1991-92", London.

Economist Intelligence Unit (1992) "Country Report: Kuwait", No. 1, London.

Economist Intelligence Unit (1992) "Country Report: Kuwait", No. 1, London.

Elyasiani, E. and S. Mehdian (1990a) "Efficiency in commercial banking industry, a production frontier approach", *Applied Economics*, 22, 539-551.

Elyasiani, E. and S. Mehdian (1990b) "A Non-parametric Approach to Measurement of Efficiency and Technological Change: The Case of Large US Banks", *Journal of Financial Services Research,* 4, 157-168.

Elyasiani, E. and S. Mehdian (1992) "Productive Efficiency Performance of Minority and Nonminority-Owned Banks: A Nonparametric Approach", *Journal of Banking and Finance*, 16, 933-948.

Elyasiani, E. and S. Mehdian (1995) "The Comparative Efficiency Performance of Small and Large US Commercial Banks in the Pre-and Post-Deregulation Eras", *Applied Economics*, 27, 1069-1079.

English, M., S. Grosskopf, K. Hayes, and S. Yaisawarng (1993} "Output Allocative and Technical Efficiency of Banks", *Journal of Banking & Finance*, 17, 349-366.

European Commission (1997) "The Single Market Review", Credit Institutions and Banking, Subseries II: Impact on Services, Vol. 3, (London: Kogan Page).

Evanoff, D. (1998) "Assessing the Impact of Regulation on Bank Cost Efficiency", Economic Perspective, Federal Reserve Bank of Chicago, 21-32.

Evanoff, D. and R. Israilevich (1991) "Productive Efficiency in Banking", Economic Perspective, Federal Reserve of Chicago, 15 (July/August), 11-32.

Evanoff, D. and R. Israilevich (1995) "Scale Elasticity Versus Scale Efficiency in Banking", *Southern Journal of Economics*, 61, 1036-1047.

Fama, E. (1970) "Efficient Capital Markets: a review of theory and empirical work, *Journal of Finance*, 25, 383-417.

Fanjul, O. and F. Maraval (1985) "La eficiencia del sistema bancario Espanol", Research Paper, Alianza University, Madrid.

Farrell, M. (1957) "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253-290.

Favero, C. and L. Papi (1995) "Technical Efficiency and Scale Efficiency in the Italian Banking Sector: A Non-parametric Approach", *Applied Economics*, 27,385-395.

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

Fieleke, N. (1977) "The Growth of U.S. Banking Abroad: An Analytic Survey", *Economic Review*, Federal Reserve Bank of Boston, 9-40.

Fisher, A. and P. Molyneux (1996) "A Note on the Determinants of Foreign Bank Activity in London between 1980 and 1989", *Applied Financial Economics*, 6(3): 271-77.

Focarelli, D. and A. Pozzolo (2000) "The Determinants of Cross-Border Shareholding: An Analysis with Bank-Level Data from OECD Countries", Paper presented at the Federal Reserve Bank of Chicago Bank Structure Conference, May.

Forestieri, G. (1993) "Economies of Scale and Scope in the Financial Services Industry: A Review of Recent Literature", in OECD (Ed.), 63-124, *Financial Conglomerates*, Paris.

Forsund, F., C. Lovell, and P. Schmidt (1980) "A Survey Of Frontier production Functions And Of Their Relationship To Efficiency Measurement", *Journal of Econometrics*, 13, 5-25.

Frankel, A., and P. Morgan (1992) "Deregulation and Competition in Japanese Banking", Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, 78 (18), 579-93.

Fry, M. (1988) "Money, Interest, and Banking in Economic development", The Johns Hopkins University Press, Maryland, 1.

Fry, M. (1995) "*Money, Interest, and Banking in Economic Development*", Baltimore, London: The John Hopkins University Press.

Fukuyama, H. (1993) "Technical and Scale Efficiency of Japanese Commercial Banks: Anon-Parametric Approach", *Applied Economics* 25, 1191-1112.

Fukuyama, H. (1995) "Measuring Efficiency and Productivity Growth in Japanese Banking: A Nonparametric Frontier Approach", *Applied Financial Economics*, 5, 95-117.

Gallant, A. (1981) "On the Bias in Flexible Functional Forms and Essentially Unbiased Form: The Fourier Flexible Form", *Journal of Econometrics*, 15,211-245.

Gallant, A. (1982) "Unbiased Determination of Production Technologies", Journal of Econometrics, 20, 285-324.

Gathon, H. and F. Grosjean (1991) "Efficacité productive et rendement d'échelle dans les banques belges", Cahiers Économiques de Bruxelles, 145 -160.

GCC Secretariat General (http://www.gcc-sg.org).

Gilligan, T. and M. Smirlock (1984) "An Empirical Study of Joint Production and Scale Economies in Commercial Banking", *Journal of Banking & Finance*, 8 (March), 67- 77.

Girardone, C., P. Molyneux, and E. Gardener (2000) "Analysing the Determinants of Bank Efficiency: The Case of Italian Banks", *SBARD*, University of Wales, Bangor.

Goddard, J., P. Molyneux and J. Wilson (2001) "*European Banking: Efficiency, Technology and Growth*", John Wiley: London.

Goldberg, L. (1992) "The Competitive Impact of Foreign Commercial Banks in the United States", in R. Alton Gilbert, ed., "*The Changing Market in Financial Services*", Proceedings of the Fifteenth Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis, Norwell, Mass and Dordrecht: Kluwer Academic. 161-200.

Goldberg, L. (2001) "When is U.S. Bank Lending to Emerging Markets Volatile?", *Mimeo*, Federal Reserve Bank of New York.

Goldberg, L. and A. Rai (1996) "The Structure-performance Relationship for European Banking", *Journal of Banking & Finance*, 20, 745-771.

Goldberg, L. and A. Saunders (1981a) "The Determinants of Foreign Banking Activity in the United States", *Journal of Banking and Finance*, 5, 17-32.

Goldberg, L. and A. Saunders (1981b) "The Growth of Organizational Forms of Foreign Banks in the U.S.: A Note", *Journal of Money, Credit, and Banking*, 13(3), 365-374.

Goldberg, L. and Grosse (1994) "Location Choice of Foreign Banks in the United States", *Journal of Economics and Business*, 46(5), 367-79.

Goldfeld, S. and R. Quandt (1965) "Some Tests for Homoscedasticity", *Journal of the American Statistical Association*, 60, 539-547.

Gough, T.J. (1979) "Building Society Mergers and the Size Efficiency Relationship", *Applied Economics*, 11, 185-194.

Grabowski, R., N. Rangan, and R. Rezvanian (1993) "Organizational Forms in Banking: An Empirical Investigation of Cost Efficiency", *Journal of Banking & Finance*, 17, 531-538.

Grabowski, R., N. Rangan, and R. Rezvanian (1994) "The Effect of Deregulation on the Efficiency of US Banking Firms", *Journal of Economics and Business*, 39-54.

Gramely, L. (1962) "A Study of Scale Economies in Banking", Federal Reserve Bank of Kansas City.

Gramley, L. (1962) "A Study of Scale Economies in Banking", Federal Reserve Bank of Kansas City, USA.

Greenbaum, S. (1967) "A Study of Bank Cost", *National Banking Review*, June, 415-434.

Greene, W. (1980) "Maximum Likelihood Estimation of Econometric Frontier Functions", *Journal of Econometrics*, 13, 27-56.

Greene, W. (1980) "On the Estimation of a Flexible Frontier Production Model", *Journal of Econometrics*, 13, 101-115.

Greene, W. (1990) "A Gamma-distributed Stochastic Frontier Model", *Journal of Econometrics*, 46, 141-164.

Greene, W. (1993) "The Econometric Approach to Efficiency Analysis", in Fried, H., C. Lovell and S. Shmidt (Eds.), The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York, 68-119.

Greene, W. (2000) "*Econometric Analysis*", 4th ed., Prentice-Hall, New Jersey.

Greenwald, B., J. Stiglitz, A. Weiss (1984) "Informational imperfections in the capital markets and macro-economic fluctuations", *American Economic Review,* 74 (1), 194-199.

Greenwood, J. and B. Jovanovic (1990) "Financial Development, Growth, and the Distribution of Income", *Journal of Political Economy*, 98 (October), 1076-107.

Grosse, R. and L. G. Goldberg (1991) "Foreign Bank Activity in the United States: An Analysis by Country of Origin", *Journal of Banking and Finance*, 15(6), 1092-1112.

Gujarati, D. (1995) "*Basic Econometrics*", 3rd ed., McGraw-Hill, New York.

Gujarati, D. (2002) "*Basic Econometrics*", 4th ed., McGraw-Hill, New York.

Gulf Business (2002) "Making things Work", August, 58-59.

Gulf Business (2002) "Tightening the Control", August, 12-14.

Hao, J., W. Hunter, and W. Yang (1999) "Deregulation and Efficiency: The Case of Private Korean Banks", Federal Reserve Bank of Chicago.

Hardwick, P. (1989) "Economies of Scale in Building Societies", Applied Economics, 21, 1291-1304.

Hardwick, P. (1990) "Multi-product Cost Attributes: A Study of UK Building Societies", Oxford Economic Papers, 42, 446-461.

Hasan, I. and W. C. Hunter (1996) "Efficiency of Japanese Multinational Banks in the United States", in Andrew H. Chen, ed., *Research in Finance*, Volume 14, Greenwich, CT and London: JAI Press. 157-173.

Heffernan, S. (1996) "*Modern Banking In Theory and Practice*", John Wiley & Sons, Chichester, England.

Hempel, G. and D. Simonson (1999) "*Bank Management: Text and Cases*", John Wiley and Sons: New York, Fifth Edition.

Hermalin, B. and N. Wallace. (1994) "The Determinants of Efficiency and Solvency in Savings and Loans", *Rand Journal of Economics*, 25, 361-381.

Hodgkins, B. and E. Goldberg (1981) "Effects of foreign acquisitions on the balance sheet structure and earning performance of US banks", in US Comptroller of the Currency, Foreign Acquisitions of US Banks, ed., R. F. Dame (Richmond, VA), 401-30.

Horovitz, P.M. (1963) "Economies of Scale in Banking", in P.M. Horovitz et al. "Private Financial Institutions" Englewood Cliffs, N.J., Pertince-Hall, 1-55.

Houpt, J. V. (1980) "Foreign Ownership and the Performance of U.S. Banks", Staff Study 109, (July), Board of Governors of the Federal Reserve System, Washington, D.C.

Hsiao, C. (1985) "Benefits and Limitations of Panel Data", *Econometric Reviews*, 4, 121-174.

Hughes, J. and L. Mester (1993) "A Quality and Risk-adjusted Cost Function for Banks: Evidence on the 'too-big-to-fail' Doctrine", *Journal of Productivity Analysis*, 4, 293-315

Hughes, J., L. Mester, and C. Moon (2000) "Are All Scale-Economies in Banking Elusive or illusive: Evidence Obtained by Incorporating Capital Structure and Risk Taking into Models of Bank Production", *The Wharton School*, University of Pennsylvania, (WP 33/00).

Hultman, C. and R. McGee (1989) "Factors Affecting the Foreign Banking Presence in the United States", *Journal of Banking and Finance*, 13(3), 383-96.

Humphrey, D. (1985) "Cost and Scale Economies in Bank Intermediation", in Aspinwall R. and Eisenbeis (Eds.) "*Handbook for Banking Strategies*", John Wiley, New York.

Humphrey, D. (1987) "Cost Dispersion and the Measurement of Economies in Banking", Federal Reserve Bank of Richmond, *Economic Review*, 24 (May/June), 5-34.

Humphrey, D. (1991) "Productivity in Banking and Effects from Deregulation", *Economic Review*, Federal Reserve Bank of Richmond, March/April, 16-28.

Humphrey, D. and L. Pulley (1997) "Banks' Responses to Deregulation: Profits, Technology, and Efficiency", *Journal of Money, Credit and Banking*, 29(1), 73-93.

Humphrey, D.B. (1987) "Cost Dispersion and the Measurement of Economies in Banking", Federal Reserve Bank of Richmond Economic Review, May/June, 24-38.

Humphrey, D.B. (1990) "Why Do Estimates of Banks Scale Economies Differ", Federal Reserve Bank of Richmond Economic Review, Vol. 76,38-50.

Hunter, W. and S. Timme (1986) "Technical Change, organizational form and the structure of bank production", *Journal of Money, Credit and Banking*, 18(2), 152-166.

Hunter, W. and S.G. Timme (1991) "Technological Change in Large US Banks", *Journal of Business*, 64, 339-362.

Hunter, W., S. Timme, and W. Yang (1990) "An Examination of Cost Sub Additivity and Multiproduct Production in Large U.S. Banks", *Journal of Money, Credit and Banking*, 22(4), 504-525.

Hymer, S. (1970) "The Efficiency (Contradictions) of Multinational Corporations", *American Economic Review*, 441-448.

Intarachote, T. (2000) "*Financial Liberalisation in Thailand*", PhD, University of Wales, Bangor.

Isik, I. and K. Hassan (2002) "Technical, Scale and Allocative Efficiencies of Turkish Banking Industry", *Journal of Banking & Finance*, 26, 719–766

Jbili, A., V. Galbis, and A. Bisat (1997) "Financial Systems and Reform in the Gulf Cooperation Council Countries", Middle Eastern Department, International Monetary Fund.

Jondrow, J., C. Lovell, I. Materov, and P. Schmidt (1982) "the Estimation of Technical Inefficiency in The Stochastic Frontier Production Function Model", *Journal of Econometrics*, 19, 233-238.

Kalish, L. and R. Gilbert (1973) "An Analysis of Efficiency of Scale and Organisation form in Commercial Banking", *Journal of Industrial Economics*, 21 (July), 293-307

Kaparakis, E., S. Miller, and A. Noulas (1994) "Short-run Cost Inefficiency of Commercial Banks: A Flexible Stochastic Frontier Approach", *Journal of Money, Credit, and Banking*, 26, 875-93.

Kaparakis, E., S. Miller, and A. Noulas (1994) "Short-run Cost Inefficiency of Commercial Banks: A Flexible Stochastic Frontier Approach", *Journal of Money, Credit and Banking*, 26, 875- 893.

Keeton, W. (1995) "Multi-Office Bank Lending to Small Businesses: Some New Evidence", Federal Reserve Bank of Kansas City, *Economic Review*, 80(2): 45-57. 39

Keynes, J. (1963) "*General Theory of Employment, Interest and Money*" Macmillan, London.

Kolari, J. and A. Zardkoohi (1987) "Bank Cost, Structure and Performance", Lexington, Mass, Lexington Books.

Kraft, E. and D. Tirtiroglu (1998) "Bank Efficiency in Croatia: A Stochastic-Frontier Analysis", *Journal of Comparative Economics*, 26, 282–300

Krugman, P. (1998) "Crises: The Next Generation?", < http://econ.tau.ac.il/research/sapir/Krugman.pdf>

Krugman, P. and Obstfeld, M. (1997) "*International Economics: Theory and Application*", 4th ed., Addison Wesley Longman, US.

Kumbhakar, S. (1997) "Modelling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship", *Journal of Econometrics*, 76, 351-356.

Kumbhakar, S. and C. Lovell (2000) "Stochastic Frontier Analysis" Cambridge University Press http://www.cup.cam.ac.uk.

Kumbhakar, S., S. Ghosh, and J. McGukin (1991) "A Generalized Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms", *Journal of Business and Economic Statistics*, 9, 279-286.

Kwan, S. and R. Eisenbeis (1994) "An Analysis of Inefficiencies in Banking: A Stochastic Cost Frontier Approach", Federal Reserve Bank of San Francisco, USA, Working paper (December).

Kwan, S. and R. Eisenbeis (1995) "An Analysis of Inefficiencies in Banking: A Stochastic Cost Frontier Approach", Journal of Banking & Finance, 19, 733-734.

Landi, A. (1990) "Dimensioni, costi e profitti delle banche italiane", Bologna, Il Mulino.

Lang, G. and P. Welzel (1996) "Efficiency and Technical Progress in Banking: Empirical Results for a Panel of German Cooperative Banks", *Journal of Banking & Finance*, 20, 1003-1023.

Lawrence, C. and R. Shay (1986) "Technology and Financial Intermediation in a Multiproduct Banking Firm: An Econometric Study of I.S.S. Banks, 1979-1982"; in C. Lawrence and R. Shay (eds), *Technological Innovation, Regulation and the Monetary Economy* (Cambridge, Mass.: Balinger).

Leibenstein, H. (1966) "Allocative Efficiency vs. X-efficiency", *The American Economic Review*, 56, 392-415.

Leightner, E. and C. Lovell (1998) "The Impact of Financial Liberalisation on the Performance of Thai Banks", *Journal of Economics and Business*, Vol. 50, 115-131.

Levine, R. (1996) "Foreign Banks, Financial Development, and Economic Growth", in Claude E Barfield, ed., *International Financial Markets: Harmonisation Versus Competition. Washington*, D.C.: AEI Press.

Levine, R. (2002) "Bank-Based or Market-Based Financial Systems: Which Is Better?", Journal of Financial Intermediation, 11(4), 398-428

Levonian, M. and J. Soller (1995) "Small Banks, Small Loans, Small Business", *Mimeo,* Federal Reserve Bank of San Francisco.

Levy-Garboua, L. and F. Renard (1977) "Une etude statistique de la rentabilite des banques en France en 1974", Cahiers Economiques et Manetaires, 5.

Lloyd-Williams, D., P. Molyneux and J. Thornton (1994) "Market Structure and Performance in Spanish Banking", *Journal of Banking & Finance*, 18(May), 433-443.

Longbrake, W. and J. Haslem (1975) "Productive Efficiency in Commercial Banking", *Journal of Money, Credit and Banking*, 7, 317-330.

Lovell C. and J. Pastor (1997) "Target Setting. An Application to a Bank Branch Network", *European Journal of Operational Research*, 98, 290 - 299.

Lozano, A. (1997) "Profit Efficiency for Spanish Savings Banks", *European Journal of Operational Research*, 98, 382-392.

Mahajan, A., N. Rangan and A. Zardkoohi (1996) "Cost Structures in Multinational and Domestic Banking", *Journal of Banking & Finance*, 20, 283-306.

Mahajan, A., N. Rangan, and A. Zardkoohi (1996) "Cost Structures in Multinational and Domestic Banking", *Journal of Banking and Finance*, 20(2): 238-306.

Mankiw, G. (1994) "*Macroeconomics*", Worth Publishers, New York.

Martin, F. and M. Sassenou (1992) "Cost Structure in French Banking: A Re-examination Based on a Regular CES-quadratic form", Caisse des Depots de Consignation, May, Paris.

Martinez, P., M. Soledad, and S. Schmukler (1999) "Do Depositors Punish Banks for 'Bad' Behaviour? Market Discipline in Argentina, Chile and Mexico", *Policy Research Working Paper,* No.2058, February, Washington: World Bank.

Maudos, J., J. Pastor, and J. Quesada (1995) "A Decade of Technical Change in Spanish Savings Banks: A Summary of Different Measurement Techniques", Istituto Valenciano de Investigaciones Economicas (IVIE) discussion paper, 4.

Maudos, J., J. Pastor, F. Perez and J. Quesada (1999) "The Single European Market and Bank Efficiency: The Importance of Specialisation", 'Paper read at the Annual Conference of the European Association of University Teachers in Banking and Finance, Universidade Nova de Lisboa, Lisbon, 3-4 (September)'.

Maudos, J., J. Pastor, F. Perez, and J. Quesada (2002) "Cost and Profit Efficiency in European Banks", *Journal of International Financial Markets,* Institutions and Money 12, 33–58

McAllister, R. and D. McManus (1993) "Resolving the scale efficiency puzzle in banking", *Journal of Banking & Finance,* 17, 389-405.

McKillop, D. and C. Glass (1994) "A Cost Model of Building Societies as Producers of Mortgages and Other Financial Products", *Journal of Business, Finance and Accounting,* 21(7), October, 1031-1046.

McKinnon, R. (1973) "Money and Capital in Economic Development", Washington, D.C.: Brookings Institution.

MEED (1999) "A mixed year", June 25th, 9-11.

Meeusen, W. and J. van den Broeck (1977) "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error", *International Economic Review,* 18, 435-444.

Mester (1987a) "Efficient Production of Financial Services: Scale and Scope", *Business Review,* Federal Reserve Bank of Philadelphia, January/February, 15-25.

Mester (1987b) "A Multiproduct Cost study of Savings and Loans", *Journal of Finance,* 42, 423-445.

Mester, L. (1993) "Efficiency in the savings and loan industry", *Journal of Banking & Finance,* 17, 267-286.

Mester, L. (1994) "Efficiency of Banks in the Third Federal Reserve District", *The Wharton School,* University of Pennsylvania, (WP 13/94).

Mester, L. (1996) "A Study of Bank Efficiency Taking into Account Risk- Preferences", *Journal of Banking & Finance,* 20, 389-405.

Mester, L. (1996) "Measuring Efficiency At U. S. Banks: Accounting for Heterogeneity is Important", European Journal of Operational Research, 98, 230-243.

Mester, L. (1997) "Measuring Efficiency At US Banks: Accounting For Heterogeneity Is Important", *European Journal of Operational Research*, 98, 230-243.

Mester, L.J. (1992) "Traditional and Non-traditional Banking: An Information-theoretic Approach", *Journal of Banking and Finance*, 16 (3), 545-566.

Miller, S. and A. Noulas (1996) "The Technical Efficiency of Large Bank Production", *Journal of Banking & Finance*, (20), 495-509.

Miller, S. and A. Parkhe (1998) "Patterns in the Expansion of U.S. Banks' Foreign Operations", *Journal of International Business Studies*, 29(2), 359-390.

Miller, S. and A. Parkhe (1999) "Home-Country Environment as a Source of International Competitiveness: An Analysis of the Global Banking Industry", *Mimeo,* Michigan State University.

Mishkin, F. (1998) " *The Economics of Money, Banking, and Financial Markets*", Addison Wesley, US.

Mishkin, Frederic S. (2001) "Financial Policies and the Prevention of Financial Crises in Emerging Market Countries," NBER Working Paper No. 8397 forthcoming in Martin Feldstein, ed., Economic and Financial Crises in Emerging Market Countries (University of Chicago Press: Chicago).

Mitchell, K. and N. Onvural (1996) "Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier-Flexible Functional Form", *Journal of Money, Credit and Banking*, 28(2), 178-199.

Molyneux, P. and J. Thornton (1992) "Determinants of European bank profitability: A note", *Journal of Banking & Finance*, 16, 1173-1178.

Molyneux, P., Altunbas Y., and E.P.M. Gardener (1996) "*Efficiency in European Banking*", John Wiley & Sons, Chichester, England.

Mullineaux, D. (1975) "Economies of Scale and of Financial Institutions", *Jounral of Monetary Economics*, 1, 233-240.

Mullineaux, D. (1978) "Economies of Scale and Organisational Efficiency in Banking: A Profit-Function Approach", *Journal of Finance*, 33, 259-280.

Mullineaux, D.J. (1975) "Economies of Scale of Financial Institutions", *Journal of Monetary Economics*, 1, 233 - 240.

Murphy, N. (1972) "Cost Of Banking Activities: Interaction Between Risk And Operating Cost: A Comment", *Journal of Money, Credit and Banking*, 4 (August), 614-615.

Murray, J. and R. White (1983) "Economies Of Scope In Multi-Product Financial Institutions: A Study Of British Columbia Credit Unions", *Journal of Finance*, 38, 887-901.

Nakamura, L. (994) "Small Borrowers And The Survival Of The Small Bank: Is Mouse Bank Mighty Or Mickey?", *Business Review*, issue Nov, 3-15

Newbold, P., W. Carlson, and B. Thorne (2003) "*Statistics For Business And Economics*", 5th ed., Prentice Hall, New Jersey.

Nicholson, W. (1995) "*Microeconomic Theory: Basic Principles and Extensions*", The Dryden Press, Texas, US.

Nigh, D., R. Kang, and S. Krishnan (1986) "The Role of Location-Related Factors in U.S. Banking Involvement Abroad: An Empirical Analysis", *Journal of International Business Studies*, (Fall), 59-72.

Nolle, D. (1994) "Are Foreign Banks Out-Competing US Banks In The US Market?", Office of the Comptroller of the Currency, Economic and Policy Analysis, *Working Paper*, May, 94-5,.

Noulas, A., S. Ray, and S. Miller (1990) "Returns To Scale And Input Substitution For Large U.S. Banks", *Journal of Money, Credit and Banking*, 22(1), 94-108.

Oral, M. and R. Yolalan (1990) "An Empirical Study On Measuring Operating Efficiency And Profitability Of Bank Branches", *European Journal of Operational Research*, 46, 282-294.

Pallage, S. (1991) "An Econometric Study of the Belgian Banking Sector in terms of Scale and Scope Economies", Cahiers Economiques de Bruxelles, 130, pp. 126-143.

Panzar, J. and R. Willig (1977) "Economies Of Scale In Multi-Output Production", *Quarterly Journal of Economics*, 91, 481-494.

Parkan, C. (1987) "Measuring the Efficiency of Service Operations: An Application to Bank Branches", Engineering Costs and Production Economics, 12, 237 - 242.

Pastor, J., F. Pérez, and J. Quesada (1995) "Efficiency Analysis in Banking Firms: An International Comparison", *Working Paper,* EC 95-18, Istituto Valenciano de Investigaciones Economicas, Valencia, Spain.

Pastor, J.T., A. Lozano, and J.M. Pastor (1997) "Efficiency of European Banking Systems: A Correction by Environmental Variables", *Working Paper* EC 97-12, Istituto Valenciano de Investigaciones Economicas.

Peek, J. and E.S. Rosengren (1995)."Bank Regulation and the Credit Crunch", *Journal of Banking and Finance,* 19, 2-4, pp. 679:92.

Peek, J., E. Rosengren, and F. Kasirye (2000) "The Poor Performance of Foreign Bank Subsidiaries: Were the Problems Acquired or Created?", *Journal of Banking and Finance*, 23(2-4): 579-604.

Peristiani, S. (1997) "A Critique of Empirical Studies of Relations Between Market Structure and Profitability", *Journal of Money, Credit and Banking*, 29(3), 326-337.

Peterson, E. (1988) "*The Gulf Cooperation Council: Search for Unity in a Dynamic Region*", Westview Press, Colorado, US.

Pi, L., and S. Timme (1993) "Corporate Control And Bank Efficiency", *Journal of Banking and Finance*, 17, 515-530.

Porier, D.J. (1976) "The Econometrics of Structural Change with Special Emphasis on Spline Functions", (North-Holland, Amsterdam).

Pulley, L. and D. Humphrey (1993) "The Role of Fixed Costs and Cost Complementarities in Determining Scope Economies and the Cost of Narrow Banking Proposals", *Journal of Business*, 66(3), 437-462.

Pulley, L., A. Berger, and D. Humphrey (1994) "Do Consumers Pay for One-Stop Banking? Evidence from a Non-Standard Revenue Function", The Wharton Financial Institution Centre, (WP 01/94).

Qatar Central Bank (2000) "*Executive Instruction Until March 2000*", Second edition, Qatar.

Qatar Monetary Agency (1992) "*The Banking And Financial System In Qatar*", Department of Research and Statistic, Qatar.

Ramchander, S., A. Reichert, and S. Jayanti, (1999) "The Determinants Of Foreign Bank Involvement In U.S. Banking Markets", *Managerial Finance*, 48-65.

Rangan, N., R. Grabowski, H. Aly, and C. Pasurka (1988) "The Technical Efficiency of U.S. Banks", *Economic Letters*, 28, 169 - 175.

Resti, A. (1997) "Evaluating The Cost-Efficiency Of The Italian Banking System: What Can Be Learned From The Joint Application Of The Parametric And Non-Parametric Techniques", *Journal of Banking & Finance*, 2, 221-250.

Richmond, J. (1974) "Estimating the Efficiency of Production", *International Economic Review*, 15, 515-521.

Rodriguez J.R.O., Alvarez A.A., and P.P. Gomez (1993) 'Scale and scope economies in banking: A study of savings banks in Spain' Universidad De La Laguna, Tenerife, Spain.

Rogers, K. (1998) "Nontraditional Activities And The Efficiency Of US Commercial Banks", *Journal of Banking & Finance*, 22, 467-482.

Rutherford, D. (2000) "*Routledge Dictionary of Economics*", Routledge, London.

Rybcynski, T. (1988) "The Internationalisation Of Finance And Business", *Business Economics*, 23, 14-20.

Sabi, M. (1996) "Comparative Analysis of Foreign and Domestic Bank Operations in Hungary", *Journal of Comparative Economics*, 22, 197-188.

Saudi Arabia Monetary Agency (2001) "Annual Report", Kingdom of Saudi Arabia.

Savory, R. (1980) "The Ancient Period" Pages 3-13 in Alvin J. ottrell (ed.) "*The Persian Gulf States: A General Survey*", Johns Hopkins University Press, Baltimore.

Schaffnit C., D. Rosen, and J.C. Paradi (1997) "Best Practice Analysis of Bank Branches: An Application of DEA in a Large Canadian Bank", *European Journal of Operational Research*, 98, 269 - 289.

Schmidt, P. (1986) "Frontier Production Functions", *Econometrics Reviews*, 4, 289-328.

Schweitzer, S. (1972) "Economics Of Scale And Holding Company Affiliation In Banking", *Southern Economic Journal*, 39(October), 258-266.

Schweitzer, S. and J. McGee (1961) "Chicago Banking", *Journal of Business*, 34(3), 203-366.

Sealey, C. and J. Lindley (1977) "Inputs, Outputs And A Theory Of Production And Cost At Depository Financial Institutions", *Journal of Finance*, 32, 1251-1266.

Secretariat General (2001) "*The economic Bulletin*", Volume XVI, Riyadh.

Secretariat General (2002), http://www.gcc-sg.org.

Seth, R. (1992) " Profitability Of Foreign Banks In United States", Federal Reserve Bank of New York, *Research Paper*, no. 9225, December.

Seth, R., D. Nolle, and S. Mohanty (1998) "Do Banks Follow Their Customers Abroad?", *Financial Markets, Institutions, and Instruments*, 7(4), 1-25.

Shaffer S. (1991) "Potential Merger Synergies Among Large Commercial Banks", *Working Paper*, Federal Reserve Bank of Philadelphia, October.

Sharpe, W. (1964) "Capital Asset Prices: A Theory of Market Equilibrium", *Journal of Finance*, September, 425-442.

Shaw, E. (1973) "*Financial Deepening in Economic Development*, Oxford University Press, New York.

Sherman, H.D. and F. Gold (1985) "Bank Branch Operating Efficiency: Evaluation with Data Envelopment Analysis", *Journal of Banking and Finance*, 9, 297-315.

Sinkey, J. (1992) "*Commercial Bank Financial Management In the Financial-Services Industry*", 4th ed., Macmillan Publishing Company, New York.

Solon, G. (1989) "The Value of Panel Data in Economic Research" in D. Kasprzyk, G Duncan, G. Kalton and M.P, Sing, eds., Panel Study (John Wiley, New York).

Spong, K., R. Sullivan and R. DeYoung (1995) "What Makes A Bank Efficient? A Look At Financial Characteristics And Bank Management And Ownership Structure", *Financial Industry Perspective*, Federal Reserve Bank of Kansas City.

Srivastava, P. (1999) "Size, Efficiency and Financial Reforms in Indian Banking", Indian Council for Research on International Economic relations.

Standard and Poor's Creditweek (2002) "Capital Market Trends in The Gulf Cooperation Council Countries", Standard and Poor's Creditweek, October 16, London

Stevenson, R. (1980) "Likelihood Functions for Generalised Stochastic Frontier Estimation", *Journal of Econometrics*, 13,57-66.

Stevenson, R. (2001) "Likelihood Functions For Generalized Stochastic Frontier Estimation", *Journal of Econometrics*, 13, 57-66.

Stiglitz, J. (1989) "Financial Markets and Development", *Oxford Review of Economic Policy*, 5(4), 55-68.

Stiglitz, J. (1985) "Credit Markets And The Control Of Capital", *Journal of Money, Banking, and Credit*, 17(12), 133-152.

Stiglitz, J. (1994) "The Role of the State in Financial Markets", In Proceedings of the World Bank Annual Bank Conference on Development Economics 1993. Edited by Michael Bruno and Boris Pleskovic, World Bank, 19-52, Washington.

Stiglitz, J. (1998) "The Role of Financial System in Development", Presentation at the fourth Annual Bank Conference on Development in Latin America and the Caribbean. The World Bank, http://www. worldbank.org/html/extdr/extme/jssp062998.htm.

Summers, L. and V. Summers (1989) "When Financial Markets Work Too Well: A Cautious for a Securities Transaction Tax" , 'In Regulatory Reform of Stock and Futures Markets: A Special Issue for the Journal of Financial Services research,' Boston, Mass.: Kluwer Academic Publishers.

Taylor, W., R. Thompson, R. Thrall, and P. Dharmapala (1997) "DEA/AR Efficiency and Profitability of Mexican Banks - A total Income Model", *European Journal of Operational Research*, 98, 346-363.

Terrell, H. S., (1993) "U.S. Branches and Agencies of Foreign Banks: A New Look", Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, 79 (10), 913-25.

Thakor, A. (1996) "The Design of Financial Systems: An Overview", Journal of Banking and Finance, 20, 917-948.

The World Factbook (2002), (fhttp://www.cia.gov/cia/publications/factbook).

Thomas, R. (1997) "*Modern Econometrics: an introduction*", Addison Wesley Longman Limited, England.

Tobin, J. (1984) "On the Efficiency of the Financial System", *Lloyds Bank Review*, (153, July), 1-15.

Tschoegl, E. (1983) "Size, Growth, and Transnationality among the World's Largest Banks", *Journal of Business*, 187-201.

Tulkens, H. (1993) "On FHD Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit", *Journal of Productivity Analysis*, 4, 183-210.

Turati, G., (2001) "Cost Efficiency and Profitability in European Commercial Banking", 7th European Workshop on Efficiency and Productivity Analysis, Oviedo (Spain).

UAE Central Bank (2001) "*Annual Report*", Abu Dhabi, UAE.

UAE Central Bank (2001) "Dynamic Growth of The UAE Monetary & Banking Sector", <http://www.cbuae.gov.ae/DynamicGrowtha.htm>.

Udell, F. (1987) "Designing The Optimal Loan Review Policy", *Proceedings*, Federal Reserve Bank of Chicago, 507-531

Ursacki, T. and I. Vertinsky. (1992) "Choice of Entry, Timing, and Scale by Foreign Banks in Japan and Korea", *Journal of Banking and Finance*, 16, 405-21.

Varian, H.L. (1990) "*Microeconomic Analysis*", 2nd ed., W.W. Norton, New York.

Vassiloglou, M. and D. Giokas (1990) "A Study of the Relative Efficiency of Bank Breaches: An Application of Data Envelopment Analysis", *Journal of the Operational Research Society*, 41, 591-597.

Venett, R.V. (1993a) "Concentration, Efficiency and Entry Barriers as Determinants of EC Bank Profitability", *Working Paper*, University of Ghent, Belgium, July.

Vennet, R. (1998) "Cost And Profit Dynamics In Financial Conglomerates And Universal Banks In Europe", paper presented at the SUERF/CFS Colloquim, Frankfurt, 15-17 October .

Vennet, R. V. (1993a) "Cost Characteristics of Credit Institutions in the EC", Paper presented at the 20th Annual Meeting of the European Finance Association, Copenhagen Business School, 26-28 August, Published in section II-D of the Proceedings, 1-38

Wachtel, P. (2000) "Globalization of Banking: Why does it matter?" Presente y Futuro del Sistema Financiero en Paises Emergentes, Caracas, Venezuela, October 25-26,2000

Wald, A. (1943) "Tests Of Statistical Hypotheses Concerning Several Parameters When The Number Of Observations Is Large", Transactions of the American Mathematical Society, 54, 426-482.

White, H. (1980) "A Heteroscedasticity Consistent Covariance Matrix Estimator And A Direct Test Of heteroscedasticity", *Econometrica*, 48, 817-818.

White, H. (1982) "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, 50, 1-26.

Williamson, O.E. (1975) "*Markets and Hierarchies: Analysis and Antitrust Implications*", New York, Free Press.

Williamson, S. (1986) "Costly Monitoring Financial Intermediation, and Equilibrium Credit Rationing", *Journal of Monetary Economics*, 18(2), 159-179.

World Bank (1999/2000) "*Entering the 21st Century; World Development Report 1999/2000*", Oxford University Press.

World Bank (2001) "*World Development Indicators Database*", July.

Worthington, A. (1988) "Efficiency in Australian Building Societies: An Econometric Cost Function Approach Using Panel Data", *Applied Financial Economics*, vol. 8, issue 5, 459-67

Yue, P. (1992) "Data Envelopment Analysis and Commercial Bank Performance: A Primer with Applications to Missouri Banks", Federal Reserve Bank of St. Louis, January/February, 31-45.

Zaim, O. (1995) "The Effect Of Financial Liberalisation On The Efficiency Of Turkish Commercial Banks", *Applied financial Economics*, 5, 257-264.

Zimmerman, G. (1989) " The Growing Presence Of Japanese Banks In California", Federal Reserve Banks of San Francisco Economic Review, Summer, 3-17.

Review of scale and scope economies studies in US banking	Author	Year	Data	Methodology	Findings
	Alhadeff	1954	Data on 210 California banks studied for the period 1938-1950. Data obtained from Federal reserve data of San Francisco	Simple cost function	Both small and large sized banks exhibited increasing returns to scale, while constant return to scale reported for medium sized banks.
	Horvitz	1963	Data obtained form a sample of banks that are members in FDIC for the period 1940-1960	Simple cost function	Very large banks operate at more scale economies than small banks; however, magnitudes of scale economies were declining.
	Schweiger and McGee	1961	Data on 6,233 banks member in Federal reserve member in the year 1959	Multiple regression approach	Small banks realise scale economies by increasing their size up to $50 million deposits. Branch banks have more cost savings than unit banks of the same size.
	Gramely	1962	Data from Tenth Federal Reserve District of 270 small unit banks covered in the period 1956-1959	Multiple regression approach	Average cost decline sharply with size increases among small banks, but the cost curves become flatter afterwards as banks size increase. However, the results did not indicate if further increase in bank size further decrease cost. Generally, negative coefficient of the size in relation to cost indicated that banks in sample were enjoying scale economies.
	Benston	1965a	Data are from the FCA programme of the Federal Reserve bank of Boston. The data cover the period 1959-1961 for 80 to 83 banks	Cobb-Douglas cost function	The coefficient of deposit and mortgage loans reported negative and significant, suggesting economies of scale; while time deposits and instalment were negative and significant, implying diseconomies of scale. Moreover, economies of scale were also found in branch banks.

Review of scale and scope economies studies in US banking

Author	Year	Data	Methodology	Findings
Benston	1965b	Same sample as in Benston 1965a	The Same as that in Benston 1965a	Economies of scale is found for branch banks
Greenbaum	1967	Data drawn from Fifth and Tenth Federal Reserve Districts for 413 and 745 banks	Weighted output index	Declining average cost for small banks, and increasing average cost for large banks. Branch banks operating cost were higher than unit banks.
Bell and Murphy	1968	Data are from the FCA programme of the Federal Reserve bank of Boston, New York and Philadelphia. The data cover the period 1963-1965 for 210 to 283 banks	Cobb-Douglas cost function	Authors found evidence of economies of scale for most bank services such as demand deposits and business and mortgage loans. They also found that branching operations have higher cost than unit banking operation
Schweitzer	1972	Data from the banks in the 9th Federal Reserve District for 1964	Cobb-Douglas cost function	Evidence found for U-shaped cost curve. Banks with assets less than $3.5 million exhibited scale economies. Banks with assets between $3.3 million and $25 million show constant returns to scale. Banks with assets over $25 million have diseconomies of scale.
Murphy	1972	Functional Cost Analysis (FCA) data for the years 1968 on 967 banks	Cobb-Douglas cost function	Generally, banks included in the sample found to exhibit constant returns to scale.
Kalish and Gilbert	1973	Functional Cost Analysis (FCA) data for the years 1968 on 989 banks	Cobb-Douglas cost function	Cost curve found to be shaped as a U-curve. Unit banks found to have the least cost, followed by affiliated banks and branch banks

Review of scale and scope economies studies in US banking

Author	Year	Data	Methodology	Findings
Mullineaux	1975	Functional Cost Analysis (FCA) data for the year 1970	Cobb-Douglas cost function	Economies of scale found to be larger for unit banks than branch banks
Longbrake and Haslem	1975	Functional Cost Analysis (FCA) data for the years 1968 on 989 banks	Cobb-Douglas cost function	Unit banks show having the lowest cost. However, as number of accounts and average deposit increase, unit banks show the highest cost.
Benston, Hanwek, and Humphrey	1983	Functional Cost Analysis (FCA) data from 1975 to 1978	Translog cost function	They found evidence of U-shaped cost curves. That is banks with deposit more than US$ 50 million experienced decreasing scale returns while bank in branching states enjoy small scale economies. The findings suggest that the optimal size of banks ranges between US$ 10 and 25 million in deposits.
Murray and White	1983	Canadian credit unions in 1977	Translog cost function	Most of credit union show existence of economies of scale, in which returns on scale increases as output increase. Evidence is also available on the strong existence of scope economies between mortgage and other lending activities. Credit union that are large and multiproduction are found more cost efficient than small and single product credit unions.
Gilligan and Smirlock	1984	US Federal Reserve Functional Cost Analysis (FCA) data for 714 banks observed in the year 1978. They also use data from Federal Reserve Bank of Kansas City on 2700 unit banks for the years 1973-1978.	Translog cost function	Scale economies exist for banks with less than $25 million in deposit, but scale diseconomies for banks beyond USD 100 million deposits. The results for the data from the Federal Reserve Bank of Kansas City indicate that banks with less that $10 million exhibit scale economies and those with $50 million show diseconomies of scale. Generally, banks output found to exhibit scope economies

Review of scale and scope economies studies in US banking	Author	Year	Data	Methodology	Findings
	Lawrence and Shay	1986	FCA programme from 1979 to 1982	Translog cost function	Banks show constant returns to scale. When analysis is based on quartile, banks show significant scale economies. Branch banks exhibit scale economies at all sizes, and unit bank showed siginifican scale economies in general. Scale economies were found between products such as loans and deposits and deposits and investments. However, significant scope diseconomies were found between loans and investments
	Kolari and Zardkoohi	1987	US Federal Reserve Functional Cost Analysis (FCA) data for the period 1979-1983	Translog cost function	Unit banks show flat cost curve (constant returns to scale), while branch bank exhibited U-shaped cost curve but tend to be upward in general. Generally, evidence of scope economies is present. However, cost benefit from joint production appeared greater for small banks than large banks. Moreover, unit banks exhibited greater scope economies than branch banks
	Berger, Hanweck, and Humphrey	1987	Functional Cost Analysis (FCA) data	Translog cost function	Scale economies are evident for banks up to $50 million, and any size in assets above $50 million show diseconomies of scale. Moreover, scope economies are only evident for small banks.
	Berger et. al.	1987	1983 Functional Cost Analysis (FCA) data for 413 branching state banks and 214 unit state banks.	Deterministic Translog cost function	Evidence is available for scale economies. In general, slight scale economies at branch level and slight scale diseconomies at the bank firm level. Scale diseconomies are shown at large unit state banks. Diseconomies of scope

Review of scale and scope economies studies in US banking

Author	Year	Data	Methodology	Findings
Humphrey	1987	Data include 13,959 US banks observed over 1980, 1982, and 1984	Translog cost function	By two to four times variation is found greater in average costs between banks that have the highest cost and those with the lowest cost compared to the observed differences in the average cost across bank size classes. Moreover, the result on cost economies did not show strong evidence of competitive advantage for large banks over small banks.
Cebenoyan	1988	Functional Cost Analysis (FCA) data on 413 branch state banks and for 214 unit state banks, all observed in the year 1983	Translog cost function	Branch banks show slight scale economies and slight scale diseconomies at the level of banking firm. Unit state banks showed large scale diseconomies for large banks. The study's finding indicated that there were scope diseconomies in banking, suggesting other consideration of joint production such as risk diversification, customer convenience and joint demand of products. Expansion path sub-additivity test indicates that production of a given bundle of output is more cost effective by two banks than one bank for banks with deposit amount of $10 million up to $1 billion
Hunter, Timme, and Yang	1990	Call and Income Report data on 311 largest US commercial banks	Translog cost function	Cost gains for large banks are more if they separate their production bundles into groups of specialist banks. Moreover, the finding shows no strong evidence on cost subadditivity.
Noulas, Ray, and Miller	1990	Call and Income Report data on 309 banks with assets over $1 billion	Translog cost function	Banks with assets between $1 billion and $3 billion exhibit scale economies. However, diseconomies of scale are found for banks beyond $3 billion in assets.

Review of scale and scope economies studies in US banking

Author	Year	Data and study period	Methodology	Findings
Mester	1993	Data form reports on Condition and Income on 328 branch banks with assets exceeding $1 billion	Translog cost function	Global scale economies exist in all banks sizes. Moreover, scope economies is also evident.
Mester	1996	Data come form the Third Federal Reserve District on 214 banks for the years 1991-1992	Translog cost function	Evidence on the flat frontier indicating constant returns to scale
Pulley, Berger, and Humphrey	1994	Data consist of panel of 683 US banks with assets over $100 for years 1987-1984 and located in state that allow state branching during 1980s. The sample also extended to include 626 banks covering the year 1990	A combination of a quadratic structure multiple outputs + log-quadratic Translog for input prices.	Revenue complementarities are not evident. Moreover, no significant evidence of scope economies among the set of revenue efficient banks. Weak evidence is found of significant revenue ray scale economies for small banks.
Mitchell and Onvural	1996	1986 and 1990 Call and Income Report data for 306 banks	Fourier Flexible cost function	Minimum cost or efficiency scale is found at $500 million. Diseconomies of scale beyond this level is quite small.

Review of scale and scope economies studies in US banking	Author	Year	Data and study period	Methodology	Findings
	Chem, Mason, and Higgins	2001	Data US banks in different categories by quarter from 1988 to 1997.	SFA and Regression model	Generally, economies of scale do exist in the banking industry. No economies of scale in the banks with agricultural loan specialization. Scale economies dominate the benefits of geographic diversification in community bank mergers.

Notes: SFA is the Stochastic Frontier Approach.

Sources: Berger and Humphrey (1997), Intarachote (2000), Casu (2000), Girardone (2000), Al-Jarrah (2002), and author's own updates.

Review of scale and scope economies studies in European banking

Author	Year	Data	Methodology	Findings
Levy-Garboua and Renard	1977	Data on French banks for 94 banks observed in 1974	Cobb-Douglas cost function	Finding show banks exhibiting Increasing returns to scale
Gouch	1979	Data on UK building societies for the period 1972-1976	Linear average cost function	No evidence of scale economies is found
Cooper	1980	Data on UK building societies covering the year 1977	Cobb-Douglas cost function	Scale economies are evident for societies with sizes under £100 million, and diseconomies of scale is found for larger societies.
Barnes and Dodds	1983	UK building societies in the years 1970-1978	Linear average cost function	No evidence of scale economies were found
Fanjul and Marvell	1985	Data on 83 Spain commercial banks and 54 savings banks for 1979	Cobb-Douglas cost function	Significant cost economies with respect to accounts per branch and deposits per account; and constant return to scale are found relating to the number of branches.
Fanjul and Maravall	1985	Sample of 83 commercial and 54 savings Spanish banks in the year 1979	Cobb-Douglas functional form	Significant cost economies and constant return to scale relating to the number of branches.
Dietsch	1988	243 French banks observed in the year 1986	Translog cost function	The results suggest that scale economies exist for the banks under study
Hardwick	1989	97 UK building societies in the year 1985	Translog form	Economies of scale exist for societies with assets under £280 million and diseconomies of scale for those with assets beyond £1500 million. No evidence found for scope economies

Review of scale and scope economies studies in European banking

Author	Year	Data	Methodology	Findings
Baldini and Landi	1990	Sample of 294 Italian banks studied in the year 1987	Translog cost function	Scale economies increase at plant level as bank size increase, however, scale economies become smaller and tend to decrease at the firm level as banks size increase
Hardwick	1990	97 UK building societies in the year 1985	Translog form	Economies of scale exist for societies with assets under £5500 million, where no evidence of scope economies found
Kolari and Zardkoohi	1990	Cooperative and saving banks in Finland during 1983 and 1984	Translog model	Cost curve of cooperative and saving banks tend to be L-shaped at plant level and U-shaped at firm level Diseconomies of scope in the joint production of Advances and bills. The author implied that merger among smaller banks are preferred than merger between larger banks
Martin and Sassenou	1992	Data on French banks covering the year 1987	CES-quadratic function	Scale economies exist in small banks, and large banks exhibit diseconomies of scale Scope economies exist in small banks production
Drake	1992	76 UK building societies in the year 1988	Translog form	Mild scale economies for societies in the £120-500 million asset size, but no evidence of scope economies
Glass and McKillop	1992	Data form the Bank of Ireland, on of the larges Irish banks, for the period 1972-1990	Hybrid translog model	Apart form the sub-period 1976-1978, banks were exhibiting diseconomies of scale, and decreasing product-specific scale economies for investment but increasing for loans Diseconomies of scope

Review of scale and scope economies studies in European banking

Author	Year	Data and study period	Methodology	Findings
Dietsch	1993	Data on 343 French banks observed in the year 1987	Translog cost function	Economies of scale were found for banks of all output size ranges Scope economies were not evident
Alevarez and Gromes	1993	64 Spanish banks in 1990	Hybrid translog form	Scale and scope economies for medium-sized saving banks and diseconomies of scale and scope for larger banks
McKillop and Glass	1994	Data obtained from annual returns for a sample of 89 UK building societies in the year 1991	Hybrid translog form	Evidence existed for augmented scale economies of national and local societies, but constant returns to scale for societies with regional base No scope economies and cost complementarities are found between the provision of mortgage and non-mortgage products.
Drake	1995	76 UK building societies in the year 1988	Translog form	No scale economies is found when expense-preference behaviour is take into considertion. No evidence of scope economies
Lang and Welzel	1996	Sample of 700 German cooperative banks	Standard translog cost function	Scope economies are evident in the largest banks
European Commission	1997	Sample on 10 EU countries covers the years 1987 (with 295 banks) to 1994 (with 1451 banks) obtained from the IBCA Bankscope database	SFA and DEA	Generally, in all countries the analysis show evidence of economies and diseconomies of scale. Small banks, in particular in Germany and France, show exhibiting increasing returns to scale. There is strong evidence indicating that largest banks realise large economies of scope.

	Author	Year	Data and study period	Methodology	Findings
Review of scale and scope economies studies in European banking	Casu and Girardone	1998	Data cover 32 Italian banking groups and 78 bank parent companies and subsidiaries for the year 1995	Translog Cost Function	Moderate evidence on the existence of scale economies; but the analysis shows strong evidence of scope economies especially with banking groups
	Cavello and Rossi	2001	Sample contains 442 European banks studies over 1002-1997 period	Translog function	Scale economies are evident but more pronounced for small banks.
	Altunbas, Gardener, Molyneux, and Moore	2001	Data cover the period 1989-1997 obtained for banks from 15 European countries	Fourier Flexible functional form and stochastic cost frontier	Scale economies (inefficiency) ranges from 5% and 7%. Smallest banks and those banks with size ranging from ECU 1 billion to ECU 5 billion enjoy more scale economies

Notes: SFA is the Stochastic Frontier Aproach; DEA is the Data Envelopment Approach.

Sources: Berger and Humphrey (1997), Intarachote (2000), Casu (2000), Girardone (2000), Al-Jarrah (2002), and author's own updates.

Review of X-efficiency studies in US banking

Author	Year	Data	Model	Findings
Sherman and Gold	1985	Data on saving bank branches with 14 offices observed in 1982	DEA	The mean inefficiency of the sample is 96%. Six banks out of fourteen appeared to be relatively inefficient.
Parkan	1987	Sample contains 35 branches of Canadian banks	DEA	11 out of 35 banks found to be relatively inefficient
Rangan, Grabowski, Aly and Pasurka	1988	Data on 215 US independent banks	DEA	The mean efficiency for the sample is 81%
Elyasiani and Mehdian	1990a	144 US banks taken randomly for the years 1980-1985	DEA	Mean efficiency of the sample is 90%
Elyasiani and Mehdian	1990b	Data on 191 of US large banks observed in the years 1980-1985	DEA	The mean efficiency found for the sample is 88%, with a note of an inward shift in the estimated frontier because of the technological advancements effect
Aly, Grabowski, Pasurka, and Rangan	1990	Data acquired form the call reports on 322 independent banks for the period 1996	DEA	Generally, high level of overall efficiency, with more allocative efficiency (81%) than technical efficiency (75%)
Ferrier and Lovell	1990	Sample contains 575 banks for the year 1984	DEA and SFA	Inefficiency found by the DEA reported 21% which is less than that of 26% reported by the SFA
Berger and Humphrey	1991	All US banks 1984	TFA	Inefficiencies dominate scale and product mix economies. Inefficiency mean found to be 19% where this inefficiency is operational in nature, comes from overuse of inputs, rather than financial, involving overpayment of interest

Review of X-efficiency studies in US banking

Author	Year	Data	Model	Findings
Yue	1992	Data on 60 Missouri banks covering the period 1984-1990	DEA	Overall inefficiency is 20%, providing that scale inefficiency was proven not to be the major sources of inefficiency
Elyasiani and Mehdian	1992	Sample contains 80 banks of minority-owned observed in the year 1988	DEA	Mean efficiency is found to be at the level of 89%
English et al.	1993	Sample contains 442 banks of the year 198	DEA	Mean efficiency estimated is 75 and 76%
Pi and Timme	1993	Sample covers the years 1988-1990 and contains 112 banks	SFA	Mean efficiency estimated was 87%
Bauer, Berger and Humphrey	1993	Panel data on 683 US branching state banks observed in the period 1984-1989	SFA and TFA	The mean efficiency found was 87%, providing the mean efficiency of both SFA and TFA found to be consistent
Eluasiani and Mehdian	1995	Sample contains 150 US banks observed in the years 1979 and 1986	DEA	The mean efficiency estimated ranged between 95% and 97%
Grabowski et al.	1994	Sample of 669 banks observed in the uears 1979, 1983, and 1987	DEA	The efficiency estimates ranged between 73 and 67%

Review of X-efficiency studies in US banking

Author	Year	Data	Model	Findings
Karapakis, Miller, and Noulas	1994	Data cover 5548 US banks for assets over US$50 million, observed in the year 1986	SFA	Overall mean inefficiency is found to be 10%
Berger, Leusner, and Mingo	1994	Data cover 760 US banks for the period 1989-1991	DFA	Mean efficiency is reported 90% for the intermediation approach and 66% for the production approach
Wheeloch and Wilson	1994	Data on 269 banks participating in the FCA programme for the year 1993	DEA	Mean inefficiency found to be large, about 50%
Hunter and Timme	1995	Sample covers the year 1985-1990 for 317 banks with assets greater than US$1 billion	DFA	Inefficiency found to range between 23 and 36%
Kwan and Eisenbeis	1995	Data cover simi-annual period 1986 to 1991 and contain 254 banks holding companies	SFA	Larger banks reported efficiency level of 92%, which is more than small size banks with efficiency level of 81%. Moreover, the mean X-inefficiency is found to be declining over time.
Clark	1996	Data contain 440 banks observed over the years 1988-1990	TFA	The mean efficiency estimated is found to be at the level of 73% and 90%
DeYoung and Nolle	1996	Data contain 1812 of US banks observed over the years 1985-1990	DFA	The annual mean efficiency found to be 56% and 73%
Mahajan et al.	1996	Data on multinational US banks for the years 1987-1990	TFA	The annual mean efficiency found to be 77% and 88%

Review of X-efficiency studies in US banking

Author	Year	Data	Model	Findings
Mester	1996	Data comes form the Third Federal Reserve District on 214 banks for the years 1991-1992	SFA	The mean cost X-inefficiency found to be in the order of 6 to 9%
Spong, Sullivan, and DeYoung	1996	Data cover 143 state banks for the year 1994	SFA	The efficiency of banks in the least efficient group report a level of 71%, and those in the most efficient group have an efficiency level of 94%
Miller and Noulas	1996	Sample of 201 banks studied over the period 1984-1990	DEA	Mean inefficiency found is 97%
Berger et al.	1997b	Sample of 832 banks branches studied over the period 1989-1991	SFA	Generally, the average efficiency found for the sample is 94 and 79%
Berger and DeYoung	1997	Sample covers the period 1985-1994	SFA	Generally, the average efficiency found for the sample is 92%
Mester	1997	Sample contains 6630 Us banks for different US districts studied over the period 1991-1992	SFA	The efficiency estimates raged between 85 and 93%
Perstiani	1997	Sample covers the period 1980-1990	DFA	The efficiency estimates ranged between 77 and 81% for different US districts

Review of X-efficiency studies in US banking

Author	Year	Data	Model	Findings
Humphrey and Pulley	1997	Sample contains a panel of 683 banks with assets greater than US$100 million in 1988. The data cover the years 1977-1980,1981-1984, 1985-1988	TFA	The efficiency rages between 81 and 85% for the three periods, which indicate that deregulation caused better competitive environment
Berger and Mester	1997	Sample contains 6000 commercial banks covering the period 1990-1995	DFA	Average efficiency estimate is 87%
DeYoung et al.	1998	Sample covers the year 1992 and contains 3997 banks	SFA	Mean efficiency found to be at the level of 66%
Rogers	1998	Sample covers the period 1991-1995 and contains more than 10000 commercial banks	SFA	For restricted and unrestricted models, cost efficiency ranged from 71 to 76% and from 65 to 66%, revenue efficiency ranged from 41 to 44% and 50 to 51%, and profit efficiency raged from 69 to 71 and 65 to 68%.
Chem, Mason, and Higgins	2001	Data US banks in different categories by quarter from 1988 to 1997.	SFA and Regression model	large banks dominate in efficiency, but medium banks are not more efficient than small banks. Those banks without agricultural loan specialization still have economies of scale in cost efficiency.

Notes: SFA is the Stochastic Frontier Approach; DEA is the Data Envelopment Approach; DFA is the Distribution Free Approach; TFA is the Thick Frontier Approach; and FCA is the Functional Cost Analysis.
Sources: Berger and Humphrey (1997), Intarachote (2000), Casu (2000), Girardone (2000), Al-Jarrah (2002), and author's own updates.

Review of X-efficiency studies in European banking

Author	**Year**	**Data**	**Model**	**Findings**
Vassiloglou and Giolias	1990	Sample of 20 Greek banks	DEA	Mean efficiency found is 91%
Berg, Forsund, and Jansen	1992	Sample consists of 264 Norwegian banks for the period 1980-1989	DEA	The smallest quartile in the sample found to be the most efficient banks in the sample
Fukuyama	1993	Sample of 143 Japanese banks studied for the year 1990	DEA	Sample efficiency mean is 86%
Drake and Howcroft	1993	Data cover 190 branches belong to six largest clearing banks in the UK	DEA	The mean efficiency found at the level 92% although standard deviation of 51% indicate a considerable deviation of inefficiency around this mean efficiency
Berg, Clauseen, and Forsund	1993	Sample consists 502 Finish banks, 141 Norwegian banks, and 120 Swedish banks, all observed for the year 1990	DEA and Malmquist Index	Overall efficiency estimates reported 58% for Finland, 78% for Norway, and 89% for Sweden banks
Berg, Forsund, Hjalmarsson, and Suominen	1993	Data cover 502 Finnish, 126 Swedish, and 150 Norwegian banks for the year 1990	DEA	Overall efficiency estimates reported 55% for Finland, 57% for Norway, and 78% for Sweden banks

Review of X-efficiency studies in European banking

Author	Year	Data	Model	Findings
Tulkens	1993	Sample consists of 773 branches of large Belgian banks studied for the month of January 1987	FDH	136 braches out of 773 are found to be efficient, where small branches appeared to be more efficient than large branches
Berg and Kim	1994	Sample of 173 Norwegian banks ovserved in the year 1988	TFA	Annual mean efficiency found to be at the level of 81%
Altunbas, Molyneux, and DiSalvo	1994	Data consists of 515, 452, and 483 Italian credit co-operative banks for the years 1990, 1991, and 1992, respectively	SFA	Mean inefficiency estimates found to be at the level of 13.1% in 1990, 15.9% in 1991, and 17% in the year 1992
Grifell-Tatje and Lovell	1995b	Spanish banks data nearly on all savings banks for the period 1986-1991	DEA and Malmquist index	The average annual efficiency is between 78% and 83%. The Mamquist index yielded an average of 97% efficiency
Grifell-Tatje and Lovell	1995c	Spanish banks data nearly on all savings banks for the period 1986-1991	DEA and Malmquist index	The average annual efficiency is between 75% and 8%. The Mamquist index yielded an average of 95% efficiency
Berg, Forsund, and Bukh	1995	Data obtained for the year 1993 and consists of 714 banks in 4 Nordic countries	DEA	The most efficient banks are the largest banks operating in Danish and Swedish banking industries

Review of X-efficiency studies in European banking

Author	Year	Data	Model	Findings
Maudos, Pastor, and Quesada	1995	Data on savings banks in Spain and cover the period 1985-1994	SFA	Technical change is found to have an average of 68% on the average cost
Pastor, Perez, and Quesada	1995	Data cover 168 US banks, 45 Australian, 59 Spanish, 22 German, 18 UK, 17 Belgian, and 67 French banks for the year 1992	DEA and Malmquist Index	Mean efficiency found at level of 81% for banks of US, 89% of Spain, 93% of German, 92% of Italy, 92% for Austria, 54% of the UK, 95% of France and 92% of Belgium
Favero and Papi	1995	Sample of 174 Italian banks of the year 1991	DEA	Mean efficiency found to be at the level of 96% for the production approach and 95% for the intermediation approach
Allen and Rai	1996	Data on 194 banks in 11 OECD countries (only 9 countries) for the period 1988-1992	DFA and SFA	Inefficiency is at average of 27.5%
Resti	1997	Data on 270 Italian banks obtained for the period 1988-1992	SFA and DEA	Mean efficiency found at the level of 69% using SFA and 74% using the DEA

Review of X-efficiency studies in European banking

Author	**Year**	**Data**	**Model**	**Findings**
Altunbas et al.	1997	Data on German banks over the period 1988-1995	SFA	Estimates of efficiency level ranged between 93% and 95%
Altunbas and Molyneux	1997	Data on European banks for the period 1988-1995	SFA	Estimates of efficiency level ranged between 72% and 76%
Lovell and Pastor	1997	Data on 545 Spanish branch offices observed in the first semester of the year 1995	DEA	Mean efficiency found to be at the level of 92%
Athanassopoulos	1997	Data on 60 commercial branches of large Greece commercial banks	DEA	Mean efficiency found to be at the level of 90%
European Commission	1997	Sample on 10 EU countries covers the years 1987 (with 295 banks) to 1994 (with 1451 banks) obtained from the IBCA Bankscope database	SFA and DEA	Average efficiency ranges between 71% and 77% over the 5 years using SFA. DEA efficiency estimates ranged between 93% and 96% but have been decreasing for the last four years of the study period
Pastor, Lozano and Pastor	1997	Data obtained for 24 Belgian banks, 150 French, 203 German, 29 Danish, 26 Italian, 68 Luxembourgian, 28 Spanish, and 45 British banks for the year 1993	DEA	Mean efficiency for these countries are found as follows: 78% Belgian banks, 71% Denmark, 37% French, 51% German, 85% Italian, 59% Luxembourgian, 71 % Netherlands, 85% Portugal, 82% Spanish, and 56% British banks

Review of X-efficiency studies in European banking

Author	Year	Data	Model	Findings
Leightner and Lovell	1998	Sample of 31 Thai banks for the period 1989-1994	SFA	The mean sample efficiency estimate for the period range between 42% and 69%
Casu and Girardone	1998	Sample consists 32 Italian banking groups and 78 bank parent companies and subsidiaries for the year 1995	SFA and DEA	The mean efficiency using SFA is 92% for banking group, and 94% for banks parent and companies; for the DEA, the mean inefficiency found is 88% for banking group and 90% for bank parent and companies
Maudos, Pastor, Perez, and Quesada	1998	Sample of 879 European banking firms observed over the period 1993-1996	DFA	For 11 European countries, efficiency results show an average of 91% with 5% truncation
Dietsch and Weill	1998	Sample covers 661 commercial, mutual, and savings banks of 11 EU countries studied over the period 1992-1996	DEA, Malmquist Index	Both profit and cost efficiency increased over the study period. Moreover, improvement of technical progress lead to an increase in total productivity
Maudos, Pastor, Perez, and Quesada	1999	Sample contains 879 banks from 11 European countries for the period 1993-1996	Translog model	Mean cost efficiency reported at the level of 44%; when bank specialisation is taken into account, mean cost efficiency is reported at the level of 74%
Casu and Molyneux	1999	A sample of 750 banks from Germany, Franc Italy, Spain, and UK, covering the period 1993-1997	DEA	The results suggest low levels of efficiency, ranging between 61% and 69%.

Review of X-efficiency studies in European banking	Author	Year	Data	Model	Findings
	Altunbas, Gardener, Molyneux, and Moore	2001	Data cover the period 1989-1997 obtained for banks from 15 European countries	Fourier Flexible functional form and stochastic cost frontier	Relative inefficiency of various banking industry within the sample have increased overtime, but averaging at the level of 25% of the total cost
	Maudos, Pastor, Perez, and Quesada	2002	Data on ten countries of the European Union banks, using IBCA information for the period 1993–1996.	SFA and DFA	Overall, the authors find high levels of efficiency in costs and lower levels in profits

Notes: SFA is the Stochastic Frontier Approach; DEA is the Data Envelopment Approach; DFA is the Distribution Free Approach; TFA is the Thick Frontier Approach; and FCA is the Functional Cost Analysis.

Sources: Berger and Humphrey (1997), Intarachote (2000), Casu (2000), Girardone (2000), Al-Jarrah (2002), and author's own updates.

Review of X-efficiency studies in other banking systems

Author	Year	Data	Model	Findings
Oral and Yolalan	1990	Bank branches in Turkey	DEA	Efficiency found to be ranging from 53% and 87%
Fukuyama	1995	Data on 462 Japanese banks covering the years 1981 and 1990	DEA	Mean efficiency ranges between 44% and 46%
Zaim	1995	Data on 95 Turkish banks observed in the years 1981 and 1990	DEA	Annual mean efficiency is 83% and 94% for the studied years
Schaffinit, Rosen, and Paradi	1997	Data obtained for the year 1993 on 291 branches of large Canadian banks based in Ontario	DEA	Efficiency is found to be at the level of 72% for DEA; however, a refined DEA model yielded efficiency at the level of 54%
Bhattacharyya, Lovell, and Sahay	1997	Data cover 70 Indian commercial banks over the period 1986-1991	DEA	Efficiency is found to be at the level of 80% for the sample; Publicly owned banks reported higher efficiency level (87%) than privately owned banks (75%) and foreign owned counterparts
Taylor, Thompson, Thrall, and Dharmapala	1997	13 Mexican commercial banks studied over the period 1989-1991	DEA	The mean efficiency found to range between 69 and 75 per cent

Review of X-efficiency studies in other banking systems

Author	Year	Data	Model	Findings
Kraft and Tirtiroglu	1998	43 Croatian commercial banks in 1994 and 1995	SFA	Cost X-efficiency in their estimates range from 54.7 to 87.9 per cent
Worthingron	1998	Sample includes 22 Australian building societies studied over the period 1992-1995	SFA	The results indicate that building societies are 20 per cent cost inefficient
Hao, Hunter, and Yang	1999	Data on 19 Korean private banks observer over the period 1985 to 1995	SFA	Efficiency ranges form 85% to 91%
Srivastava	1999	85 Indian commercial and public banks studied over the period 1994-1995	SFA	The mean cost efficiencies of private and public banks are 98.18 and 98.11 per cent respectively. The mean efficiency of recent entrants reports higher level than current banks. Moreover, highest cost efficiency is, generally, reported for middle-sized banks, followed by small and large sized banks.
Altunbas, Liu, Molyneux, and Seth	2000	Data on Japanese banks (130 banks in 1993 and 1994, and 121 in 1995)	Fourier Flexible and SFA	Inefficiencies in two models are about similar and range between 5% and 7%
Intarachote	2000	Sample on 15 Thai banks, 14 foreign banks and other finance and specialized institutions	DEA	Mean inefficiency ranges from 26% to 48% for national banks, 33% to 50% for the foreign banks, and 6% to 14% for the finance and specialized institutions

Review of X-efficiency studies in other banking systems

Author	Year	Data	Model	Findings
Isik and Hassan	2002	Data consist of 39 banks from 1988, 54 banks from 1992 and 56 banks from 1996 of Turkish banks over the 1988–1996 period.	DEA and parametric approach	Overall cost and profit efficiencies for the Turkish banks are 72 and 83 per cent respectively
Al-Jarrah.	2002	Data on 82 banks operating in Jordan, Egypt, Saudi Arabia, and Bahrain over the period 1992-2000	Fourier Flexible and SFA	Cost efficiencies are found to be 95 per cent, standard profit 66 percent, and alternative profit efficiencies 58 per cent. Islamic banks found to be the most cost and profit efficient banks. Geographically, Bahrain banks are the most efficient, while Jordanian banks are found to be the least efficient.

Notes: SFA is the Stochastic Frontier Aproach; DEA is the Data Envelopment Approach; DFA is the Distribution Free Approach; TFA is the Thick Frontier Appraoch; and FCA is the Functional Cost Analysis.

Sources: Berger and Humphrey (1997), Intarachote (2000), Casu (2000), Girardone (2000), Al-Jarrah (2002), and author's own updates.

List of GCC Banks Included in the Sample (1995-2000)

Qatar Banks	Foreign Banks	1. HSBC 2. Arab Bank Limited 3. Grindlays Bank 4. P.N.B Paribas 5. Standard Chartered Bank 6. Mashreq Bank PSC 7. United Bank Limited 8. Saderat Iran Bank
	National Banks	9. Qatar National Bank 10. Qatar Islamic Bank 11. Doha Bank 12. Commercial Bank of Qatar 13. Qatar International Islamic Bank 14. Al-Ahli Bank of Qatar

UAE Banks	Foreign Banks	1. Ahli Bank of Kuwait 2. ABN AMRO Bank 3. Arab Bank 4. Arab African Bank 5. Bank Melli Iran 6. Bank of Baroda 7. Bank Saderat Iran 8. Cairo Bank 9. Banorab Bank 10. Indosuez Bank 11. Lebanese Bank 12. Pariba Bank 13. Barclays Bank 14. Citibank 15. Standard Chartered Grindlays Bank 16. Habib Bank Zurich 17. Habib Limited Bank 18. Janata Bank 19. Lloyds Bank 20. National Bank of Bahrain 21. Standard Chartered Bank 22. HSBC

UAE Banks	National Banks
	23. Abu Dhabi Commercial Bank 24. Abu Dhabi Investment Company 25. Arab Bank for Investment & Foreign Trade 26. Arab Emirates Investment Bank 27. Bank of Sharjah 28. Commercial Bank International 29. Commercial Bank of Dubai 30. Dubai Islamic Bank 31. Emirates Bank International 32. First Gulf Bank 33. Investbank 34. Mashreqbank 35. Middle East Bank 36. National Bank of Abu Dhabi 37. National Bank of Dubai Public Joint Stock Company 38. National Bank of Fujairah 39. National Bank of Ras Al-Khaimah 40. National Bank of Sharjah 41. National Bank of Umm Al-Qaiwain 42. Union National Bank 43. United Arab Bank

Saudi Arabia Banks	1. Al Bank Al Saudi Al Fransi 2. Arab National Bank 3. Bank Al-Jazira 4. National Commercial Bank 5. Riyad Bank 6. Saudi American Bank 7. Saudi British Bank 8. Saudi Hollandi Bank 9. Saudi Investment Bank

Kuwait Banks	1. Alahli Bank of Kuwait 2. Bank of Kuwait & The Middle East 3. Burgan Bank 4. Commercial Bank of Kuwait 5. Gulf Bank KSC 6. National Bank of Kuwait 7. Kuwait Finance House 8. Gulf Investment Corporation 9. Industrial Bank of Kuwait. 10. Kuwait Investment Company

Bahrain Banks	1. Al-Ahli Commercial Bank. 2. Arab Banking Corporation 3. Bahraini Saudi Bank 4. Bank of Bahrain and Kuwait. 5. Commercial Bank of Bahrain 6. Gulf International Bank 7. National Bank of Bahrain 8. Bahrain Islamic Bank 9. TAIB Bank 10. Bahrain International Bank 11. United Gulf Bank

Oman Banks	1. Bank Dhofar Al-Omani Al-Fransi 2. Bank Muscat 3. National Bank of Oman 4. Oman Arab Bank 5. Oman International Bank 6. Oman International Development and Investment Co.