

**The Impact of Financial Liberalisation on the
Efficiency of Malaysian Banks: An Empirical Analysis
using Frontier Measurements**

by

Azrie Tamjis

A Doctoral Thesis

**Submitted in Partial Fulfilment of the Requirements for the Award of
Doctor of Philosophy of Loughborough University**

© Azrie Tamjis, September 2014

Abstract

The Asian financial crisis in 1997–98 left a severe impact on Malaysia’s economy and banking system. This has forced the Malaysian government to undertake financial restructuring initiatives to restore market and public confidence, and to meet the ongoing challenges associated with market structure, financial innovation and globalisation. Therefore, Bank Negara Malaysia (BNM) introduced a ten-year Financial Sector Master Plan (FSMP) to strengthen domestic banks and the regulatory structure, and to promote the banks’ efficiency by stimulating a competitive banking industry through financial liberalisation. The crisis for banks in Malaysia and the region has been extensively studied in the past (Sufian, 2010). However, empirical studies of the post-crisis period, and the implementation of the FSMP, remain limited. Hence, a data set of all banks in Malaysia, which covers the period 2000–2011, was employed to examine the effect of the FSMP’s initiatives on Malaysian banks’ efficiency between 2000 and 2011. To measure this efficiency, this study employs both parametric and nonparametric models: namely, stochastic frontier analysis (SFA) and data envelopment analysis (DEA). Economic functions such as, cost-, standard profit- and alternative profit-efficiency were used in a 1-stage SFA model, which includes control variables (e.g. capital adequacy, asset quality and liquidity) and environmental variables (e.g. ownership, size, specialisation, deregulation periods and market structure) in the model specifications. In addition, this study employs SFA as the main measurement method, while the DEA model was used to cross-check consistency (Resti, 1997; Bauer et al., 1998). Both SFA and DEA demonstrated that, in most cases, the consistency was moderate.

The level of cost efficiency of Malaysian banks worsened over the years 2000–2011, with average cost efficiency during this period was at 76.5%. Despite the various liberalisation measures introduced to the banking industry – particularly during the three phases of the FSMP; 2000–2003; 2004–2007; 2008–2011 – cost efficiency trended downward, due to the effects of consolidation by domestic banks, deregulation of interest rates, the introduction of foreign Islamic banks, and the global credit crisis. Banks in Malaysia were forced to adjust their inputs and outputs to the rapid changes in the banking industry, which might have made a negative impact on cost efficiency. On the other hand, the banks demonstrated a steadily increasing profit efficiency trend, which fluctuated with the introduction of interest

rate liberalisation (early second phase of the FSMP (i.e. 2004)) and during the global credit crisis (early third phase of the FSMP (i.e. 2008)). The average profit efficiency for 2000–2011 was 93.3%. The profit efficiency exhibited an increasing trend in the first (2000-2003) and second (2004-2007) phases of the FSMP, suggesting that the effect of consolidation by domestic banks had resulted in higher market concentration and greater market power among the remaining banks. However, the profit efficiency average scores fell in 2004, 2008 and 2011. This is attributed to the deregulation of interest rates, the deleveraging of the inflow of foreign funds, and the rapid increase in policy interest rates. At a more granular level, domestic banks were found to be more cost efficient, but marginally less profit efficient, when compared to foreign banks. In terms of bank specialisation, conventional banks were more cost- and profit-efficient than Islamic banks. With regard to economies of scale, the majority of Malaysian banks revealed scale economies, illustrated by a U-shape, with medium-sized banks being more scale efficient than small and large banks.

These results suggest that, to enhance Malaysian banks' efficiency, the government must maintain competitive pressure on the large domestic banks that were consolidated during the first phase of the FSMP (2000-2003). Policymakers may want to further open up banking markets, improve risk management and governance, encourage financial innovation, and support expansion of smaller banks. The implementation of deregulation initiatives during periods of uncertainty (e.g. the global credit crisis) have also resulted in decreasing trend of cost and profit efficiency. Hence, monitoring initiatives, using tools such as frontier measurement is important for regulator's macro- and micro-prudential surveillance.

Keywords: Cost efficiency, Profit efficiency, Stochastic Frontier Analysis, Data Envelopment Analysis, Malaysian Banks, Financial Liberalisation, Financial Sector Master Plan.

Table of Contents

Abstract	i
Table of Contents	iii
List of Figures	vi
List of Tables	viii
Abbreviations	ix
Acknowledgements	xi
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.2 Aims and Motives of Research	4
1.3 Research Methodology and Data	6
1.4 Research Outline	7
CHAPTER 2 THEORY AND MEASUREMENT OF PRODUCTION EFFICIENCY	11
2.1 Introduction	11
2.2 Production Efficiency and the Theory of the Firm	12
2.2.1 Economic Theory and Efficiency	14
2.2.2 Competitive Model and Efficiency	18
2.2.3 Market Imperfections	21
2.2.4 Market Structure and Performance of the firm hypotheses	24
2.2.5 Causes of Inefficiency of the Firm	28
2.3 The Measurement of Efficiency of the Firm Using Frontier Efficiency	36
2.3.1 Technical and Allocative Efficiency	38
2.3.1.1 Input-orientated Measures of Technical and Allocative Efficiency	39
2.3.1.2 Output-orientated Measures of Technical Efficiency and Allocative Efficiency	42
2.3.2 Scale Efficiency	44
2.3.3 Scope Efficiency	49
2.4 The Measurement of Frontier Efficiency	51
2.4.1 Nonparametric Frontier Approaches	53
2.4.1.1 Data Envelopment Analysis (DEA)	54
2.4.1.2 Free Disposal Hull (FDH) Analysis	56
2.4.2 Parametric Frontier Approaches	57
2.4.2.1 Deterministic Frontier Model	58
2.4.2.2 Stochastic Frontier Analysis (SFA)	59
2.4.2.3 Thick Frontier Approach (TFA)	61
2.4.2.4 Distribution Free Approach (DFA)	61
2.5 Functional Forms	62
2.5.1 Cobb-Douglas Functional Form	63
2.5.2 Translog Functional Form	64
2.5.3 Fourier Flexible Functional Form	66
2.6 Is There a Best Frontier Approach?	67
2.7 Conclusions	69
CHAPTER 3 EMPIRICAL STUDIES OF BANKING EFFICIENCY	72
3.1 Introduction	72
3.2 The Role of the Banking Sector in Economic Growth and Development	73
3.3 Government Intervention in the Banking Sector	76
3.4 Empirical Studies of Banking Efficiency	84
3.4.1 Empirical Studies of Banking Efficiency in Developed Countries	85
3.4.2. Empirical Studies of Banking Efficiency in Developing/Transition Countries	102
3.4.3 Empirical Studies of Banking Efficiency in Malaysia	115

3.5 Conclusion	121
CHAPTER 4 THE MALAYSIAN BANKING SECTOR, 2000–2011.....	123
4.1 Introduction.....	123
4.2 Brief Overview of the Malaysian Economy.....	124
4.3 Overview of the Malaysian Banking System between 2000 and 2011	126
4.4 The Structure of Malaysian Banks.....	150
4.4.1 The Ownership of Malaysian Banks.....	150
4.4.2 Malaysian Banks by Specialisation	152
4.4.3 Assets and Liabilities Structure of Malaysian Banks	155
4.4.4 Income Statements Structure of Malaysian Banks	166
4.5 Financial Ratio Analysis of Malaysian Banks	175
4.6 Competition Indicators of the Malaysian Banking Sector	178
4.7 The Development of Research Hypotheses	180
4.8 Conclusion	184
CHAPTER 5 METHODOLOGY AND DATA	187
5.1 Introduction.....	187
5.2 Methodology	189
5.2.1 Stochastic Frontier Analysis (SFA).....	191
5.2.1.1 Cost- and Profit-Efficiency	192
5.2.1.2 Scale-Efficiency	201
5.2.2 Data Envelopment Analysis (DEA).....	202
5.2.2.1 Cost- and Profit-Efficiency	206
5.2.2.2 New DEA Cost- and Profit-Efficiency.....	208
5.2.2.3 Scale-Efficiency	210
5.2.3 Consistency Tests between SFA and DEA Models.....	210
5.3 Data and Variables	212
5.3.1 Input and Output Variables.....	213
5.3.1.1 Input Prices, Outputs and Output Prices.....	216
5.3.2 Control and Environmental Variables	221
5.3.2.1 Control Variables	223
5.3.2.2 Environmental Variables.....	230
5.4 Stochastic Frontier Model(s) Specification.....	238
5.4.1 Performing 1-Stage Analysis	241
5.4.2 Identifying the Preferred Model.....	244
5.5 Conclusion	256
CHAPTER 6 EMPIRICAL RESULTS: ANALYSIS AND DISCUSSION	258
6.1 Introduction.....	258
6.2 Examining the Coefficients of the Preferred SFA Models	260
6.2.1 Coefficients of the Preferred Cost Frontier.....	261
6.2.2 Coefficients of the Preferred Standard Profit and Alternative Profit Frontiers	265
6.3.1 Testing of Hypotheses 2 to 5	275
6.3.2 Testing of Hypotheses 7 to 9	280
6.4 Impact of Liberalisation via the FSMP on the Efficiency of Malaysian Banks.....	284
6.4.1 SFA Efficiency Scores.....	284
6.4.1.1 SFA Cost Efficiency Scores	284
6.4.1.2 SFA Standard Profit- and Alternative Profit-Efficiency Scores.....	289
6.4.2 DEA Efficiency Scores.....	293
6.4.2.1 DEA and NDEA Cost Efficiency Scores	294
6.4.2.2 DEA and NDEA Profit Efficiency Scores.....	295
6.5 Consistency Condition Tests of SFA and DEA	298
6.5.1 Comparison of Efficiency Distributions	300
6.5.2 Correlations of Efficiency Rankings.....	301

6.5.3 Identification of Best-Practice and Worst-Practice Banks.....	303
6.5.4 Stability of Efficiency Scores over Time.....	305
6.5.5 Efficiency and Accounting-based Performance Measures	307
6.6 Characteristics of Efficiency Scores of Malaysian Banks.....	311
6.6.1 The CAMEL Rating system	311
6.6.2 Ownership Structure	316
6.6.3 Bank Specialisation.....	323
6.6.4 Bank Size	327
6.7 Scale Economies and Technological Change.....	330
6.7.1 Scale Economies	330
6.7.2 Technological Change	335
6.8 Conclusions.....	338
CHAPTER 7 CONCLUSIONS	344
7.1 Introduction and Summary of Findings	344
7.2 Policy Implications	350
7.3 Limitations of Present Research and Future Research.....	352
BIBLIOGRAPHY	354

List of Figures

Figure 1.1 Structure of the Thesis	8
Figure 2.1 Information Flow of Chapter 2	12
Figure 2.2 Productivity and Technical Efficiency: A Single Input and a Single Output	18
Figure 2.3 Structure-Conduct-Performance Paradigm	27
Figure 2.4 Technical and Allocative Efficiency (Input-oriented)	40
Figure 2.5 Technical and Allocative Efficiency (Output-orientated)	43
Figure 2.6 Economies of Scale and Average Cost Curves	46
Figure 2.7 Economies of Scale	47
Figure 2.8 Effects of Scale Production	48
Figure 2.9 The Concept of Economies of Scope	50
Figure 2.10 Conceptual Efficiency Framework	53
Figure 2.11 Data Envelopment Analysis	55
Figure 2.12 Free Disposal Hull	57
Figure 3.1 Information Flow of Chapter 3	73
Figure 4.1 Information Flow of Chapter 4	124
Figure 4.2 The Structure of the Malaysian Financial System as of December 2011	128
Figure 4.3 Overview of Malaysian Banking Sector Development, 1997-2011	132
Figure 4.4 Restructuring Plan for the Malaysian Banking Sector	134
Figure 4.5 The three Phases of the Financial Sector Master Plan, 2001-2010	136
Figure 4.6 Total Assets of Domestic and Foreign Banks in Malaysia	152
Figure 4.7 Total Assets of Conventional and Islamic Banks in Malaysia	155
Figure 4.8 Total Assets, Total Liabilities and Capital in Malaysian Banks	156
Figure 4.9 Percentage of Capital, Amounts Owing to Other Banks and Deposits Accepted relative to Liabilities and Capital, 2000–2011	158
Figure 4.10 Annual Growth of Deposits Accepted, Amounts Owing to Other Banks and Capital of Malaysian Banks, 2000–2011	160
Figure 4.11 Annual Growth of Loans, Investments and Amounts Due from Other Banks, 2000–2011	163
Figure 4.12 Percentage of Loans, Investments and Amounts Due from Other Banks Relative to Total Assets, 2000–2011	164
Figure 4.13 Growth and Structure of Income of Malaysian Banks, 2000–2011	167
Figure 4.14 Composition of Income of Malaysian Banks, 2000–2011	167
Figure 4.15 Growth and Structure of Expenses in Malaysian Banks, 2000–2011	168
Figure 4.16 Composition of Expenses in Malaysian Banks, 2000–2011	169
Figure 4.17 Trends of Profit before Tax, Total Income and Total Expenses of Malaysian Banks, 2000–2011	171
Figure 4.18 Annual Growth in Total Income, Total Expenses and Profit before Tax in Malaysian Banks, 2000–2011	172
Figure 4.19 Trend of Interest Margin in Malaysian Banks, 2000–2011	173
Figure 4.20 Trends of Profit Performance Indicators in Malaysian Banks, 2000–2011	176
Figure 4.21 Trends of Cost Performance Indicators in Malaysian Banks, 2000–2011	177
Figure 4.22 Trends of Concentration and Competitive Indicators in Malaysian Banks, 2000–2011	179
Figure 5.1 Information Flow for Chapter 5	188
Figure 5.2 Overview of Frontier Efficiency Estimations	190
Figure 5.3 Accounting for Environmental Variables in Stochastic Frontier Analysis	223
Figure 5.4 Stages Performed in Deriving Preferred Model	242
Figure 6.1 Information Flow of Chapter 6	260
Figure 6.2 Average SFA Cost Efficiency Scores of Malaysian Banks, 2000–2011	286

Figure 6.3 Average SFA Standard Profit and Alternative Profit Efficiencies Scores of Malaysian Banks, 2000–2011.....	289
Figure 6.4 Average DEA and NDEA Cost Efficiency Scores of Malaysian Banks, 2000–2011	294
Figure 6.5 Average DEA and NDEA Profit Efficiency Scores of Malaysian Banks, 2000–2011....	296
Figure 6.6 Cost Efficiency Scores of Domestic and Foreign Banks, 2000–2011	319
Figure 6.7 Profit Efficiency Scores of Domestic and Foreign Banks, 2000–2011	322
Figure 6.8 Cost Efficiency Scores for Conventional and Islamic Banks, 2000–2011	325
Figure 6.9 Profit Efficiency Scores for Conventional and Islamic Banks, 2000–2011	326
Figure 6.10 Trends of Scale Economies for Malaysian Banks, 2000–2011	333

List of Tables

Table 3.1 Efficiency Studies in US and Developed Countries Banking System	98
Table 3.2 Efficiency Studies in Transition and Developing Countries Banking System.....	112
Table 4.1 Malaysian Banks as of December 2011	129
Table 4.2 Liberalisation Measures in Malaysia, 2000-2011	149
Table 4.3 Total Assets of Malaysian Banks by Ownership	151
Table 4.4 Total Assets of Malaysian Banks by Specialisation	154
Table 4.5 Sources of Funds of Malaysian Banks (Liabilities and Capital), 2000–2011	161
Table 4.6 Uses of Funds of Malaysian Banks (Total Assets), 2000–2011	165
Table 4.7 Income Statement Structure for Malaysian Banks, 2000–2011	174
Table 5.1 Descriptive Statistics of Banks’ Inputs and Outputs, 2000–2011	220
Table 5.2 Descriptive Statistics for Control Variables, 2000–2011	230
Table 5.3 Descriptive Statistics for Environmental Variables, 2000–2011	238
Table 5.4 Stages Involved in Searching for Preferred Model	243
Table 5.5 Estimated Cost Frontier Models	253
Table 5.6 Estimated Standard Profit Frontier Models.....	254
Table 5.7 Estimated Alternative Profit Frontier Models.....	255
Table 6.1 The Maximum Likelihood Parameters for the Preferred Cost Frontier	263
Table 6.2 The Maximum Likelihood Parameters Estimation for the Preferred Standard Profit.....	268
Table 6.3 The Maximum Likelihood Parameters Estimation for the Preferred Alternative Profit...	271
Table 6.4 Testing of Research Hypotheses 2–9	284
Table 6.5 Average SFA Cost Efficiency Scores, 2000–2011	285
Table 6.6 Average SFA Standard Profit- and Alternative Profit-Efficiency	293
Table 6.7 Average DEA and NDEA Cost Efficiency Scores of Malaysian Banks, 2000–2011.....	295
Table 6.8 Average DEA and NDEA Profit Efficiency Scores of Malaysian Banks, 2000–2011	297
Table 6.9 Descriptive Statistics of Efficiency Scores of SFA, DEA and NDEA.....	301
Table 6.10 Spearman Rank-Order Correlation between SFA, DEA and NDEA efficiency Scores	303
Table 6.11 Correspondence of Best Practice and Worst Practice of Cost Efficiency across Techniques	305
Table 6.12 Correspondence of Best Practice and Worst Practice of Profit Efficiency across Techniques	305
Table 6.13 Stability of Efficiency Scores – The Correlations of k-Year-Apart Efficiencies.....	307
Table 6.14 Correlations between Frontier Cost Efficiencies and Non-Frontier Performance Measures	309
Table 6.15 Correlations between Frontier Profit Efficiencies and Non-Frontier Performance Measures	310
Table 6.16 Descriptions of CAMEL Rating System.....	313
Table 6.17 CAMEL indicators for Best and Worst Practice Banks	316
Table 6.18 Cost Efficiency Scores of Domestic and Foreign Banks, 2000–2011	320
Table 6.19 Profit Efficiency Scores of Domestic and Foreign Banks, 2000–2011.....	322
Table 6.20 Cost Efficiency Scores for Conventional and Islamic Banks, 2000–2011.....	325
Table 6.21 Profit Efficiency Scores for Conventional and Islamic Banks, 2000–2011	327
Table 6.22 Cost Efficiency Scores of Malaysian Banks According to Asset Size, 2000-2011	329
Table 6.23 Profit Efficiency Scores of Malaysian Banks According to Asset Size, 2000-2011	330
Table 6.24 Scale Economies for Malaysian Banks, 2000–2011	332
Table 6.25 Scale Efficiency for Malaysian Banks, 2000–2011	335
Table 6.26 Technological Change of Malaysian Banks, 2000–2011	337
Table 6.27 Decomposition of Technical Changes	338

Abbreviations

AE	Allocative Efficiency
APE	Alternative Profit Efficiency
ATM	Automated Teller Machine
BAFIA	Banking and Financial Act
BIS	Bank for International Settlements
BLR	Base Lending Rate
BNM	Bank Negara Malaysia (Central Bank of Malaysia)
CBA	Central Bank Act
CDRC	Corporate Debt Restructuring Committee
CE	Cost Efficiency
CEO	Chief Executive Officer
CIR	Cost Income Ratio
COLS	Corrected ordinary least squares
CRS	Constant Returns to Scale
DEA	Data Envelopment Analysis
DFA	Distribution Free Approach
DMU	Decision Making Units
DRS	Decreasing Returns to Scale
EU	European Union
FDH	Free disposal hull
FDI	Foreign Direct Investment
FF	Fourier-flexible
FSMP	Financial Sector Master Plan
GDP	Gross Domestic Products
HHI	Herfindahl-Hirschman index
IBS	Islamic Banking Scheme
IMF	International Monetary Fund
IRS	Increasing Returns to Scale
KLSE	Kuala Lumpur Stock Exchange
LAC	Long-run average cost curve
LR	Log likelihood ratio
Max.	Maximum Value
MC	Marginal Cost
MEPS	Malaysian Electronic Payment System
MGS	Malaysian Government Securities
Min.	Minimum Value
ML	Maximum Likelihood
MOLS	Modified Ordinary Least Square
MR	Marginal Revenue
MYR	Malaysian Ringgit
NDEA	New Data Envelopment Analysis
NEP	New Economic Policy
NPLs	Non-Performing Loans
OLS	Ordinary Least Square
OPR	Overnight Policy Rate
RAM	Rating Agency Malaysia
RE	Revenue Efficiency

ROA	Return on Assets
ROE	Return on Equity
RWCR	Risk Weighted Capital Ratio
S.D.	Standard Deviation
SAC	Short-run average cost curve
SE	Scale Efficiency
SFA	Stochastic Frontier Analysis.
SMEs	Small and medium enterprises
TA	Total Assets
TC	Total Costs
TE	Technical Efficiency
TFA	Thick Frontier Approach
TFP	Total Factor Productivity
TP	Total Profits
USA	United States of America
VRS	Variable Returns to Scale
WTO	World Trade Organisation

Acknowledgements

Alhamdulillah, all praise to Allah, the most gracious and the most merciful.

I wish to express my utmost gratitude to my supervisor, Dr Robert Hamilton for his efforts in giving endless guidance, continuous commitment, and motivation throughout this study. Without his thorough approach and our in-depth discussions, I would certainly not have the same confidence about my work. Particularly, his patience in reviewing my writing and giving constructive criticism on many previous drafts has enabled me to complete this thesis. He has been a constant source of inspiration and his contribution to my personal development extends far beyond this thesis. I would also like to express my gratitude to my co-supervisor, Dr Keith Pond, for all his tireless support, motivation and administrative assistance given during the process of my research. His encouragements, thoughtful advice, generosity in giving considerable time, and priceless suggestion have been of great assistance for this research.

A special thanks to Professor Tom Weyman-Jones and Dr Yizhe Dong with whom I have had the chance to consult on theoretical and methodological problems during the research process.

I would also like to acknowledge the financial support from my sponsor, Bank Negara Malaysia, without which this thesis could not have become a reality.

Most of all, I am very grateful to my family, especially my late mother, Rokiah Salleh, who passed away during my study, my father, Tamjis Tahir for his consistent encouragement to acquire knowledge, my brother, my parents-in-law, my wife, Azmiah Abdul Razak and my children, Muhammad Haziq, Alya Qistina, Adibah Humaira and Adilah Humaira for their love, understanding, encouragement and patience. Last but not least, thanks to all my friends and family, who in different ways have all inspired me throughout this journey.

May Allah bless us. Ameen.

Chapter 1 Introduction

1.1 Background

Developing countries need efficient intermediation of funds from savers to borrowers to enable productive allocation of resources, which then contributes to higher economic growth (McKinnon, 1973; Shaw, 1973). Thus, the development of the financial system is essential for the general enhancement of productivity and economic growth in developing countries. However, financial systems in developing countries generally demonstrate excessive government intervention in mobilising and allocating resources, which inhibits competition and efficient resource allocation. As a result, many developing countries have progressively implemented financial liberalisation and reforms to promote competition and efficiency in resource allocation (Kumbhakar and Sarkar, 2003).¹ Such liberalisation has increased competitiveness among banks significantly by inducing them to adjust their strategies in key areas (e.g. in the development of new delivery systems and service quality) (Casu and Girardone, 2006). A competitive environment also encourages banks to be more efficient by reducing operating costs, enhancing overall management, improving risk management, and offering new products and services (Denizer et al., 2000). So, higher competition can cause banks to reduce their costs and therefore increase their efficiency, leading to more efficient allocation of resources and higher economic growth (McKinnon, 1973; Shaw, 1973; Fry, 1995).² As a consequence, issues regarding banks' efficiency and financial liberalisation in developing countries have become more important for micro- and macro-stability of the banking industry (Kolari and Zardkoohi, 1987).

As mentioned earlier, the banking system seeks to enhance resource allocation and aims to improve the efficiency of banks (Demirguc-Kunt and Levine, 2008). However, developments in the banking industry may suffer distortions as a result of financial crises. In a deep recession, an economy may experience a sharp decline in economic growth. For

¹ The financial liberalisation programme includes, but is not limited to, a deregulation of exchange and interest rate controls, reduction in government directed lending, privatisation of government-owned banks, a reduced entry barrier for foreign banks and a reduction in the statutory reserve ratio (Caprio et al., 1997).

² For instance, greater market competition in the banking system could reduce market power and therefore should reduce the price of financial services, resulting in welfare gains for the public and customers (McKinnon, 1973; Shaw, 1973).

instance, Malaysia was hit by the Asian financial crisis in 1997–98: the Malaysian Ringgit fell by 40.0% against the US Dollar; the stock market plunged by over 70.0%; resulting in extreme volatility in financial markets; and the country's sovereign rating being downgraded (Jomo and Chin, 2001; Sufian, 2009). The scenario worsened as economic activity declined: GDP contracted by 7.5% with weak regional export demand; and companies were in distress and unable to service debt and over leveraging. In the banking system, the number of non-performing loans (NPLs) increased sharply, which caused capital erosion due to over-concentration of risk (mainly in the large corporate sector) (Bank Negara Malaysia, 1997, 1998). At the same time, the intermediation process was also inefficient due to tight liquidity, with loan growth declined sharply (Rajoo, 2008).

This crisis negatively impacted the banking sector in Malaysia, resulting in many problems, such as an increase in NPLs. So, Malaysia initiated a comprehensive crisis management arrangement to restore stability, expedite recovery and restart growth without relying on International Monetary Fund (IMF) assistance. Due to this early intervention, Malaysia has managed a speedy recovery at minimal costs to the government (less than 5% of GDP) (Bank Negara Malaysia, 1999, 2000). Realising the vulnerabilities of the financial sector, the central bank then introduced a plan outlining the strategic focus, common goals and action to be taken through the Financial Sector Master Plan (FSMP). The FSMP is a sequence of measures to accomplish the desired financial sector. It has goals for maintaining financial stability and makes 119 recommendations in six broad clusters (Bank Negara Malaysia, 2001b). The FSMP spans over three phases: in the first phase, the focus was on building capacity in domestic financial institutions and strengthening the regulatory framework.³ Subsequently, phase two was a vision to level the playing field for incumbent foreign banks, in which some restrictions on foreign banks were removed. Phase three will see Malaysia open up its banking industry to foreign players, in line with the World Trade Organisation (WTO)'s liberalisation program (Bank Negara Malaysia, 2001b; Hon et al, 2011).

Prior to the Asian financial crisis, the banking system in Malaysia was fragmented with 33 domestic banking institutions. Banks at that time had little capital and were unable to face

³ This capacity building is imperative in laying a solid foundation on which financial institutions can strengthen and compete effectively amid further liberalisation without accumulation of risk factors that may adversely affect financial stability and growth (Bank Negara Malaysia, 2001b; Hon et al., 2005).

the macro pressures arising from economic vulnerability. The banks also over-relied on corporations for their finance, where less attention was given to small and medium enterprises (SMEs) and retail consumers, which caused gaps in access to finance. Companies relied too heavily on banks for their finance, and the bond capital market was not sufficiently developed to support the need for funds of large firms (Bank Negara Malaysia, 1999b). Before the Asian crisis, banks were lacking in terms of effective risk management and corporate governance, leaving them highly fragile. Furthermore, the market conditions were very rigid, where rules-based regulation and micro-based supervision were implemented in the financial sector. The pricing mechanism of banking products and services was also rigid (e.g. government controlled interest rates), which did not encourage competition among financial players (Ang and McKibbin, 2007; Said et al., 2008). In terms of Islamic banking, the prominence of Islamic finance was limited and not given full attention with regard to its possibilities and potentials (Abdul Majid et al., 2011). In general, the banking industry in Malaysia was distorted by interest rate controls and selective credit policies, a lack of competition, and weak supervision by the regulators.

The FSMP initiatives changed the financial landscape of the banking industry. As of 2011, the banking industry was consolidated and rationalised, from 33 domestic financial institutions into 10 banking groups (Abdul Majid et al., 2011). Banks were also found to be diversifying and improved their efficiency in delivery channels for financial products and services by enhancing access to financing, particularly for SMEs and consumers (Hon et al., 2011). These changes diversified the financial sector, with a deep and liquid debt securities market and a better focus on investment banks assisting corporations to get alternative finance in the bond market. Banks are now more focused on corporate governance and risk management, particularly with the implementation of principles-based regulations, coupled with an adequate supervisory and surveillance framework (Abdul Ghani, 2010). Moreover, the market structure has improved, with an increased emphasis on market orientation, supported by greater regional cooperation, increased competitive pressure from new and current foreign banks, and freedom in the pricing of lending and deposits (Tahir et al., 2008). Similar to initiatives in other developing countries, the objectives of these reforms and liberalisations, via the FSMP, are to promote diversity, efficiency and productivity, and to facilitate a competitive banking system by improving resource allocation and building a stronger economy (Fry, 1995). As a consequence, banking efficiency received even more attention in the aftermath of the financial crisis, with structural reforms and liberalisations,

rendering the examination of this banking efficiency an important issue for both the public and policymakers (Berger and Mester, 1997).⁴

1.2 Aims and Motives of Research

The aims of this research are: first, to carry out an extensive cost- and profit-efficiency analysis of Malaysian banks for the years 2000–2011 using frontier approaches and examining how changes in the financial services affected efficiency, productivity and the market structure of the banking industry in Malaysia. Second, to examine the impact of market liberalisation initiatives, via the FSMP, on efficiency and productivity in Malaysian banks. And third, to investigate the determinants of efficiency in Malaysian banks, and their significance, by examining bank-specific (i.e. inherent risks (e.g. capital adequacy, asset quality and liquidity), ownership structure, bank specialisation, and bank size) and market environmental (i.e. market concentration, FSMP's phases and the global credit crisis) factors.

There are very few empirical efficiency studies of Malaysian banking, compared with the vast number of such studies in the United States (US) and European banking systems. Additionally, this research is motivated by areas that still need to be fully addressed in relation to the empirical studies of Malaysian banks. First, this study covers the entire period of implementation of the FSMP. Although Hon et al. (2011) attempted to examine the plan's impact, the data set they employed did not span the entire timeline of the FSMP and may not reflect the competitive aspects of the banking system as it only compared domestic banks. Second, most Malaysian bank studies focus on the banks' cost efficiencies; a limited number of empirical studies look into the profit efficiency function. There are even fewer studies that conduct and compare both cost- and profit-efficiency using the same Malaysian bank data set. Third, to the author's knowledge, there are no studies in Malaysia that perform a comprehensive consistency analysis between nonparametric and parametric methods for cost- and profit-efficiency. Bauer et al. (1998) suggested six consistency conditions rather than only concentrating on rank-order correlation for consistency. Fourth,

⁴ Improved banking efficiency could result in better resource allocation, which benefits society by intermediating greater amounts of funds, providing more products with better prices and service quality for customers, improving bank profitability and achieving greater safety and soundness in banking sector (Berger and Mester, 1997). Therefore, the study of efficiency could assist banking regulators to design policies by evaluating the impacts of financial liberalisation, consolidation and market structure on efficiency.

the timeline of data used for most Malaysian banking efficiency studies is usually short (on average, between three and eight years) with a smaller sample of banks, which may not represent the entire banking system, particularly as frontier measurement encompasses the relative measurement concept. Fifth, studies of explanatory factors, or the causes of inefficiency, are also lacking, which may result in a failure to explain the factors that affect the inefficiency of Malaysian banks. Sixth, at the time of writing this thesis, to the author's knowledge, there are no studies on the efficiency of Malaysian banks that cover the impact of the global credit crisis, which occurred during the implementation of the FSMP. Further to the areas discussed above, this present research is important for a number of other reasons.

First, this study addresses a contemporary policy issue in relation to market structure. One of the primary objectives of the FSMP is to liberalise and increase market competition in the banking sector. Thus, this study examines how banking structure, efficiency and productivity change affect bank performance, which will provide evidence for policy changes to market competition. This would also assist policymakers to assess the level of achievement they have made and anticipate potential risks within the market. Additionally, this study also investigates scale economies and scale efficiency, which may help policymakers determine the appropriate size of banks and ascertain the most suitable market structure for the Malaysian banking market.

Second, frontier efficiency models could be used to supplement the central banks' risk assessment tools, particularly in identifying the causes of inefficiency. Hence, it is useful for regulators and policymakers, such as central banks, to employ frontier efficiency models for banking industry assessment, and consequently exercise early interventions and maintain market order through their micro- and macro-prudential regulatory policies.

Third, this study will extend and complement the existing literature of international banking efficiency. It will be the first to cover the impact of the FSMP on the Malaysian banking industry and will expand the literature on financial liberalisation, particularly within developing countries; which is currently dominated by studies of developed nations.

Fourth, this study is important as it employs a dataset that comprises all commercial and Islamic banks (including foreign and domestic-owned) in Malaysia, providing a comprehensive overview of the Malaysian banking sector over the period 2000–2011.

Moreover, different characteristics of the banks are examined regarding CAMEL ratings system, ownership structure (i.e. foreign vs. domestic banks), by specialisation (i.e. Islamic vs. conventional banks) and size.

In particular, this study seeks to address the following questions:

1. How cost- and profit-efficient are Malaysian banks? What is the general level of cost- and profit-efficiency of the Malaysian banking sector and how has it varied over time?
2. Has the financial liberalisation of the banking sector via the FSMP improved the efficiency and productivity of the banking industry in Malaysia?
3. What is the impact on Malaysian banks' cost- and profit-efficiency with regard to capital, non-performing loans, and liquidity?
4. Does cost- and profit-efficiency vary across bank ownership (i.e. foreign vs. domestic), specialisation (i.e. Islamic vs. conventional) and size (i.e. in terms of asset size)?
5. What was the impact of the global credit crisis on the cost- and profit-efficiency of Malaysian banks?
6. Do the parametric stochastic frontier analysis (SFA) and nonparametric data envelopment analysis (DEA) methods provide consistent results?
7. Are there any economies and diseconomies of scale in the Malaysian banking sector? And do the returns of scale differ across bank size?
8. Has there been any technological change in the Malaysian banking sector?

1.3 Research Methodology and Data

In this study, frontier measurement is employed to measure the efficiency of Malaysian banks for the years 2000–2011 and to answer the above questions. For better estimation of cost- and profit-efficiency, and taking into account the effect of heterogeneity (e.g. ownership structure, banks specialisation, inherent risks, and size), this study uses Battese and Coelli's (1995) SFA 1-stage approach, which may have an impact on the efficiencies. In

this one-stage approach, a set of control variables (e.g. capital adequacy, asset quality and liquidity) and environmental variables (e.g. ownership, specialisation, financial liberalisation and size) are included into the specification of cost- and profit-efficiency functions. Failure to include these variables in the 1-stage approach can result in biased estimation of the parameters of the cost- and profit-frontier and scores (Coelli et al., 2005). These different sets of control and environmental variables are tested in several stages using statistical testing (i.e. the log-likelihood ratio test), searching for the best fitting model that is later utilised for the estimation of efficiency scores in Malaysian banks (discussed later in Chapter 5).

As suggested by Bauer et al. (1998), when more than one technique is used for efficiency studies, comparing different techniques for consistency could assist various parties for better approximation of efficiency measures, which is useful for policymakers. Therefore, this study employs two DEA models for cost- and profit-efficiency and tests their consistency with SFA. For consistency testing, following Bauer et al. (1998), a variety of consistency tests are applied to assess the robustness of the results derived from these two models (i.e. parametric vs. nonparametric).

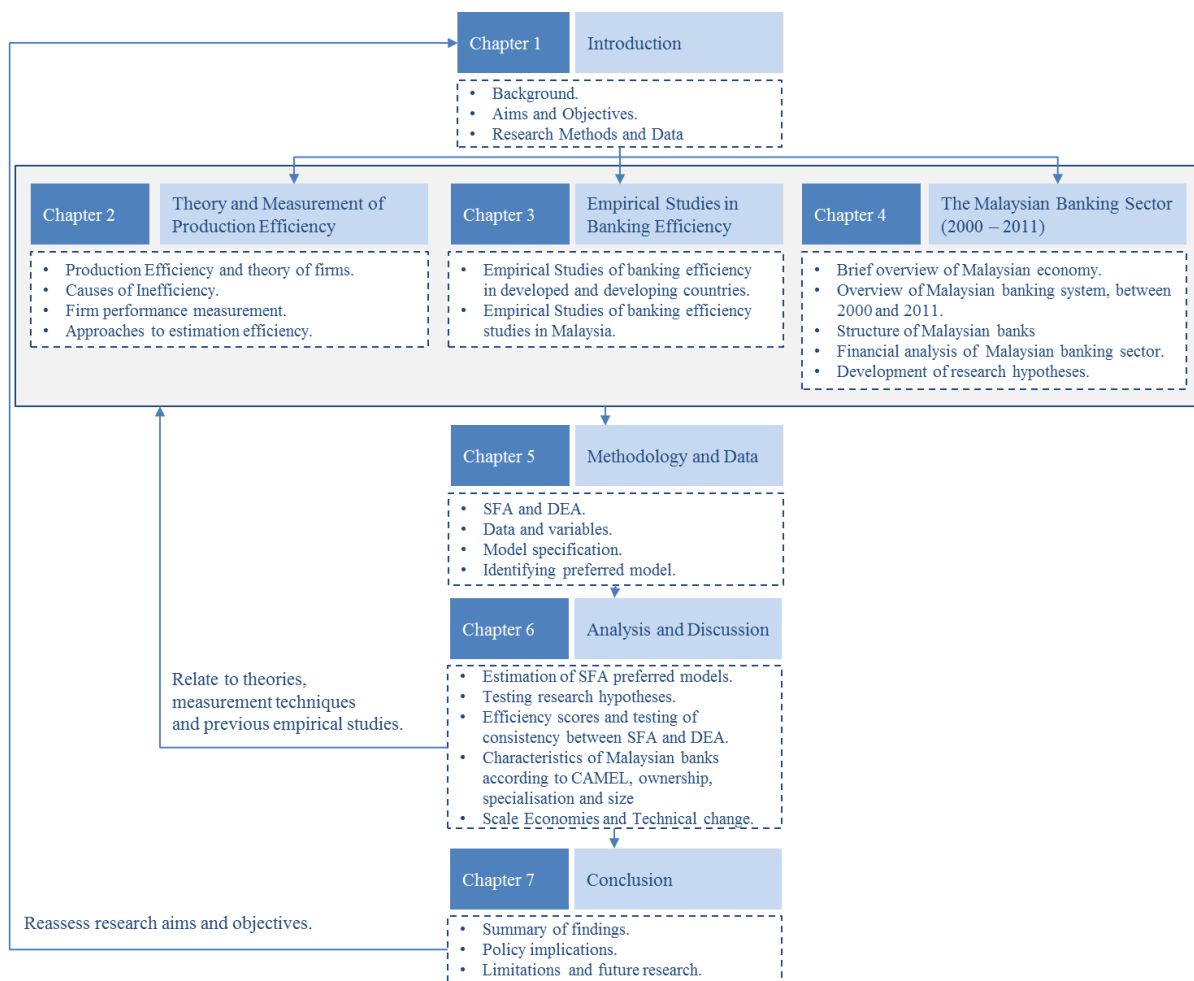
This research employs a confidential dataset of the Malaysian banking industry from 2000 to 2011 from a Malaysian financial organisation. The data comprise of all banks operating in Malaysia and also covers the period of implementation of the FSMP, allowing us to test the impact of liberalisation on the efficiency of Malaysian banks. The sample is an unbalanced panel comprising 32 banks in 2000 followed by 39 banks in 2011 and a total of 354 observations.⁵

1.4 Research Outline

Figure 1.1 presents the structure of the present research and its organisation into seven chapters:

⁵ This includes all domestic conventional, foreign conventional, domestic Islamic and foreign Islamic banks in Malaysia. Investment banks are excluded from this study due to differences in business operations, strategies and target customers.

Figure 1.1 Structure of the Thesis



Chapter 2: Theory and Measurement of Production Efficiency

This chapter introduces the theoretical framework related to production efficiency and discusses two different types of efficiency: absolute and relative. The discussion continues with the concept of production function, which is the main underlying principle for productive efficiency. As the production function derived from the neoclassical theory of the firm, this chapter reviews briefly on traditional neoclassical theory, which views firms as rational and having the objective of maximising profits. It seeks to explain how the market works, but treats the firm as a black box which transforms resources into goods and services. Since neoclassical theory does not take into account the internal structure of firms; managerial and behavioural theories highlight the possibility of internal inefficiency in firms' decision-making processes, and rationalise why firms might not always operate efficiently. Consequently, this chapter introduces the frontier efficiency measurement, using the concept of the optimisation process as its underlying production activity. The main types of

concept in frontier efficiency measurement are reviewed: for example, technical, allocative, cost, profit and scale efficiencies. This chapter also reviews the parametric and nonparametric frontier techniques and functional forms (for parametric models), enabling estimation of a firm's efficiency relative to other firms in the industry. At the end of this chapter, the question of which is the best technique between parametric and nonparametric models are discussed.

Chapter 3: Empirical Studies of Banking Efficiency

This chapter reviews the literature on banking efficiency in developed and developing countries, and Malaysia. It begins by discussing the significance of the banking sector in supporting economic development and growth. The discussion continues by bringing attention to the effects of government intervention on the banking industry through regulation and liberalisation. A well-balanced regulation of the banking system is strived for to stimulate economic growth. Also, discussions on empirical literature of banking efficiency is presented, focusing on different contextual areas such as the effect of deregulation, ownership, market structure, size, and risk appetite on the degree of efficiency of banks.

Chapter 4: The Malaysian Banking Sector, 2000–2011

This chapter provides a brief overview of the Malaysian banking sector and outlines the deregulation measures implemented by the FSMP, which was introduced to strengthen and gradually liberalise the sector. It presents the contextual background required when assessing the empirical results obtained during analysis in later chapters. The effect of financial liberalisation is analysed using the banks' balance sheets and income statements for the years 2000–2011. Financial ratios and competitive indicators (e.g. Herfindahl-Hirschmann index and interest spread) are also presented.

Chapter 5: Methodology and Data

This chapter outlines the methodology employed to measure the cost- and profit-efficiency of the banks. The chapter begins with a discussion of SFA as the main method of study, and followed by DEA, and new DEA (NDEA) (as cross-checking methods). As this study

involves a set of control and environmental variables, the SFA model's specifications – with different combinations of control and environmental variables – are tested to derive the preferred model, prior to measuring the banks' efficiency scores. For the preferred SFA models, three economic efficiencies are estimated to a single data set: cost, standard profit, and alternative profit efficiencies. Following Bauer et al. (1998), this chapter introduces five consistency conditions to assess the robustness of the efficiency estimates obtained from the SFA and DEA models. It also discusses the variables employed in the empirical analysis and brief descriptive statistics of the sample data.

Chapter 6: Empirical Results – Analysis and Discussion

This chapter empirically analyses the cost- and profit-efficiency of the Malaysian banking sector for the years 2000–2011, using the preferred SFA, DEA and NDEA models. Based on the parameter estimation, the coefficients and their significance (i.e. t-ratios) are analysed to test the hypotheses (e.g. ownership, specialisation and size) developed in earlier chapters. One of the main hypotheses is to measure the impact of financial liberalisation via FSMP on the efficiency scores of Malaysian banks and their trends from 2000 to 2011. The trends of efficiency scores may provide indications of the impact (positive or negative) of deregulation measures implemented through the FSMP. This chapter also assesses the consistency of result estimation in the SFA and DEA models. It compares distribution statistics (e.g. mean, standard deviation, skewness and kurtosis), ranking correlations of efficiency scores across different techniques, identification of best-practice and worst-practice banks across different techniques, stability of efficiency scores over time, and ranking correlation between efficiency scores and standard financial ratios. It looks into the diversity of efficiency scores among the different characteristics of the banks, based on certain characteristics (e.g. foreign banks vs. domestic banks, Islamic banks vs. conventional banks, bank size and CAMEL ratings system). This chapter also examines scale economies, scale efficiency and technological change in Malaysian banks.

Chapter 7: Conclusions

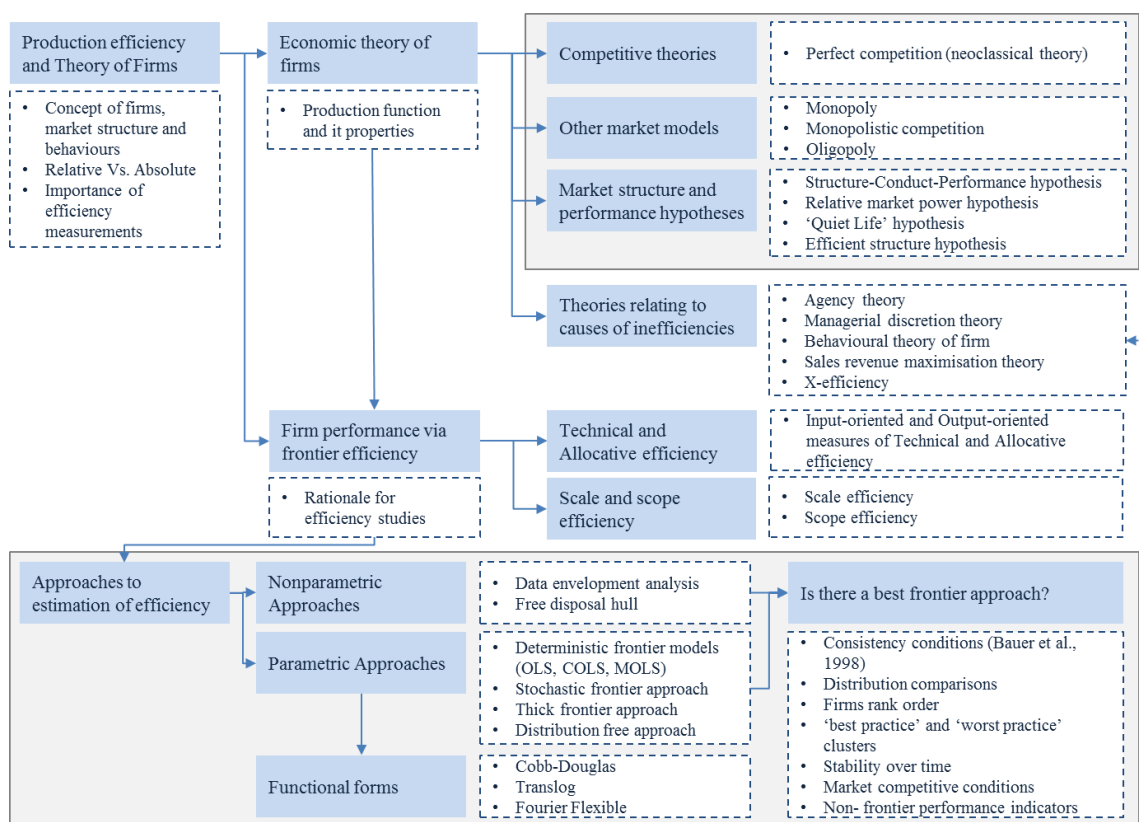
This chapter summarises and concludes the main findings of the study. Policy implications, limitations and future research are also presented.

Chapter 2 Theory and Measurement of Production Efficiency

2.1 Introduction

The objective of this chapter is to provide an introduction to the theoretical framework regarding productive and frontier efficiency. Additionally, the chapter explains various approaches used to estimate the efficiency of the firm. In Section 2.2, the basic description of productive efficiency is explained and the section provides the key concepts for measuring absolute efficiency versus relative efficiency. The section describes neoclassical theory, highlighting the firm as an economic agent with the single aim of maximising profit. Several alternative types of market structures are reviewed to explain why some firms have different objectives and therefore do not exhibit profit maximising behaviour. This section provides alternative theories which question the neoclassical theory and the other market models that focus more on the internal structure and behaviour of those firms which do not maximise profit. Section 2.3 introduces the underlying economic theory of frontier efficiency, concentrating on the estimation of productive activity as an optimisation process. The section also presents the concepts of technical efficiency, allocative efficiency, scale efficiency, and scope efficiency which are used as tools to measure the performance of the firm. In Section 2.4, the estimation methods of frontier efficiency are discussed. This part focuses on two common methods of estimating frontier efficiency; the parametric and the nonparametric approaches, and explains their properties and the differences between them. It presents a discussion of various functional forms used in the parametric approach and the distinctive features of both parametric and nonparametric approaches are examined and critically reviewed. A discussion of key advantages and disadvantages of parametric and nonparametric approaches is presented at the end of the section. Figure 2.1 provides a summary of the content of this chapter.

Figure 2.1 Information Flow of Chapter 2



2.2 Production Efficiency and the Theory of the Firm

Firms need markets to function properly. In his seminal paper entitled 'The Nature of the Firm', Coase (1937) points out that there is no definite theory to determine the market or the boundaries of a firm. The boundaries or markets are described as a place for a variety of exchanges, which brings together the suppliers and buyers of a particular product. Markets are usually suppressed, and the resource allocation of the firm is accomplished through demand and supply. Therefore, market conditions can influence the efficiency of the firm through the use of limited resources to produce maximum output (Goddard et al., 2001). Baumol and Blinder (1988) defined efficiency as the absence of wasted resources. The efficiency of the firm can be described as internal efficiency and as allocative efficiency (Shepherd, 1985). Internal efficiency affects the managers of the firm; it therefore focuses on the effective management of the firm and how the firm transforms resources (inputs) into various outputs. It includes practices such as cost control, and asset and human resource

management. On the other hand, allocative efficiency⁶ refers to a general equilibrium between production and market demand. Allocative efficiency occurs when marginal costs are equal to the price offered by all firms in the market (Goddard et al., 2001). Shepherd (1985) and Goddard et al. (2001) state that an optimum level of both internal efficiency and allocative efficiency should be achieved in order to attain overall efficiency.

The efficiency of a firm can be measured using two different approaches: absolute⁷ and relative efficiency. Koopmans (1951) defines absolute efficiency as a condition of optimum inputs and outputs where technology is unable to increase any output without reducing the inputs. That is, when a firm achieves 100 percent efficiency of input and output, it cannot be further improved without the balance between inputs and outputs being compromised. However, the empirical application of absolute efficiency is not known and has not been implemented. Thus, using the work of Koopmans (1951), Farrell (1957) introduces the concept of relative efficiency, which measures the performance of a firm relative to the best performance firms of a particular market. Relative efficiency (100 percent in comparison to other firms) is attained when other firms cannot further improve their efficiency without compromising their inputs or outputs (Zhu, 2002). It resembles benchmarking procedures where an individual firm's production is compared with the best-performing firm of the industry (Berger and Humphrey, 1997). Relative efficiency is widely discussed in the literature because of its implications for managerial performance and public policy; the results from relative efficiency can provide quantitative evidence, substantiate performance differentials that are predicted qualitatively by economic theories and identify sources of inefficiencies of the firms (Lovell, 1993; Cummins and Weiss, 1998).

As mentioned in the previous paragraph, the estimation of the relative efficiency of a firm has implications for public policy and managerial performance. Kolari and Zardkoohi (1987) and Lovell (1993) state that performance estimation of firms could assist policy makers to formulate appropriate policies for an industry. For instance, firms in an industry can be

6 Goddard et al. (2001) define allocative efficiency as meeting the demands of consumers through the production of goods and services. In the context of the present research, allocative efficiency will observe the firms' approaches using the best combination of input and output; relative to their prices and magnitude, rather than the allocation of resources due to demand from the market.

7 Vilfredo Pareto and Tjalling Koopmans introduced this concept, but no empirical application was implemented until Farrell (1957) utilised it to arrive at relative efficiency estimation.

benchmarked and ranked (from best-performing to worst-performing firms) using the estimation of relative efficiency, where the findings from such a measurement can help policy makers to analyse and identify ways to encourage inefficient firms to reach best practice. By employing the measurement of relative efficiency, policy makers can also explore their hypotheses regarding the sources of inefficiency or productivity differential of firms in an industry (Lovell, 1993). Once the sources of inefficiency of the firms are identified, the policy makers may use these findings to implement appropriate policies to improve the performance of the firms, as well as determining the market structure⁸ that best serves the public (Molyneux et al., 1996; Berger and Humphrey, 1997). Further, the economic performance of a country relies on the performance of micro economic units (e.g. firms), where well-performing firms can further contribute to economic growth (Lovell, 1993).

At the managerial level, the identification of best and worst practice provides necessary information and feedback for managers to improve their management practice and performance. Berger and Humphrey (1997) state that managers can adopt the internal policies, procedures and controls performed by ‘best-practice’ firms and avoid practices used by ‘worst-practice’ firms.⁹ Moreover, better management practice would result in a higher level of production efficiency, implying better chances of survival, improved profitability, and lower prices for consumers (Kolari and Zardkoohi, 1987; Berger et al., 1993).

2.2.1 Economic Theory and Efficiency

The basic concept of productive efficiency can be described through a production function, which is derived from the microeconomic theory of the firm (see Section 2.2.2 Competitive Model and Efficiency). The production function is used to describe marginal products and

⁸ Goddard et al. (2001) argue that the performance of firms can be influenced by the market structure in which they operate. Hence, it is crucial for a policy maker to determine the market structure that suits both the firms and the general public. For example, a highly concentrated market with a number of inefficient firms might indicate to policy makers that they should reduce any barriers to entry. By doing so, the level of market competition is expected to increase.

⁹ That is, management can achieve a higher level of efficiency through enhancing positive determinants and through eliminating negative determinants experienced by both ‘best-practice’ and ‘worst-practice’ firms.

the law of diminishing returns,¹⁰ which can be viewed from the purely technical relationship between inputs and outputs at any period (Koutsoyiannis, 1975). A production function can be defined as the maximum quantity of output produced by given amount of inputs for a given technology (Cyert and George, 1969; McGuigan and Moyer, 1989). In other words, a production function can be seen as a series of processes of transforming inputs (resources) into maximum outputs (goods and services), subject to technological conditions. Thus, an efficient firm should be able to maximise output based on given input and technology. Following Coelli et al. (2005), the production function of a firm can be expressed as follows:

$$y = f(x) \quad (2.1)$$

where y represents output and $x = (x_1, x_2, \dots, x_N)$ is a vector of inputs. For this production function, the decision makers are assumed to have control over the inputs and include them in the general structure of the function $f(\cdot)$. In addition, there are a number of properties related to the function $y = f(x)$ which are (Coelli et al., 2005):

1. Non-negativity: where the value of $f(x)$ is a finite, non-negative, and real number.
2. Weak essentiality: where the production of positive output is impossible without the use of at least one other input. The function should pass through the origin.
3. Non-decreasing in b : which means additional units of an input will not decrease output. If $x^0 \geq x^1$ then $f(x^0) \geq f(x^1)$. If the production function is continuously differentiable, the monotonicity property supports non-negative marginal products.
4. Concave in x : this implies that any linear combination of the vector x^0 and x^1 will produce an output not less than the same linear combination of $f(x^0)$ and $f(x^1)$.

¹⁰ The marginal product can be defined as the extra amount of output that can be produced when the firm uses one additional unit an input. It can be expressed as $MP_b = \Delta a / \Delta b$, where MP_b is the marginal product measured in terms of actual units of output per unit of input; Δa is additional changes in input; and Δb is the changes in output (McGuigan and Moyer, 1989). Additionally, the law of diminishing returns states that if an input in the production of a product is increased, while all other inputs are held constant, a point will eventually be reached at which additions of input yield progressively smaller, or diminishing unit of output (McGuigan and Moyer, 1989). The marginal products and the law of diminishing returns are reflected in one of the properties of a production function.

Concavity implies that all marginal products are non-increasing when the production function can be differentiated.

Production efficiency is strongly associated with cost and profit elements. The relationships among the cost, profit, and production functions can be explained by duality theory.¹¹ The duality theory is described as the possibility of making appropriate transformations of the production function into maximising behaviour (i.e. the profit function) and minimising behaviour (i.e. the cost function) of the firm by differentiating a function which characterises an optimal solution for production (Weitzman, 1973; Diewert, 1974; Lopez, 1982; Lusk et al., 2002). Therefore, a firm can create the optimal input and output mix based on the objective of the firm either to minimise cost or to maximise profit. For the cost function, the key assumption is to minimise cost in a perfectly competitive market where the firm has no influence on prices (Coelli et al., 2005). Based on Coelli et al., (2005); the cost function can be expressed as:

$$C = f(w, q) \tag{2.2}$$

where C is cost, $w = (w_1, w_2, \dots, w_N)$ is a vector of input prices, and q is the output vector. Like the production function, the cost function has its own set of properties that must be satisfied (Coelli et al. 2005):

1. Non-negativity: cost is always non-negative
2. Non-decreasing in w : cost will not be decreased when input prices increase.
3. Non-decreasing in q : it costs more to produce additional output.
4. Homogeneity: multiplying all input prices proportionately will cause a proportionate increase in costs.
5. Concave in w : this may not be intuitive, but it implies that input demand functions do not slope upwards.

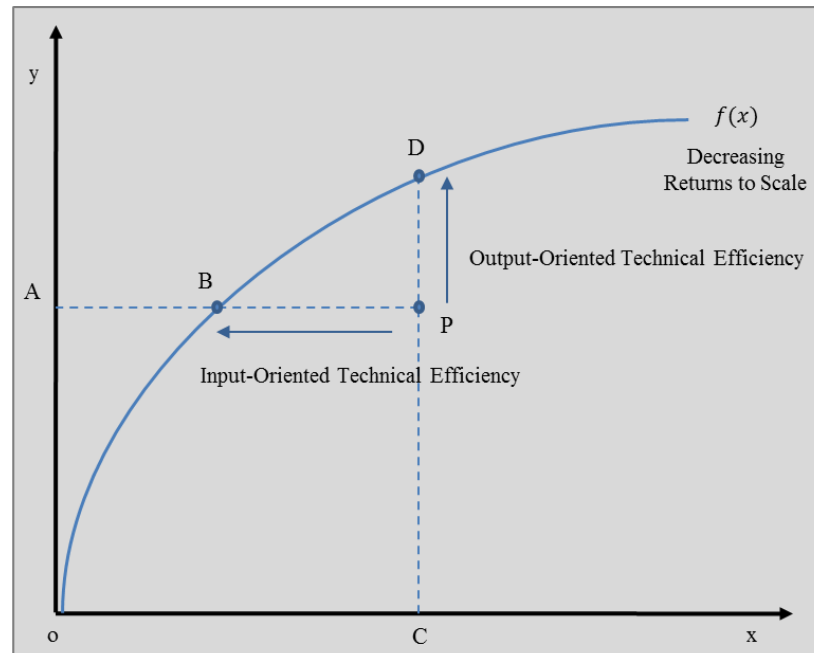
¹¹ Shephard (1953) introduced the theoretical and mathematical foundation of duality. Duality theory can be broadly explained as a mathematical technique of translating concepts and theorems into other concepts or theorems in a one-to-one manner.

Comparably, for the profit function, the firm is assumed to be making a decision to maximise profits for a given production technology. The profit function can be expressed as follows (Coelli et al., 2005):

$$P = f(p, w) \tag{2.3}$$

where P is profit, p is the output vector and $w = (w_1, w_2, \dots, w_N)$ is a vector of input prices. Like the cost function, the profit function should satisfy a number of properties. The properties are similar to those of the cost function. These are, non-negativity, non-decreasing in p , non-decreasing in w , and homogeneity; the only exception is that the profit function is convex in w rather than concave, as was demonstrated above for the cost function (Coelli et al., 2005).

In production efficiency, productivity is commonly defined as the relationship between inputs and outputs, and can be regarded as a natural performance measure (Coelli et al., 2005). A productivity increase can be accomplished by a firm using either the minimum input to produce a given quantity of output, or by producing greater output from a given quantity of the input. Generally, from the production function $f(x)$, the productivity of a firm can be measured using the ratio of outputs to inputs; and efficiency, on the other hand, can be measured from the distance between the observed value (or the actual value) and the potential value; that is, the maximum value of the production function (a detailed concept of technical efficiency is discussed in Section 2.3.1). A firm is considered as technically efficient when the output is at its maximum, based on the given inputs and technology. According to Sherman and Zhu (2006), the terms ‘productivity’ and ‘efficiency’ have been used interchangeably in various contexts. Coelli et al. (2005) demonstrated that these terms are not precisely the same; where productivity and technical efficiency can be described from the example shown in Figure 2.2 using the single input (x) and a single output of production (y) (Kumbhakar and Lovell, 2000; Coelli et al., 2005).

Figure 2.2 Productivity and Technical Efficiency: A Single Input and a Single Output

Source: Coelli et al. (2005, p. 55)

where $f(x)$ is the production function, P is the observed value of a firm represented by the observed input C and observed output A. D represents the potential value on the production function which can be produced at C. B is the potential value on the production function, based on the minimum input that can be used to produce output A. Thus, productivity of the firm at point P is defined as the ratio of output to input, CP/AP , and technical efficiency can be defined as the ratio of the potential input to the actual input (an input-orientated measure), or actual output to potential output (an output-orientated measure). The input-orientated measure of technical efficiency is equal to AB/AP , while the output-orientated measure of technical efficiency is represented by CP/CD , which is always smaller than or equal to 1.

2.2.2 Competitive Model and Efficiency

As discussed already, the relationships between inputs and outputs can be demonstrated using the production, cost, and profit functions. Based on these relationships, the firm will make decisions to produce the maximum output based on the allocated input (or, in the case of the cost function, the optimisation behaviour of the firm will be based on minimising cost

to produce a given output). This system derives from neoclassical theory and stems from the static equilibrium concept. Neoclassical theory assumes a market with a large number of firms whose single objective is to maximise profit (Cyert and Hedrick, 1972). Neoclassical theory treats firms as a black-box.¹² It focuses on the input and output of the firm rather than on how the transformation is done.¹³ The theory also describes firms as contenders in a perfectly competitive market; they are price takers and rely on freely available information to maximise profit (Demsetz, 1997). The main characteristics of perfect competition are as follows (Mc Guigan and Moyers, 1989):-

1. There are a very large number of buyers and sellers, each of whom buys or sells such a small proportion of the total industry output that a single buyer or seller's actions cannot have a noticeable impact on the market price.
2. A homogeneous product is produced by each firm and there is no product differentiation.
3. There is free entry and exit from the market with minimal barriers to entry and exit.
4. Buyers and firms have perfect knowledge of prices and information is available freely.

Should all of the above conditions be satisfied, a market can be in competitive equilibrium in which all firms can earn a normal profit and remain in business (McGuigan and Moyer, 1989). In a perfectly competitive market, market demand and market supply determine the equilibrium price and quantity traded. No single buyer or seller can make a significant impact on market price, as all economic agents in a competitive market are price takers (i.e. each takes market prices as given and beyond their control) (McGuigan and Moyer, 1989). The firm has no control over prices and can sell as much as it wants at the going market price and it faces a horizontal demand curve which reflects the perfect elasticity of demand. As mentioned earlier, the main objective of the firm is to maximise profit. Profit maximisation is achieved at the point where marginal revenue (MR) equals marginal cost

¹² 'Black box' can be described as a process where inputs and outputs (and the relationship between them) are known, but the internal structure or working is not fully understood or considered (Berger and Mester, 1997; Castelli et al., 2010).

¹³ Demsetz (1997) asserts that the objective of neoclassical theory is to understand price-guided resource allocation rather than management-guided resource allocation. This is consistent with Campus (1987), who defines neoclassical economics as an approach that pays attention to prices, output, and income distribution based on market demand and supply.

(MC) (i.e. where an increase in revenue from selling an extra unit is equal to the cost of producing that extra unit). Hence, at the point where $MR=MC$, the firm is considered to achieve economic efficiency (Griffith and Wall, 2004). Economic efficiency can be broken down into two parts, the productive efficiency and allocative efficiency. A firm is described as achieving productive efficiency when it produces a given level of output with the lowest cost of production. Additionally, allocative efficiency is realised when price corresponds to or equals marginal cost. A deviation of market price from marginal cost can result in allocative inefficiency (Griffiths and Wall, 2004).

In the long run scenario, firms are unable to make abnormal profits. However, in the short run it is possible for some firms to earn abnormal profit. The existence of abnormal profit will attract new competitors and drive down the market price until all firms earn normal profit again in the long run. This shows that the demand curve of a firm is perfectly elastic; any movement of price below or above the market price could result in firms going into deficits in the long run.¹⁴ If firms in a perfectly competitive market are inefficient and unable to earn normal profits, they are at risk of either being acquired or driven out of the market (McGuigan and Moyer, 1989). In other words, firms that are unable to allocate resources efficiently to produce the maximum level of output might fail to reach the efficient frontier, and could be forced out of the market (Machlup, 1967).

However, past literature suggests that inefficient firms that are not on the estimated efficiency frontier can continue to survive in the market due to many factors (this is discussed later in the chapter). As mentioned before, neoclassical theory is not designed to consider the managerial behaviour of the firm (Machlup, 1967). Its sole purpose is to explain the function of the market (Cyert and George, 1969; Demsetz, 1997). For instance, Williamson (1981) argues that the optimality concept in neoclassical theory is arbitrary because it does not recognise transaction costs¹⁵ and the rationality of producers. In reality

¹⁴ For example, if a single firm tries to sell its product above the market price, rational buyers will move to other sellers. On the other hand, if the firm sells below the market price, the quantity demanded will approach infinity and the firm will encounter financial loss.

¹⁵ Robins (1987) describes transaction cost as the cost associated with an economic exchange which is independent of the competitive market price of goods produced. Examples of transaction cost include search and information costs, monitoring costs and enforcing contractual performance. Despite being independent of the competitive market price, the value of transaction cost is determined by the types of economic exchanges conducted.

however, transaction costs (e.g. the cost of information) can influence the economies of the production process, and decision making by the producers can sometimes be irrational and imperfect. This could also indicate that market competition is imperfect.

2.2.3 Market Imperfections

In the previous discussion, the perfect competition model described a large number of firms, where no single firm has market power over prices. However, in certain industries, a few large firms dominate the market, have control over prices, and earn abnormal profits. Thus, several different market structures arise from violations of the perfect competition model's assumptions, and might require an industry with a small number of firms to operate. Such markets are imperfectly competitive and the level of imperfect competition relies on several conditions: the number and size of the firms, product homogeneity, and the degree of independent decision making by individual firms.

The first market structure to be discussed which is in direct contrast to the perfect competition model is monopoly. Monopoly can be characterised as (1) a single firm producing a specific commodity, (2) no close substitute products exist, (3) substantial barriers to entry exist (through cost advantages, product differentiation, and scale economies) and, (4) there are large capital requirements (McGuigan and Moyer, 1989). Sraffa (1926) accepts the above description of monopoly and argues that competition is being eliminated from the market for two basic reasons: barriers to entry and cost advantages from economies of scale. Because of the existence of entry barriers, a monopolist has control over price and can set price anywhere along its downward sloping demand curve.¹⁶ As a result, the marginal revenue of a monopolist is different from the marginal revenue of the firm in the perfectly competitive market (McGuigan and Moyer, 1989). In comparison to a perfectly competitive market (where firms face a horizontal demand curve and the marginal revenue curve is identical to its demand curve (price=MR)), the monopolist faces a downward sloping demand curve, which results in a different relationship between marginal revenue

¹⁶ The downward sloping demand curve of the monopolist market is in accordance with the 'law of demand', in which there is an inverse relationship between price and quantity demanded (McGuigan and Moyer, 1989). Moreover, the demand curve of the firm is also identical to the industry's demand curve, because the firm is the industry (Koutsoyiannis, 1975).

(MR) and price, where MR is less than price ($MR < \text{price}$). The monopolist maximises its profits when the two following conditions are met: (1) marginal cost (MC) is equal to MR; and (2) the slope of MC is greater than the slope of MR at their point of intersection (Koutsoyiannis, 1975). Once these conditions are fulfilled, the monopolist continues to maximise profits when the price is higher than its marginal costs. Unlike firms in the perfectly competitive market (which is a price taker and only requires a decision on output), the monopolist faces decisions regarding both price and output. The monopolist can either produce the output at $MR=MC$ and sell at the corresponding price, or it may set its price and sell the corresponding output that the market is willing to take (Goddard et al., 2001). Nevertheless, the decision-making process (e.g. setting higher prices) for the monopolist is simpler than for firms in the perfect competition model. This is due to the existence of entry barriers (e.g. a sole producer of products with no close substitutes). In the absence of competition, the monopoly could still receive abnormal profits despite being inefficient in its resource allocation (Demsetz, 1982, Baumol and Blinder, 1988; Berger and Hannan, 1997).

Some market structures are neither perfect nor monopolist, but somewhere in between. Chamberlin (1933) introduced two additional market structures monopolistic competition and oligopoly. The conditions of monopolistic competition are similar to those of perfect competition model. Monopolistic competition can be described as a market with a large number of firms, where each one sells a differentiated product, makes their own decisions, and enjoys easy entry and exit from the market (McGuigan and Moyer, 1989). The most distinguishing characteristic for monopolistic competition is the differentiation of the output (heterogeneous products) of each firm. Using Sraffa's (1926) downward sloping demand curve, Chamberlin (1933) explains that the demand curve is influenced by product differentiation for the pricing policy of the firm, style and services associated with the product as well as the firm's selling strategy. Furthermore, product differentiation allows firms to determine the price of the product. For instance, firms in monopolistic competition are not price takers (i.e. facing a downward sloping demand curve) and have some degree of market power; but with competition of close substitutes offered by other firms, their discretion over price is limited (McGuigan and Moyer, 1989). Additionally, the demand curve is influenced by (1) the pricing and selling strategy of the firm, (2) the pricing, output

and selling policies of competitors, and (3) the income, pricing and selling policies of product from other industries. Thus, any changes in these things will shift the demand curve in monopolistic competition, which in turn could affect the profit maximisation of the firms in the industry (Demsetz, 1959).

The oligopoly model, on the other hand, describes a market with a few large firms which are interdependent in price, output, product quality, and terms of sale (McGuigan and Moyer, 1989). In this market structure, any actions by an individual firm in terms of pricing would have a noticeable impact on the sales of other firms in the market. The number of firms in the market should be sufficiently small for there to be conscious interdependence. Each firm is setting their future according not only to their firm's policies but to those of their competitors as well (Griffiths and Wall, 2004). Mazzeo (2002) posits that due to the small number of competing firms, the market structure affects the degree of price competition where earnings are based on production choices made by the firms. In other words, the price and output of a firm in an oligopolistic market can be determined by consumer preferences, product substitutability, and the responses on prices by competitors. Thus, this non-collusive oligopoly model is different from other market structures (e.g. perfect competition and monopoly) due to the characteristic of assuming a certain degree of interdependence among firms in the industry. However, the uncertainty arising from oligopolistic interdependence can be avoided by the firms in this market structure through collusive agreements, such as cartels (e.g. price fixing) and price leadership (e.g. leading the industry by setting prices). A collusive agreement is generally in the form of implicit agreements; it could commonly be illegal in most countries if it were announced in the open market (Koutsoyiannis, 1975). An outcome of the collusive oligopoly model can be that firms behave like a monopolist, which restricts the supply of the industry output at higher prices to enable them to earn abnormal profits. That is, collusive firms in oligopoly can exercise market power by fixing their price above marginal cost and produce a smaller quantity of output; hence, it is an inefficient market structure in terms of resource allocation (McGuigan and Moyer, 1989).

The preceding discussion briefly describes the four basic structures found in neoclassical economic theory: perfect competition, monopoly, monopolistic competition and oligopoly. These market structures can be defined based on several conditions: the number and size of

the firms, types of products, degree of control over prices, and the extent of entry and exit barriers (McGuigan and Moyer, 1989). This indicates that the structure of the market can influence the performance of firms. The relationship between market structure and firms' performance is normally examined using several hypotheses, such as the structure-conduct-performance (SCP) hypothesis, relative-market-power (RMP) hypothesis, the quiet life hypothesis, and the efficient-structure hypothesis.

2.2.4 Market Structure and Performance of the firm hypotheses

The market-power-related hypothesis can be explained by the structure-conduct-performance hypothesis, the relative-market-power hypothesis and the quiet life hypothesis. These hypotheses describe the impact of market structure on the behaviour and performance of the firm. Conversely, the efficient structure hypothesis tends to explain the efficiency of the firm (e.g. superior management and advanced technology) which, in turn, influences the performance of the firm. Thus, the two different approaches enable views from different angles on firms' performances and the market structure (Berger, 1995).

The efficient-structure (ES) hypothesis¹⁷ states that the performance of the firm in a given market structure is influenced by the efficiency of the managers, rather than by collusive behaviour in a more concentrated market. The ES hypothesis suggests that the superior efficiency of particular firms can result in greater market concentration, enabling them to earn abnormal profits. Thus, the performance of the firm is strongly influenced by their specific efficiency (i.e. good management and technology) instead of market concentration conditions. The ES hypothesis differs from the market-power-related hypotheses (i.e. the SCP hypothesis, the RMP hypothesis and the quiet life hypothesis), in which it views efficiency as a factor for a firm to gain market share at the expense of less efficient firms, which consequently increases market concentration. Market-power-hypotheses, on the other hand, describe the market structure (i.e. market power) as the key factor that influences the performance of the firm (Catena, 2000).

¹⁷ Demsetz (1973) introduces the efficient-structure hypothesis: he asserts that market concentration is indicated by the differential efficiency of some superior firms instead of the collusive behaviour of a concentrated market. These superior firms are assumed to have a large market share that may result in a greater market concentration and the earning of supernormal profit (Berger, 1995).

Berger (1995) stated that there is a positive statistical relationship between measures of market structure (concentration and market share) and profitability. In this regard, the positive relationship between profitability and market structure elements can be explained by two market power hypotheses, namely the traditional SCP hypothesis and the relative-market power (RMP) hypothesis. In the structure-conduct-performance¹⁸ hypothesis, the performance of the firm is assessed with respect to the interaction between market structure and market conduct. Figure 2.3 illustrates the SCP paradigm, which explains the relationship between market structure, market conduct, and market performance (Mason, 1939; Bain, 1951). This hypothesis implies that firms are allowed to set prices if the market is concentrated as result of imperfect competition (Berger, 1995). For instance (in the context of banking), banks can set prices (e.g. lowering deposit rates and increasing their loan rates) that are less favourable to consumers in a concentrated market due to imperfect competition within the market which, in turn, could earn greater profits. Another theory to describe market power is the relative-market-power (RMP) hypothesis.¹⁹ This hypothesis explains that only firms with a large market share and differentiated products are able to set prices and earn abnormal profits (Shepherd, 1982). The effect of a concentrated market on a firm's performance can be further presented by the quiet life hypothesis.²⁰ Berger and Hannan (1998) argue that managers of the firm may exercise the quiet life in a more concentrated market and thus benefit from higher prices and pay less attention to keeping costs under control. Berger and Hannan (1998, p. 455) stated:

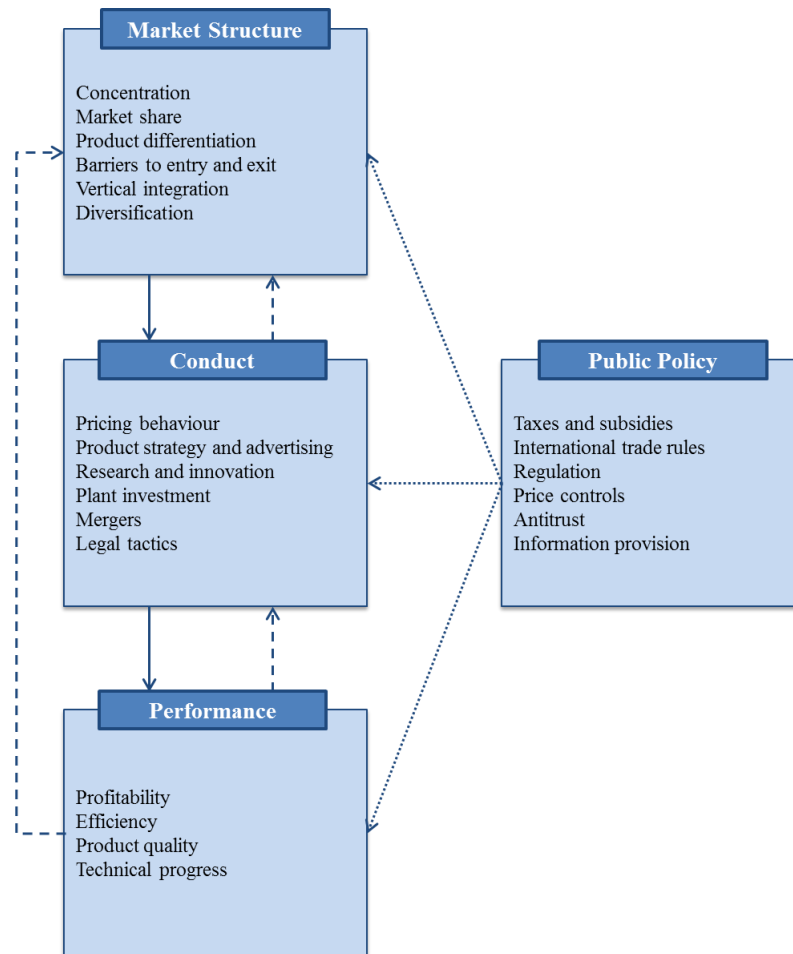
“Market structure may influence for one of the several related reasons. First, if high levels of market concentration allow firms to charge prices in excess of competitive levels, then managers may take part of the benefits of the higher prices not as higher

¹⁸ The structure-conduct-performance paradigm was introduced by Mason (1939) and Bain (1951), who describe the relationship between market structure and market performance. They suggest that market structure can influence the behaviour and performance of the market. For SCP acronyms, S represents market structure, which describes the degree of concentration, market share, market entrance barrier, and market integration. Subsequently, C is market conduct and can be described by the behaviour of buyers and sellers, pricing policies, investment in research, and advertising strategies. Finally, P represents market performance. This can be measured by comparing the profitability and efficiency of firms, and industry's product quality, and technical progress (Waldman and Jensen, 2001)

¹⁹ The market power hypothesis was developed by Ravenscraft (1980) and Mueller (1983). According to the RMP hypothesis, market power (due to product differentiation and barriers to entry) allows firms to capture market share, to set higher prices, and earn abnormal profit (Catena, 2000).

²⁰ The quiet life hypothesis was developed by Hicks (1953). He asserted that firms with greater market power are more risk averse. He further added that firms with greater market power are capable of gaining a combination of higher profit with lower risk compared to firms with less market power.

profits, but in the form of a “quiet life”, in which they don’t work as hard to keep costs under control. The difference between the actual price and the competitive price may provide a “cushion” or comfort zone. In the absence of other disciplining mechanism, managers may allow unit cost to rise to consume part of this cushion and still allow owners to earn economic rents without the full effort of cost minimisation. Second, market power may allow managers to pursue objectives other than firm profits or managerial leisure. Third, the price cushion provided by market power may simply allow inefficient managers or practices to persist without any intention to pursue goals other than maximising firm value. The lack of market discipline in concentrated markets may simply blunt the economic signals that would normally force changes in management to keep costs low, leaving managers in a position for which they do not have comparative advantages. Thus, market power may allow managerial incompetence to persist without any wilful shirking of work effort, pursuit of other goals, or efforts to defend or obtain market power.”

Figure 2.3 Structure-Conduct-Performance Paradigm

Source: Waldman and Jensen (2001, p. 7). The solid arrows show the primary relationships of this paradigm, in which market structure influences market conduct and consequently determines the market performance. The dashed arrows indicate the feedback effects of market performance on market conduct and market structure. Government policies have direct influence on market structure, market conduct, and market performance, as shown by the dotted arrows.

The above discussion emphasises the impact of the market structure on firm performance using empirical hypotheses, namely the efficient structure hypothesis, structure-conduct-performance hypothesis, the relative-market power hypothesis, and the quiet life hypothesis. The importance of this relationship between market structure and firm performance was explained in Berger and Hannan (1998), who postulate that the market structure can influence the performance of the firm and the behaviour of the managers. They further argue that a concentrated market structure could result in a shift of managers' behaviour to a quiet life of pursuing personal objectives; this allows inefficient managers to persist within the

firm without any intention to pursue goals other than profit maximisation. Additionally, this implies that the managers are unable to act with global rationality (i.e. unlimited information, unlimited time at their disposal and an unlimited ability to compare all possible alternative actions and choose the one that maximises profit), as postulated by neoclassical theory. In reality however, the managers face bounded rationality with limited and(or) distorted information, limited time and limited abilities to assess all alternatives to select the best approach to maximise firm value (discussed in the next section) (Simon, 1955). Considering the above, several managerial and behavioural theories surfaced to challenge the neoclassical theory and the market-structure hypotheses; these include agency theory, managerial discretion theory, sales revenue maximisation theory, growth maximisation theory, the behavioural theory of the firm and X-efficiency; that arises from the complexity in managing modern enterprises. Accordingly, these managerial and behavioural theories provide the necessary clues as to why some firms survived while not being efficient. These theories seek to explain the causes of the inefficiency of firms.

2.2.5 Causes of Inefficiency of the Firm

Neoclassical theory assumes that firms are making rational decisions with their single objective being to maximise profit. Moreover, the theory suggests that inefficient firms will be forced out of the market. However, as mentioned earlier, this has not always been the case. Some firms continue to survive despite being inefficient. Neoclassical theory has earned criticism for being too unrealistic and narrow for the current economic environment (i.e. global rationality and a single objective to maximise profit). For example, neoclassical theory assumes that a firm is a single ownership entity (no separation between ownership and management). The owner is responsible for making decisions and is assumed to have unlimited information, time and ability in making decisions to maximise profit. However, past literature shows that firms are unable to achieve their goal in maximising profit due to limited knowledge, limited and(or) distorted information and limited ability in decision making. There are arguments that firms may not want to pursue a single goal to maximise profit, due to having a multitude of goals, of which profit is just one of them (Koutsoyiannis, 1975). In the modern world, however, the firm is commonly organised by the separation of owners and managers (i.e. managers, who have discretion to attain goals other than profit

maximisation). Therefore, neoclassical theory is challenged by two different theoretical views: institutional economics (i.e. transaction costs, taxes, computational limitations and other friction) and behavioural economics (i.e. non-rational and systematically uniformed behaviour of the controlling group or agent) (Merton and Bodie, 2004). The introduction of managerial theory (i.e. maximisation of the managerial utility) and behavioural theory (i.e. satisficing behaviour and bounded rationality),²¹ such as principal-agency theory (Jensen and Meckling, 1976), managerial discretion theory (Williamson, 1964), sales revenue maximisation theory (Baumol, 1959), growth maximisation theory (Marris, 1964), the behavioural theory of the firm (Simon, 1955; Cyert and March, 1964) and X-efficiency (Liebenstein, 1966) seeks to explain the behaviour of the controlling group in pursuing objectives other than maximising the firm's profit.

The first managerial theory discussed is agency theory, which attempts to solve the principal-agent conflicting goals and the difficulties faced by the principal in monitoring and verifying the actual tasks performed by the agent (Eisenhardt, 1989). Jensen and Meckling (1976) define the agency relationship as a contract between principal (owner) and agent (manager). In the contract, the owner hires the manager to perform the operational tasks on behalf of the principal by delegating some authority to the manager. To some extent because the ownership is divorced from the day-to-day operations of the business there is good reason to assume that the manager will not act in the best interest of the owner if the manager is a utility maximiser (Fama and Jensen, 1983). This independence of action might be due to their superior knowledge of the firm as well as to their ability to disguise their actions from the principals; where the managers or agents may not always act towards achieving the objectives of the principals (Griffiths and Wall, 2004). However, the managers' discretion in setting the goals is not unlimited. The managers have to ensure that a minimum level of profit is reached for a dividend policy that is acceptable to the members of the firm (i.e. shareholders), avoid a relative fall in firm value (i.e. reduced share prices in

²¹ For example, the basic feature of managerial theory is that the managers will maximise their own utility, subject to a minimum profit constraint (which is set at a level adequate to satisfy the owners of the firm) necessary for the job security of the high-level managers. On the other hand, the behavioural theory of the firm describes the complex nature of the firm, which includes in the theory of the firm the factor of realism. The firm is seen to have multiple goals, with different groups connected with its activity in various ways (e.g. managers, workers, customers, suppliers, bankers and so on); each has its own goals or demand (Koutsyiannis, 1975; McGuigan and Moyer, 1989; Griffiths and Wall, 2004).

the stock exchange and risk of a take-over) and safeguard investments for the satisfactory operations of the firm. If these conditions are not satisfied, the managers may be made redundant by the owners (Koutsoyiannis, 1975). Thus, this conflict between principal and agent could in return reduce the profitability and efficiency of the firm. To overcome this principal-agent problem, a firm has to incur an agency cost,²² which is established from monitoring mechanisms and which sets appropriate incentives to control any deviant activities of the agent. Another approach to minimise conflicting objectives between principal and agent is through corporate governance solutions, for the managers' self-discipline, such as introducing a board of directors (Fama and Jensen, 1983) and capital and labour market disclosure (Fama, 1980).

As a result of principal-agent segregation, managers may seek to maximise their own utility over the firm's productivity (Williamson, 1963). Managerial-discretion theory deals with the degree of freedom held by the managers in utilising a certain portion of the firm's resources (Mique and Belanger, 1974). Williamson (1963) describes managerial-discretion theory as any acts by which managers allocate excessive resources of the firm to maximise their own utility instead of maximising the profit of the firm. Nevertheless, a manager's discretion is constrained by the firm's profit that is expected by the market and shareholders (in the form of a paid dividend and the firm's share price); and if these expectations are not fulfilled, the job security of the manager can be at risk (Yarrow, 1976). As mentioned above, the managers have discretion in setting policies that maximises their own utility instead of the firm's profit. The manager's utility can be in the form of salary, power, status, professional excellence, and prestige, in which these can be achieved at their discretion. In this case, Williamson (1963) iterates that managers do not display neutral behaviour towards cost and they have preferences in the expenses of the firm such as staff expenditure,

²² Jensen and Meckling (1976) state that the agency costs can be separated into monitoring cost, bonding cost, and residual loss. The monitoring cost is expenditure incurred from checking the activities of the managers. Monitoring cost ranges from the expenses related to Directors' remuneration, internal and external auditing, budget restrictions, compensation policies, and operational rules. On the other hand, bonding cost refers to the costs incurred by providing incentives to ensure that the agent will act in the interest of the principal. Those incentives come with sacrifices for the manager, who in turn provides a guarantee not to deviate from the owners' objectives. Finally, residual loss is the loss sustained after monitoring and bonding cost. This can be described as the losses suffered by the principal as a result of decision making by the agent which decreases the principal's profit (Jensen and Meckling, 1976; Fama and Jensen, 1983; Berger and di Patti, 2006).

managerial emoluments (i.e. slack payments²³) and discretionary investments (using the firm's fund) to satisfy their utility (i.e. $U = f(S, M, I_D)$, where U is a manager's utility, S is staff expenditure, including managerial salaries (administrative and selling expenditure), M is managerial emoluments, and I_D is discretionary investment). Moreover, managers may exercise their discretion in using funds (to which they have access) to satisfy their own needs (e.g. office renovation, private jet hire, or staying in a lavish hotel). But in order for managers to continue to enjoy this discretion, they have to ensure that the owners' needs are satisfied by returning acceptable profits (Yarrow, 1976). If unproductive expenses and managerial slack²⁴ continue to exist, firms will incur unnecessary expenses which can result in higher overall inefficiency of the firm (Williamson, 1963).

Third, instead of profit maximising being the primary goal of the firm (as outlined in the neoclassical theory), Baumol (1958) suggested an alternative objective of the firm which is to maximise sales revenue. The sales-revenue maximisation theory highlights the objective of some firms to maximise sales revenues, subject to the constraint of returning a satisfactory level of profit to shareholders as well as retaining profit for expansion or future growth (Baumol, 1958; Hawkins, 1970). The sales maximisation goal is reflected in the Williamson (1963) managerial-discretion theory, but in a much broader sense; managers seek to increase their utility through greater expenses on staffing levels and emoluments, as well as investing in projects, all made possible from increased sales revenue. Moreover, funds for greater expenditure are easier to accumulate from sales revenue than from profits or new external finance (since profits have to be distributed to shareholders and new external finance would require greater accountability) (Griffiths and Wall, 2004). In addition, Baumol (1958) states that the salaries and other perks of managers are more closely related to sales revenue than to profit. For example, the managers' salaries and incentives are simpler to measure through sales revenue. The positive growth in sales revenue will strongly reflect the reputation of managers. The managers are more interested in measuring sales than profits and are willing to forgo additional profit when the minimum

²³ Slack payments are payments to the factors of production over and above the price required to keep them in their present employment (Carter, 1971; Nohria and Gulati, 1996).

²⁴ Agency theorists often suggest that managerial slack is a source of an agency problem which results in the firm becoming inefficient, risk averse, and lacking in performance (Jensen and Meckling, 1976; Fama, 1980). For example, managerial slack such as redundant employees, unused capacity, and unneeded capital expenditure can affect the input and output of the firm (Nohria and Gulati, 1996).

target level is achieved; they reject the opportunity to increase the profit of the firm at the expense of sales (Baumol, 1958). Managers prefer a steady profit performance rather than impressive profit over a period of time. Thus, to a certain extent, managers are seen as risk avoiders who intend to maintain stable performance and are reluctant to pursue promising projects that are risky; and such attitudes may prevent a firm's economic growth (Koutsoyiannis, 1975).

So far, the discussions assumed that the owner's goal is in conflict with the goal of the manager. Marris (1964) however, suggested that both owners and managers can achieve a common goal (i.e. via the firm's growth rate rather than its sales revenue or profit) which is reflected in growth-maximisation theory. Growth maximisation theory derives from a model of the managerial enterprise, which explains that the maximisation of a firm's growth rate can be achieved through the maximisation of the growth rate of demand for the firm's products and the growth of its supply of capital (i.e. $g = g_D = g_C$, where g is balanced growth of the firm, g_D is the growth in demand for the products of the firm, and g_C is the growth of the supply of capital) (Marris, 1964). The objective of growth maximisation is to maximise jointly the growth rates of demand and capital, which in turn, this maximises managers' utility as well as the owners' utility.²⁵ For instance, managers strive for growth in the demand for the products and services of the firm to raise their power and status; on the other hand, owners seek growth in the capital value of the firm to increase their personal wealth (Griffiths and Wall, 2004). Hence, a balanced rate of growth of the firm is required to satisfy both parties, subject to the managerial and job-security constraints faced by the managers. Marris (1964) incorporate financial policies such as a financial security constraint (as part of the job-security constraint factor which can be represented by the leverage, liquidity and retention ratios) in the decision making process of the firm. For example, the retention ratio policy (i.e. retained profits over total profits) can influence the utility of managers and owners. If the profits that are distributed to shareholders are low (a high retention ratio), the retained profits can be used for reinvestment and to stimulate the growth of the firm; but this will jeopardise managers' job security where the shareholders may

²⁵ The manager's utility can be described as $U_M = f(g_D, s)$, where U_M is the utility of manager, g_D is rate of growth of demand for the products of the firm, and s is a measure of job security, which can be decomposed into a managerial constraint and a job security constraint. The owner's utility can be defined as $U_O = f(g_C)$, where U_O is the utility of the owner, and g_C is the growth rate of capital.

decide to terminate their position due to the fall in share prices and increased potential take-overs. On the other hand, if the distributed profit is high (a low retention ratio), the firm will have less funds for economic growth; and shareholders might be more content and the price of the shares will be sufficiently high to deter take-overs. This example implies that despite having a solution which maximises the utility of both managers and owners, the conflicts of achieving balanced utility between managers (who seek the firm's growth) and owners (i.e. shareholders, who seek higher dividends) can still occur, and can result in greater inefficiency in the firm (Griffiths and Wall, 2004).

Unlike the managerial theories, the behaviour theory of the firm describes the firm as an organisation with various groups such as owners, managers, suppliers, customers, workers and etc., each of which has their own goal, or set of goals. Hence, the firm is not seen as a single-decision unit, as outlined in neoclassical theory but as a multi-goal organisational coalition (Cyert and March, 1963). The single goal of the firm is seen as being too remote from the complexity of modern firms (particularly large firms), where these firms are established as a coalition of different groups which are connected based on different activities, and each group within the firm has its own set of goals or demands (e.g. production, inventory, sales, market share and profit goals) (Koutsoyiannis, 1975; Griffiths and Wall, 2004). Some of the objectives of these coalitions of groups might be in conflict with the ultimate goals of the firm. Senior management has to decide the ultimate goals for the firm by performing continuous bargaining and setting an aspiration level in resolving the inevitable conflicting goals of the groups within the coalition (Cyert and March, 1963) which could result in organisational slack.²⁶ Therefore, the goals of the firm (as a result of a continual process of bargaining) will be in the form of aspiration levels instead of maximising process; that is, the behavioural theory seeks to 'satisfice'²⁷ rather than optimise

²⁶ Organisational slack is described as the difference between total resources and total payment. For example, wages paid in excess of those required to maintain labour and supplies purchased above the average price (Cyert and March, 1963; Carter, 1971). Organisational slack can result in increased unnecessary costs and thus may reduce the efficiency of the firm. Additionally, the conflicting goals between groups that arise during the process of the goal-setting at the top management level can be resolved through using monetary payments, side payments (in the form of policy commitments), slack payments, prioritising and sequencing attention given to demands and decentralising decision-making. Such means are used to ensure compliance or to satisfy individuals in some groups (Cyert and March, 1964; Koutsoyiannis, 1975).

²⁷ The word satisfice was introduced by Herbert A. Simon; it is a combination of the words 'satisfy' and 'suffice'. The word satisfice reflects bounded rational behaviour where firms will try to achieve a satisfactory rather than the maximum level of profit (Simon, 1955, 1959).

profits, sales, or other magnitudes (Simon, 1955). Simon (1955) introduced the concept of bounded rationality²⁸ (as opposed to global rationality in neoclassical theory) given the limitations in information, time and decision making to support the satisficing behaviour of large firms. Consequently, should the top managers attain the goals of the firm (which take the form of aspiration levels), the firm's performance can be considered as satisfactory; that is, it is possible for firms to survive in the market by not maximising profit, due to facing complex decision-making processes and organisational slack within the firm.

The final theory discussed is X-efficiency, which provides the rationale for why firms are not achieving maximum efficiency in their production process. The concept of X-efficiency was first introduced by Leibenstein in 1966. The 'X' was initially introduced because Leibenstein was unclear on how to describe this concept of inefficiency (Frantz, 1992). Leibenstein (1966, 1975, 1979) identifies X-efficiency as one of two main sources of inefficiency in the firm (another form of inefficiency mentioned by Leibenstein is allocative efficiency).²⁹ X-efficiency seeks to evaluate intra-firm behaviour and relationships as well as interactions among individuals within the firm, instead of the working of the price system. The X-efficiency theory hypothesises that: 'Neither individuals nor firms work as hard, nor do they search for information as effectively as they could' (Leibenstein, 1966, p.407). For this hypothesis, Leibenstein (1979) identified non-maximising behaviour as the key idea of X-efficiency, where the level of X-inefficiency is mainly determined by the level of effort of individuals in the firms. The degree of X-inefficiency can be measured as the difference between the maximum effective utilisation of inputs and the actual utilisation of inputs (Leibenstein, 1975). Therefore, X-inefficiency can be attributed to several mechanisms: first, the inefficiency of the firm can occur due to a lack of competitive pressure. For instance, firms may imitate each other and cooperate, rather than drive for aggressive competition, particularly in an uncertain market with competitive interdependence. Second, the inefficiency could stem from asymmetric information which mostly favours at least one

²⁸ Simon (1955) constructed the concept of bounded rationality which is based on the actual decision process in the behaviour of human individuals or a group of individuals. Bounded rationality explains that the rationality of individuals is restricted by cognitive limitation, relying on limited time and information in making complex decisions.

²⁹ Leibenstein (1966, 1975 and 1979) found two possible sources of inefficiency, which are losses from allocative inefficiency and X-inefficiency. Allocative inefficiency derives from the divergence between price and marginal cost such as monopoly, special tariffs, and barriers to competitive output pricing. X-inefficiency on the other hand makes reference to non-allocative inefficiencies.

party and is not available to others. Third, inefficiency occurs due to a principal-agent problem; in view of the imperfect contract between managers and owners, the situation allows managers to exercise their discretion regarding the amount and quality of work (Leibenstein, 1979; Frantz, 1992). Fourth, not all factors of the production function are identified. The most widely employed factors are capital, labour, and land. Nonetheless, there could be other input factors that affect the production of goods and services. Accordingly, Leibenstein (1979) also provides some direction in identifying X-inefficiency, namely via pressures arising from (1) the environment in which the firm operates, (2) the internal structure of the firm and, (3) the traditional rules that have been established in the firm. From the X-efficiency theory, Leibenstein (1975) criticises the assumption that firms have the single goal to maximise profit (as outlined by neoclassical theory), which he claimed that firms are unable to maximise profit due to maximisation of managerial utility (Demsetz, 1995).

In summary, neoclassical theory assumes that firms are making rational decisions with their single objective to maximise profit that focuses on the input and output of the firm, rather than on how the transformation is made. Moreover, neoclassical theory fails to explain why some firms survived regardless of not being efficient. Thus, managerial and behavioural theories focus on the organisational structure, managerial discretion and the behaviour of individuals within the firm to explain the inability of a firm to pursue goals other than profit maximisation. This also explains why firms continue to survive despite being inefficient in managing their inputs and outputs.

In relation to this research, the neoclassical, managerial and behavioural theories discussed earlier have implications towards the efficiency of Malaysian banks. For example, the market structure of the Malaysian banks can be seen as oligopolistic, where the banking industry in Malaysia is a concentrated market (as a result of consolidation of domestic banks after the 1997-98 Asian financial crisis) (Batchelor et al., 2005). Moreover, a few large banks have been leading the banking industry in setting the prices of the banking products which are more likely to be determined based on their internal cost structure, liquidity, profitability and risk appetite. Hence, this demonstrates that market concentration allows banks to apply market power to their products and enjoy a quiet life, which is strongly

related to banks being inefficient (Berger, 1995). In addition, the governance and organisational structure of the Malaysian banks have improved over time (since the Asian financial crisis in 1997-98), where a chairman of a bank is not allowed to become its chief executive officer (CEO) (as stipulated in BNM's Guidelines on Corporate Governance for Licensed Institutions). As a result of the separation between owners and managers, it is expected that higher agency costs, greater managerial discretion and the satisficing behaviour of senior management will exist within the banks in Malaysia. For instance, a CEO may want to increase sales and invest in risky financial products to generate higher income and accumulate additional funds, where they can use such funds to satisfy their managerial utility. Hence, theories such as agency, managerial-discretion, growth maximisation and the behavioural theory of the firm can help to explain why some banks in Malaysia are inefficient, particularly where this is due to the separation of ownership from the day-to-day operations of the banks.

2.3 The Measurement of Efficiency of the Firm Using Frontier Efficiency

The theory of the production function showed that the optimisation process is one of the criteria used in determining the efficiency of the firm. Maximum production, maximum profit and/or minimum cost are means that are used in the optimisation process (Lovell, 1993).³⁰ This is consistent with the theoretical concept according to which the firm optimises inputs and outputs to maximise profit. Additionally, apart from the production function the neoclassical, managerial, and behavioural theories that were discussed in the previous section explain why firms are operating in an inefficient manner. However, in order to determine the degree of inefficiency of the firm, empirical analytical methods such as frontier estimation are required to measure inefficiencies based on relative efficiency (as mentioned in Section 2.2). This section introduces frontier-analysis approaches that are used to measure the degree of efficiency of firms. These approaches derive mainly from the work of Farrell (1957), who introduces a foundation for estimating relative efficiency using the economic theory of production and its optimisation process.

³⁰ Lovell (1993) states that optimum productivity can be described in terms of production possibilities and behavioural goals of the production unit, where firms aim to have optimum cost, revenue, and/or profit subject to applicable constraints on quantities and prices.

As discussed earlier, the measurement of efficiency can assist policy makers and other interested parties to identify sources of inefficiency and to improve the competitive market conditions. The empirical measurement provides qualitative as well as quantitative evidence when theories are unable to offer guidance, or they give contradictory signals on the impact of some phenomenon on the performance of the firms. It is necessary to quantify discrepancies that are envisaged qualitatively by theory, using appropriate measurement techniques (Lovell, 1993). Lovell (1993) and Berger and Humphrey (1997) suggest that frontier estimation is useful because it: (1) provides information to policy makers regarding the policies implemented, (2) identifies causes of inefficiency, (3) ranks firms in the industry, and (4) evaluates best-practice and worst-practice firms. Moreover, frontier estimation is often regarded as superior to non-frontier methods (i.e. ratio analysis and ordinary least squares regression) (Lovell, 1993; Berger and Humphrey, 1997). One of the simplest non-frontier methods for measuring the performance of firms is financial-ratio analysis. This approach consists of using conventional performance ratios which examine the financial ratios of the firm, such as return on assets (ROA), return on equity (ROE), liquidity ratio, cost-to-income ratio and capital ratio; all of these are commonly utilised by regulators, analysts and bank managers. However, financial-ratio analysis is unable to control the impact of input prices and output prices as well as other exogenous factors, which restrict the performance ratios from approaching a closer estimation to the true performance of the firms (Lovell, 1993). That is, the financial ratios are limited in scope due to their one-dimensional view of products or processes, in which frontier measurement is superior as it involves programming and statistical techniques that consider the interactions of the key variables mentioned above (Berger and Humphrey, 1997; Barr et al., 2002). Consequently, another commonly used non-frontier approach is the ordinary least squares (OLS) regression method, which is used to estimate the production function. OLS is a method of estimating the unknown parameters of a function in a linear regression model. A line is generated for a linear approximation that minimises the sum of squared vertical distances between the observations in the dataset; this measures the average performance of the firms (Lovell, 1993). Hence, the result of average performance (via OLS) of the production function used to construct production frontier is inconsistent with the economic theory of the optimisation process; as it does not demonstrate the minimising or maximising behaviour of the firm (Lovell, 1993).

In view of the usefulness of frontier estimation and its common use in the literature, the discussion in this section begins with the efficiency measurement introduced by Farrell (1957). Based on Debreu (1951) and Koopmans (1951),³¹ Farrell (1957) introduces and employs a production function as an empirical method to estimate the productivity frontier. Frontier estimation is based on a relative measurement rather than an absolute measurement. Coelli et al, (2005) further iterate that frontier estimation focuses on comparing observations and evaluating optimal values of inputs and outputs. For example, as mentioned and described earlier in Section 2.2.1 and Figure 2.2, firms that are on the estimated frontier are said to be efficient. However, other firms that are positioned below the frontier are identified as inefficient. The relative measurement of inefficiency is determined by the distance between the production frontier (isosurface) and inefficient firms (Kumbhakar and Lovell, 2000). The production frontier efficiency can be categorised into technical efficiency, allocative efficiency, scale efficiency, and scope efficiency, which are discussed in the following subsections.

2.3.1 Technical and Allocative Efficiency

Farrell (1957) suggests that productivity efficiency can be decomposed into technical efficiency (TE) and allocative efficiency (AE). His work brought forward the proposition of the empirical treatment of a production function as a ‘frontier’, and proposed that productive efficiency consists of both TE and AE. In general, TE is described as the ability of a firm to produce maximum output based on a given input, or as using the minimum input for a given output. Mester (1997) asserts that a firm is said to be technically inefficient if it is producing less than maximum output from a given set of inputs, or is using more than the minimum input required for a given level of output. On the other hand, Lovell (1993) defines AE as the ability of a firm to combine inputs and outputs in optimal proportions, provided that the prices of inputs and outputs are known. Allocative inefficiency can be described as utilising an incorrect mix of inputs, given their prices and production technology, which may be

³¹ Debreu (1951) proposed the first measure of technical efficiency using a ‘coefficient of resource utilisation’, which uses the smallest quantity of inputs to produce a certain level of outputs. Additionally, Koopmans (1951) describes a producer as technically efficient when an increase of any output requires an increase of at least one input or a reduction in at least one output, and if a reduction in any input requires a reduction of at least one output or at least an increase in another input.

driven by factors such as the excessive or poor regulations made by policy makers (Mester, 1997; Isik and Hassan, 2003).

TE and AE can be analysed from two perspectives, namely input-orientated TE and AE and output-orientated TE and AE (Lovell, 1993). For input-orientated efficiency, TE and AE are viewed from a cost-minimisation standpoint, addressing a firm's optimal mix of inputs to produce outputs. Conversely, output-orientated TE and AE refer to a revenue maximisation viewpoint, measuring the optimal mix of outputs based on given inputs (Kumbhakar and Lovell, 2000).

2.3.1.1 Input-orientated Measures of Technical and Allocative Efficiency

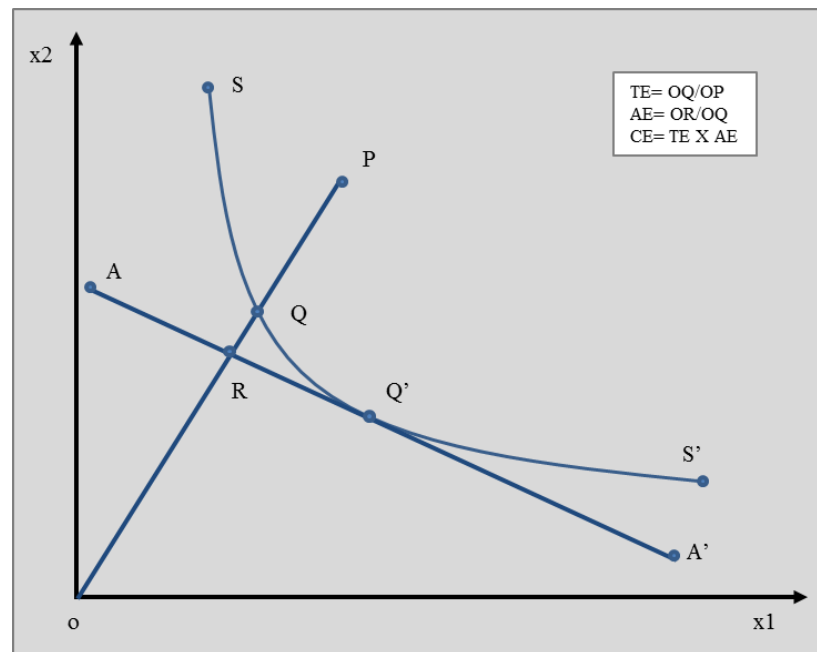
Farrell's (1957) original study was discussed with reference to the relationship between input/input space and its focus on the input-reduction process. This method describes the input-orientated measures of TE and AE (Coelli et al., 2005). Farrell (1957) suggests that the efficiency of a firm depends on its ability to produce maximum output from a given input; and in this context, TE can be explained using the frontier isoquant from the Figure 2.4. Figure 2.4 shows the isoquant SS' ; and under the assumption of constant returns to scale and a given level of technology, the firm produces one output (y) and two inputs (x_1 and x_2) ($y=f(x_1, x_2)$), using the best available technology. The isoquant SS' is the minimum combinations of the two inputs required to produce a given level of output. Every possible combination along the isoquant (for example at point Q and Q') is considered technically efficient; any point above or to the right of it can be described as technical inefficient because they can still reduce the use of inputs without reducing the output level. For example, assuming that a firm is producing at point P using two inputs (x_1 and x_2) to produce a given level of output, the firm is considered technically inefficient because by moving from P to Q it can produce the same quantity of output with fewer inputs. Thus, the TE of the firm can be measured by the ratio of OQ/OP . TE is represented by an indicator which takes a value from zero to 1. For instance, if OQ/OP is closer to 1, it indicates that a level of output of the firm for a given input is closer to the most technically efficient point for a given combination of inputs. The further the firm moves away from the frontier, the

more the performance of the firm deteriorates, which can result in TE moving closer to zero. The technical efficiency (TE) equation can be expressed as follows (Coelli et al., 2005):

$$TE = OQ/OP \quad (2.4)$$

where TE is technical efficiency, OQ is the potential inputs being used to produce a given output, and OP is the actual inputs being used to produce the same level of output that could be produced using fewer inputs.

Figure 2.4 Technical and Allocative Efficiency (Input-oriented)



Source: Farrell (1957, p. 254)

However, TE does not consider the price and the optimal allocation of inputs. Thus, AE was introduced by Farrell (1957), and this is estimated based on optimal proportions of inputs using known prices. The relative prices of production factors are incorporated in the isocost AA' in Figure 2.4, which is tangential to the isoquant SS' . The isocost AA' represents the ratio of the two input prices and measures the minimum cost to produce a given output. In other words, the isocost AA' shows the various combinations of inputs that require the same level of expenditure (i.e. cost). For instance, if the firm's overall production is efficient (i.e. TE and AE), it will occur at Q' (but not Q), which is the cost efficient point. At point Q' , the

SS' and AA' slopes are the same and this intersection indicates that the combination of inputs and output is fully efficient (where both TE and AE are attained simultaneously). Consequently, by moving Q to Q', the firm could produce the same output at a lower cost by adjusting its input levels. At point Q, the firm is considered allocatively inefficient (but technically efficient), where the firm made an inefficient choice of a combination of inputs at the given prices, and therefore incurred greater cost than it would have done if it had produced at point Q'. Thus, AE is measured by the ratio OR/OQ. The distance QR represents the potential reduction of production cost that can be made to be considered allocatively efficient. Similar to TE, an AE ratio closer to 1 indicates an output combination that is closer to the allocatively efficient point for a given level of output. Mathematically, AE is given by (Coelli et al., 2005):

$$AE = OR/OQ \quad (2.5)$$

where AE is allocative efficiency, OR is the observed combination of input prices that is associated with the cost-minimising input at Q', and OQ represent the observed combination of inputs prices that are associated with technical efficiency at point Q.

As mentioned earlier, the tangency point Q' represents the point of economic (overall cost) efficiency, in which both TE and AE are simultaneously efficient. In Figure 2.4, if firm at point P wants to be economically efficient, it must also be technically and allocatively efficient at point Q'. Thus at point P, the cost efficiency of the firm is the fraction OR/OP. In the case of input-orientation, this ratio (i.e. OR/OP) is termed by Farrell (1957) as the overall cost efficiency of the firm, measured by the product of technical efficiency (TE) and allocative efficiency (AE) (Coelli et al., 2005).

$$CE = TE \times AE = OQ/OP \times OR/OQ = OR/OP \quad (2.6)$$

where CE is overall cost efficiency, TE is technical efficiency; and AE is the allocative efficiency of the firm.

2.3.1.2 Output-orientated Measures of Technical Efficiency and Allocative Efficiency

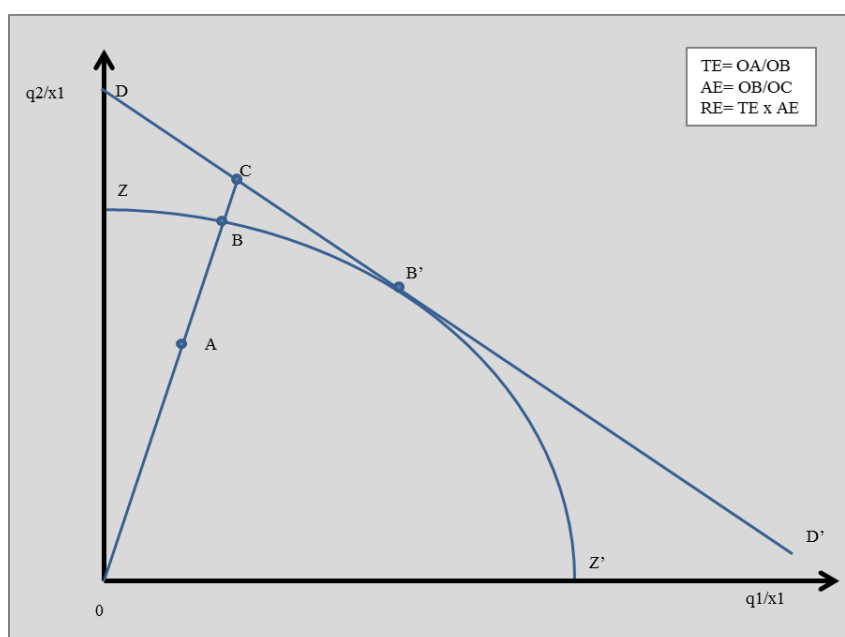
The idea of Farrell (1957) discussed previously is shown in the input/input space ratio from a cost-minimisation perspective. Aigner and Chu (1968) replicated the Farrell (1957) approach by estimating a parametric function for the input/output space; this provides a revenue-maximisation perspective. The differences between input-orientated efficiency and output-orientated efficiency can be described by addressing a couple of questions. Coelli et al. (2005) state that cost minimisation or input-orientated efficiency seek to address the question of ‘By how much can input quantities be proportionally reduced without changing the quantities or output produced?’ Conversely, the revenue maximisation or output-orientated efficiency is addressed in the question of ‘By how much can output quantities be increased without changing the quantities of the inputs?’

The latter question may be addressed using Figure 2.5, which illustrates a firm with two outputs (q_1/x and q_2/x) and one input (x) and assumed conditions of constant returns to scale and a given technology. Following Coelli et al. (2005), ZZ' is the possibility curve of different potential methods of producing two outputs with a given level of input. ZZ' represents the upper boundary of the production possibilities. Point A is inefficient, as the combination of the two outputs is inside the production possibility curve ZZ' . Point B represents an efficient firm, as it lies on the upper boundary of the production possibility curve ZZ' . The ratio of the distance between OA and OB is used to estimate technical efficiency (TE). Technical efficiency can be represented with an indicator which takes a value from zero to 1. A value of OA/OB closer to 1 indicates that the level of output of the firm for a given input is closer to the most technically efficient point for a given combination of outputs. If a firm is technically efficient, the ratio of OA/OB should be 1. The distance AB indicates the extent to which the two outputs can be proportionally increased without requiring extra inputs. Thus, to measure output-orientated technical efficiency, it will be the ratio of (Coelli et al., 2005):

$$TE = OA/OB \quad (2.7)$$

where TE is technical efficiency, OA represents the actual production of two outputs using one input, and OB represents the potential production of two outputs without requiring extra input.

Figure 2.5 Technical and Allocative Efficiency (Output-orientated)



Source: Coelli et al. (2005, p. 255)

AE is introduced to consider the optimal proportions of output based on prices, which is not reflected in TE. AE derives from an output perspective that consists of selecting the mix of outputs that maximises revenue with a given quantity of input. In Figure 2.5, the iso-revenue line DD' shows every possible combination of two outputs that generate the same total revenue, based on a given input and the available technology. The iso-revenue line DD' is tangential to the production possibility curve ZZ' when the slopes of both DD' and ZZ' are equal at point B' . At point B' the firm is considered to have optimal revenue efficiency (RE), in which both TE and AE are attained simultaneously. This point of tangency B' also indicates the maximum revenue efficiency that a firm can achieve. Although B is technically efficient, due to its location on the production possibility curve, it is however not allocatively efficient, in which could result in AE of less than 1. The distance CB is the

distance that can increase the revenue of the firm should it moves from point B to point B'. Since the revenue from production at B' is only a fraction of that at B, the ratio of OB/OC is used to measure AE. AE takes a value between zero to 1. The closer AE is to 1 implies that the output combination is closer to the allocatively efficient point for a given level of input. AE can be written as follows (Coelli et al., 2005):

$$AE = OB/OC \quad (2.8)$$

where AE is the allocative efficiency, OB is the observed combination of output prices that is associated with technically efficient production at point B, and OC is the observed combination of output prices that is associated with revenue efficiency at point B'.

AE and TE have a revenue-increasing interpretation which is similar to the cost-reduction interpretation in the input-orientated measures. To achieve overall revenue efficiency, firms should have the optimal mix of outputs to maximise revenue. This can be demonstrated in Figure 2.5, where at point B' both TE and AE are attained simultaneously. Hence, in the case of output-orientation, the overall revenue efficiency (RE) is a function of technical efficiency and allocative efficiency (Coelli et al., 2005).

$$RE = TE \times AE = OA/OB \times OA/OC = OA/OC \quad (2.9)$$

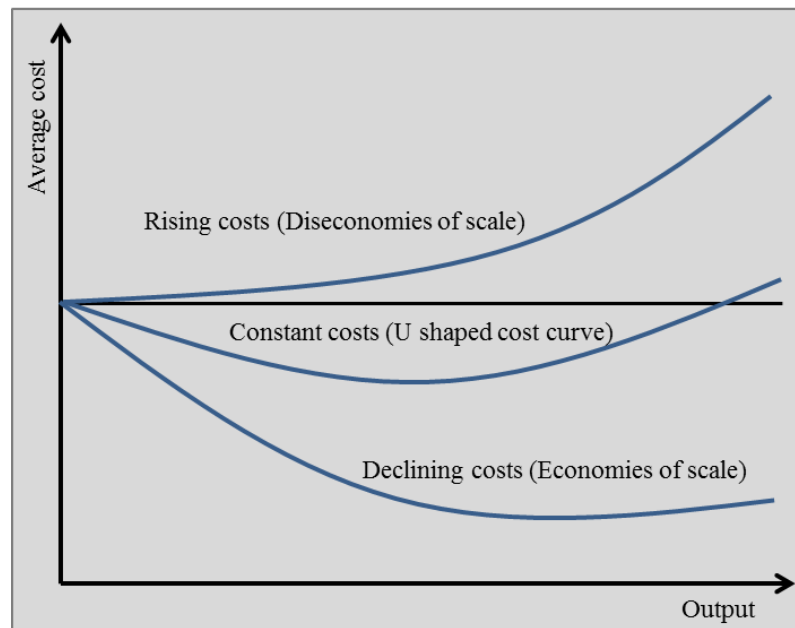
where RE is the overall revenue efficiency, TE is the technical efficiency and AE is the allocative efficiency.

2.3.2 Scale Efficiency

Section 2.3.1 reviewed the measurement of efficiency based on inputs and outputs, prices and production technology. The efficiency of a firm can be classified as technical efficiency, allocative efficiency and economic efficiency (i.e. cost and revenue efficiency). A firm may be technically efficient, allocatively efficient and economically efficient (i.e. cost, revenue, or profit efficient), but not necessarily scale efficient. Economies of scale are normally associated with the size of a firm and its cost advantage, which manifests as increased

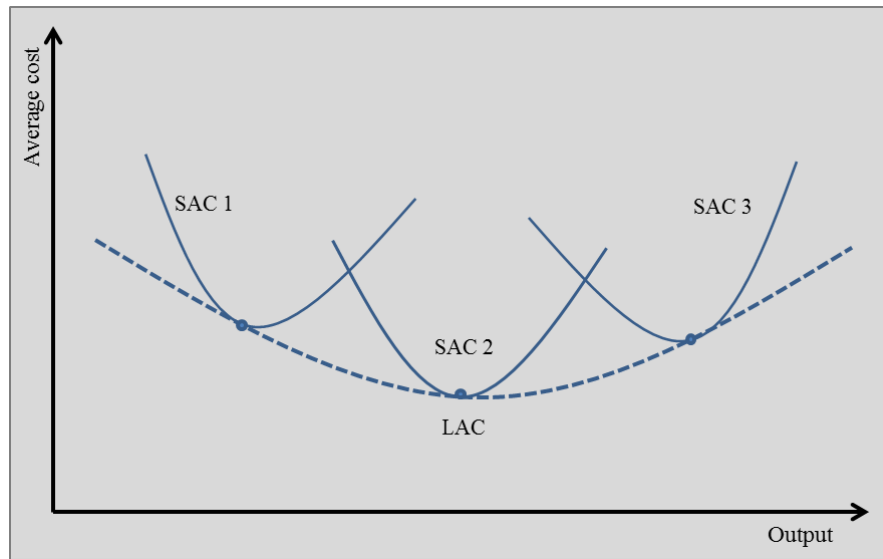
output of products (Clark, 1988). Molyneux et al. (1996) describe economies of scale as a rate at which output changes, as all factor quantities are varied; they measure whether firms with similar production and managerial technologies are operating at optimal size. That is, economies of scale are measured by the ratio of a proportionate change in output to a given proportionate change in all inputs (McGuigan and Moyer, 1989). This is measured by the ratio of the percentage change in the inputs relative to the percentage change in the outputs. Economies of scale exist when a proportionate increase in output is associated with a less than proportionate increase in inputs. On the other hand, diseconomies of scale materialise when inputs increase proportionately more than the output does as the latter increases. Constant returns to scale refer to the proportional change of output being equal to the proportional change in input (Clark, 1988).

Economies of scale are based on the shape of the average cost curve illustrated in Figure 2.6. The figure shows the U-shaped cost curve which represents a cost function exhibiting all three of these scale-economy characteristics. For empirical measures of scale economies, Clark (1988) suggests that the ratio of the proportion of output in relation to the proportion of input determines different types of scale economies. Given the total cost function defined by $TC=f(Q)$ where Q is output, the average total cost can be derived as $ATC= f(Q)/Q$ and marginal cost is $\partial TC/\partial Q$. The average cost will decline as long as marginal cost lies below average cost. Hence, the scale economies (SE) are given as ATC/MC , which is the elasticity of cost with respect to output (Altunbas et al., 1996). If the scale economies are equal to 1 ($SE=1$), it represents constant returns to scale; if the scale economies are less than 1 ($SE<1$), this reflects increasing return to scale; and conversely, if the scale economies are greater than 1 ($SE>1$), the firm is considered inefficient and decreasing returns to scale are present.

Figure 2.6 Economies of Scale and Average Cost Curves

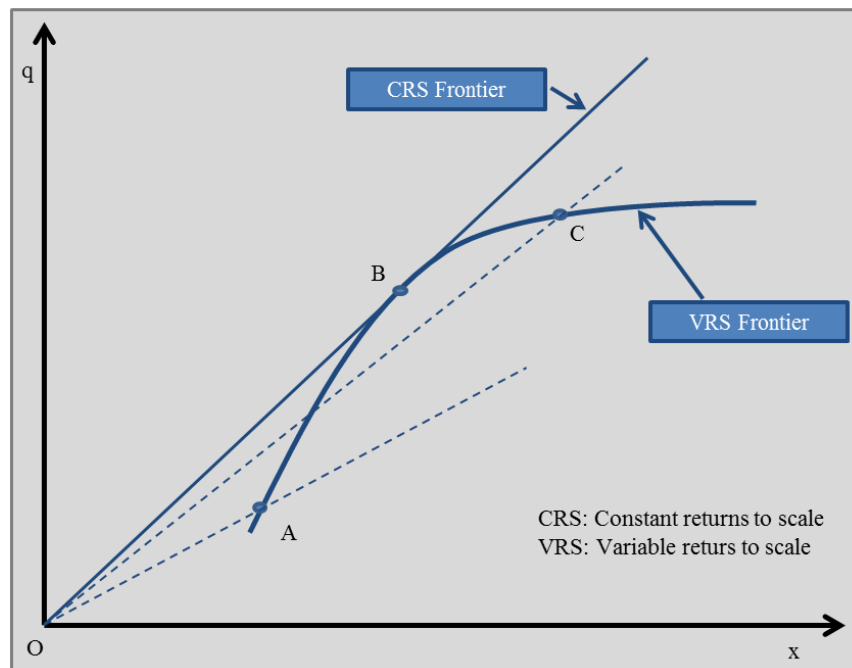
Source: Sinkey (1992, p. 306)

In addition, the firm's long run average cost describes what is happening to average cost when the firm expands, and is at a tangent to the series of short run average cost curves. Each short run average cost curve relates to a separate stage or phase of expansion (McGuigan and Moyer, 1989). Figure 2.7 shows a series of short-run average cost curves (SAC1, SAC2, and SAC3), with each producing different levels of output and a long-run average cost curve (LAC). In this figure, the firm will choose the size that yields the lowest average cost for their particular level of output. A curve of long-run average cost (LAC) is drawn from the series of short-run average cost curves, where each point of the LAC is to a point of tangency with a corresponding short-run average cost curves (SAC1, SAC2 and SAC3), and it shows the least-cost method of production for any level of output. Scale economies appear as the slope of the long-run average cost curve indicating how costs vary with output (Humphrey, 1990). A downward sloping LAC suggests economies of scale, due to average costs of production declining as output increases. Conversely, an upward sloping LAC suggests diseconomies of scale, indicated by higher costs incurred as output increases. A firm experiences constant returns to scale at the lowest point of the LAC, where a firm achieves its lowest average cost.

Figure 2.7 Economies of Scale

Source: Humphrey (1990, p. 38)

Many firms can be technically and allocatively efficient but in terms of productivity they may not be equally optimal. This inconsistency is due to scale effects. A firm can improve its cost and profit efficiency by producing at constant returns to scale as shown in Figure 2.8; the firms operating at points A, B and C are all technically efficient. This is because all of them are operating on the production frontier. However, the productivity of each firm differs due to the effect of scale. These firms are assumed to use one input (x) and produce one output (q). Firm A operates in the increasing returns to scale region of the production frontier. Firm A can increase its scale of production by being more productive and moving towards firm B. Firm B would be more scale efficient than Firm A, due to a higher ratio of output compared to input. Firm B is able to produce more output due to its lower average cost. On the other hand, firm C is operating in the decreasing returns to scale part of the production frontier. For firm C to become more productive, it has to move towards point B by reducing its scale of production. By doing so, it can reduce its average cost. Point B is the ideal region, because it is said to be operating at the most productive size, or equivalently, at the technically optimal productive scale (TOPS). This is also the point on the production frontier at which a ray from the origin is tangential to the production frontier (i.e. the slope of the ray displays a proportional change of output, which has a similar proportional change in input, where the ray angle from the origin is 45°).

Figure 2.8 Effects of Scale Production

Source: Coelli et al. (2005, p. 59)

An input-orientated measure of scale efficiency for a firm operating with a given input and output can be described as the ratio of technical efficiency of constant returns to scale over the technical efficiency of variable returns to scale (Coelli et al., 2005). That is,

$$SE = \frac{TE_{CRS}}{TE_{VRS}} \quad (2.10)$$

where SE is scale efficiency, TE_{CRS} is its technical efficiency with constant returns to scale, and TE_{VRS} is its technical efficiency under variable returns to scale. When $SE \geq 1$, the firm is experiencing increasing returns to scale, $SE = 1$ refers to constant returns to scale, $SE \leq 1$, denotes decreasing returns to scale.

According to Akhavein et al. (1997), scale efficiency can increase a firm's profit and improve cost in several ways. First, scale efficiency improves the cost per unit of output as the size of the firm increases. Second, scale efficiency can improve the profit efficiency

from a superior combination of inputs.³² Third, a firm that is scale efficient could have greater market power. Firms with greater market power can set prices and can charge higher prices for its products and thus raise profits for the firm. Additionally, De Young (1997) states that firms can improve their cost efficiency by moving closer to the cost frontier; through economies of scale (or growing larger), firms can reduce their per unit average cost and achieve cost savings by spreading their costs over large quantities of output.

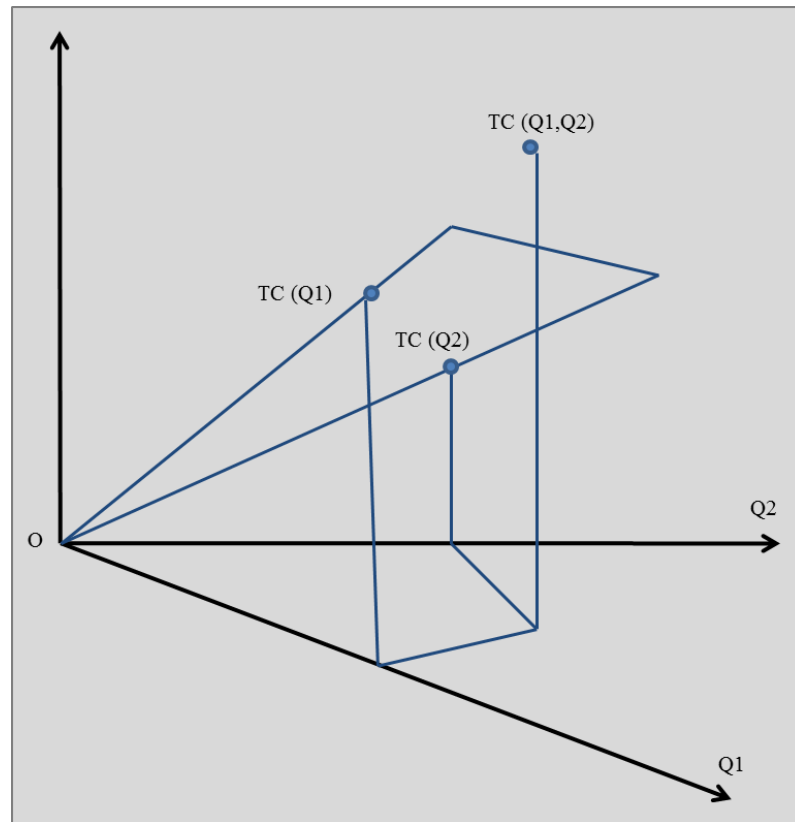
2.3.3 Scope Efficiency

Another type of efficiency that improves cost and increases profit efficiency is economies of scope.³³ A firm could experience cost savings and greater profit from joint production of multiple products. Economies of scope exist when two or more products can be jointly produced by a single firm with a lower cost compared to independent production by many firms (Panzar and Willig, 1981; Berger et al., 1987; Clark, 1988; Molyneux et al., 1996). Clark (1988) suggests two types of economies of scope, namely global and product-specific. Global economies of scope refers to the comparison of costs between joint production and by separate production by a number of firms. The economies of scope occurs if the joint production cost is less than the total cost of producing each product independently. On the other hand, product-specific economies of scale derive from adding different products with joint production to produce a different product. It exists when costs are reduced from joint production from various product mixes. However, in order to ascertain the best product mix, the cost complementarities between various pairs should be estimated.³⁴

³² For example, the profit efficiency of a firm could be improved without improving its cost efficiency through mergers and acquisitions. The newly formed firm could enjoy revenue growth greater than the increase in cost.

³³ Samuelson (1966) defines scope efficiency as joint-ness in production as the capability of a firm to produce multiple outputs at a lower cost at a given level of input, rather than a series of separate firms, each of which specialises in the production of single output.

³⁴ Cost complementarities refer to the extent to which the cost of producing a particular product may vary from the output levels of other products. Synergised cost complementarities can result in lower cost. Nonetheless, the best mix production combination has first to be determined by the firm.

Figure 2.9 The Concept of Economies of Scope

Source: Baumol et al. (1988, p. 72)

Following Baumol et al. (1988), the concept of economies of scope can be explained geometrically as seen in Figure 2.9. If the outputs are produced independently, the cost functions are $TC(Q_1)$ and $TC(Q_2)$. However, when the outputs are produced jointly, the joint cost for producing is $TC(Q_1, Q_2)$. The economies of scope exist if the two outputs are jointly produced and result in lower cost when compared with the cost of producing the same quantity of output independently, that is $TC(Q_1, Q_2) < TC(Q_1) + TC(Q_2)$. A measure of scope can be expressed by the Equation 2.11.

$$SCOPE = \frac{TC(Q_1) + TC(Q_2) - TC(Q_1, Q_2)}{TC(Q_1, Q_2)} \quad (2.11)$$

where $SCOPE$ is the scope efficiency, $TC(Q_1)$ is the total cost in producing output 1 independently, $TC(Q_2)$ is the total cost in producing output 2 independently, and $TC(Q_1, Q_2)$ is the total cost to produce outputs 1 and 2 jointly. If $SCOPE > 0$, it indicates the

firm's overall economies of scope; on the other hand, if the $SCOPE < 0$, it implies diseconomies of scope.

As discussed above, scale and scope efficiencies can increase profit and the cost efficiency of the firm. Therefore, a firm is often required to identify the sources that would result in economies of scale and scope. The sources of economies of scale and scope can be derived from various sources which arise when a firm has extra capacity in its inputs (Clark, 1988). A firm can increase its output without proportionately increasing the input in the production activity. The excess in input would reduce the output cost per unit when output increases. Berger et al. (1987), Clark (1988) and Humphrey (1990) state that the potential sources of economies of scale and scope can come from several areas such as better information technology, specialised labour, organisational flexibility, and information distribution. For example, banks are often required to distribute asymmetric information to borrowers and lenders. Hence, information distribution costs can be reduced when the bank expands in size (Berger et al., 1987). Information such as the financial result, industry information, and economic condition is used by the bank to grant loans. This information is usually reused for either the same or different borrowers. Thus, economies of scale arise when information is used for different borrowers. On the other hand, economies of scope exist when information is used for the same customer for different products (Clark, 1988; Humphrey, 1990). However, the estimation of economies of scope in the banking sector is difficult, due to the insufficiency of cost data for each output (Berger et al., 1987). Hence, this research does not analyse the economies of scope of Malaysian banks.

2.4 The Measurement of Frontier Efficiency

Frontier analysis provides an empirical and functional measurement of efficiency that is better than financial ratios (Farrell, 1957). Financial ratios are nonetheless important, although they are being limited in scope due to a one-dimensional view of products or processes, and also ignore any interactions, substitutions, and trade-offs between key variables (Siems and Barr, 1998).³⁵ Therefore, a more inclusive multiple-input and multiple-

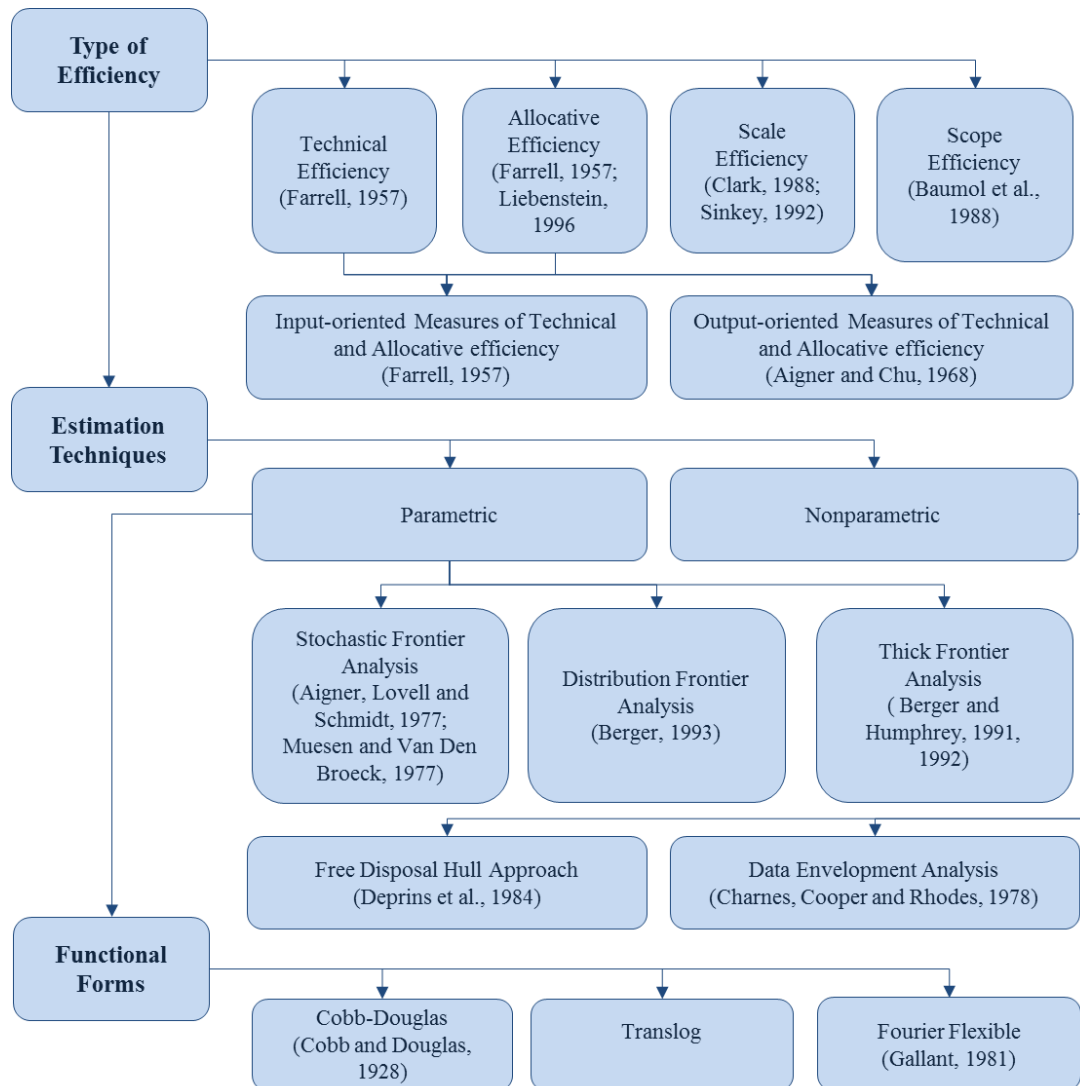
³⁵ According to Berger et al. (1993), financial ratios are regarded as misleading indicators of efficiency for three reasons: First, financial ratios do not control for product mix or input prices. Second, financial ratios such

output framework is required to evaluate the productive efficiency that provides benchmarking information. Berger and Humphrey (1997) and Barr et al. (2002) state that frontier measurement approaches are superior to financial ratios because they involve programming and statistical techniques that remove the differences in input prices and other exogenous factors. Nevertheless, frontier analysis relies on accounting measures from financial data, which in turn provides ratio information such as prices of input and prices of output (Berger and Humphrey, 1997).

The frontier-efficiency measures focus on relative instead of absolute estimation and are concerned with comparing potential and actual performance (Kumbhakar and Lovell, 2000). They measure efficiency using the observed sample of firms in the industry (Farrell, 1957). As mentioned in Section 2.3, firms located on the estimated frontier are said to be efficient, while others (those not on the frontier's isosurface) are inefficient. Efficiency (e.g. technical efficiency and allocative efficiency) is therefore being measured based on the distance to the frontier curve. The measurement of efficiency widely used in the literature can be divided into two groups: the parametric and nonparametric approaches (see Figure 2.10). The approaches are similar, as they attempt to benchmark the relative performance of the firms by estimating their best frontier, but the underlying assumptions of these two approaches are very different.

According to Berger and Humphrey (1997), the differences in the underlying assumptions between nonparametric and parametric models can be identified in terms of functional forms and random noise. The functional form is apparent in the parametric approach as it requires functions or restrictive structures for its specification to estimate the best practice frontier. On the other hand, the nonparametric approach requires relatively little structure. In terms of assumptions regarding random noise, the parametric approach assumes two-component error terms. The two components are the inefficiency term (managerial weakness) and the random error term. Conversely, the nonparametric approach assumes that the deviation from best practice is due to managerial error (inefficiency) only. Hence, random noise is not identified, due to whole error term being considered as inefficiency.

as cost-to-asset ratio assume that all assets are equally costly to produce. Third, the use of simple ratios would not distinguish between X-efficiency gains and scale and scope efficiency.

Figure 2.10 Conceptual Efficiency Framework

Adapted from Ahmad Mokhtar et al. (2006, p. 5)

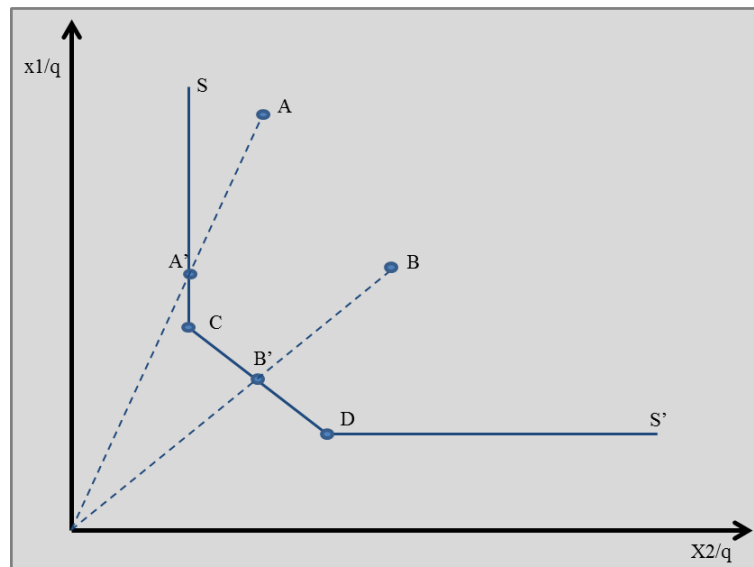
2.4.1 Nonparametric Frontier Approaches

The nonparametric frontier was introduced by Charnes, Cooper and Rhodes (1978), who used linear programming techniques to construct a production frontier and to measure the efficiency of the firms relative to their estimated frontier. Thus, this method envelops a data set of observations, but makes no accommodation for noise or random error (Lovell, 1993). Hence, the nonparametric method assumes (1) no measurement error in constructing the frontier, (2) no luck or chance that temporarily gives a production unit better performance,

(3) no inaccuracies created by accounting rules that would make measured outputs and inputs deviate from economic outputs and inputs; and (4) no sampling error caused by the selection of a sample instead of conducting a census of the entire population. Any of these errors that appear in an inefficient unit's data may be reflected as part of inefficiency; the approach assumes that random error is zero (Kumbhakar and Lovell, 2000). The common nonparametric approaches such as data envelopment analysis (DEA) and free disposal hull (FDH) put relatively little structure on the specification of the best frontier.

2.4.1.1 Data Envelopment Analysis (DEA)

DEA is built upon the work of Farrell (1957). It was introduced by Charnes, Cooper and Rhodes (1978), who developed an application of mathematical programming to locate a frontier based on the observed output and input quantities used by each firm to evaluate their efficiency level. DEA is described as a nonparametric approach and not dependent on functional forms (Avkiran, 1999; Brown, 2003; Attaulah et al., 2004; Drake et al., 2006). It relies on linear programming techniques for constructing a nonparametric envelopment (piecewise linear envelopment) to a set of observed outputs and inputs (Berg et al., 1993). Figure 2.11 illustrates the concept of DEA with four firms, utilising two inputs (x_1/q and x_2/q) to produce one output (q). Firms C and D are on the efficient frontier, because they use the least input combinations to produce output. The curve SS' connects the points C and D which represents the efficient frontier that is enveloping all observable data. Firms A and B are considered inefficient. However, firms A and B could reduce the quantity of their inputs to produce the same output on the frontier, as shown at points A' and B' (Coelli et al., 2005).

Figure 2.11 Data Envelopment Analysis

Source: Coelli et al. (2005, p. 165)

Casu and Girardone (2002) assert that the efficiency scores are generated from the actual data of firms under assessment where the score is relative (not an absolute estimate) to other firms within the sample. The best performance firms in the sample are utilised as a benchmark (i.e. firms that are located on the estimated frontier). The scores can be generated by either measuring all inputs used in the production process and determining the minimum level of inputs required to produce the same quantity of output or by measuring all inputs used in the production process and determining the levels of output that are required to be increased, based on the given inputs (depending on minimisation or maximisation constraint imposed in the linear programming). Firms with fewer inputs used to produce the same output or producing greater outputs based on a given input are considered more efficient and generate higher scores (Brown, 2003). The efficiency scores do not imply any explicit relationship between inputs and outputs (i.e. no algebraic relationship between inputs and outputs) because DEA does not impose any functional forms in its nonparametric approach (Ray, 2010).

There are two common variants to DEA: the Charnes, Cooper and Rhodes (1978) (CCR) model and the Banker, Charnes and Cooper (1984) (BCC) model. Charnes, Cooper and Rhodes (1978) introduced the CCR model, which is based on constant return to scale (CRS)

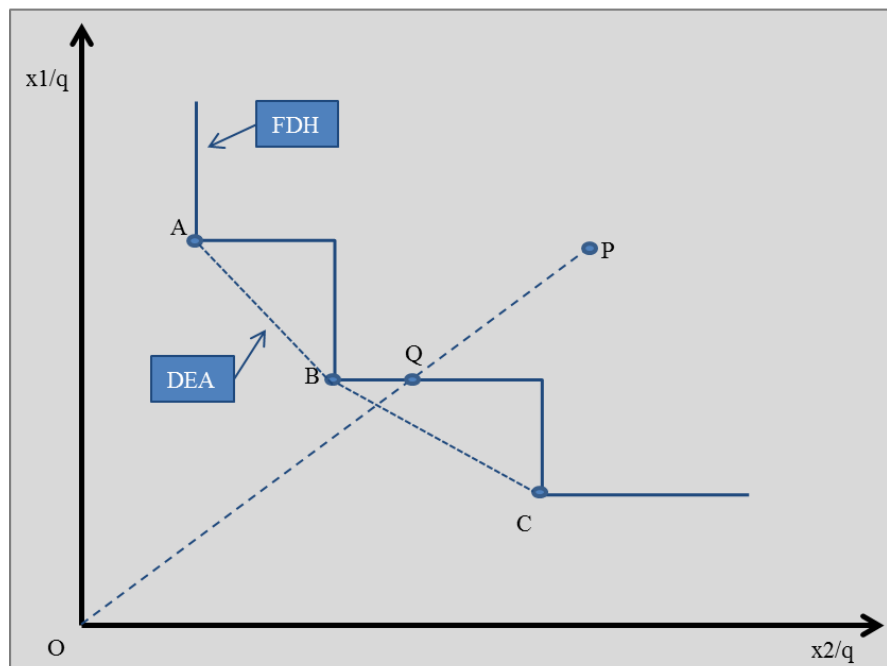
when enveloping the observed data to determine the shape of the frontier. The CCR model assumes that all firms are operating at an optimal scale. The assumption of CRS is particularly inappropriate when firms are experiencing economies and diseconomies of scale. Therefore, Banker, Charnes and Cooper (1984) argue that there are many factors that can cause firms to fail to operate at their optimal scale such as regulations, imperfect competition, macroeconomic shocks, and internal structural deficiency. The BCC model drops the CRS assumption by introducing variable returns to scale (VRS) in an attempt to identify simultaneously the most efficient scale for each firm in determining both technical efficiency and the envelopment surface. Nevertheless, there are various extensions and new concepts of DEA such as the additive model, the slack-based measure, the Russell measure, a non-radial model, and a multi-stage model.³⁶ Economic efficiencies such as cost and profit efficiency had also been extended from the basic DEA model. These are discussed in Chapter 5 (Methodology and Data).

2.4.1.2 Free Disposal Hull (FDH) Analysis

The FDH is nonparametric frontier model developed by Deprins et al. (1984) and extended by Tulkens (1993). FDH is a special case of the DEA model where the points connecting the DEA vertices are not included in the frontier. Instead, the FDH production possibility set is composed of only the DEA vertices and the free disposal points interior to these vertices (Berger and Humphrey, 1997). The FDH approach relaxes the assumption of convexity and assumes that there is no linear substitution possible between input or output combinations on a piecewise linear frontier. Unlike the convexity shape of the DEA frontier, the FDH has a staircase shape. Because the FDH frontier is either congruent with or interior to the DEA frontier, it therefore generates larger estimates of the average efficiency score than does DEA. The differences between FDH and DEA is that FDH evaluates firms' efficiency based on observed data, while DEA measures efficiency based on firms on the frontier (Tulkens, 1993). One of the shortcomings of FDH is that the approach is unable to measure economic efficiency (e.g. cost and profit efficiency), but only measures technical efficiency (Cook and Seiford, 2009).

³⁶ For more details of the extension of DEA models, see Cook and Seiford (2009).

Figure 2.12 Free Disposal Hull



Source: Deprins et al. (1984, p. 287)

To illustrate the FDH concept, Figure 2.12 shows the staircase-shaped frontier based on observed data with five production points A, B, C, P and Q. Points A, B and C describe efficient firms; P is inefficient as it does not lie on the frontier that tightly envelops the observable data. However, firms producing at point P could reduce the quantity of their inputs to become efficient (Deprins et al., 1984). Unlike DEA where input slack can be compared with the efficient frontier, FDH allows comparisons to be made only on actual observations, which makes slack more difficult to estimate when compared to DEA (Borger et al., 1998).

2.4.2 Parametric Frontier Approaches

In comparison to the nonparametric approach, the parametric approach requires functional forms (e.g. Cobb-Douglas, transcendental logarithms and Fourier Flexible) and a distribution assumption (e.g. half-normal, truncated normal and gamma) to disentangle residual (noise) into random error and inefficiency. The parametric approach can be divided into the deterministic and stochastic frontier methods. The deterministic frontier model is

normally measured using the ordinary least squares (OLS) estimation to estimate efficiency. On the other hand, the stochastic frontier model has been widely employed in the past literature due to its ability to distinguish inefficiency from random noise. The stochastic frontier approaches that are discussed in this section are the stochastic frontier approach (SFA), the thick frontier approach (TFA), and the distribution frontier approach (DFA).

2.4.2.1 Deterministic Frontier Model

One of the earlier models for the parametric approaches is the deterministic frontier model. This model uses a one-sided error term with non-negative properties in which the whole residual is considered as inefficient. The deterministic frontier utilises the OLS and OLS variants such as corrected ordinary least squares (COLS) and modified ordinary least squares (MOLS). The model can be written as follow:

$$\ln TC_i = f(w_{ni}, y_{ki}, \beta_i) + \ln \varepsilon_i \quad (2.12)$$

where $\ln TC_i$ is the logarithm of costs of the i -th firm, w_{ni} is the vector of prices of the variable inputs of the i -th firm, y_{ki} is the vector of the output levels of the i -th firm, β_i represents the set of parameters to be estimated, and ε_i is the error term of the firm's cost/profit inefficiency and is assumed to be non-negative and greater than zero.

The level of inefficiency of firms is indicated by the error term ε_i . If ε_i is equal to zero, inefficiency does not exist, but if it is greater than zero, inefficiency exists within the firm. However, the model has its weakness in that it is found to be inconsistent with the theoretical notion of the optimisation process, where the focus of this model had turned into data envelopment through the construction of a cost frontier (Kumbhakar and Lovell, 2000). The problem however can be corrected by using the COLS method introduced by Winsten (1957). This is done by shifting the intercept downwards (in the case of cost frontier but shift upwards for production and profit frontiers) by the highest residuals and ensuring that the firm with the lowest cost is utilised as the benchmark with which to estimate other firms' cost frontiers. Nonetheless, this technique is vulnerable to outliers. Thus, Afriat (1972) and Richmond (1974) proposed the MOLS method, which assumed that the error term ε_i is a

one-sided distribution, similar to the half-normal assumption; where the intercept parameter is shifted downwards (for cost frontier) by the mean of the assumed distribution.

Both COLS and MOLS require the parameters to be first estimated by OLS, and subsequently followed by the adjustment of the intercept parameters as mentioned above. The main drawback of COLS and MOLS is that they only change the estimation of the intercept parameter, making the frontier line parallel to the OLS regression line, which leads to having two lines with the same structure. As a result, the frontier does not meet the aim to bound the data as closely as possible (Kumbhakar and Lovell, 2000). Additionally, neither COLS nor MOLS separate inefficiency from the random error ε_i which treats all residuals as inefficiency (Coelli et al., 2005). In view of this inherent weakness, stochastic frontiers were introduced to overcome the limitations of deterministic frontier models. As mentioned earlier, the stochastic frontier models are capable of separating the error term into noise and inefficiency, which is discussed in the next section.

2.4.2.2 Stochastic Frontier Analysis (SFA)

In contrast to deterministic models, the stochastic frontier allows for the composite error term ε_i to be split into an inefficiency component and a random error (Berger and Humphrey, 1997). SFA was independently introduced by Aigner et al. (1977) and Meusen and Van den Broeck (1977) in the same year. Similarly to the frontier efficiency principle, SFA assumes that a firm's output is influenced by technology and a given input. Thus, the production of the firm is influenced by the parametric function of known inputs, output, control factors, and random error. The economic functional forms such as cost or profit functions are widely used to reflect the optimisation process of the firm. For illustration of SFA, the single-equation stochastic cost function model is shown below:

$$\ln C_{it} = f(Y_{nit}, W_{kit}, Z_{jit}) + \ln \varepsilon_{it} \quad (2.13)$$

Where C_{it} is the observed total cost of production for the i -th firm at time t . Y_{nit} is a vector of outputs, W_{kit} is an input price vector, and Z_{jit} is a vector of exogenous factors. Following Aigner et al. (1997), the assumption of the composed error term is as below:

$$\varepsilon_{it} = V_{it} + U_{it} \quad (2.14)$$

where V_{it} and U_{it} are independently distributed, V_{it} represents random error and is assumed to be normally distributed with zero mean and variance, $\sigma_v^2 \geq 0$ is drawn from a one-sided distribution that is assumed to capture the effect of inefficiency. U_{it} is assumed to be a half-normal distribution with mean zero and variance σ_u^2 . Inefficiency U_{it} , can be estimated by using the conditional mean of the inefficiency term from the error term. Jondrow et al. (1982) proposed that log likelihood for inefficiency is explained in terms of the two variance parameters, $\sigma^2 = \sigma_v^2 + \sigma_u^2$, which capture the variance of composed error term ε_{it} and $\lambda = \sigma_u / \sigma_v$, which measures the fraction of inefficiency relative to statistical noise.

SFA avoids problems associated with nonparametric frontiers by explicitly considering the stochastic properties of the data and separates the composite error term into inefficiency and statistical noise (Kaparakis et al, 1994). The composite error term can be disentangled by making distribution assumptions. The most common composite error model is the half normal-normal distribution which assumes that a half normal distribution is used to capture inefficiency and a standard normal distribution is used for capturing the random error (Aigner et al. 1977; Berger 1993; Berger and Mester 1997). Apart from the half-normal assumption, there are several alternative assumptions for the composite error, such as gamma, exponential, truncated-normal and deterministic kernel distribution (Greene, 1990; Berger, 1993; Bos et al. 2009).³⁷ The assumption imposed on the composite error term prevents inefficiency from generating a negative value (Fries and Taci, 2005). Finally, the structured composite error term can trace whether the firm's inefficiency is derived from either measurement problems or managerial deficiency (Fries and Taci, 2005).

³⁷ Greene (1990) argues that in comparison to a gamma distribution, the half-normal assumption for inefficiency is inflexible and imposes an arbitrary restriction in which firms are clustered near full efficiency. Stevenson (1980) and Berger and De Young (1997) specified a more general truncated normal distribution and found minor but significant improvement compared to the half-normal model. Additionally, Venet (1998) employs both half-normal and exponential distributions in his study, but he found little difference between these distributions.

2.4.2.3 Thick Frontier Approach (TFA)

The thick frontier approach (TFA) was developed to avoid the restrictive assumptions on the error term found in SFA. TFA has little structure and generates less information (Kumbhakar and Lovell, 2000). TFA also requires functional specifications such as cost and profit functions to estimate the efficiency of the firm. The TFA model compares the efficiency of groups of firms, rather than trying to estimate the frontier of production functions. In determining the groups, firms are ranked and divided into quartiles based on cost per unit of assets. The highest quartile is represented by firms that are presumed to have higher efficiency, and at the other end, the lowest quartile is assumed to have firms with lower efficiency. A cost function is estimated for both quartiles. The deviations in the error terms, and predicted cost between the highest and the lowest quartile of the observations are assumed to represent random error and inefficiency, respectively (Berger and Humphrey, 1992). That is, the estimation of random error and inefficiency in TFA relies on differences between the highest and lowest quartiles rather than distribution assumptions. In addition, Berger and Humphrey (1991) state that TFA requires less specificity in the maintained assumptions; therefore it is less likely to be violated substantially by the data. However, it does not generate cost efficiency estimates for each firm; instead, it produces one cost efficiency estimate for the highest-cost quartile relative to the lowest-cost quartile. Thus, TFA may not provide accurate estimates of the general level of overall cost efficiency, which makes it unlikely to be used by managers and policy makers (Kumbhakar and Lovell, 2000).

2.4.2.4 Distribution Free Approach (DFA)

The distribution-free approach (DFA) was introduced by Schmidt and Sickles (1984) and Berger (1993). Similar to TFA, DFA avoids using distributional assumption for the error component. Nevertheless, DFA requires a functional form for the construction of the efficient frontier similar to SFA, but inefficiency is separated from the random error differently. DFA makes no strong assumptions regarding the distribution of inefficiency and the random error. Instead, DFA uses stability over time to separate inefficiency from random error. It assumes that inefficiency is constant through time and that random error

tends to average out over time (Bauer et al., 1998). Hence, DFA is applied to panel but not cross-section data. By using a panel data set (i.e. repeated observation of each firm over a period of time), the estimation of efficiency is determined as the difference between each firm's residual and the average firm on the frontier (Fries and Taci, 2005). DFA can be estimated as follows (Kumbhakar and Lovell, 2000):

$$\ln c_{it} = f(y_{nit}, w_{kit}, \beta_{it}) + \ln \varepsilon_{it} \quad (2.15)$$

where $\ln c_{it}$ is the logarithm of cost of the i -th firm at time t , y_{nit} is the vector of output levels of the i -th firm at time t , w_{nit} is the vector of prices of variable inputs of the i -th firm at time t , β_{it} represents the set of parameters to be estimated, ε_{it} is the error term of i -th firm at t .

De Young (1997) states that the statistical properties of DFA are intuitive, apply few arbitrary assumptions, and are easy to apply. It also avoids possible endogeneity³⁸ of prices while providing efficiency estimates. DFA also supports economic interpretations that are comparable to those from a cost and profit frontier (Adams et al., 1999). However, the primary drawback of DFA is the requirement of time-invariant cost efficiency (i.e. constant through time and no consideration of technological change) where this assumption becomes less plausible as time increases. Hence, it is possible that inefficient firms turn into efficient firms as time passes (Kumbhakar and Lovell, 2000).

2.5 Functional Forms

As discussed above, a functional form is required for parametric models, particularly for cost and profit efficiency. The functional form is used to estimate the relationship between the dependent and explanatory variables. For example, cost is expressed as a function of input prices and outputs (i.e. $tc=f(w,y)$, where tc is total cost, w is the input price and y is output), and profit is expressed as function of input price and output price (i.e. $tp=f(w,q)$; where tp is total profit, w is input prices and q is output prices). Coelli et al. (2005) suggest

³⁸ A variable is said to be endogenous should there be a correlation between the variable and error term, in which arise from several factors such as measurement errors and omitted variables.

that several factors are considered when choosing functional forms, namely flexibility, linear in parameters, regularity and parsimony³⁹. Thus, the estimations from parametric models could lead to different results as a consequence of the choice of the functional form. The most common functional forms such as Cobb-Douglas, transcendental-logarithm (translog), and the Fourier-Flexible functional form are discussed in the following subsections.

2.5.1 Cobb-Douglas Functional Form

In 1928, Charles Cobb and Paul Douglas published a study on American economic growth between 1899 and 1926. The authors considered a simplified view of an economy in which output is determined by the amount of labour involved and the capital invested. The Cobb-Douglas function can be described in the following logarithm form:

$$\ln C_i = \ln \alpha_0 + \sum_{n=1}^N \beta_n \ln w_{ni} + \sum_{k=1}^K \delta_k \ln y_{ki} \quad (2.16)$$

where $\ln C_i$ is the logarithm of total cost, w_{ki} is the vector price of input of the i -th firm, y_{ki} is the vector price of the i -th output level, and $\alpha_0, \beta_n, \delta_k$ are the parameters to be estimated, representing the cost elasticities of the output and input prices. $\sum_{i=1}^n \beta_i = 1$ indicates that the Cobb-Douglas functional form is homogeneous of degree 1 in input prices. This homogeneity restriction in the cost function suggests that an increase in input prices can result in a proportional increase in total cost (this homogeneity restriction is consistent with the production function's properties discussed in section 2.2.1.)

This function is easy to estimate and interpret compared to other functional forms (e.g translog and Fourier-Flexible functional forms). However, the primary drawback of the

³⁹ Coelli et al. (2005) state that, first, a functional form can be divided into first-order and second-order flexible. The second-order flexible is normally preferred as it provides second order approximation, which can be used in measuring scale economy and changes in technology. Nonetheless, increased flexibility has its drawbacks, such as more parameters to estimate, which can contribute to issues such as multicollinearity and reduced degrees of freedom (Chambers, 1988). Second, the parameters should be agreeable to the linear regression techniques, particularly to measure causality via coefficients of the unknown parameters. Third, the functional form should satisfy the economic regularity properties, and fourth, the functional form should be parsimonious. That is, the functional form is adequate for the estimation of efficiency (Fuss, McFadden and Mundlak, 1978).

Cobb-Douglas function is that it is a first-order approximation; hence a constant value for elasticity of scale exists, making it impossible to test different values of economies of scale (Coelli et al., 2005).⁴⁰ Consequently, this inflexible property of the Cobb-Douglas functional form has caused it to be used rarely to estimate cost and profit efficiency (Coelli, et al., 2005).

2.5.2 Translog Functional Form

One of the most widely employed functional forms is the translog (transcendental logarithmic) production function (Berger and Mester 1997; Bauer et al. 1998; Bos and Kool 2006; Fu and Heffernan 2007). Unlike the Cobb-Douglas cost function (i.e. linear homogeneity restrictions are implemented on the first-order approximation), the translog functional form provides a second-order approximation and allows economies of scale to vary with the output level (Coelli et al., 2005). This specification allows for the necessary flexibility when estimating the frontier function (Bos and Kool 2009). The translog function imposes few restrictions on the first-order and second-order effects at the same time. It provides a second-order logarithmic approximation to an arbitrary continuous transformation surface (Coelli et al. 2005; Kaparakis et al. 1994). Kumbhakar and Lovell (2000) state that the translog function provides a better fit to the frontier which envelops the Cobb-Douglas specification. The general model of the translog cost function can be expressed as:

⁴⁰ The Cobb-Douglas function assumes that any change in output will result in a change in total cost by a constant number δ_i , which indicates the constant elasticity of total cost, as shown by first-order derivatives $\frac{\partial \ln C_i}{\partial \ln y_j} = \delta_i + \sum_{i=1}^n \beta_i \ln w_i$. Additionally, the time trend is implicitly assumed to be constant in the Cobb-Douglas functional form which may not reflect the effect of technological advances, which often cause economic relationships (production function) to change over time (Coelli et al., 2005).

$$\begin{aligned}
\ln C_{it} = & \alpha_0 + \sum_{j=1}^J \beta_j \cdot \ln y_{jit} + \sum_{m=1}^M \delta_m \cdot \ln w_{mit} \\
& + \frac{1}{2} \sum_{j=1}^J \sum_{k=1}^J \beta_{jk} \cdot \ln y_{jit} \ln y_{kit} + \frac{1}{2} \sum_{m=1}^M \sum_{n=1}^M \delta_{mn} \cdot \ln w_{mit} \ln w_{nit} \\
& + \sum_{j=1}^J \sum_{m=1}^M \gamma_{jm} \cdot \ln y_{jit} \ln w_{mit} + u_{it} + v_{it}
\end{aligned} \tag{2.17}$$

where $\ln C_{it}$ is the logarithm of total cost of i -th firm at time t , $\ln y_{jit}$ is the logarithm of the outputs for the i -th bank at time t ; $\ln w_{mit}$ is the natural logarithm for the input price vectors for i -th bank at time t ; and $\alpha, \beta_j, \delta_m, \beta_{jk}, \delta_{mn}$, and γ_{jm} are parameters to be estimated. Since the duality theorem requires that the cost function is linearly homogenous in input prices, and that parameters of second-order partial derivatives are symmetric, the following restrictions have to be imposed on the parameters of the translog function form:

$$\sum_{m=1}^M \delta_m = 1, \sum_{m=1}^M \delta_{mn} = 0 \text{ and } \sum_{j=1}^J \gamma_{jm} = 0 \tag{2.18}$$

However, as noted in some studies (for example Berger and Mester 1997), the translog function is not free from weaknesses. McAllister and McManus (1993) and Mitchell and Onvural (1996) show that there are ill-fitting data in the translog function, particularly in measuring scale economies.⁴¹ They further iterate that the translog functional form may not fit well with data that are far from the mean in terms of output size, which can result in a specification bias that contributes a false conclusion concerning scale economies (Berger and Mester, 1997). Despite its limitations, the translog functions are widely used due to their flexibility and their forecasting capability when compared to other functional forms (Altunbas and Chakaravarty, 2001).

⁴¹ Translog form uses the Taylor series of expansion method (a mathematical technique used to expand a function in a point) that is approximated from a Cobb-Douglas function (Mitchell and Onvural, 1996). The expansion point in the translog function is the mean of the variables to be extended (i.e. C , y and w). Hence, the translog functional form is a locally flexible form, where the cost function does not fit all the data to the greatest possible extent.

2.5.3 Fourier Flexible Functional Form

The Fourier-Flexible (FF) functional form was first proposed by Gallant (1981). It derived from two main components: the first is the standard translog function, and the second the trigonometric Fourier series. These two components are dependent on each other, where the linear combination of sine and cosine of the variables is used to estimate the translog functional form. The FF functional form has mathematical and statistical properties with an infinite Fourier series capable of representing any function in its entire range. When using the FF form, one avoids holding any maintained hypothesis by allowing the data to reveal the true cost function through a large value of fitted parameters (Berger and Mester, 1997). Therefore, FF is independent of arbitrary and restrictive functional forms (Huang and Wang, 2002).

As mentioned earlier, the translog function may not fit well with data far from the mean of output size. Thus, the FF functional form overcomes this problem with better approximation across a broad range of outputs by incorporating additional parameters for the Fourier trigonometric terms (Girardone et al., 2004).⁴² However, additional parameters and trigonometric terms can give rise to a couple of problems. First, with trigonometric terms, for the FF function to hold it requires a larger sample of observations (i.e. higher degree of freedom) and, second, the additional parameters can give rise to multicollinearity issues in estimating the parameters of the model specification (Chambers, 1988; Coelli et al., 2005, Fu and Haffernan, 2007). Huang and Wang (2002) state that the FF form has not attracted much attention compared to the translog form in efficiency studies, due to the technical difficulties in constructing and estimating it.⁴³ They further argue that most of the studies utilising the FF form failed to incorporate allocative inefficiency into their statistical model, which can increase the potential for specification error.

⁴² Wheelock and Wilson (2001) argue that there is an unresolved statistical problem for the FF functional form regarding the supplement of the translog function, as to whether to employ either trigonometric terms or orthogonal polynomials. Additionally, the number of terms that should be included for estimation is an unresolved problem for the FF functional form.

⁴³ Nevertheless, Berger and Mester (1997) discovered that few differences are noted between the FF and translog functional forms. The improvement in goodness of fit from using the FF functional form appears to be small. That is, both functional forms yield a similar average level and dispersion of measured efficiency. These functional forms rank the individual firms in almost the same order.

2.6 Is There a Best Frontier Approach?

As seen from the earlier discussion, the measurement of efficiency through frontier estimation can be divided into two main techniques: parametric and nonparametric. These techniques differ from one another in terms of functional form and random noise (Berger and Humphrey, 1997). The key advantage of a parametric approach is its ability to separate random error into inefficiency and noise (Coelli et al., 2005). However, parametric requires the specification of functional form and distribution assumptions on random error, which could influence the shape of the efficient frontier and the distribution of the residuals. Nevertheless, the restriction properties imposed on the functional form might result in misspecification errors. Hence, any misspecification of the functional form will result in an overestimation or under estimation of the efficiency score measurement (Berger and Mester, 1997; Ramanathan, 2003). Apart from the functional form utilised in the parametric approach, the arbitrary assumption on the distribution probability of the error term could also affect the efficiency scores. Greene (1990) found that different assumptions imposed on the error term can generate different efficiency scores.⁴⁴ Therefore, the nonparametric approach prevents misspecification errors by not placing *a priori* functional form in the measurement of efficiency. However, the nonparametric approach is unable to separate residuals (noise) into random errors and inefficiency components, where all deviation from the frontier is considered as inefficiency (Coelli et al., 2005). With the random error being assumed as managerial inefficiency, any errors in one of the observations that lie on the efficient frontier may influence the structure of the frontier (i.e. alter the measured efficiency of all the observations that are compared to this observation) and consequently result in biased efficiency scores (Coelli et al., 2005). In terms of inefficiency scores, since the nonparametric approach assumes the random error as inefficiency, it is expected that inefficiency for the nonparametric approach will be greater relative to parametric models (Coelli et al., 2005). In the case of the US banking industry, Berger and Mester (1997) found that nonparametric approaches have lower efficiency means compared to the

⁴⁴ Greene (1990) examines 123 US electric utility companies using the stochastic frontier analysis mainly to test the gamma distribution assumption; they compare it against three different types of distribution assumptions, namely the half-normal, exponential and truncated normal distribution. From the results, Greene found that the different assumptions generate different means of inefficiency scores: gamma 0.1051, half-normal 0.1234, truncated normal 0.1039, and exponential 0.0989.

parametric approaches. Nevertheless, Ferrier and Lovell (1990), on the other hand, reports that higher efficiency is exhibited by the DEA (nonparametric approach) compared to the SFA translog model (parametric approach) due to DEA being sufficiently flexible to envelop the data more closely than the SFA translog, which has a second-order approximation flexibility.

As discussed earlier, both approaches have relative advantages and disadvantages and it is difficult to reach agreement on which approach generates better efficiency scores (Bauer et al., 1998). Berger and Humphrey (1997) also state that, there is no consensus as to the preferred method for determining the best-practice frontier against which relative efficiencies are measured. Thus, Bauer et al. (1998) suggest that consensus on which is the best single frontier approach to measure efficiencies should not be sought. Instead, they recommend a set of consistency conditions as a checking device when two or more methodologies are employed.⁴⁵ These consistency conditions are deemed to be useful for both regulators and decision makers. The consistency conditions proposed by Bauer et al. (1998) are as follows:

1. The efficiency scores from different approaches should have comparable means, standard deviations and other distributional properties.
2. The different approaches should rank the firms in the same order based, on their efficiency scores.
3. The different approaches should be able to cluster similar firms in both a best-practice and worst-practice category.
4. The efficiency approaches should remain reasonably stable across different time intervals.
5. Efficiency scores from different approaches should reflect the market's competitive conditions.

⁴⁵ Resti (1997) suggests that between parametric and nonparametric models, the parametric model should be considered as an 'ideal benchmark'. On the other hand, the nonparametric model should be employed as 'empirical observable alternative', because of its high tolerance towards scant data populations and extreme sizes.

6. The measured efficiencies from different approaches should be consistent with non-frontier performance indicators, for example return on assets or a cost to earnings ratio.

It was suggested by Bauer et al. (1998) that, if all six conditions are met for the different methodologies applied, one can be more confident in drawing up regulatory policy conclusions based on the results.⁴⁶

2.7 Conclusions

In this chapter, the relationship between economic theory and the firm's internal efficiency was discussed using various theoretical ideas presented in the literature. The internal efficiency of the firm can be influenced by different market structures. The basic underpinning assumption regarding the market has been that of firms which are operating to maximise profits. In addition, firms are expected to operate efficiently. If not, they will be eliminated from the market. Market theory emphasises that market structures such as monopoly, monopolistic competition, and oligopoly can be imperfect due to market power, particularly in setting their prices above the competitive level. The firms with greater market power may exhibit a lower level of efficiency due to not working as hard to maximise profits. The market structure theory intends to explain how the market works, rather than the internal operation of the firm. Hence, alternative theories such as agency, managerial-discretion, sales revenue maximisation, growth maximisation, the behavioural theory of firm, and X-efficiency are introduced to challenge neoclassical theories. These alternative theories are helpful to this chapter, as it presents the rationale why some firms operate inefficiently and do not seek to maximise profit.

Additionally, this chapter provides the methods for estimating the efficiency of the firm. The relationship between inputs and outputs within the production function is discussed, particularly regarding optimising behaviours, such as minimising cost and maximising profits. The relative (not absolute) estimation of efficiency is discussed. From relative

⁴⁶ Nevertheless, Bauer et al. (1998) generate mixed evidence and are unable to draw a policy conclusion based on the six consistency conditions proposed. For the first three conditions, they found that parametric and nonparametric approaches were not mutually consistent. However, within the group of parametric and nonparametric approaches, the results displayed greater consistency with each other. Nonetheless, a possible consistency is found in the final three conditions.

production efficiency, the input and output orientation of production efficiency, technical efficiency and allocative efficiency were also discussed. The overall (economic) efficiency is a measure of technical efficiency and allocative efficiency; and when both technical efficiency and allocative efficiency are optimised simultaneously, economic efficiency is also maximised, which is consistent with the underlying economic theories which assume that firms act to maximise profits. Additionally, this chapter emphasised different types of efficiency: technical, allocative, scale, and scope efficiencies. This chapter presents various approaches to measuring relative efficiency to estimate the best-practice frontier, namely the nonparametric (e.g. Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH)), and parametric approaches (e.g. Deterministic Frontier Models, Stochastic Frontier Analysis (SFA), Thick Frontier Analysis (TFA), and Distribution Free Approach (DFA)). The parametric approach utilises econometric specifications and requires a functional form and an assumption about the error term to measure economic (i.e. cost and profit) efficiency. On the other hand, the nonparametric approach relies on mathematical linear programming to envelop the actual data, where random error or deviation from the frontier is considered as inefficiency. The discussion continued with commonly used functional forms employed in parametric models, such as the Cobb-Douglas, translog and Fourier-Flexible functional forms.

In the final part of this chapter, a question relating to nonparametric and parametric techniques is presented: “Is there a best frontier method?” This question is valid due to the differences in methodology and the assumptions of the approaches. An answer to the question of which method is the best has been arguably difficult despite the extensive use of both nonparametric and parametric approaches in measuring the efficiency levels of the banking industry, in which Bauer et al. (1998) argue that there is no consensus within banking research findings as to which approach is better in estimating efficiency scores. Thus, the above question is addressed by introducing a number of consistency conditions to mediate the differences between parametric and nonparametric approaches.

The next chapter introduces the importance of frontier analysis in the banking industry. The Frontier analysis is used by these policy makers as a tool to benchmark the performance of an individual bank against the best-practice banks, to identify sources of inefficiency, and to

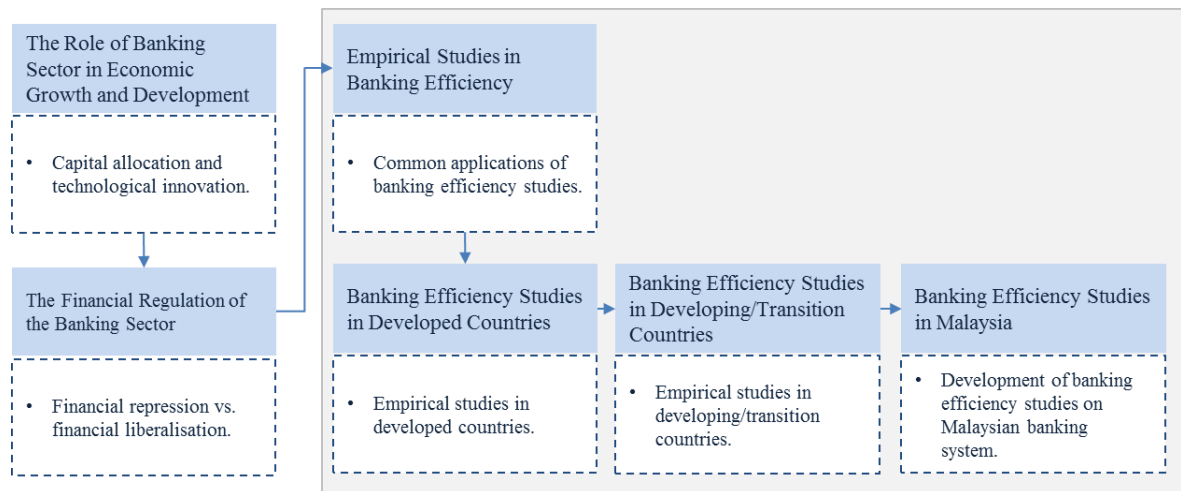
provide feedback on policies implemented in the industry, such as capital regulation, deregulation of interest rates, removal of geographic restrictions on branching, and mergers and acquisitions. The next chapter also discusses and reviews previous banking efficiency studies in developed, developing countries and Malaysia in relations to various contextual issues such as financial liberalisation, ownership, market structure, size and inherent risks.

Chapter 3 Empirical Studies of Banking Efficiency

3.1 Introduction

In the previous chapter, the relevant theories of firms and estimation of frontier efficiency were discussed. Chapter 2 provided brief information on different approaches that are commonly used in estimating efficiency of the firms. The importance of frontier efficiency was also discussed in the previous chapter, particularly the benefits received by various users such as bank managers and policy makers. As policy makers are interested in the performance and conduct of certain industries, this chapter highlights relevant information on the importance of the banking sector's efficiency and related discussions on regulation and financial liberalisation. Consequently, this chapter also reviews a number of empirical studies on banking efficiency in developed countries, developing countries and Malaysia. A number of contextual applications of banking efficiency literature are also discussed; relating to financial liberalisation, ownership, market structure, size and risks.

Section 3.2 introduces the importance of banks for economic development and growth. This section discusses the role of the bank as an intermediary that accumulates capital to be channelled towards productive uses for encouraging economic growth. In Section 3.3, the discussion continues on the relevant theories of financial repression and financial liberalisation and common approaches taken by the regulators in liberalising their banking industry. Section 3.4 discusses empirical studies of banking efficiency in countries with varying economic status (developed and developing countries). Banking efficiency studies in Malaysia are also discussed; for example in the context of the Asian financial crisis and financial liberalisation (e.g. the Financial Sector Master Plan). Based on the above, Figure 3.1 displays the information flow of Chapter 3 (Empirical Studies of Bank Efficiency).

Figure 3.1 Information Flow of Chapter 3

3.2 The Role of the Banking Sector in Economic Growth and Development

A bank is a financial intermediary whose main activity is to provide loans to borrowers and collect deposits from savers. In this case, financial intermediation refers to borrowing by the deficit units (i.e. borrowers) from banks rather than directly from the surplus units (i.e. savers) themselves. That is, financial intermediation is a process that consists of savers depositing funds with the financial intermediary (i.e. the banks) who then lend it to borrowers (Hellwig, 1991; Matthews and Thompson, 2005; Casu et al., 2006). The utility function of borrowers and lenders differs in several ways (e.g. size transformation, maturity transformation and risk transformation) (Casu et al., 2006). For instance, borrowers often require large quantities of funds and on the other hand, lenders normally have smaller amounts of surplus funds; which indicates that the capacity of lenders is generally lower than the size required by borrowers. Thus, the banks can bridge the gap and reconcile the needs and objectives of both borrowers and lenders by performing the ‘size transformation’ function, where the bank pools a number of deposits and parcels them to be lent in larger sums (Arestis and Demetriades, 1997; Matthews and Thompson, 2005; Casu et al., 2006).

In another example, savers may also want to have access to their funds during an emergency, and are wary of being short of liquidity. Since banks normally face a maturity mismatch in their assets and liabilities (i.e. they borrow short and lend long), the bank has to carry a

‘maturity transformation’ function. In overcoming the savers’ concern, banks can offer savers safety and liquidity for their deposited funds; and place risks where they are best yielded, such as in loans and investments (Rajan and Zingales, 2001; Matthews and Thompson, 2005; Casu et al., 2006). In addition, the bank also faces default and pricing risks from funds lent to borrowers. In ensuring that savers’ funds are safe, the bank has to perform a ‘risk transformation’ function; this can be done by diversifying investments, controlling risks, and screening and monitoring borrowers, as well as holding adequate capital and reserves as a buffer against unexpected losses (Beck and Levine, 2004; Matthews and Thompson, 2005; Casu et al., 2006).

Therefore, with a ‘coalition’ of savers and borrowers, the banking institutions help to minimise the costs associated with direct lending, particularly on transaction costs and other costs derived from information asymmetries⁴⁷ through economies of scale or economies of scope in the transaction technology (Freixas and Rochet, 2008). By reducing transaction costs and asymmetric information, banks continue to allocate resources to productive uses within an uncertain environment (Stoughton, 1993; Merton and Bodie, 1995). As mentioned before, the variety of services provided by the banks is in contrast to the direct finance system (Hellwig, 1991), where the role of banks has now expanded and become more complex by performing additional roles such as brokerage services, trade financing and securitisation (Casu et al., 2006).⁴⁸

Consequently, the banking sector is identified as one of the major contributors to a country’s Gross Domestic Product (GDP) (Beck and Levine, 2004; Hassan et al, 2007). The banking sector promotes the productivity and output of a country in several ways; for example, the banking sector increases the savings rate and capital allocation, improves the liquidity of investments, facilitates portfolio diversification, supports the real economy and eases the impact of external shocks (Beck and Levine, 2004). Levine (1997) and Zingales and Rajan

⁴⁷ Asymmetric information exists in a situation where one party in a transaction has superior information compared to another. Asymmetric information normally results in adverse selection and moral hazards. Nevertheless, with increasing advances in technology, asymmetric information is on declining trend, where users can access relevant information using various communication technologies such as the Internet.

⁴⁸ Levine (1997) grouped these various functions of banks into: (1) savings mobilisation, (2) risk management, (3) acquiring information on investment opportunities, (4) monitoring borrowers and corporate control, and (5) facilitating transactions of goods and services.

(2001) assert that, as intermediary institutions, banks promote economic growth through capital accumulation.⁴⁹ The banks accumulate savings to be channelled into productive activities such as lending to deserving entrepreneurs to introduce new products, and promote technology as a catalyst for economic growth (Schumpeter, 1912; McKinnon, 1973). At the same time, banks can benefit from the lower cost of funds from lenders (savers) and transfer this low cost to borrowers (e.g. entrepreneurs and innovators) (McKinnon, 1973; Shaw, 1973).

Savers and bank customers require investment information to make decisions. The information if collected individually entails a higher cost, a longer time, a lengthier process and requires good analytic skills. Hence, the banks can address these problems by collecting and processing the investment information; subsequently, the cost is distributed to/among customers (Millon and Thakor, 1985; Stoughton, 1993). As an intermediary, the information collected by the banks can address the issues of information asymmetries such as adverse selection and moral hazard.⁵⁰ The information held by the banks can also be utilised to determine lending and investment decisions. Consequently, banks will monitor the performance of the projects as well as exercising corporate control (Bencivenga and Smith, 1991). Banks can also perform portfolio diversification from their scale of operation; in which investors/savers are supplied with a wider range of returns that meet their risk appetite. Moreover, by diversifying the loan and investment portfolio, the bank can make the cost of delegation and monitoring as small as possible and offer close to risk-free deposits (Diamond, 1984; Freixas and Rochet, 2008). Hence the existence of banks could satisfy the marginal utility of savers and stimulate real investment in a country (Arestis and Demetriades, 1997). Another key activity of the bank is to facilitate economic transactions. Banks provide faster and cheaper mechanisms for receiving and making payments, such as

⁴⁹ This is consistent with Levine (1997), who states that the banking sector affects economic growth in two ways, namely through capital allocation and through technological innovation. The banks will pool all the savings for capital accumulation and reallocate savings into different capital-producing technologies. Thus, allocating capital to innovations can result in a competitive price of supply, which in return will give rise to greater competition in the market structure (Merton and Bodie, 2004).

⁵⁰ Adverse selection refers to a market process of producing an undesirable result due to imperfect information between lenders and borrowers. In addition, moral hazard refers to the risk that a party to a transaction has not entered into the contract in a good faith. For example, a person may provide misleading information about its assets, liabilities or credit capacity to secure a loan or advances from the bank (Boyd et al., 2002).

payment systems, remittance services and automated teller machines, to promote productivity gains, technological innovation and economic growth (Levine, 1997).

The discussion above summarises the significant roles of the banking system in promoting economic growth. In general, the banking system could lead to a more efficient utilisation of funds and improve the availability of funds to higher-risk and innovative ventures in order to promote economic growth and development in a country (Casu et al., 2006). Apart from mobilising funds, banks are imperative to economic growth by executing monetary policy, assessing markets' emerging risks and providing payment services. The importance of banking institutions for economic growth and stability suggests that banking institutions should be regulated (McKinnon, 1973; Shaw, 1973). Nevertheless, past literature shows that heavy regulation of banking institutions may impair economic growth. Thus, suitable and effective liberalisation initiatives are needed to address the financial repression of the banking sector. Moreover, financial liberalisation can also result in the higher efficiency of banks (Girardone et al., 2004), which, subsequently can result in greater economic growth (an increase in GDP) (Hassan et al, 2007; Ferreira, 2012). Therefore, the next section will briefly discuss the impact of government intervention on the banking sector.

3.3 Government Intervention in the Banking Sector

As mentioned earlier, the financial intermediation of banks provides capital accumulation and the allocation of funds to promote economic growth. The banking sector influences the social, economic and political landscape of a country in which the government and the private sector have substantial roles to play (McKinnon, 1973). McKinnon (1973) argues that the government should provide adequate regulation of financial services to support the private sector's operations. The government should also conduct constant reviews of current regulation and deregulation in order to achieve the ideal regulation of the market. Deregulation or financial liberalisation does not mean all legal restrictions should be removed; rather it should result in a comprehensive and stable set of policies and regulations that promote market efficiency (Shaw, 1973).

According to Bhattacharya et al. (1997), theories of modern banking focus on contractual obligations of banks as intermediary institutions, relating to areas such as monitoring lending, resolving market imperfections, liability contracts with depositors, and problems in coordinating contracts (e.g. illiquidity or bank runs). This indicates that the relationship between external parties and banks is important for banking sector stability and could influence the performance of the banks. Unregulated banking is prone to financial instability and can lead to situations such as depositor panics, domino-like collapse, and asset-price spirals (Schooner and Taylor, 2010). Therefore, Benston and Kaufman (1996) assert that government-imposed regulations are required due to the existence of potential market failures. Moreover, if the banking sector fails, it could result in a detrimental impact on the real economy (Levine, 1991). Hence, government intervention through regulation can help to ensure that the market works efficiently and/or to adjust market outcomes to achieve social objectives (Schooner and Taylor, 2010). Schooner and Taylor (2010) and Gart (1994) further suggest that there are four market failures that generally form the basis for regulatory intervention: (1) the existence of a monopoly, (2) the existence of externalities, (3) the existence of information asymmetries, and (4) protecting public goods. These market failures are discussed as follows:

1. The existence of a monopoly

The existence of a monopoly might promote the market power of one or a few banks, which restricts market competition. These banks are likely to raise prices, restrict supply, offer poorer services and restrict innovation (Gart, 1994).

2. The existence of externalities

The existence of externalities can arise in a situation where the impact of producing goods or services can result in costs or benefits to a third party that are not reflected in the price. In other words, the costs/benefits are the consequence of economic activity which is subsequently experienced by a third party. Externalities can either be positive or negative. For example, positive externalities include the production of a commonly accepted medium of exchange and storing of value in that exchange. Negative externalities, on the other hand, cover activities such as action to prevent bank runs, to

avoid economic distress from bank failure, and to reduce costs for prudent banks (Khemani, 1990).

3. The existence of information asymmetries

Information asymmetries occur when it is assumed that in a well-functioning market, buyers and sellers have all the information needed to evaluate competing products and services, and consequently make well-informed decisions. But, in reality, due to bounded rationality, information is limited and can be priced, just like any other commodity. For example, a party involved in a transaction may deliberately seek to mislead another party by giving false information or failing to disclose key facts (Boyd et al., 2002).

4. Protecting Public Goods

Public goods are described as products that one individual can consume without reducing their availability to another individual. Thus, public goods are referred to as 'non-rivalrous' and 'non-excludable'. Due to the nature of non-rivalry and non-excludability, it is difficult for producers to charge for public goods. In the context of banking, financial stability is seen as a public good and consequently adequate regulations and a safety net (e.g. deposit insurance) are required (Gart, 1994).

In many developing countries, banking institutions face the prudential regulations and restrictions of regulators who are mandated to promote stability, resilience and a sound financial system. The soundness of the banks can influence the financial landscape and infrastructure as well as macroeconomic performance and monetary policy in many developing countries. Thus, regulators are paying increasing attention to monitoring the health and efficiency of banking institutions and financial markets, and macroeconomic developments that may threaten a country's financial stability (Bertus et al., 2007). Many developing countries rely heavily on prudential regulation to control banks' excessive risk taking, and mitigate moral hazard and bank failure. Such failure is a costly event, which increases contagion risk⁵¹ and can collapse an economic system (Pettway, 1980). In view of

⁵¹ Contagion risk in this context refers to the financial difficulties faced by a bank that spill over to a larger number of banks in the financial system or even to the financial system as a whole.

this significant effect, the government intervenes in banks' operations and imposes 'financial repression'⁵² through banking regulations, on areas relating to the minimum capital requirement, the establishment of new banks, an interest rate ceiling/floor, lending direction, standards of risk management, directed lending, the qualifications of directors and managers, and organisational governance (McKinnon, 1973; Shaw, 1973; Fry, 1995; Kroszner and Strahan, 1999; Bertus et al., 2007). These regulations and restrictions are implemented to protect banking consumers from the consequences of potential bank failures. Apart from mitigating bank failures, regulations are employed to prevent monopoly power, which can restrict competition and could result in excessive pricing (Schooner and Taylor, 2010).⁵³ Additionally, regulations imposed on banking institutions can also minimise the potential economic turmoil from economic shocks and financial crises (Caminal and Matutes, 2002).

Excessive financial repression or regulation can result in inefficient capital allocation, high cost of financial intermediation and lower rates of returns to savers, which can impede economic growth (Roubini and Sala-i-Martin, 1992). Heavy governmental intervention in banking activities is also seen as one of the factors that influences lower efficiency and poor performance in banks (Kumbhakar and Sarkar, 2003). Therefore, financial liberalisation is introduced to free certain market activities from financial repression (McKinnon, 1973; Shaw, 1973).

Financial liberalisation refers to the government's initiatives to reduce regulatory restrictions in the financial sector and thus increase the participation of private entities. The primary reason behind financial liberalisation is to improve market competition, increase the

⁵² The concept of financial repression can be inferred from McKinnon (1973), Shaw (1973) and Fry (1995). Financial repression normally refers to controls on interest rates, directed lending to government and restrictions on entry into financial industry. Financial repression is deemed as preventing proper capital allocation and impairing economic growth.

⁵³ However, Stigler (1971) and Barth et al. (2006) state that banking regulation can also be seen as 'restricting' as well as 'promoting'. For example, the government can implement a set of policies to promote competition in the banking system and concurrently, impose some restrictions preventing large banks from manipulating market power to drive smaller banks from the market. In addition, the government can also restrict larger banks from exercising excessive pricing, which consequently protects the basic needs of the consumer and the public. Therefore, the government needs to perform a balancing act between 'promotions' and 'restriction' to ensure financial stability, soundness and steady economic growth in banking sector (Caminal and Matutes, 2002).

efficiency of banks (i.e. effective resource allocation and output expansion) and improve technology and innovation (i.e. Internet banking, payment systems, securitisation and financial derivatives) (Hamilton, 1989; Lozano-Vivas, 1997; Kumbhakar and Sarkar, 2003). Additionally, financial liberalisation is also required for banks to cope with the emergence of new challenges in the banking system, such as cross-border banking, a universal banking environment, diversified banking activities via fee-based services, competition with financial markets, Internet banking, and innovative financial products (e.g. derivative products) (Schooner and Taylor, 2010).

Financial liberalisation requires several preconditions to be successful (Fry, 1995). These preconditions are:

1. Adequate prudential regulation and supervision to support the financial market structure;
2. A stable monetary policy (a reasonable level of price stability);
3. Fiscal stability (i.e. avoid excessive government borrowings that can increase inflationary expansion and foreign capital inflow);
4. Highly competitive behaviour in banks, and;
5. A fair tax system on the banking industry.

If these conditions are not met, financial liberalisation can result in excessive risk taking by banks, high real interest rates, bankruptcies of firms and bank failures (Arestis et al., 2003). Apart from the preconditions mentioned above, Dunham and Kelegama (1994) suggest that appropriate 'sequencing' processes are also required. They suggested three key aspects of financial liberalisation (Dunham and Kelegama, 1994): (1) the speed of implementation, (2) the stages involved, and (3) the order of liberalisation initiatives.⁵⁴ In addition, Abiad and

⁵⁴ First, In terms of implementation speed, the government has to consider either implementing liberalisation on a gradual basis or as a 'one-shot' programme. Second, the speed of financial liberalisation depends on the political consequences, microeconomic situations, income distributions and protection of the banking industry. Second, Financial liberalisation involves several key stages, for example, the removal of protection requires robust banking regulation and the strengthening of weaker banks. Third, Financial liberalisation requires an appropriate order of liberalisation of varying markets such as commodities, labour and financial markets, to attain a fully liberalised financial industry. A proper 'sequencing' process can support market order for a

Moody (2005) suggest six policy dimensions for implementing financial liberalisation on the banking system, namely (1) credit controls, (2) interest rate controls, (3) entry barriers, (4) operational restrictions, (5) privatisation and (6) international financial transactions;⁵⁵ and based on these six broad dimensions, Fry (1995) suggests that several specific liberalisation initiatives can be implemented by the government. For example, the initiatives include but are not limited to removing the interest rate ceiling on deposits and loans, removing barriers to foreign exchange transactions, allowing foreign banks' entry, promoting branch expansion and removing government-directed lending policy.

Financial liberalisation can result in increased market competition and improved operational practice and efficiency of the banks; Molyneux et al. (1996) state that the general public will benefit from increased competition in the banking system. Greater market competition can result in lower interest costs and a wider choice of banking providers as well as a broader range of banking products and services for the general public to enjoy. Common methods of achieving increased competition in the banking system via financial liberalisation are: (1) the deregulation of interest rates and (2) the removal of entry barriers to foreign banks.

First, McKinnon (1973) and Shaw (1973) argue that a banking system that is undergoing financial repression from government-controlled interest rates can discourage savings and investments. Similar to product prices, the interest rate performs a rationing function by allocating a limited supply of credit to borrowers who demand and compete for it. Recognising the importance of banks for intermediation and monetary policy transmission functions, McKinnon (1973) and Shaw (1973) posit that financial repression can distort actual interest rate equilibrium, and may not reflect the demand and supply of funds of the market players. They further argue that, if the government-controlled interest rate ceiling results in low or negative real interest rates, savings in banking institutions will also decline

smooth transition to a more liberalised market. In other words, liberalisation initiatives should be managed rather than adopting a 'laissez-faire' or free market approach (Arestis, 2005).

⁵⁵ For example, first, the policy relating to credit controls refers to lending directed by the government towards several sectors and high capital adequacy requirements. Second, interest rate controls refer to government controls on the interest rate's ceilings, floors or bands. Third, entry barriers are described through licensing requirements and limiting the entry of foreign banks. Fourth, operational restrictions can be from the perspectives of staffing, branching and advertising. Fifth, privatisation refers to specific privatisation of government-owned financial institutions, and sixth, international transactions refer to capital and current account inflows (Abiad and Moody, 2005).

(as savers and investors will move to alternative products: for example, the stock market, property or direct lending), which will subsequently lead to fewer loans being mobilised through the banking system (Gart, 1994). Thus, by addressing this issue the government can therefore encourage savings and investments and increase market competition by removing the government-controlled interest rate and delegating power to the market players (McKinnon, 1973; Shaw, 1973; Fry, 1995). For example, McKinnon (1973) and Shaw (1973) examined the effect of deregulating the interest rate on individual financial systems in developing countries and reported that interest rates in these countries shifted towards a more competitive market equilibrium (this positive effect exists mainly because the power of setting interest rates is given to the market players). They believed that deregulation of interest rates is required for a country's economic growth and development, in which the banking system would become more competitive and inefficiency could be removed from the banks (Berger and Mester, 2003). Moreover, by removing the government-controlled interest rate ceiling, the government could encourage savings, improve monetary policy transmissions, enhance product innovativeness, and improve the allocation of resources in the banking system, subsequently supporting the real growth of the national economy (McKinnon, 1973; Shaw, 1973). However, the drawback of removing the interest rate ceiling is that it can also result in unhealthy pricing competition, inefficiency in banks, asset and liability mismatches, and the possibility of innovative products that are not easily understood (e.g. asset-backed securities and derivatives products) (Demirguc-Kunt and Detragiache, 1999; Kraft and Tirtiroglu, 1998).

Second, financial liberalisation can be made by removing entry barriers against foreign banks entering the domestic market. Unite and Sullivan (2003) suggested that the removal of restrictions on the entry of foreign parties could be made through (1) allowing foreign-controlled banks to enter the domestic banking system and (2) promoting and allowing foreign ownership of equity in domestic banks. The removal of restrictions on foreign banks and foreign ownership in domestic banks could provide a positive impact on the banking industry. First, it could reduce the monopoly of certain domestic banks. Second, domestic banks could improve their cost management and allocate resources efficiently with the introduction of competitive pressure from foreign banks (De Young and Nolle, 1996). And third, foreign banks, particularly from developed countries, would bring in and introduce

advanced technologies, the latest operational methods, financial innovations and new managerial practices. Eventually, this new knowledge would spill over into the domestic market, either through competitive observations or the mobilisation of human capital within the banking system (Levine, 1996; Goldberg, 2003). De Young and Nolle (1996) suggest that, with the removal of entry barriers to foreign banks, domestic banks have to strive harder, realising that they are no longer protected by local regulators, which results in an increasingly efficient banking system. Despite many positive outcomes, there are also several drawbacks that can impact the domestic market with the introduction of foreign competitors: (1) a lower franchise value in domestic banks; (2) imbalanced concentration of lending sectors by foreign and domestic banks, (3) increased potential instability of the domestic banks; and (4) different objectives of foreign banks at odds with domestic market priorities (Hellmann et al., 2000; Bayraktar and Wang, 2004).⁵⁶

An increasingly competitive market structure stemming from financial liberalisation may result in the higher efficiency of firms, reduced managerial slack and improved productivity (see Chapter 2). For example, Gart (1994) states that the positive impacts of financial liberalisation on banks has induce management to reduce their operating expenses by making improvements in computer technology and has led management to replace labour with automation. With this automation, banking institutions focus more on portfolio management and shift from ‘relationship’ banking to ‘risk-pricing’ banking, which can improve the time management of banks. Additionally, financial liberalisation also forces the banks to seek new ways to increase revenues; managers have to expand and diversify business into non-asset-based fee income and may also charge for services that used to be free (Gart, 1994, Goldberg, 2003). Nevertheless, financial liberalisation can also lead to an

⁵⁶ First, the disadvantages of opening the market to foreign competitors might include a lower franchise value in domestic banks, in which domestic banks may have the incentive and appetite to take on higher risks to compete with foreign banks. For example, large foreign international banks with greater reputations and stronger network linkages would generally have better access to cheaper funds (from retail deposits, the interbank market or the global money market) compared to domestic banks. On the other hand, domestic banks would have to incur higher costs for funds compared to foreign banks, which could result in a lower franchise value. Second, armed with more advanced products and services, foreign banks may practise ‘cherry picking’, by selecting the most profitable segments of the market and leaving behind the riskier sectors to be served by the domestic banks. Third, the domestic banking market will also face potentially greater instability from foreign banks pulling out, particularly when foreign banks face economic turmoil either in the host or in their home country. Fourth, the objectives of the domestic banking market may not be achieved due to differences in priorities and the business focus of foreign banks, in which their lending pattern tends to ignore domestic priorities (Hellmann et al., 2000; Bayraktar and Wang, 2004).

adverse impact on banking institutions' profitability. For instance, the average cost of bank retail deposits may increase after the deregulation of the interest rate due to industry competition for cheaper funds, which in turn can narrow their profit margins and slow growth (Gart, 1994, Hellmann, Murdock and Stiglitz, 2000). Consequently, with greater freedom in the market and the aim of maximising profits, banks are more likely to invest in higher-risk products (e.g. derivative products and high yield notes) (Kraft and Tirtiroglu, 1998).

3.4 Empirical Studies of Banking Efficiency

As mentioned in chapter 2, neoclassical theory assumes that all firms have the objective of maximising profit. However, several alternative theories and hypotheses of the firm (e.g. agency theory, managerial theory, behavioural theory and the relative market power hypothesis) argue with this view, as profit maximisation is not reflected in the behaviour of firms. These alternative theories provide a background for empirical analysis of potential sources of inefficiency, particularly internal bank-specific characteristics (e.g. ownership, risk and bank type) and external environmental factors (e.g. market structure, financial liberalisation and economic growth). These explanatory factors also become the foundation for investigating various contextual topics, which inform and provide feedback on government policies. This can be evidenced from Berger and Humphrey (1997), who performed a survey of 130 frontier analysis studies in 21 countries. The motivation of their study was to find the common themes and issues relating to banking efficiency studies, one of which is to inform government policy. Consequently, in terms of economic status, the banking efficiency literature is dominated by studies from developed countries such as the US and Western Europe. On the other hand, studies of banking efficiency from developing and transitional countries, including Malaysia, are limited. Nevertheless, the increasing interest in developing countries is attributable to various changes occurring in the banking industry over the last two decades, such as the globalisation of financial markets, financial liberalisation, market reform, privatisation, financial crisis and technological advancement. Hence, banking efficiency studies in developing countries may show different results compared to developed countries, providing greater insight into the banking systems of developing countries. In view of the dissimilarities in the results from countries of different

economic status, the following sections review empirical banking efficiency studies in: developed countries, developing countries and Malaysia. These studies were applied in various contexts; for example deregulation, ownership, market structure, size and inherent risk.

3.4.1 Empirical Studies of Banking Efficiency in Developed Countries

Earlier studies of banking efficiency concentrated on US banks. However, most of these studies focused on the scale and scope of banks' efficiency, rather than on the efficiency frontier estimations. These methods were questioned. Berger and Humphrey (1991) asserted that the scale and scope of efficiency is a secondary concern. They further suggested that the efficiency frontier should be employed primarily to reflect the banks' cost and profit-efficiencies.⁵⁷ As a result, since the early 1990s frontier analysis has increasingly been employed to measure the efficiency of banks (see Ferrier and Lovell, 1990; Berger, 1991, 1992; Berger et al., 1993). Frontier analysis is also extensively used to explain the progression of the economy, growth in the banking sector, policy implementations, competition in the banking market, as well as the specific characteristics of banks. Therefore, a number of common themes/areas associated with government policy in developed countries are reviewed in this study, such as (1) financial liberalisation/deregulation, (2) ownership, (3) market structure, (4) bank size, and (5) inherent risk.⁵⁸

First, earlier studies of financial liberalisation and the performance of banks were applied using US banking data, which recorded several deregulation events that have taken place since the early 1980s. The deregulation initiatives in the US include the lifting of the interest rate ceiling (i.e. Marquette vs. First of Omaha, Depository Institutions Deregulation and

⁵⁷ This is evidenced by Berger et al. (1993), who surveyed the results of previous studies on scale economies of the banking industry, and found that the average cost curve has a relatively flat U-shape. Previous studies surveyed by Berger et al. (1993) are more interested in investigating the presence of scale economies rather than economic efficiency (e.g. cost/profit efficiency), which assumes that all banks are equally efficient. This assumption however does not hold, due to differences in the cost structure among banks (Berger and Humphrey, 1991).

⁵⁸ Earlier, the discussion on market competition explained the need for financial liberalisation, which can give impact to the banks' risk-taking behaviour, the performance of domestic and foreign banks and market concentration by big banks. In the context of this research, explanatory factors (e.g. financial liberalisation, market concentration, ownership, bank type, size, risks and time) will be employed as control and environmental variables; these variables will be discussed in detail in chapter 5 ('Methodology and Data').

Monetary Control Act of 1980), allowing interstate branching (i.e. Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994) and allowing thrifts (term used in US for savings banks) to conduct commercial lending and compete in money market mutual funds (i.e. Garn-St. Germain Depository Institutions Act of 1982) (Humphrey, 1993). Hence, studies of the relationship between deregulation and bank efficiency in the US that employ frontier analysis were on an increasing trend in the 1990s (e.g. Berger and Humphrey, 1991, 1992; Humphrey and Pulley, 1997; and Berger and Mester, 2003). For example, Berger and Humphrey (1991) employed a TFA model on 13,951 insured commercial banks in the US using one year's cross-sectional data (1984) to measure and analyse the inefficiencies of all US banks.⁵⁹ They found that the inefficiencies of 19.0% occurred in the state banks were operational rather than financial. Based on the decomposition made on the cost components, the technical inefficiency derives mainly from the overuse of labour and capital (Berger and Humphrey, 1992; Humphrey and Pulley, 1997). They further suggested that the impact of deregulation increases market competition and pressures banks to cut costs substantially in a very short period of time.

The immediate impact of deregulation on banks in a short period of time was also evidenced in Berger and Humphrey (1992), who use a parametric TFA model to examine the general level of efficiency of large US banks reported in the Federal Reserve's Functional Cost Analysis Program for the years 1980, 1984 and 1988. In this study, 1980 is regarded as the pre-deregulation period, 1984 as the mid-deregulation period and 1988 as the post-deregulation period, by which time adjustments to deregulation may or may not have been completed. During this period (1980–1988), US banks underwent substantial deregulation, which included the lifting of the interest rate ceiling on certain deposits. Banks therefore entered into a phase of significant disequilibrium, where they attempted to adjust their interest rates pricing according to market changes and increased competition. The efficiency for state banks in 1980, 1984 and 1988 were 74.7%, 80.8% and 73.8%, respectively. The efficiency of banks showed some improvement in 1984 but worsened in 1988, indicating

⁵⁹ In this study, Berger and Humphrey (1991) introduced TFA to examine the impact of deregulation on the efficiency of US banks. They argued that TFA has a benefit over SFA due to less specificity in the maintained assumptions, which are less likely to be violated by the data. However, TFA may not provide accurate estimates of the general level of overall cost efficiency, which makes it unlikely to be used by management and policy makers (Kumbhakar and Lovell, 2000).

that the adjustment process to the new regulated interest rate equilibrium have yet been completed. Berger and Humphrey (1992) also found that deregulation in the 1980s had little impact on technical change or productivity growth due to the slow response by banks in minimising costs in response to the deregulation initiatives.

This suggests that the banks needed time to adjust to deregulation. For example, Humphrey and Pulley (1997) utilised data from 683 US banks between 1977 and 1988 using a parametric TFA model to analyse the impact of deregulation on the profitability of US banks. They found that banks adjusted both their outputs and inputs by modifying the prices of loans and deposits (i.e. the price of outputs) as well as input utilisation of labour and capital to minimise the negative impact of deregulation, such as increased funding costs. US Banks responded to the deregulation of interest rates in the early 1980s by increasing the interest rate of core deposits. This reaction by US banks was made to increase the level of core deposits, rather than relying on purchased funds as their sources of funds. In order to mitigate the higher cost in core deposits, US banks increased fees for deposit services, shifting the asset mix towards a floating ratio, and taking greater asset risk in search for higher revenues. Based on their observations, banks responded to deregulation in three main stages: cost offsetting and reduction, cost shifting, and revenue augmentation.⁶⁰ Their findings suggest that the profit efficiency of banks bore a minimal result in the early phase of deregulation between 1977 and 1981 at 12.0% and 1.0% for large and small banks, respectively. Consequently, profit efficiency improved after 1981 to 45.0% and 39.0% for large and small banks respectively. They concluded that adjustment by banks following deregulation can take up to four years to complete (Humphrey and Pulley, 1997).

As mentioned above, increased profit efficiency can be achieved in a few years after the introduction of deregulation. Nevertheless, an increase in profit efficiency may not necessarily improve cost efficiency. For example, Berger and Mester (2003) examined the performance of US banks in terms of profit and cost efficiency for the years 1984, 1991 and 1997, which are the years of deregulation of interest rates. Berger and Mester (2003)

⁶⁰ For example, cost offsetting and reduction refers to the activities of banks in charging more for previously low-cost or free services to offset higher deposit interest rates and reducing costs such as branch operating costs. Consequently, cost shifting refers to banks' approach in shifting higher funding costs and interest rates to borrowers (via floating rate loans). Finally, revenue augmentation refers to expanding assets portfolios with riskier assets that have an expectation of higher returns (Humphrey and Pulley, 1997).

employed three different economic functions namely, cost, standard profit and alternative profit on the TFA model. The standard profit function considers revenue and cost effects in both outputs and inputs, which reflects the operation of the banks. Profit maximisation can also describe the economic objectives of managers, who take into account both revenues and costs (Berger et al., 1993). The standard profit function uses input and output prices for the estimation of profit efficiency. On the other hand, alternative profit efficiency is computed similarly to profit efficiency except that the output is taken as a total amount rather than output prices. The concept of alternative profit is helpful when some of the underlying assumptions of cost and standard profit functions are not met (e.g. for firms operating within an imperfect competitive market structure with inaccurate output prices) (Berger and Mester, 1997).⁶¹ Based on the TFA model using the Fourier Flexible specification, they report that between 1991 and 1997, US banks' cost productivity worsened, but profit efficiency increased significantly. The cost inefficiency annual rate fell by 0.3% between 1984 and 1991 but increased by 2.7% between 1991 and 1997. On the other hand, standard and alternative profit efficiency registered an increase of annual growth by 13.7% and 16.5% respectively. To explain this perplexing finding, Berger and Mester (2003) argued that, in a more liberalised environment, banks have been more innovative and have offered a variety of products and services to customers, intending to maximise banks' revenue. By doing this, banks are also affected by higher costs, but concurrently increase revenue in higher proportion. Additionally, Berger and Mester (2003) also include control and environmental variables⁶² in a one-stage approach, and investigate explanatory factors such as bank size, inherent risks, market power and mergers. This was performed to include the effect of heterogeneity of banks. They found first that larger banks demonstrate higher revenues but also suffer from higher operational costs. Second, inherent risks are reduced when the performance of profit and cost efficiency is increased, due to shifts in the asset mix during the economic boom during the mid-1990s. The shifting of asset mixes from securities to loans and other off-balance sheet products has resulted in a greater proportion

⁶¹ These three economic efficiency functions are discussed in detail in Chapter 5.

⁶² There are two types of explanatory factors that affect inefficiency estimation, namely control variables and environmental variables. Control variables are introduced into the specification of a stochastic model to test for potential biased results on the output measures of the estimation measurement. On the other hand, environmental variables are imputed into the model specification to test the sources of efficiency (Battese and Coelli, 1995; Puig-Junoy and Ortun, 2002, Coelli et al., 2005). Failure to control for control and environmental variables may result in biased estimates of production frontier and inefficiency scores (Sherlund et al., 2002, Bos et al., 2009) (This is discussed in detail in chapter 5).

of portfolio returns, which in turn increases the profit- and cost-efficiencies of banks. Third, market power has little effect on the performance of banks. This also supports the conclusion that the market power from the consolidation exercise (i.e. mergers and acquisitions) may not lead to higher efficiency (Berger and Mester, 2003).

In summary, studies of US banking efficiency show that banks can benefit from banking deregulation (e.g. through greater competition, a flexible cost structure, increased profits and higher productivity and efficiency). Nevertheless, Berger and Humphrey (1997) state that banks require time to adjust their cost and profit structures in the early phase of the deregulation. At this phase, banks usually have lower efficiency but subsequently improve in later stages as market competition increases stemming from financial liberalisation.

In another developed country, Spain, Grifell-Tatje and Lovell (1996) found that deregulation has little impact on the cost efficiency of the bank. The cost efficiency of Spanish banks did not improve as a result of deregulation. Furthermore, Vivas (1997) investigates the impact of deregulation on the profit efficiency of Spanish savings banks from 1986 to 1991. The results from the TFA model showed no significant reduction in the profit inefficiency of Spanish banks, which were recorded at 46.0% to 42.0% in 1986 and 1991, respectively. Vivas (1997) also asserted that Spanish savings banks are favourable to forgoing their short-term profit for a further expansion in market share.⁶³ In Australia, Avkiran (2000) examines the changes in productivity in the retail banking sector from a sample of ten Australian banks during the deregulation period 1986–1995. He employed the Malmquist productivity index⁶⁴ and found that the total productivity rise was on average 3.2% per year. During the deregulation period, the decomposition of total productivity change is driven more by technological progress than technical efficiency. This can be explained by the greater competition faced by Australian banks from building societies, credit unions and mortgages originators during the deregulation period. In view of the greater competition,

⁶³ Vivas (1997) stated that the impact of deregulation on the banking industry could result in a greater level of inherent risk such as a higher potential of loan loss provision, when banks are eager to capture a higher market share from liberalisation initiatives.

⁶⁴ Malmquist (1953) introduced a quantity index that refers to the amount by which one consumption bundle must be radially scaled in order to generate the same utility level provided by some base consumption bundle. The Malmquist quantity index has been applied to the measurement of productivity change, and is normally employed using nonparametric linear programming techniques (Grifell-Tatje and Lovell, 1996).

Australian banks responded to increased competition through such practices as securitisation, product innovation, increase fee-based income and new delivery channels. For instance, a re-engineering of retail delivery channels was aimed at improving cost efficiency and customer services. These self-service technologies have become tools in minimising service delivery costs and were particularly promoted in rural areas, which reduce banks' costs (e.g. costs reduced from not setting up branches in rural areas). Avkiran (2000) further suggests that government and regulators can provide incentives to banks, for example through tax incentives to encourage banks to innovate and at the same time improve the degree of efficiency of the banks.

Girardone et al. (2004) employed data from 1,958 Italian banks from 1993 to 1996 to examine the effect of heterogeneity in Italian banks as well as to identify the characteristics of efficient banks. Using the SFA model with a two-stage analysis,⁶⁵ they found that the inefficiency of Italian banks ranges between 13.0% and 15.0% of their total costs and the inefficiency level reduces over time, indicating that deregulation initiatives introduced in 1992 by the European Union had produced a positive impact in improving the overall cost of the Italian banking industry. Additionally, the impact of deregulation by the European Union, which aimed to increase competition by removing the entry barriers to European banks, is also analysed by Casu and Girardone (2006). They investigated the impact of competition on bank efficiency. The result from Panzar and Rosse model⁶⁶ (*H*-stats) is regressed against efficiency estimation and other explanatory variables. Casu and Girardone (2006) found that an efficient banking system might not necessarily result in higher competition. They argued that the relationship between competition and efficiency is not straightforward as increased competition may force banks to be more efficient but higher efficiency may not result in higher competition in the EU banking system.

⁶⁵ In the two-stage approach, the first stage estimates the scores of the frontier efficiency using traditional inputs and outputs. Consequently, the second stage involves the regression of scores from the first stage with the explanatory variables (Cebenoyan et al., 1993; Berger and Hannan, 1997; Bonin et al., 2005). Additionally, the one-stage approach includes explanatory variables into the estimation of frontier efficiency (Battese and Coelli, 1995; Coelli et al., 2005) (This is discussed in detail in chapter 5).

⁶⁶ Panzar and Rosse (1987) developed a test statistics to distinguish competitive market structure (i.e. perfect competition, monopoly, monopolistic competition and oligopoly). The *H*-statistic test is the sum of elasticities of total revenue with respect to input prices. An *H*-statistic result can be interpreted as follows: if *H* equals 0 or is negative it implies oligopoly or monopoly, or if *H* equals 1, indicates perfect competition, and between 0 and 1 ($0 < H < 1$) implies the monopolistic competition.

Therefore, studies in deregulation in the developed countries provide mixed findings. While some show a negative impact, there are also many studies that found a positive impact on the banking system where bank efficiency improves following the introduction of deregulation initiatives (see Girardone et al., 2004; Avkiran, 2000). In addition, deregulation in developed countries, for example in the EU, has improved the competition structure in the banking system, which in turn has increased the efficiency of the banks (see Casu and Girardone, 2006). Nonetheless, deregulation may also cause greater risks to the productivity of banks and unhealthy market competition; Vivas (1997) reports that banks forgo a certain level of profit in order to gain a bigger market share. This has resulted in a higher level of inherent risks, which can lower the profit efficiency of the banks.

Second, empirical studies regarding ownership (e.g. foreign versus domestic banks) and the efficiency level of banks are also discussed in developed countries. The introduction of foreign banks is expected to increase competitive pressure on domestic banks. Claessens et al. (2001) found that foreign banks are superior to domestic banks in developing countries, but the opposite is true in developed countries. Both foreign and domestic banks have their own competitive advantages when operating in the banking industry. These circumstances can be explained by two alternative hypotheses proposed by Berger et al. (2000), namely the ‘global advantage hypothesis’⁶⁷ and ‘home field advantage hypothesis’. The ‘global advantage hypothesis’ refers to the comparative advantages of foreign banks over domestic banks, such as superior managerial competency, the lower cost of capital and operations, greater risk management skills, advanced technology, enhanced international knowledge, a wider range of services, and superior investment expertise. On the other hand, the ‘home advantage hypothesis’ refers to the comparative advantages of domestic banks over foreign banks in their home country, such as higher-scale economies and greater business

⁶⁷ Berger et al. (2000) proposed two types of global advantage hypothesis, namely the general and the limited form. These forms are based on the countries of origin of foreign banks. The general form refers to the comparative advantage held by foreign banks regardless of their country of origin. On the other hand, the limited form suggests that only certain foreign banks from a particular set of countries can outperform domestic banks.

knowledge (e.g. domestic banks are conversant with local regulations, culture and language).⁶⁸

The results of previous empirical studies concerning the relationship between foreign and domestic banks and bank efficiency present mixed findings. The pattern of previous studies' findings however, shows that most foreign banks display higher efficiency than the domestic banks in the developing countries, whereas, in developed countries (e.g. the US and Europe), domestic banks normally exhibit higher efficiency than foreign banks (see DeYoung and Nolle, 1996; Hasan and Hunter, 1996; Mahajan et al., 1996; Sathye, 2001). For example, Berger and De Young (2001) employed a parametric DFA model on 7,000 US banks, and reported that foreign banks in the US are less efficient than domestic banks, suggesting that international barriers such as language, culture, currency and regulatory structures exist within the operations of foreign banks, supporting the 'home advantage hypothesis'. Additionally, De Young and Nolle (1996) studied the US market from 1985 to 1990, utilising 62 foreign banks using the DFA model. They found that in the US banking industry, US-owned banks are more efficient than foreign-owned banks. The profit efficiency gap between US banks and foreign banks varies between 3.0% and 10.0%.

Similarly, Berger et al. (2000) examined the efficiency levels of foreign and domestic banks in five countries, namely Germany, Spain, France, the UK and the US. They found foreign banks display lower cost and profit efficiency than domestic banks in Germany, France and the UK. However, in Spain and the US, either cost or profit efficiency is greater in domestic banks. In Spain, domestic banks display higher cost efficiency than foreign banks by 2.1% but lower profit efficiency by 5.4%. On the other hand, in the US, domestic banks are more profit-efficient than foreign banks by 25.1%, but less cost-efficient by 2.9%. This implies that in Spain, the Spanish banks have a slight cost advantage over their foreign rivals; but

⁶⁸ There are two perspectives of what makes domestic banks more efficient than foreign banks. First, the headquarters of foreign banks may face difficulties in monitoring and evaluating the performance of extended entities abroad from a distance. Additionally, due to diseconomies of operations, foreign banks face adversities in terms of attracting and sustaining retail customers. Second, domestic banks are protected by inherent market barriers such as local culture, language, markets and regulations, making it difficult for foreign banks to compete (Berger et al., 2000).

having lower costs may hurt Spanish banks by earning less revenue.⁶⁹ In the US, higher profit by the domestic banks is likely due to extra spending on efforts to produce a variety of financial services that generates substantially greater revenues (Berger et al., 2000).

Third, as mentioned earlier, greater competition in the banking system can result in a higher level of efficiency and financial innovation. In chapter 2, relevant hypotheses regarding market structure and behaviour were discussed: the structure–conduct–performance hypothesis, the relative market power hypothesis, the ‘quiet life’ hypothesis and the efficient structure hypothesis. These hypotheses suggest that banks with greater market concentration possess the power to set higher prices and exercise market power (Berger, 1995). With insufficient levels of competition, the banks’ management might not control the costs of their operations to maximise profits. That is, by possessing such market power, this could allow managers to pursue goals other than profit maximisation, and subsequently result in cost and profit inefficiencies (Williamson, 1963; Berger and Hannan, 1997). For instance, Berger and Hannan (1997) investigated the impact of market structure on the efficiency level of banks. They analysed US banks survey data from the Federal Reserve for the years 1980 to 1989. They found that the relationships of X-efficiency and loan rates, deposits, and market concentration were statistically significant (5% significance level), registering coefficients of 0.18, -0.47 and -0.12, respectively. The negative relationship between concentration and X-efficiency supports the existence of a structure–conduct–performance hypothesis. The result also suggests that concentration is negatively related with X-efficiency and indicates that the ‘quiet life’ hypothesis occurs within the market power hypothesis where managers enjoy the current state of affairs and work less than they should in order to maximise profits. As discussed earlier in chapter 2, there are two types of market structure hypothesis: market-related hypotheses (e.g. relative market power hypothesis, structure–conduct–performance hypothesis and quiet life hypothesis) and efficient structure hypothesis. These hypotheses have policy implications. On the other hand, if market concentration is attributed to market power, anti-trust policies may be socially beneficial which could force the prices towards competitive levels. If market concentration is driven by high efficiency (the efficient-structure hypothesis), anti-trust actions might result in

⁶⁹ This may be attributed by poor investment choices, poor risk diversification (low risks may result in low expected returns), or/and poor service quality (e.g. skimping on expenditures necessary to monitor and service customers) (Berger et al., 2000).

greater inefficiency, coupled with prices that are less favourable to consumers (Berger and Humphrey, 1997).

In past banking literature, size has been analysed extensively and is often argued to be an important factor to explain variations in the efficiency of banks. Hence, the fourth application to discuss on studies of bank efficiency in developed countries regards bank size. The analysis of the relationship between size and bank efficiency provides important information for policy makers as well as bank managers to evaluate the optimal scale required for the operation of banks (Casu and Girardone, 2002). Past literature suggests that large banks have several advantages over small banks, product diversity, larger outreach networks, greater technology, superior managerial performance, better cost control, higher profitability and greater economies of scale and scope (i.e. from both growth and joint production) (Evanoff and Israilevich, 1991; Casu and Girardone, 2002; Tsionas et al., 2003). On the other hand, large banks face greater bureaucratic problems that are complex to manage, which could result in lower efficiency levels (Delis and Papanikalou, 2009).

The results from past banking efficiency studies in developed countries reveal that the relationship between size and efficiency varies. For example, Elyasiani and Mehdian (1990) employed data from 144 US banks from the Call and Income Report, 1985, to examine the technical efficiency and scale-efficiency of US banks. They found that large banks are more efficient than small banks. Furthermore, Berger, Hancock and Humphrey (1993) examined the efficiency estimate of the profit function and applied it to US commercial banks for the period 1984 to 1989. They focused on the standard profit function using the DFA model and intermediation approach⁷⁰ to measure the general efficiency level of US banks. They also found that large banks are more efficient than medium and small banks. The profit inefficiency of large, medium and small banks was 0.4%, 7.1% and 11.6% respectively. Likewise, Kaparakis, Miller and Noulas (1994), Kwan and Eisenbeis (1996), Clark (1996)

⁷⁰ The intermediation approach was proposed by Sealy and Lindley (1977), and perceives banks as financial intermediary between savers and borrowers. In this approach banks are assumed to use capital and labour to collect, deposit and transform them into loans and investments. Another common approach is the production approach, which views banks as producing loans and deposits accounts using capital and labour as inputs. Outputs are normally measured by the number of deposit and loan accounts, and by the number and type of transactions performed or documents processed (Molyneux et al., 1996) (This is discussed in detail in Chapter 5).

and Hasan and Marton (2003) found that larger banks are more efficient. Tsionas et al. (2003) investigated the performance of 19 Greek banks for the period 1993–1998. They employed the DEA model and reported that large and small banks were more efficient than medium banks, which displayed a cost efficiency of 97.0%, 95.0%, and 94.0%, respectively. They suggest that larger banks have the ability to adopt innovative technologies, whereas smaller banks use efficient market exploitation to increase their market share. On the other hand, Akhigbe and McNulty (2003) used the SFA model on US bank data (1990 to 1996) to investigate profit efficiency based on smaller banks in a less competitive market. Based on the single frontier evaluation, they found that small banks (74.7%) are more profit-efficient than large banks (72.5%). Following the banking deregulation measures introduced by the regulators during early 1990s, large banks require major restructuring and cost-cutting initiatives more than the small banks. Other studies that found similar results, which suggest that smaller banks possessed higher efficiency than large banks, are Elyasiani and Mehdiian (1995), Clark (1996), Casu and Girardone (2002) and Koetter (2006).

As mentioned earlier, a higher level of market competition can result in higher risk-taking behaviour by the banks. Thus the final contextual application of empirical studies in developed countries is on the inherent risks faced by banks. Based on empirical studies of banking efficiency in developed countries, the relationship between risk and bank efficiency is found to be inverse (Kaparakis et al., 1994; Wheelock and Wilson, 1995; Kwan and Eisenbeis, 1996; Berger and De Young, 1997; Fiordelisi et al., 2011). The Basel capital adequacy accord⁷¹ series from Basel I to Basel III (introduced between 1998 and 2011) emphasised the need for robust prudential regulations to tackle various risks within the banks such as market, credit, operational and liquidity risk. These risks are reflected in financial capital via risk-weighted assets and various methodologies. In general, banks with greater risk are required to maintain larger stores of financial capital to buffer against potential losses. According to Berger and De Young (1997), who investigated US commercial banks using the data of the Annual Reports of Conditions and Income between

⁷¹ The Basel capital adequacy accord is a set of agreements set by the Basel Committee on Banking Supervision (BCBS), which provides recommendations on banking regulations regarding capital risk, market risk, operational risk and liquidity through Basel I, II and III. The purpose of the accords is to ensure that financial institutions have enough capital to meet obligations and absorb unexpected losses (Basel Committee on Banking Supervision, 1999, 2005, 2010).

1985 and 1994, from the 1-stage SFA model, the coefficient for non-performing loans was negative (-0.06) and significant at 1%. Thus, they found that the relationship between loan quality and cost efficiency supports the ‘bad luck’ hypothesis, where an increase in problem loans can result in lower cost efficiency (e.g. an increase in problem loans will lead to higher expenses in monitoring and managing the bad loans portfolio). Berger and De Young (1997) introduced several hypotheses relating to problem loans and their relationship with cost efficiency and financial capital. These hypotheses use mnemonics such as ‘bad luck’, ‘bad management’, ‘skimping’ and ‘moral hazard’. They explained the source of problem loans and how they occur. First, ‘bad luck’ indicates problem loans that derive from external sources. That is, extra expenses incurred by the banks when loans turn bad and cannot be controlled by management. Second, ‘Bad management’ refers to poor handling of the loan portfolio and operating expenses, and consequently reflects the low cost efficiency of banks. Third, ‘Skimping’ reflects the effort spent on loan origination, which in turn will impact the loans’ performance in future. ‘Skimping’ is a trade-off between a loan’s problems and its operating costs. For example, banks that focus on earning good long-run profits could ‘skimp’ by reducing loan screening, which leads to problem loans in the future. And fourth, the ‘Moral hazard’ hypothesis explains the problem of the entire banking industry where the negative behaviour of a single bank taking excessive risks is considered excusable due to the importance of the industry.

Kaparakis et al. (1994) examined the productive efficiency of 5,548 US banks for the year 1986 using the SFA model and simultaneously tested the effect of external factors (e.g. state population density, state branching regulations, portfolio riskiness) on the cost efficiency scores. The average inefficiency of the US banks in their sample is 9.0%. They employed a 2-stage analysis to test the external factors and found that banks become inefficient when they move to a more competitive market, with an increase in operational risks, a high ratio of non-performing loans, increased purchased funds, and low equity to capital ratio. Likewise, Wheelock and Wilson (1995) used quarterly data of US banks from the Call Report from 1984 to 1993. They examined the determinants of bank failure of US banks through an estimation of managerial inefficiency using the nonparametric DEA model. They reported that banks with lower capital, higher loan loss provisions and a higher loan–deposit ratio have a higher risk of failure. Similarly, Altunbas et al. (2000) examined around 130

Japanese banks for the period 1993 to 1995 using the SFA model and investigated the impact of risk and quality factors on bank costs. The results display a stable inefficiency level, between 6.0% and 7.0% from 1993 to 1995. In this study, they included the loan loss provisions, financial capital and liquidity ratio to control for output quality and risks. When testing the coefficient of the determinants, they found that efficient banks had fewer non-performing loans, lower liquidity and higher financial capital. As discussed above, there are various types of inherent risk that can affect the efficiency level of the banks. For this present research, the explanatory factors for inherent risk that will be employed in the analysis are capital adequacy, asset quality and liquidity.⁷²

This section has provided an overview of empirical studies in banking efficiency from developed countries. Most of the earlier studies that pioneered the development of frontier efficiency originated from developed economies, such as the US and the EU. The findings from these empirical studies have been used by regulators to measure the efficiency of banks within developed economies relating to macroeconomic events (e.g. financial liberalisation and policy affecting market structure) and the characteristics of banks (e.g. ownership, size and inherent risks). Furthermore, the same methodology is also applied to banking systems in developing countries. The differences in market condition, regulatory structure, language and culture in developing economies may generate different findings to those in developed countries. Hence, in the next section, empirical studies of banking efficiency in developing countries are discussed.

⁷² Inherent risk factors such as capital adequacy, asset quality and liquidity will be used as control variables in the SFA model. Details of these control variables are discussed in Chapter 5 ('Methodology and Data').

Table 3.1 Efficiency Studies in US and Developed Countries Banking System

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Elyasiani and Mehdiian (1990)	144 US banks from the Call and Income Report, 1985	Intermediation/cost	Deterministic	Inputs: labour, capital, deposits Outputs: loans, investments	Smaller banks are more efficient than large banks during the early years of the sample. The gap reduces from year 5 onwards. Market conditions affect small and large banks differently. Small banks struggled to keep up with market changes compared to bigger banks.
Berger, Hancock and Humphrey (1993)	US commercial banks from the Report Condition and Income, 1984–1989	Intermediation/profit	DFA	Inputs: labour, purchased funds Outputs: business loans, commercial and industrial loans, real estate loans, instalment loans	Inefficiencies in the US banking industry are quite large. These inefficiencies derived from poor revenues rather than excessive costs. Larger banks are more efficient than smaller banks.
Mahajan, Rangan and Zardkoohi (1996)	US multinational and domestic banks from Call Report, 1987–1990	Intermediation/cost	TFA	Inputs: labour, purchased deposits, capital Outputs: total loans, demand deposits, government securities	Domestic banks enjoy a cost advantage over multinational banks. There were significant diseconomies of scale for all the banks including both domestic and multinational banks. The scale diseconomies have more impact on domestic banks compared to multinational banks. There is no evidence that extraordinary size is required to compete in the international market.
De Young and Noelle (1996)	62 US foreign-owned banks (US based subsidiaries of foreign banks), 1985–1990	Intermediation/cost	DFA	Inputs: price of labour, price of fund, core deposits, physical capital Outputs: total loans, total securities	They found that in the US banking industry, US-owned banks are more efficient than foreign owned banks. Foreign owned banks are less efficient due to input inefficiency where they have to utilise higher variable inputs compared to US-owned banks, particularly due to a reliance on expensive financing using purchased funds.

Note: DFA: Distribution Free Analysis, TFA: Thick Frontier Analysis

Table 3.1 Efficiency Studies in US and Developed Countries Banking System (Continued)

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Berger and Hannan (1997)	US banks, Fed's survey of selected deposits and other accounts, Fed's survey of the term of bank lending to business, 1980–1989	Intermediation/cost	DFA	Inputs: labour, physical capital Outputs: demand deposits, retail time and savings deposits, real estate loans, commercial and industrial loans, instalment loans Control variables: Herfindahl Index, Return on Assets (ROA), Return on Equity (ROE)	The average efficiency levels of banks from deposit and loan survey were 56.0% and 57.0% respectively. Authors performed a two-stage analysis to determine the impact of the market structure and from the regression, they found that the deposit and loan rates together with the coefficients of concentration were statistically significant and supported the structure–conduct–performance hypothesis. Concentration is negatively related with efficiency and indicates that the 'quiet life' hypothesis occurs within the market power hypothesis.
Akhavain, Berger and Humphrey (1997)	57 bank mergers from Call Report, 1980–1990	Intermediation/cost and profit	DFA	Inputs: price of loans, price of securities, price of deposits, price of labour Outputs: equity capital, loans, securities Environmental variables: pre- and post-mergers	They found that the mega-mergers in 1980s significantly increased the degree of profit efficiency in US banks. The average improvement was 16.0% from 74.0% pre-mergers to 90.0% post-mergers. One significant observation made was that most of the banks were able to further diversify their portfolio by shifting their output mix from securities to loans.
Berger and Mester (1997)	6,000 US commercial banks from Consolidator Reports of Condition and Income	Intermediation/cost, standard profit and alternative profit	DFA and SFA	Inputs: labour, purchased funds Outputs: consumer loans, business loans, securities Control variables: Non Performing Loan (NPL) ratio, financial capital	The standard profit of distribution free analysis using both translog and Fourier Functional form registered 55.0% and 54.0% respectively and for alternative profit, both estimations displayed 46.0% and 45.0% for translog and Fourier Functional form functions respectively. Cost efficiency and market power are negatively related but market power and profit efficiency are positively related. This simply indicates that in less competitive markets, banks can charge higher prices but at the same time do not sense the pressure to keep costs down.

Note: SFA: Stochastic Frontier Analysis, DFA: Distribution Free Analysis

Table 3.1 Efficiency Studies in US and Developed Countries Banking System (Continued)

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Berger and Mester (2003)	US banks from the Call and Income Report, 1984 and, 1991–1997	Intermediation/cost, standard profit and alternative profit	TFA	Inputs: purchased funds, core deposits, labour Outputs for alternative profit and cost: consumer loans, business loans, real estate loans, securities Outputs for standard profit: price of consumer loans, price of business loans, price of real estate loans, price of securities	Banks' cost productivity has worsened, but profit efficiency has increased significantly. To explain these perplexing findings, Berger and Mester (2003) argued that over time, banks have been innovative and offer a variety of products and services to customers, intending to maximise banks' revenue and in consequence, higher costs can be expected when such moves are initiated. The wider and additional products and services offered might have raised the cost, but at the same time they have increased revenues in higher proportion than cost; this can result in worsened cost efficiency.
Allen and Rai (1996)	194 banks from 11 OECD countries, 1988–1992	Intermediation/cost	SFA and DFA	Inputs: price of labour, price of funds, price of capital Outputs: total loans, investment securities	The DFA overestimates the magnitude of inefficiencies relative to SFA. There were significant differences between stochastic cost frontier and distribution free approaches, which registered at 18.0% and 68.0% respectively.
Bergendahl (1998)	48 banks from 4 Nordic countries (Denmark, Finland, Norway, Sweden), 1992–1993	Production/cost	DEA	Inputs: personnel cost, cost of materials, credit loss cost Outputs: lending, deposits, gross revenue	The average efficiency of all banks in the Nordic countries was 0.74. Based on this study, it was found that banks on the frontier have certain qualities in the output compared to the others below the frontier curve. The qualities are lending cost per unit, term cost per unit and gross revenues per unit.
Altunbas, Evans and Molyneux (2001)	German banks in 1000s, 1989–1996	Intermediation/cost	SFA	Inputs: price of labour, price of funds, price of physical capital Outputs: total loans, securities, off balance sheet items	Public savings banks and mutual cooperatives were slightly more cost- and profit-efficient compared to the private banks. Public and mutual banks were slightly better due to their ability to source a cheaper cost of funds from depositors and investors. Additionally, there were no significant differences in the efficiency levels of these three groups of banks.

Note: DEA: Data Envelopment Analysis, SFA: Stochastic Frontier Analysis, DFA: Distribution Free Analysis, TFA: Thick Frontier Analysis

Table 3.1 Efficiency Studies in US and Developed Countries Banking System (Continued)

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Koetter (2006)	32,322 observations of banks data from Deutsche Bundesbank, 1993–2003	Intermediation/cost and profit	SFA	Inputs: fixed assets, labour, borrowed funds, equity Outputs: interbank loans, commercial loans, securities Control variables: total assets, off-balance sheet, market concentration, asset risk	Two new models were introduced to capture the effect of heterogeneity in the market. The result reported that the alternative inputs showed approximately 5% lower efficiency level than the traditional input (traditional model 0.92, model 2 (market regions), 0.86 and model 3 (large banks within the federal market), 0.88). Profit efficiency on the other hand shows a similar result at 0.65 for all three models. Banks with a larger size, riskier assets, and higher market concentration exhibit positions higher than the overall cost frontiers.
Bos and Kool (2006)	401 small cooperative Dutch (Rabo) banks, 1998–1999	Intermediation/cost and alternative profit	SFA	Inputs: public relation cost, price of labour, net housing cost, admin cost and write off, price of financial capital Outputs: retail loans, wholesale loans, mortgages, insurance-brokerage-travel provisions Environmental variables: bank's strategic choices, local banking market variables	Exogenous input prices (0.97) are more cost-efficient than endogenous input prices (0.91). However, in terms of profit efficiency, the result is marginal. Market factors can only explain 10.0% of efficiency on local banks. Due to a high level of homogeneity in the market such as banks operating in the same country with a set of similar legal and institutional frameworks, the result could not explain the variations in efficiency levels.
Fiordelisi and Mare (2013)	4200 observations of Italian cooperative banks, 1997–2009	Intermediation/cost, revenue, operating, cost-income and interest efficiencies	SFA	Inputs: capital adequacy, credit orientation, liquidity, asset quality Outputs: Size, employment growth, GDP growth Environmental variables: Population, concentration, NPL, Entrepreneurial, no. of ATMs, bank branches, criminalities, solidarity	Results indicated that lower risk is related to higher survival time for cooperative banks. This means that prudent and skilful managerial abilities could increase the survival time of cooperative banks. Additionally, banks with greater efficiency have higher chances in facing higher probability of survival. It is also observed that the higher capital level could reduce the probability of default of a bank, which supports that higher capital levels provide greater buffer for loss absorbency and reduce banks' moral hazards.

Note: SFA: Stochastic Frontier Analysis

3.4.2. Empirical Studies of Banking Efficiency in Developing/Transition Countries

As discussed earlier, there is growing interest in conducting empirical research regarding banking efficiency in developing/transition countries. To date, empirical studies have dealt mainly with contextual applications rather than the introduction of new methods or approaches of frontier efficiency. Thus, the discussions in the next subsection are attributed to the application of frontier efficiency to developing/transition countries, relating to (1) financial liberalisation/deregulation, (2) market structure, (3) ownership, (4) size and (5) risk.⁷³

First, following the successful implementation of financial liberalisation in developed countries, developing countries also initiated deregulation to (1) improve the efficiency of their banks, (2) increase market competition, and (3) achieve greater savings mobilisation and a better allocation of resources and expansion in real economic growth. For example, Kraft and Tirtiroglu (1998) examined the impact of banking liberalisation on cost and scale efficiency using the parametric SFA with a translog specification model on 43 Croatian banks from 1994 to 1995. The number of Croatian banks has doubled since the initiation of banking liberalisation in 1990. Based on a stochastic frontier approach, the newly established Croatian banks (75.0%) are found to have lower cost efficiency than older banks (83.0%). The higher cost efficiency of the older banks is a result of measures taken to meet the anticipated competitive threat of new banks. Hence, the notion that free entry can result in more competition should be treated cautiously, particularly when the banking environment is influenced by free-riding opportunities for distressed borrowers, unrehabilitated state banks and difficulties faced by new banks in starting operations (Kraft and Tirtiroglu, 1998).

⁷³ Bos et al. (2009) state that the banking industry consists of different players with variations of activities and aims. Hence, heterogeneity among banks' attributes such as equity, bank type, regional location and bank size should be considered. Failure to consider these explanatory factors could result in an underestimation of frontier efficiency.

Hao et al. (2001) investigated the productive efficiency of 19 private nationwide and regional South Korean banks for the period 1985 to 1995. They also analysed the determinants of efficiency using a two-stage analysis, following the deregulation initiated by the Korean government in the early 1980s and 1990s. Based on the results from the stochastic frontier model, it was noted that both nationwide and regional banks had improved their efficiency score after the commencement of the deregulation in 1991 where banks displayed lower bad loan ratios and higher foreign and 'chaebol'⁷⁴ equity ownership percentages. The average efficiency of these banks was 87.0% in 1985 and increased to 91.0% in 1993 following the implementation of financial deregulation. This finding is consistent with Gilbert and Wilson (1998) who indicated that the gradual liberalisation of the Korean banking industry has contributed to an improvement in the productivity of Korean banks.

Al-Jarrah and Molyneux (2007) investigated cross-country banking efficiency in the Arabian banking industry from 1992 to 1990. They examined 82 banks from Jordan, Egypt, Saudi Arabia and Bahrain using stochastic frontier analysis coupled with three different economic functions namely, cost, profit and alternative profit efficiency. The objective of their study was to explore the degree of banking efficiency and the impact of financial reforms on these countries. Al-Jarrah and Molyneux (2007) argue that by implementing a different type of efficiency economic function, the results of each method provide some informational value. The cost efficiency of Arabian banks registered 95.0% but lower alternative profit and standard profit efficiency, displaying 66.0% and 58.0% respectively. This shows that Arabian banks can be cost-efficient but not profit-efficient due to the input prices (e.g. lower cost of funds for Islamic banks) and banks' optimisation behaviour. In addition, Cook et al. (2001) investigated the deregulation initiatives (i.e. the liberalisation of interest rates, reduction of the minimum treasury bills limit, removals of entry restrictions on foreign banks) taken by the Tunisian government within the period 1992–1998. The result suggested that deregulation made no impact on the efficiency level of the banks.

Attaullah et al. (2004) compared the technical efficiency of commercial banks in India and Pakistan from 1988 to 1998 and the significance of financial liberalisation during the period

⁷⁴ 'Chaebol' is a form of business conglomerate in South Korea. They are typically global multinationals that own a number of international enterprises.

of study. This study adopted the DEA model and found that the overall technical efficiency of Indian and Pakistani banks improved with the implementation of financial liberalisation. Both India and Pakistan registered an increase in overall technical efficiency, from 67.0% and 41.1% in 1988 to 82.1% and 57.4% in 1998, respectively. Concurrently, Hardy and dePatti's (2001) findings are also in accordance with those of Attaullah et al. (2004) who found that, based on the SFA model for both profit and cost efficiency, the financial reforms between 1988 and 1992 in Pakistan resulted in higher cost and profit efficiency during the post-deregulation period. Additionally, Sensarma (2006) tested the deregulation effect on cost efficiency in Indian banks under different types of ownership from 1986 to 2000. From the SFA with translog specification model, cost inefficiencies in Indian banks were on a declining trend during the sample period but the rate of decline was slower during the implementation of deregulation in 1992, indicating that banks were adjusting to the events of deregulation that were taking place. The cost efficiency of public and foreign banks improved from 85.0% and 18.0% in 1986 to 94.0% and 31.0% in 2000, respectively. Hence, India had achieved the objectives of deregulation, which aimed to improve productivity and reduce intermediation costs. Kumbhakar and Sarkar (2003) also examined the impact of deregulation and productivity of the banks in India from 1985 to 1996 using the TFP model. Unlike Sensarma (2006), Kumbhakar and Sarkar (2003) found little evidence to support that deregulation has resulted in improvements over the years. This can be attributed to the extensive over-employment of labour relative to capital during the period of study.

Drake et al. (2006) examined technical efficiency in the Hong Kong banking system from 1995 to 2001. They employed the DEA model with an environmental factor being incorporated into the two-stage analysis of efficiency. The profit efficiency of Hong Kong banks lies between 88.6% and 94.9%. Based on two-stage regression analysis, the authors argue that the environmental factors indicated by the macroeconomic and housing markets had affected Hong Kong's banking industry. Drake, Hall and Simper (2006) also asserted that the financial deregulation initiated by Hong Kong's regulators and the Asian financial crisis had not impacted the efficiency level of the banks in Hong Kong, unlike in other Asian countries such as Malaysia, Indonesia, Thailand and Korea.

Within South East Asia (SEA) (i.e. Indonesia, Malaysia, the Philippines, Singapore and Thailand), Williams and Nguyen (2005) examined the relationship between bank efficiency and governance and deregulation for the period 1990–2003. From the SFA results, they found that privatised banks demonstrate greater (alternative) profit efficiency compared to other banks with a different governance structure. The implementation of a privatisation programme in SEA countries contributed to greater efficiency in banks. The mean profit efficiency of privatised banks is higher than their competitors: for example 77.4% in Indonesia, 91.4% in Philippines and 76.9% in Korea. Additionally, Leightner and Lovell (1998) analyse the impact of deregulation on commercial banks in Thailand using sample data from 1989 to 1994. Leightner and Lovell (1998) claimed that Thai banks reacted to financial deregulation in a positive manner because: first, the deregulation phase took place during an economic boom; second, the market setting during deregulation was oligopolistic (where banks have the opportunity to gain higher profits before competition forces lower profitability); third, Thai banks have adequate international expertise to cope with forex market liberalisation; and, fourth, the foreign banks brought in international expertise.

Another common area of financial liberalisation study, in developing and transitional countries, relates to the removal of entry barriers for foreign banks and how it affects the efficiency of the banks within the domestic market. The results of previous studies on the impact of foreign banks vary from country to country. For example, Claessens et al. (2001) employed 7,900 bank observations from 80 countries for the years 1988 to 1995. They state that the efficiency of foreign banks depends on the economic development of the host countries. In developing countries, foreign banks demonstrate greater efficiency compared to domestic banks. On the other hand, foreign banks display lower efficiency when compared to the larger domestic financial conglomerates in developed countries. They also assert that the presence of foreign banks places a competitive pressure on domestic banks to improve their services and operations. This is consistent with Unite and Sullivan (2003) who reported that the presence of foreign banks has increased competitive pressure on domestic banks (in the case of Philippines) and forced domestic banks to improve their efficiency levels. On the other hand, Kraft et al. (2006) employed an SFA-translog specification model and conducted an investigation in Croatia (a developing/transition country) from 1994 to 1995, and they found that intense competition from new entrants did

not help to improve the overall efficiency of domestic banks. The result suggests that new banks are less efficient than old banks, where the average cost efficiency of new versus old banks is 75.0% and 83.0% respectively.

Therefore, empirical studies on the impact of financial liberalisation towards the efficiency of banks in developing/transition countries provide mixed findings; for instance, some studies show a positive impact while others show either a negative impact or no impact at all due to liberalisation. The result of financial liberalisation relies on the conditions of the industry and types of deregulation measures. A good banking system structure can result in the successful implementation of liberalisation whereas, a weak banking system structure can lead to instabilities and decreasing efficiency should the government continue to implement their financial liberalisation initiatives (Berger and Humphrey, 1997; Demircig-Kunt and Detragiache, 1999).⁷⁵

Second, the topic of ownership, which can affect market competition, is also applied in the banking industry of developing/transitional countries. In general, foreign banks exhibit higher efficiency levels compared to domestic banks in developing countries (Isik and Hassan, 2003; Hassan and Marton, 2003; Bonin et al, 2005; Fries and Taci, 2005; Havrylchik, 2006). Most of the findings regarding ownership (i.e. foreign vs. domestic banks) are the opposite in developed compared to developing countries. This condition can be described by the 'global advantage hypothesis' as discussed previously. For example, Havrylchik (2006) examined the cost efficiency of the Polish banks from 1997 to 2001 using the DEA model and reported that foreign banks (average TE: 73%; average AE: 88.0%) displayed greater productivity from both technical efficiency (TE) and allocative efficiency (AE) in comparison to domestic banks (average TE: 53.0%; average AE: 69.0%). Additionally, Isik and Hassan (2003) found that public and foreign banks in Turkey display more cost and technical productive efficiency than private banks. Likewise, Berger et al. (2009) employed an SFA-translog specification model on 266 annual observations of Chinese banks during the economic reforms from 1994 to 2003. They also found that banks with major foreign ownership are more efficient than domestic banks (e.g. private and state-

⁷⁵ For example, a weak banking structure can be explained by inefficient government structure and processes, a weak financial and legal infrastructure and enforcement, and a high level of corruption (Demircig-Kunt and Detragiache, 1999).

owned domestic banks). The profit and cost efficiency of foreign owned banks is higher at 69.0% and 91.0% compared to the big four banks⁷⁶, state owned banks and private domestic banks, which registered at 23.0% and 89.0%, 48.0% and 91.0%, and 56.0% and 82%, respectively. Compared to Berger et al. (2009), Kasman and Yildirim (2006) also found that foreign banks exhibit higher profit- and cost-efficiency in commercial banks from eight Central and Eastern Europe countries (Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic and Slovenia) from 1995 to 2002. From the SFA–Fourier Flexible Function model, they argued that foreign banks have the ability to generate higher profits from different value-added segments of the market. Furthermore, Hasan and Marton (2003) examine the conditions and performance of the Hungarian banking industry during the transitional period, from a centralised economy to a market-oriented system, for the years 1993 to 1998. Using both cost efficiency (CE) and profit efficiency (PE) on a parametric SFA model, banks with foreign ownership participation (CE: 73.9%; PE: 68.2%) were found to be more efficient than domestic banks (CE: 66.2%; PE: 61.9%) in Hungary. They further assert that, to build a stable and increasingly efficient banking industry during the transitional period, early reorganisation initiatives, flexible approaches to privatisation, and liberal policies towards foreign bank participation are crucial to establishing a more competitive banking system (Hassan and Marton, 2003). Therefore, unlike developed countries, foreign banks establish a stronger presence in developing countries and demonstrate superior knowledge and expertise. Nevertheless, the presence of foreign banks supports the development of the local banking sector by introducing ‘best’ international practices and know-how that may have a spillover effect on domestic banks (Levine, 1996). On the other hand, there is also empirical study within developing countries where foreign banks display lower cost efficiency. For instance, Sensarma (2006) found that foreign banks exhibit lower efficiency scores compared to domestic banks in India (using data from 1986 to 2000). The cost efficiency of foreign banks has an average of 26.7% compared to public and private banks, with an average of 87.7% and 78.7% respectively. This is attributed to the low deposit-to-total liabilities ratio and high personnel cost. Correspondingly, Bhattacharya et al. (1997) also found lower efficiency in foreign banks in India for the year

⁷⁶ Big four is colloquial name for the four main banks in China. The big four banks in China are (1) Bank of China, (2) China Construction Banks, (3) Industrial and Commercial Bank of China and (4) Agricultural Bank of China.

1986 to 1991 due to constraints imposed by Indian regulators towards foreign banks, such as high capital requirements and modest priority lending sectors.

Third, competition between foreign and domestic banks could alter the market structure of the banking industry. It is also expected that the competitive structure of the market can influence the efficiency levels of banks. For example, Fries and Taci (2005) examined 289 banks from 15 transition countries from Eastern Europe and analysed the impact of the banking industry's transformation during the post-communism transition from 1994 to 2001. From the parametric SFA-translog function model, Fries and Taci (2005) found that strong macroeconomic stability, greater competition from foreign banks and the sturdy structural development of banking institutions would result in higher cost efficiency in the banks. They also found that the majority of foreign-owned private banks exhibit a higher degree of cost efficiency than other types of private bank. Fries and Taci (2005) further suggest that to increase the efficiency of the Eastern European banking industry, state banks may offer significant ownership interests to foreign banks. Similarly, Fu and Heffernan (2007) also found that greater competition since the first phase of banking reform (i.e. the deregulation of interest rates) has resulted in better efficiency levels in commercial banks in China. However, the result of cost efficiency using the SFA model (average CE; 1985–2002: 40.7%, 1985–1992: 43.5%, 1993–2002: 39.5%) was lower in the second phase of the banking reform in China, where from the regression of 2-stage analysis, the relationship between cost efficiency and the reform variable is negative (-0.1270) at 5% significance level. Despite the banking reform initiatives, they argue that state-owned banks are still being pressured to play a special role in China's economic system, where large loans are granted to loss-making state-owned firms. Additionally, bank ownership is also applied in the 2-stage regression procedure, which suggests that Chinese banking cost efficiency could be improved by converting state banks into joint-stock ownership. Nevertheless, this 2-stage analysis is being questioned (this is discussed in detail in chapter 5). For instance Berger and Mester (1997) argue that the 2-stage analysis is valid when the additional explanatory variables in the second stage are exogenous and are not correlated with inputs and outputs in the first stage. However, Fu and Heffernan (2007) employ endogenous variables (e.g. the ratio of purchased funds to total assets, the ratio of total loans to total assets, and the ratio of total investment to total assets) that are highly correlated with inputs and outputs in the

second stage. Thus, the conclusion of correlations and their directions between cost efficiency and explanatory variables are somewhat doubtful (Fu and Heffernan, 2007). In addition, Isik and Hassan (2002) employed the SFA model to examine both profit and cost efficiency on Turkish commercial banks between 1988 and 1996. They found that Turkish banks are profit-efficient (average PE: 94.0%) but not cost-efficient (average CE: 86.0%). They also argue that there is lack of competition within the market where the number of banks is small and the demand for banking products and services is greater than the supply. Due to imperfect market competition, Turkish banks are operating at higher cost and obtain higher revenues at the same time (which has resulted in lower cost efficiency but higher profit efficiency). This can also be explained by the 'quiet life' hypothesis, where managers do not work as hard as they should to maximise profits. Similarly, the impact of higher market concentration can also be observed in the work of Grigorian and Manole (2002), who suggest that there is a positive relationship between market concentration and the efficiency level of banks from a study conducted using the DEA method on 209 banks from seventeen transition countries between 1995 and 1998.

Fourth, as earlier discussed, competition and concentration are important elements in measuring welfare-related public policy towards market structure and conduct in the banking industry. Large banks generally have the market power to charge higher prices and hence, are more cost and profit efficient than small banks. However, findings on the efficiency of banks in developing countries in terms of size display mixed results. Efficient banks in developing countries vary in terms of size. For example, Leightner and Lovell (1998) analysed commercial banks in Thailand from 1989 to 1994 during the deregulation stage using the DEA model. They found that large banks are more cost efficient than medium and small banks. A similar finding is noted in Al-Jarrah and Molyneux (2007) who utilised the SFA–Fourier Flexible specification model to examine the international cross-country banking efficiency of Arabian banking industry from 1992 to 1998. In term of size, they also found that larger banks within the Arab countries exhibit higher cost- and profit-efficiency.

On the other hand, Kwan (2006) examined the cost efficiency of Hong Kong banks between 1992 and 1999 using a SFA-translog model. Small banks registered a higher cost efficiency

score than large banks at 78.0% and 70.0% respectively. He also stated that larger banks are more inefficient as their equity is publicly traded, which results in diverse ownership and greater agency cost (for a sound corporate governance). Likewise, Isik and Hassan (2002) found that size is negatively associated with cost efficiency. In the case of the Turkish banking industry, they assert that small banks can deliver better than large banks due to better market discipline, as small banks only compete in the metropolitan areas but not in rural areas. Moreover, Bonin et al. (2005) also found an inverse relationship between cost efficiency and bank size. Based on the result of a parametric SFA model on 224 banks from eleven Eastern European transition countries between 1996 and 2000, they found that small banks are more cost and profit efficient than large banks. This finding can be explained by the early stage of consolidation schemes in the banking industry in transition countries, which caused large banks to exhibit a lower efficiency level⁷⁷. Apart from the above findings, Kasman and Yildirim (2006) found no relationship between bank size and cost and profit efficiency based on an examination of commercial banks in Central and Eastern Europe between 1995 and 2002 using the SFA-Fourier flexible specification model.

Fifth, as discussed earlier, a more competitive market structure can increase the risk-taking behaviour of banks, which consequently influences their efficiency in developing countries. For example, Grigorian and Manole (2002) employed 209 banks in seventeen transition countries from 1995 to 1998 to examine their efficiency levels and the impact of the policy framework on these banks. Based on the results from the DEA model, they found a positive association (0.074 at 5% significance level) between capitalisation and efficiency scores, suggesting that well-capitalised banks have better potential to achieve a greater efficiency level. The implementation of a tighter minimum capital adequacy ratio by the regulators has shown a positive impact where commercial banks have displayed better revenue, improved financial capacity and robust deposit-taking behaviour. From the analysis of DEA on Turkish banks, Isik and Hassan (2003) found that efficient banks have a strong level of financial capital to buffer potential losses that may arise from risky assets. On the other

⁷⁷ Bonin et al. (2005) assert that regulators of the banking sector in transition countries introduce consolidation schemes for smaller banks in order to improve the efficiency of the banks and allow them to compete with foreign banks. The regulators presume that small banks in transition countries are too small for international standards. However, Bonin et al. (2005) demonstrate through SFA that the presumption mentioned above does not hold for newly established consolidated banks in transition countries.

hand, cost efficiency is negatively related (-0.026 at 10% significance level) with the rate of asset growth indicating that rapid growth can result in outgrowing managerial skills as well as scale problems. Additionally, Fries and Taci (2005) examined the impact of bank- and country-level characteristics on the cost efficiency of commercial banks in Eastern Europe using the SFA approach. They posit that cost-efficient banks can be characterised through a higher ratio of capital to total assets and lower loan losses, which are signs of lower inherent risk in banking operations. Thus, the above discussion indicates that within the developing/transitional countries, lower inherent risk (e.g. high capital levels, low loan losses and high liquidity) would normally result in higher efficiency.

The above discussion explains how frontier efficiency is applied in developing economies. As a summary, many of these empirical studies provide mixed findings from policies implemented by regulators such as, financial liberalisation, the removal of entry barriers to foreign banks, merger and consolidation schemes in the banking industry and market concentration and competition. These policies may directly affect banks but no consensus is found on the efficiency of banks based on market structure, size and ownership. Nonetheless, one consistency in the findings was found in relation to inherent risk. It was observed that banks with a lower level of inherent risk would achieve a higher level of efficiency.

Table 3.2 Efficiency Studies in Transition and Developing Countries Banking System

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Grigorian and Manole (2002)	209 banks from 17 transition countries from Bankscope, 1995–1998	Value add/cost	DEA	Inputs: labour, fixed assets, interest expenditures Outputs: revenues, net loans, liquid assets, deposits, net loans, liquid assets Correlation variables: GDP, Inflation. M2/GDP, equity/assets, concentration, new vs. old, foreign ownership, capital adequacy, single borrower limit, forex exposure limit	There is a positive association between capitalisation and market concentration on the efficiency, where well-capitalised banks have better potential to achieve a greater efficiency level. The implementation of a tighter minimum capital adequacy ratio by the regulators has shown a positive impact: commercial banks have displayed better revenue and improved financial capacity and robust deposit-taking behaviour. Countries with a stricter policy on foreign exposure limits for banks were worse off than countries that liberalised their foreign exchange exposure policy.
Fries and Taci (2005)	289 banks from 15 transition countries from Bankscope database, 1994–2001	Intermediation/cost	SFA	Inputs: labour, physical capital Outputs: loans to non-bank entities, loans to other banks, deposits Control variables: NPL/Total Loans (TL), non-loan assets/Total Assets (TA) Environmental variables: per capita GDP, nominal market interest, assets share, ratio of TL/Total Deposits (TD), ratio of capital/assets	Smaller banks show significant scale economies while average-sized banks have constant returns to scale. Country-level factors that can increase cost efficiency level are a lower nominal interest rate, a higher market share of foreign-owned banks and a higher intermediation ratio. In terms of bank characteristic variables, to achieve higher cost efficiency level a bank should display higher ratios of capital to total assets and lower loan losses, also demonstrating signs of lower inherent risk in banking operations.
Berger, Hassan and Zhou (2009)	38 Chinese banks from Bankscope database, 1994–2003	Intermediation/cost and standard profit	SFA	Inputs: interest expense to total deposits, non-interest expenses to fixed assets Outputs: total loans, total deposits, liquid assets, other earnings assets Control variables: asset quality Environmental variables: bank ownership	Banks with majority foreign ownership (69.0%) are the most profit-efficient compared to banks with domestic private (56.0%) and state ownership ('big four', 23.0%; non-'big four', 48.0%), suggesting that a shift of resources from state-owned banks to foreign ownership would be likely to increase the efficiency of banks, in particular the 'big four' state owned banks.

Note: DEA: Data Envelopment Analysis, SFA: Stochastic Frontier Analysis

Table 3.2: Efficiency Studies in Transition and Developing Countries Banking System (Continued)

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Al-Jarrah and Molyneux (2007)	82 banks from Jordan, Egypt, Saudi Arabia and Bahrain, 1992–1990	Intermediation/cost	SFA	Inputs: price of deposits, price of labour, price of physical capital Outputs: total loans, other earning assets, off balance sheet items Control variables: loan loss reserve over total loans, capital adequacy ratio, time trends Environmental variables: market structure, organisational characteristics, geographical segmentation, bank liquidity	The cost efficiency of Arabian banks registered 95.0% and lower profit and standard profit efficiency, displaying 66.0% and 58.0% respectively. Arabian banks can be cost-efficient but not profit-efficient due to the output specifications. It was found that Islamic banks, larger banks and Bahraini banks are more efficient than their counterparts.
Isik and Hassan (2003)	Turkish banks from Banks Association of Turkey and Istanbul Stock Exchange, 1988, 1992 and 1996	Intermediation/cost	DEA	Inputs: labour, capital, loanable funds, price of labour, price of capital, price of funds Outputs: short-term loans, long-term loans, risk adjusted off balance sheet items, other earnings assets Correlate variables: efficiency structure, ownership structure, corporate governance structure, market structure, risk structure, other bank traits	Public and foreign banks show better cost and technical productive efficiency than private banks. Efficient banks are characterised with greater purchased deposit for financing risky assets and have strong financial capital to buffer potential losses from the inherently risky assets. Cost efficiency is negatively related with rates of asset growth, indicating that rapid growth can result in outgrowing of managerial skills as well as scale problems.
Bonin, Hasan and Wachtel (2005)	225 banks from 11 advanced transition countries, 1996–2000	Intermediation/cost and standard profit	SFA	Inputs: price of capital, price of funds Outputs: deposits, total loans, total liquid assets and investments other than loans and liquid assets Correlate variables: foreign strategic, other majority foreign, government owned, international participation, log assets	Participation in foreign ownership contributed to a positive impact. Foreign-owned banks are driven towards cost efficiency more than profit efficiency. Efficiencies are negatively associated with bank size as many transition countries have undertaken consolidation schemes on their banking industry due to the presumption that their banks were too small by international standards to be efficient.

Note: DEA: Data Envelopment Analysis, SFA: Stochastic Frontier Analysis

Table 3.2 Efficiency Studies in Transition and Developing Countries Banking System (Continued)

Authors/(Year)	Data	Approach/Function	Method	Variables	Results
Kwan (2006)	Hong Kong banks from Hong Kong Monetary Authority filing, 1992–1999	Intermediation/cost	SFA	Inputs: labour, physical capital, borrowed money Outputs: loans to finance trade, non-trade-related loans, other earnings assets Control variables: bank size, deposit-to-asset ratio, ratio of trade-related loans to total assets, ratio of non-trade-related loans to total assets, ratio of loans loss provision to total loans, ratio of off balance sheet to total assets, loan growth	Following the Asian financial crisis, inefficiency increased slightly as banks were having difficulties in adjusting their inputs according to market conditions, in which loan demand fell and banks struggled by incurring more costs in monitoring and addressing worsening loan portfolios. The average cost inefficiencies for Hong Kong banks were 32.0%, where large and small banks exhibit average cost-inefficiency scores of 30.0% and 22.0% respectively. In Hong Kong, smaller banks were reported to have greater cost efficiency relative to larger banks.
Kasman and Yildirim (2006)	Commercial banks from 8 Central and Eastern European Countries (Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic, Slovenia), 1995–2002	Value added/cost and alternative profit	SFA	Inputs: price of labour, physical capital Outputs: total loans, total deposits, other earnings assets Environmental variables: population, income per capita, capital ratio, concentration ratio, intermediation ratio, inflation, M2/GDP, market capitalisation	Average cost and profit inefficiency scores were found to be at 20.7% and 37.0% respectively. Using the one-stage analysis, they found no relationship between size and cost- and profit efficiency. Foreign banks on average displayed lower levels of inefficiencies, where the average cost- and profit-inefficiency of foreign banks were at 20.0% and 35.0%, whereas domestic banks registered higher average cost- and profit-inefficiency at 21.0% and 38.0% respectively. Foreign banks have the ability to generate higher profits via high value-added segments of the market and provide innovative products and services demanded in the market.
Mobarek and Kolonov (2014)	307 conventional banks (CB) and 101 Islamic banks (IB) from 18 Organisation of Islamic Countries (OIC) banks:	Intermediation/ cost	SFA and DEA	Inputs: deposits, equities, personnel expenses and fixed assets Outputs: total loans and other earning assets	CBs shows higher cost efficiency compared to the IBs, particularly with regards to the analysis of bank size. Large CBs shows higher efficiency than large IBs as shown by these two methods of frontier estimations. Moreover, the result between SFA and DEA also exhibits high degree of consistency.

Note: SFA: Stochastic Frontier Analysis, DEA: Data Envelopment Analysis

3.4.3 Empirical Studies of Banking Efficiency in Malaysia

Empirical studies of Malaysian banks' efficiency are limited, but on an increasing trend. Most such studies were written in the late 2010s and have used frontier efficiency methods to measure performance. The studies were applied in several contexts and followed the development of the economic and banking sectors. Thus, this section discusses empirical studies of banking efficiency in Malaysian banks relating to the Asian financial crisis, the banking sector consolidation scheme (i.e. the effect of mergers and acquisitions), financial liberalisation initiatives (e.g. the ten-year Financial Sector Master Plan) and market competition and structure (i.e. the introduction of new foreign banks from both conventional and Islamic banking streams). In addition, a discussion of the limitations of past literature is presented at the end of this section.

Malaysia faced the Asian financial crisis in 1997, which affected the stock market and the value of the currency and created banking turmoil. There is a limited number of studies examining the impact of the financial crisis on bank efficiency. For example, Suffian (2009) employed DEA and the Malmquist Productivity Index to analyse the impact of the Asian financial crisis on Malaysian domestic and foreign commercial banks from 1995 to 1998. Suffian (2009) reported that over this period foreign and smaller banks in Malaysia exhibited a decline in productivity. However, the medium and larger domestic banks experienced productivity growth during the period under study. Furthermore, Sufian (2010a) also compared the cost efficiency levels of Malaysia and Thailand banks from 1992 to 2003 to investigate the impact of the Asian financial crisis on these two sectors. Sufian (2010a) found that Thai banks (81.6%) display a higher level of cost efficiency compared to Malaysian banks (66.3%). The technical efficiency of banks in both countries declined abruptly 12 months after the Asian financial crisis, due to extreme measures such as capital controls and the reduction of interest rates to revive the economies. Sufian (2010a) also suggests that productivity levels are positively related to loan intensity, non-interest income and overall profits, and negatively associated with bank size, expense preference behaviour and liquidity.

Subsequent to the crisis, the initial response by the Malaysian government was to tighten fiscal and monetary policy and prevent the exchange rate from depreciating further. The government of Malaysia therefore implemented capital controls (i.e. adopting a fixed exchange rate regime by pegging the Malaysian Ringgit to US Dollars) and restructuring the financial system. The rationalising of the financial system included bank merger initiatives, which resulted in a smaller number of banks and increased capital level for the banks. Also, the government created a special-purpose-vehicle (SPV) firm to purchase bad loans (Goh et al., 2003; Sufian, 2009) (see Chapter 4 for detailed discussion on government measures during Asian financial crisis). Thus, there are several studies in Malaysia investigating the impact of mergers on bank efficiency. For instance, Khrisnasamy et al. (2004) investigated the productivity changes of ten Malaysian domestic anchor banks from 2000 to 2001 using the Malmquist Productivity Index. Within this short period of comparison, they reported that total factor productivity increased by 5.1% in the eight banks that were found to experience technological progress. Additionally, Sufian (2004) employed data of ten consolidated Malaysian domestic banks from 1998 to 2003 to analyse the impact of merger programs initiated by the Malaysian government using the DEA model. Sufian (2004) found that during the merger year, the overall cost efficiency level dropped significantly but then recovered, registering higher overall efficiency relative to the pre-merger years. Likewise, Ismail and Abdul Rahim (2009) used data from 1995 to 2005 to examine the impact of mergers on the productivity and efficiency of domestic commercial banks. They report that merger initiatives improved the technical efficiency of banks where pre-merger and post-merger technical efficiency was at 67.6% and 95.2% respectively. Moreover, Abd-Kadir, Selamat and Idros (2010) analysed nine Malaysian commercial banks from 2003 to 2007 to measure the post-merger program productivity level of Malaysian banks. They reported that average total factor productivity (TFP)⁷⁸ of Malaysian banks after the consolidation program increased by 10.1%, due to a change in technical progress rather than in technical efficiency, suggesting that mergers and consolidation program of Malaysian banking industry resulted in technological innovations and advances, such as e-banking and automated teller machine (ATM) networks, rather than operational managerial efficiency. These findings by Abd-

⁷⁸ Total Factor Productivity can be described as index of “technological progress” which is defined as combination of output per unit of labour and capital. TFP index can be written as $A = Q/(aL + bK)$, where Q is the aggregate level of output, L is labour input, K is capital input, and a and b refer to weights (Nadiri, 1970).

Kadir et al. (2010) can be supported from the empirical studies of Sufian and Habibullah (2009), Sufian (2006) and Radam et al. (2009), who found that the impact of mergers can increase the performance of banks. These empirical studies also show that most of the growth in productivity is supported by technological advances, instead of managerial superiority. By contrast, from the data of ten consolidated domestic banks from 1998 to 2004, and based on DEA estimation, Mohd Said et al. (2008) found very little difference in average efficiency scores between pre- and post-merger, registering 97.713% and 97.712% respectively. They claimed that merger activities in Malaysia provided little impact and did not enhance the productivity efficiency of Malaysian banks.

Accordingly, in 2001, Bank Negara Malaysia (BNM) introduced the FSMP. There is only one study that looks at the impact of the FSMP on the efficiency of Malaysian banks. Hon et al. (2011) investigated the impact of the ten-year plan on ten Malaysian consolidated banks from 2001 to 2005 using the DEA model and reported that the banks exhibited efficiency improvement due to the FSMP initiatives. They also suggested that Malaysian banks need to increase their efficiencies to between 11.0% and 24.0% to be fully efficient. However, deterioration in scale-efficiency was observed, which was caused by the merger programme (Hon et al., 2011). The mergers have caused the banks that are too large to be efficient, where the managers may not work as hard to maximise profit and enjoying ‘quiet life’ (Hicks, 1935). Nevertheless, this study may not fully reflect liberalisation initiatives due to the short time period under investigation (i.e. 5 years’ data from 2001 to 2005) and a limited number of banks (i.e. ten domestic banks).

One of the aims of the FSMP was to liberalise the banking industry with the introduction of new foreign banks. There are several studies examining the relationship between foreign and domestic banks and the efficiency of Malaysian banks. For example, Mohd Tahir et al. (2008, 2009 and 2010) produced a series of three empirical studies using data from 22 Malaysian domestic and foreign banks (nine domestic and thirteen foreign banks) from 2000 to 2006. In the first study, Mohd Tahir et al. (2008) investigated the cost function using the parametric SFA and reported that domestic banks (average CE: 90.9%) are more efficient than foreign banks (average CE: 74.4%). The second study (Mohd Tahir et al. 2009) utilised the DEA cost function and found that domestic banks (average CE: 88.7%) still

registered higher efficiency scores compared to foreign banks (average CE: 73.3%). The third in the series of studies involved the investigation of both cost and profit efficiency using parametric SFA. This study, (Mohd Tahir et al. 2010) reported that domestic banks were more cost-efficient, but lower in terms of alternative profit than foreign banks. The alternative profit efficiency of foreign banks (average PE: 76.9%) is greater than that of domestic banks (average PE: 63.8%). Moreover, Suhaimi et al. (2010) also found that the profit efficiency of foreign banks was higher than that of domestic banks. Sufian (2011) also employed the DEA model to examine 36 and 23 Malaysian commercial banks in 1995 and 2004 respectively, and investigated the impact of the ownership structure on banks' performance. Like Mohd Tahir et al. (2009), he found that foreign banks were less productive compared to domestic banks. Sufian (2011) argues that this finding is surprising, particularly on technology-advanced banks from developed countries that were unable to adapt and overcome the 'home field' advantage in the developing economy. Further examination suggests that foreign banks originating in Europe are less productive than their domestic and other foreign banks from US and Asia, which supports the 'limited global advantage' hypothesis. In contrast, Matthews and Ismail (2006) examined Malaysian domestic and foreign banks from 1994 to 2000 using the DEA. From their result, they found the efficiency of foreign banks was greater than that of domestic banks.

The Islamic banking development is also a key part of many initiatives within the FSMP. The key initiatives found during the implementation of FSMP are through (1) the conversion of banks in the Islamic banking scheme (IBS), which uses the 'Islamic window' banking structure,⁷⁹ into full Islamic subsidiaries under the Islamic Banking Act 1983 and (2) the introduction of new foreign Islamic banks to foster competition on domestic Islamic banks. Therefore, there are several studies that examine the efficiency of Islamic banks in Malaysia. For example, Abdul-Majid et al. (2011) analysed the efficiency levels of Malaysian banks practising Islamic banking using the parametric SFA with translog specification model. The cost efficiency results show that banks with Islamic banking operations display 34.0% greater cost efficiency than other banks without Islamic banking activities. In addition, Omar et al. (2006) examined the cost efficiency of Malaysian

⁷⁹ 'Islamic window' refers to the operation of Islamic banking within a conventional bank. The operation of an Islamic window requires the bank to establish necessary firewalls to ensure that commingling between Islamic and conventional funds does not occur (Sole, 2007).

commercial banks from 2000 to 2004, consisting of eleven Malaysian commercial banks utilising the nonparametric DEA model. Similarly, they found Islamic banks (domestic Islamic banks prior to offering licences to foreign banks) to be less efficient than conventional banks. Furthermore, Kamaruddin et al. (2008) employed DEA on twelve Malaysian banks (both domestic and foreign banks) with Islamic Banking operation window (IBS) and two full-fledged Islamic banks from 1998 to 2004 to analyse the degree of both profit- and cost-efficiency in Islamic banks in Malaysia. They reported that the average cost efficiency (97.0%) is greater than the average profit efficiency (85.3%) of Islamic banks in Malaysia, which implies that Malaysian banks are better in managing cost than profit (Kamruddin et al., 2008). Another example of domestic and foreign banks' efficiencies in Malaysia is regarding the area of Islamic banking where Sufian and Habibullah (2010) investigated the impact of foreign Islamic and domestic Islamic banks. They employed the DEA model on Malaysian foreign and domestic Islamic banks from 2001 to 2008. Sufian and Habibullah (2010) found that domestic Islamic banks perform better than their foreign bank counterparts, registering at 76.2% and 70.6%, respectively. Within the period under study, BNM issued three new Islamic bank licenses to foreign banks. These three *de novo*⁸⁰ foreign Islamic banks exhibit lower efficiency compared to established foreign and domestic Islamic banks.⁸¹

There are also several studies that fall under different contextual topics, which can range from inherent risk to market structure. For example, Abd Karim, Chan and Hassan (2010) compared the cost efficiency of banks in Malaysia and Singapore and their relationship to non-performing loans. Using stochastic frontier analysis, they reported that Singaporean banks are relatively more efficient than Malaysian banks registering an average cost efficiency of 93.5% and 86.8% respectively. They found a negative relationship between cost efficiency and non-performing loans. Dogan and Fausten (2002) investigated cost efficiency in 21 Malaysian banks from 1989 to 1998. Using DEA, the results were used to estimate the Malmquist Productivity Index and measure the productivity growth of Malaysian banks. Dogan and Fausten (2002) found that Malaysian banks' productivity

⁸⁰ For example, the Federal Reserve Bank defines a De Novo bank as a state member bank that has been in operation for five years or less.

⁸¹ De Young and Hassan (1998) assert that De Novo banks in the US will normally show reasonable profits for the first three years and require nine years to be on a level playing field with other established banks.

declined between 3.3% and 5.6% during the period of study. The deterioration of productivity was due to the insulation and protection of the banking industry from foreign bank competition. Sufian and Abdul Majid (2007) analysed the cost- and profit-efficiency of Malaysian banks listed in the Kuala Lumpur Stock Exchange (KLSE) from 2002 to 2003 by applying DEA. They reported that Malaysian banks exhibited a higher degree of cost efficiencies when compared with the level of profit efficiencies. They also found that larger banks registered a higher degree of cost, but lower profit efficiencies, than smaller banks.

As mentioned previously, most of the banking efficiency studies on Malaysian banks were published in the late 2010s and applied in several contexts such as financial crises, mergers, financial liberalisation, ownership, bank type, and risks (broadly categorised by Berger and Humphrey (1997) under ‘inform government policy’). Based on the literature reviewed, there are areas that still need to be fully addressed in relation to the efficiency of Malaysian banks. First, the impact on the efficiency of Malaysian banks during the implementation of liberalisation (i.e. FSMP) has not been investigated. Hon et al. (2011) did attempt to examine the impact of the FSMP, but the data set they employed does not span the entire timeline of the FSMP and may not fully reflect the competitive aspects of the banking system by comparing only domestic banks. Second, the majority of empirical studies on Malaysian banks focus mainly on the cost function and a limited number of empirical studies looking into profit functions. There are even fewer studies that conduct and compare both cost and profit functions using the same data set of Malaysian banks. Third, to the author’s knowledge, there are no studies in Malaysia that perform a comprehensive consistency analysis between nonparametric and parametric methods. Bauer et al. (1998) suggested six consistency conditions rather than only concentrating into rank-order correlation for consistency. The result from these consistency tests is useful for policy makers in making decisions. Fourth, the timeline of data used for most Malaysian banking efficiency studies is usually short (on average, between three and eight years) with a smaller number of banks as a sample, which may not represent the entire banking system. Fifth, studies on explanatory factors or the determinants of inefficiency are also lacking, which may result in a failure to explain the factors that affect the inefficiency of Malaysian banks. Sixth, there are few studies on the efficiency of Malaysian banks that cover the impact of the global credit crisis occurred during the implementation of FSMP. Therefore, this

research addresses some of the gaps observed from the Malaysian banking efficiency studies (discussed in detail in Chapter 5).

This present research will examine the impact of liberalisation through the ten-year FSMP, employing data of all Malaysian banks (foreign – conventional and Islamic banks and domestic – conventional and Islamic banks) from 2000 to 2011. In terms of estimation, both cost and profit (standard profit and alternative profit) efficiency will be utilised on parametric (SFA as main benchmark) and nonparametric (DEA as empirical alternative) approaches for consistency assessment. At the same time, the determinants of inefficiency using both control and environmental variables will also be tested using the one-stage approach on SFA. Consequently, the details for measurement of the efficiency estimation, data, model specification and variables such as input, output, control and environmental variables will be discussed in chapter 5 ('Methodology and Data').

3.5 Conclusion

This chapter has discussed the importance of the banking system in influencing the growth and development of the economy. In addition, it also provided relevant information regarding financial repression and financial liberalisation of the banking industry. How frontier efficiency analysis is being used in many different applications was also discussed. Based on the reviews made of banking efficiency, there exists extensive research into financial liberalisation or deregulation. Empirical studies from developed countries provide a strong foundation of frontier efficiency techniques that apply in many banking areas, most of which are of interest to regulators/policy makers. There are a limited number of studies on banking efficiency in developing countries, but these are increasing. The growing interest in developing countries is attributed to various changes occurring in the market structure, particularly in relation to financial liberalisation and globalisation. Moreover, developing countries have been investigated because of their different market structure, compared to developed countries where the banking system is usually less developed and subject to more government control.

This chapter also reviewed empirical studies of developed countries, developing countries and Malaysia regarding the impact of financial liberalisation, market structure, ownership, size and inherent risk in the banks. These factors are generally considered as the key determinants that correlate with banking efficiency. Based on the discussions in this chapter, most of the results from these studies generally exhibited mixed findings. For example, some studies found a positive impact results from financial liberalisation, which increases the efficiency of the banks, while others report little impact or negative impact on the banks' efficiency level. Similarly, a more competitive banking environment can increase the efficiency of banks; however, the results dealing with market structure and concentration also vary in both developed and developing countries. From the empirical studies reviewed, many studies found that domestic banks are more efficient than foreign banks in the developed countries. Nonetheless, in the developing/transitional countries, foreign banks display greater efficiency than domestic banks. Moreover, there are many studies that examine the factors that correlate with the efficiency of banking institutions in developing countries, of which many found that efficient banks generally have a greater market share, fewer non-performing loans, more capital, and a higher level of liquidity.

In addition, past empirical studies in Malaysian banking displayed similar trends to those above; for example, the impact of financial liberalisation on the efficiency of banks presents mixed findings. However, most studies show a positive impact from the introduction of financial liberalisation in Malaysia. Additionally, contrasting findings are shown regarding bank ownership factors. Unlike the typical findings in other developing countries, Malaysian banking exhibits greater efficiency in domestic banks compared to foreign banks, which is attributed to more regulatory controls on foreign banks operating in Malaysia.

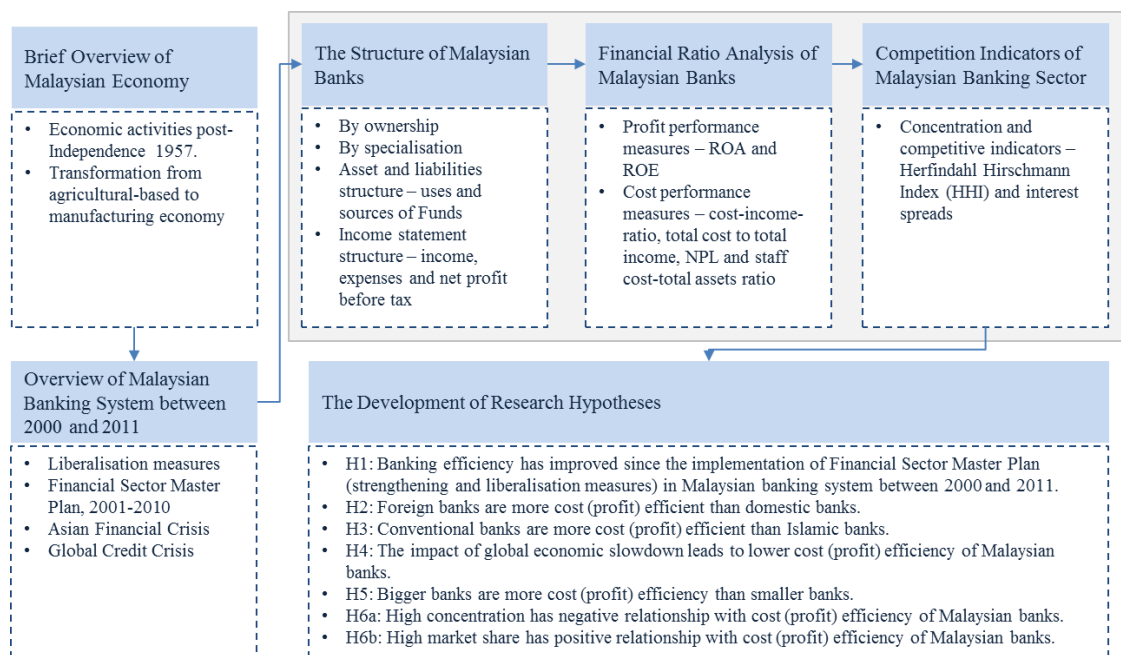
The following chapter discusses the structure of the banking system in Malaysia and the gradual financial liberalisation initiatives introduced through a ten-year Financial Sector Master Plan. Concurrently, the impact of liberalisation will also be measured using the financial items derived from the income statements and balance sheets of the banks. Additionally, research hypotheses are also formulated using the discussion presented in this chapter, chapter 2 ('Theory and Measurement of Production Efficiency') and chapter 4 ('Malaysian Banking System and Financial Liberalisation').

Chapter 4 The Malaysian Banking Sector, 2000–2011

4.1 Introduction

As discussed in the previous chapter, an efficient and competitive banking system can lead to a higher degree of economic growth, greater product innovation, easier access to financial services and improved financial stability (Molyneux et al., 1996). Therefore, this chapter explains the liberalisation measures that have been introduced in Malaysia, which have aimed to improve the competitive environment and efficiency of its banking sector. In addition, it also discusses the institutional mechanism, evolution and structure of the Malaysian banking sector. Furthermore, this chapter also provides the necessary information to understand the empirical analysis of the efficiency of the Malaysian banking sector in subsequent chapters.

Section 4.2 discusses a brief overview of the Malaysian economy and is followed by Section 4.3, which presents an overview of the Malaysian banking system and salient measures introduced to liberalise the banking sector. Section 4.4 briefly reviews the structure of Malaysian banks, particularly in terms of ownership and specialisation. Moreover, this section also provides the attributes of Malaysian banks based on their balance sheet and income statement positions, which are affected by the economic and liberalisation changes between 2000 and 2011. Section 4.5 presents the performance of the Malaysian banks based on the financial ratio analysis. Section 4.6 reviews the competitive conditions of the Malaysian banking system. At the end of this chapter, Section 4.7 presents the development of hypotheses regarding the impact of financial liberalisation on the Malaysian banking sector. Figure 4.1 outlines the visual information flow of this` chapter ('The Malaysian Banking Sector, 2000–2011').

Figure 4.1 Information Flow of Chapter 4

4.2 Brief Overview of the Malaysian Economy

Malaysia (formerly Malaya) is a relatively small country located in South East Asia. The land area of Malaysia covers approximately 330,000 square kilometres with Thailand, Singapore and Indonesia as close neighbours. Malaysia is divided into two distinct parts, which are West Malaysia (peninsular) and East Malaysia (Sabah and Sarawak in the Borneo Island). Malaysia gained its independence in 1957 after being colonised for over 400 years, since the invasion of Malacca by the Portuguese in 1511. Consequently, the British came to open up Penang in the 19th century. With colonisation by the British and other foreign countries, Malaysia has become a multiracial country⁸² with a population of 28.3 million in 2010, of which around more than 67.4% are Malays or Bumiputera (sons of the soil), 24.6% are Chinese, and 7.3% are Indians, and 0.7% are from other ethnic groups (Department of Statistics Malaysia, 2010).

⁸² Under British rule, Chinese and Indians were brought into Malaya to work in tin mining and rubber plantation estates, respectively. These two main economic activities were new to the Malays, who were not familiar with them (Jomo and Hui, 2003).

Malaysia is a small, open economy and susceptible to changes in the global economy. It is an upper-middle income economy with a USD 7,900 per capita nominal income (World Bank, 2011). The country maintains an open and liberal trading regime, which is supported by longer-term adjustment policies. For example, Malaysia implements a series of different five-year plans for its economic development, which transformed the country from an agricultural and mining-based economy to an economy based on manufacturing and services (Bank Negara Malaysia, 1999b).

After independence in 1957, Malaysia relied on its agricultural sector; with rubber as its main commodity, contributing more than 30.0% of GDP and more than 50.0% of export earnings in 1960s. Rubber and tin were the major commodities from which the British have benefited the most in the colonial era. Malaysia relies on its New Economic Policy (NEP) and its five-year national development plans for economic transformation. However, as commodity prices fell in the first half of the 1980s, the country has diversified its sources of growth to improve productivity. The country has shifted its economic strategy from an agriculture-based sector to its manufacturing sector. To materialise this economic strategy, the government of Malaysia implemented institutional reforms on the public sector, liberalised restrictive laws and offered fiscal incentives to boost business confidence and reduce the cost of operating in Malaysia (Bank Negara Malaysia, 1999b). This transformation attracted foreign direct investment (FDI) and the private sector to invest in Malaysia, which would be the main driver for Malaysian economic growth (Bank Negara Malaysia, 1999b).⁸³ In addition, Malaysia introduced ‘Vision 2020’ in 1990 with an aspiration to become a developed country by the year 2020. With such vision, from a producer of raw materials such as tin and rubber in the 1970s, Malaysia is progressing to becoming a multi-sector economy. Currently, Malaysia can be described as a leading exporter of electrical components and appliances, palm oil and natural gas (World Bank, 2011).

Malaysia’s recent development can be observed from its gross domestic product (GDP); Malaysian average annual GDP was 7.2% from 1961 to 1980; and for the years 1981 to

⁸³ To support this economic reform, the government introduced the Investment Incentive Act 1968 and Promotion of Investment Act 1986, which support a range of new tax incentives for the manufacturing sector.

1996, the Malaysian GDP has increased to an annual average of 7.4%. The strong economic growth in 1981–1996 was influenced by large domestic capital accumulation and a large influx of foreign direct investment (FDI) (Athukorala and Menon, 1995, Athukorala, 2002). The rapid increase of FDI in the period mentioned above is a result of the introduction of the Investment Incentive Act 1968, export incentive provisions, economic policy reform, and the establishment of free trade zones (Ang 2008, 2009). Moreover, sound macroeconomic management, sustained economic growth, low set-up costs, adequate human resources, stable political environment and a well-functioning financial system have also attracted large amount of FDI into Malaysia (Ang 2008, 2009). However, from 1997 to 2003, average annual GDP declined to 3.5% due to the impact of Asian financial crisis (Ang, 2007) and then recovered between 2000 and 2008 in which Malaysia registered an average annual growth in GDP of 5.5%. Consequently, the Malaysian GDP for 2009, 2010 and 2011 were at -1.7%, 7.2% and 5.6% respectively. The contraction of GDP by 1.7% in 2009 was a result of impact from global credit crisis (discussed later in this chapter). In addition, the country's poverty level has also declined from 12.3% in 1984 to 2.3% in 2009 (World Bank, 2011).

4.3 Overview of the Malaysian Banking System between 2000 and 2011

Like many other developing and emerging economies, Malaysia's financial system is a bank-based system.⁸⁴ Malaysia inherited a foreign-dominated banking system at the time of independence in 1957 after the British colonial administration left Malaysia. The Bank Negara Malaysia (BNM) (i.e. the Central Bank of Malaysia) was established in 1959. At this time there were 18 foreign commercial banks with a total of 99 branches operating in the country, as opposed to only eight domestic commercial banks with in total only 12 branches. The foreign banks held over 90.0% of the share of the market in 1957 (Detgriache and Gupta, 2006). The foreign banks (mainly British-owned

⁸⁴ A bank-based system is one in which banks take a leading role in mobilising savings, allocating capital, monitoring and exercising corporate control and providing portfolio management based on savers' appetite for risk. In contrast, the market-based system refers to a system with the securities market as the main driver alongside banks in getting society savings into firms and exerting corporate control (Levine, 2002). In Malaysia, the banking system remains the main source of finance. Nevertheless, the capital market has also gained importance; both the banking and financial market systems continue to play a role in providing sources of finance to Malaysian firms (Ibrahim, 2011).

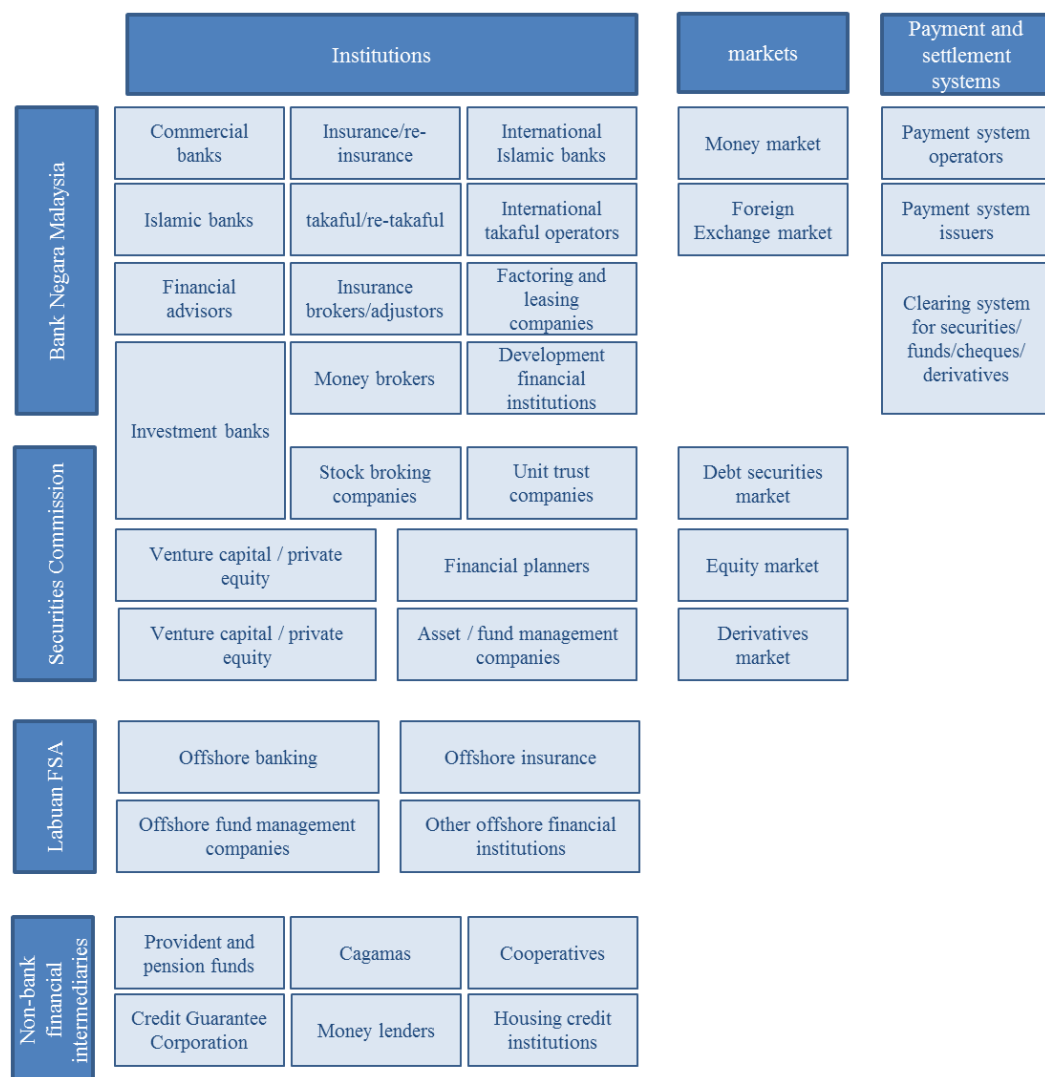
banks) established before independence specialised in foreign exchange business, finance in foreign trade, rubber plantation and tin mines (Matthews and Ismail, 2006). At the stage of initiation, the main focus of BNM was to promote and protect the domestic banks. In order to achieve this objective, regulatory restrictions were imposed on the establishment of new foreign banks and their branches (Abdul Majid et al., 2011). Foreign banks have been prohibited to open new branches since 1971 and the last foreign bank allowed to enter the market was the Bank of Nova Scotia in 1973. As a result, the number of domestic banks surpassed the number of foreign banks in 1976 (Ibrahim, 2011, Sufian, 2011).⁸⁵ However, the number of domestic banks declined by the end of 2001 following the merger and consolidation exercises encouraged and coordinated by the government and BNM after the Asian financial crisis (Matthews and Ismail, 2006; Bank Negara Malaysia, 2002, 2011a).

Banking institutions in Malaysia are regulated by the Banking and Financial Institutions Act 1989 (BAFIA 1989) (except for Islamic banks, which are regulated under Islamic Banking Act 1983). This Act was introduced primarily due to the failure of a number of deposit-taking cooperatives in the 1980s, which consequently threatened the stability and integrity of the banking system (Bank Negara Malaysia, 1999b). Therefore, this Act was increasingly necessary to support the growing competition in the banking system and to define the boundary banking institutions' operations. One of the objectives of BAFIA (1989) is to provide an integrated supervision of the banking institutions and to improve the current laws of banking operations. The Act also includes different related institutions such as discount houses, and money and foreign exchange brokers under the single supervisory and regulatory regime of the Central Bank (BNM).

⁸⁵ By 1994, all foreign bank branches were required to be locally incorporated. A locally incorporated foreign bank is a subsidiary (separate entity) owned by a foreign parent bank in the host country and required to operate under the regulation of the host country. In contrast to locally incorporated foreign banks, foreign bank branches are obliged to follow both home and host country regulation, which relies on the parent bank's capital (Ball and Tschoegl, 1982). Both foreign banks' structures have implications for the competitive structure of the host country's banking system, which also affects the price and quality of banking services. For example, a foreign subsidiary can apply direct competition with domestic banks and ring-fence their operation from parent banks. On the other hand, the operation of foreign branches is normally underdeveloped, which concentrates on certain segments of business in the host country (Cerutti et al., 2007).

At the end of 2011 (see Figure 4.2), the structure of the Malaysian financial system could be divided into three main sections: financial institutions, non-financial institutions and financial markets. Under the banking sector, the main players under the jurisdiction of BNM are the commercial banks, Islamic banks, Insurance and Takaful companies, and investment banks. However, investment banks in Malaysia are jointly supervised by BNM and the Securities Commission (SC) due to the overlapping nature of their activities. On the other hand, the main financial markets that affect the banking sector in Malaysia are the money and foreign exchange markets.

Figure 4.2 The Structure of the Malaysian Financial System as of December 2011



Adapted from Bank Negara Malaysia, 2011b

In Malaysia, the banking system is the largest component, accounting for approximately 50.0% of the total assets of the Malaysian financial system (Bank Negara Malaysia, 2011a). Malaysian banks have been the largest mobilisers of deposits, enabling them to perform intermediation functions and contribute to economic growth (Loke, 2011). The banking and financial system accounted for 115.0% of Malaysian gross domestic product (GDP) in 2011 and has grown by an average annual rate of 7.1% since 2000 (Bank Negara Malaysia, 2012a). At the end of 2011, there were 56 banks operating in Malaysia including 25 commercial banks, 16 Islamic banks and 15 investment banks (see Table 4.1).

Table 4.1 Malaysian Banks as of December 2011

Commercial Banks	
1. Affin Bank Berhad	14. Citibank Berhad *
2. Alliance Bank Malaysia Berhad	15. Deutsche Bank (Malaysia) Berhad *
3. AmBank (M) Berhad	16. HSBC Bank Malaysia Berhad *
4. CIMB Bank Berhad	17. Industrial and Commercial Bank of China (Malaysia) Berhad *
5. Hong Leong Bank Berhad	18. J.P. Morgan Chase Bank Berhad *
6. Malayan Banking Berhad	19. Mizuho Corporate Bank (Malaysia) Berhad *
7. Public Bank Berhad	20. OCBC Bank (Malaysia) Berhad *
8. RHB Bank Berhad	21. Standard Chartered Bank Malaysia Berhad *
9. Bangkok Bank Berhad *	22. Sumitomo Mitsui Banking Corporation Malaysia Berhad *
10. Bank of America Malaysia Berhad *	23. The Bank of Nova Scotia Berhad *
11. Bank of China (Malaysia) Berhad *	24. The Royal Bank of Scotland Berhad *
12. Bank of Tokyo-Mitsubishi UFJ (Malaysia) Berhad *	25. United Overseas Bank (Malaysia) Berhad *
13. BNP Paribas Malaysia Berhad *	
Islamic Banks	
1. Affin Islamic Bank Berhad	9. Public Islamic Bank Berhad
2. Alliance Islamic Bank Berhad	10. RHB Islamic Bank Berhad
3. AmIslamic Bank Berhad	11. Al Rajhi Banking & Investment Corporation (Malaysia) Berhad *
4. Bank Islam Malaysia Berhad	12. Asian Finance Bank Berhad *
5. Bank Muamalat Malaysia Berhad	13. HSBC Amanah Malaysia Berhad *
6. CIMB Islamic Bank Berhad	14. Kuwait Finance House (Malaysia) Berhad *
7. Hong Leong Islamic Bank Berhad	15. OCBC Al-Amin Bank Berhad *
8. Maybank Islamic Berhad	16. Standard Chartered Saadiq Berhad *
Investment Banks	
1. Affin Investment Bank Berhad	9. Kenanga Investment Bank Berhad
2. Alliance Investment Bank Berhad	10. Maybank Investment Bank Berhad
3. AmInvestment Bank Berhad	11. MIDF Amanah Investment Bank Berhad
4. CIMB Investment Bank Berhad	12. MIMB Investment Bank Berhad
5. ECM Libra Investment Bank Berhad	13. OSK Investment Bank Berhad
6. Hong Leong Investment Bank Berhad	14. Public Investment Bank Berhad
7. Hwang-DBS Investment Bank Berhad	15. RHB Investment Bank Berhad
8. KAF Investment Bank Berhad	

* Foreign-owned banks

Source: Bank Negara Malaysia Monthly Statistical Bulletin – December 2011

Since independence in 1957, the Malaysian banking system has been regulated by Bank Negara Malaysia (BNM), which was established on 26 January 1959 under the Central Bank of Malaya Ordinance, 1958 (CBO). The CBO was revised in 1994 as the Central

Bank of Malaysia Act, 1958 (CBA, 1958). The CBA defines BNM as the bank that governs the monetary and banking structure of the country. In 1989, the Banking and Financial Institution Act, 1989 (BAFIA, 1989) was introduced, extending BNM's powers of supervision and regulation of financial and deposit-taking institutions (Bank Negara Malaysia, 1999b). Then in 2009, the CBA (1958) was repealed with the introduction of Central Bank of Malaysia Act, 2009 (CBA, 2009), which came into force on 25 November 2009 (Bank Negara Malaysia, 2009a). In the new CBA, the mandates of BNM have been broadened, empowering the Central Bank to:

1. Formulate and conduct monetary policy;
2. Issue currency in Malaysia;
3. Regulate and supervise financial institutions based on laws enforced by BNM;
4. Provide oversight over money and foreign exchange markets;
5. Exercise oversight over payment systems;
6. Promote a sound, progressive and inclusive financial system;
7. Manage Malaysian foreign reserves;
8. Promote an exchange rate regime according to the fundamentals of the economy;
and
9. Act as a financial adviser, banker and financial agent of the Government.

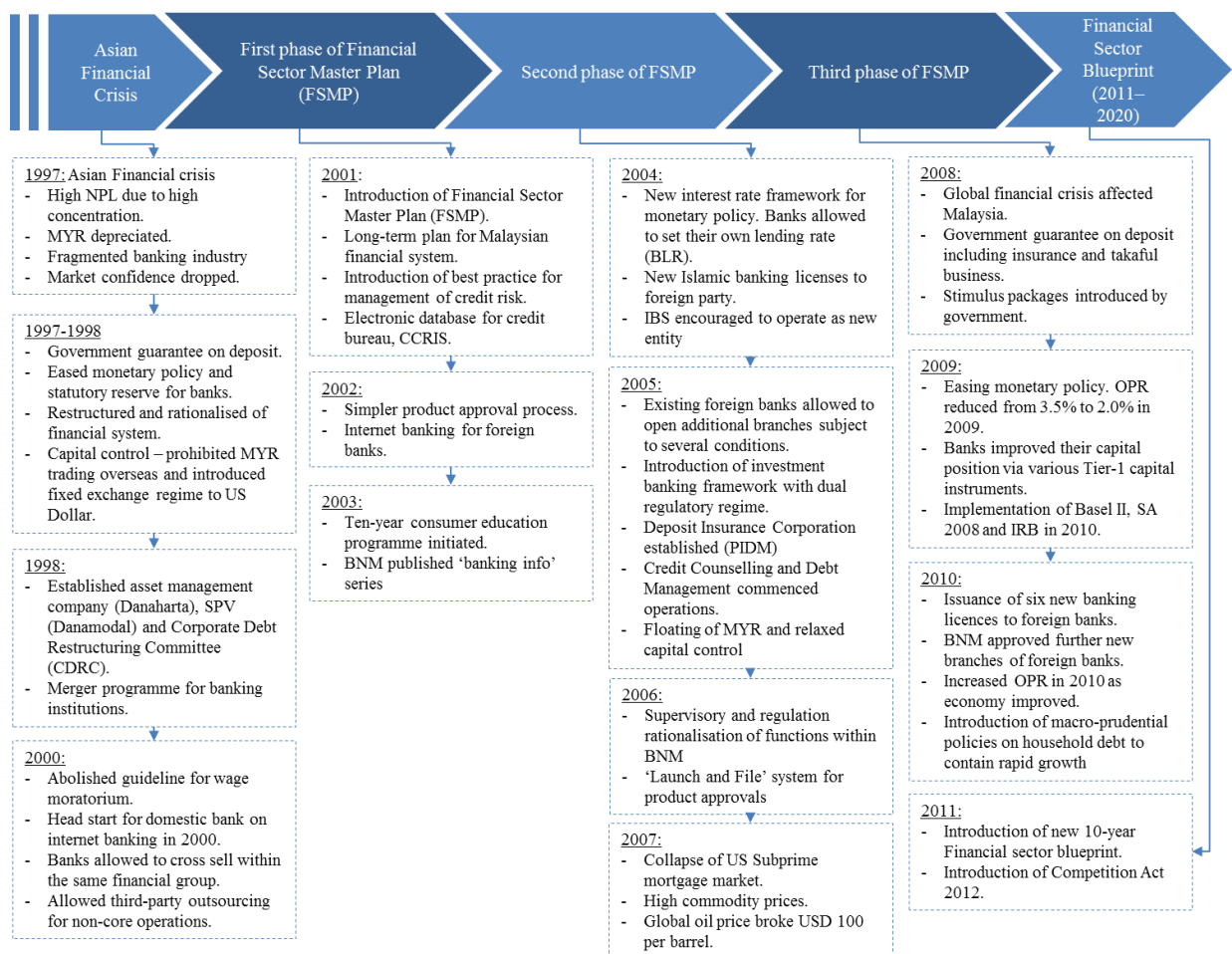
The mandates in CBA 2009 are inter-related and complementary. As the issuer of currency, BNM has to ensure that the domestic prices remain stable and the benefit of economic growth is not eroded (Bank Negara Malaysia, 1999b). At the same time, BNM is required to promote monetary stability by pursuing a monetary policy that serves the interests of the country with the primary objective of maintaining price stability while giving due regard to developments in the economy. To support price stability, a stable financial infrastructure is required to ensure that the monetary policy measures implemented are effective. This will allow the smooth functioning of the

intermediation process so that domestic savings are mobilised and transmitted to investors and savers, thereby contributing to overall growth in investment and output. In carrying out its financial stability mandate, BNM has the responsibility and powers to minimise the likelihood and impact of adverse developments that may affect financial stability including the financial intermediation process, the orderly functioning of the foreign exchange and money markets and public confidence in the financial system. Concurrently, BNM continues to act as lender of last resort and provide the necessary liquidity to contain liquidity shocks (Bank Negara Malaysia, 1999b, 2009a, 2009b). Additionally, recognising the importance of monetary policy for price stability, it is also essential that there is coordination with fiscal policy to achieve an optimal mix of both policies for price stability. Therefore, BNM continues its role to act as financial adviser, banker and financial agent to the government of Malaysia (Bank Negara Malaysia, 2009a).

The regulations and liberalisation measures introduced by BNM during the years 2000 to 2011 are based on the economic and banking market conditions (see Figure 4.3). Before the introduction of Financial Sector Master Plan (FSMP) in 2001, Malaysia faced the Asian financial crisis in 1997–1998, which started in Thailand and has been rapidly transmitted to Malaysia. The crisis initially started as a currency crisis and subsequently affected the stock market, which resulted in a sharp fall of stock indices in Kuala Lumpur Stock Exchange (KLSE) (Sufian, 2009). Hence, investors started to withdraw their positions in the stock market and dispose of Malaysian Ringgit, which consequently resulted in ‘twin crises’ (the currency and banking crises) (Jomo and Chin, 2001). This eventually escalated into an economic crisis, which affected the corporate sector and the domestic economy. At the onset of the Asian financial crisis, Malaysian banks’ credit exposure was largely from the corporate sector, the broad property sector (including the construction sector) and the purchase of securities. As the ability of the corporate sector to service its debt obligations deteriorated, non-performing loans (NPL) increased subsequently. This resulted in a deterioration in asset quality and weakened banks’ balance sheet position, creating insolvency threats (Rajoo, 2008). Additionally, due to declining public confidence, depositors also shifted their funds from smaller banks to larger banks; in view that smaller banks experienced heavy withdrawals even

though they were solvent. This resulted in liquidity problems in the smaller banks (Bank Negara Malaysia, 1997). During this period of uncertainty, banks were unable to withstand these economic shocks due to poor risk management, weak asset management, inadequate internal policies and excessive lending concentrated in large corporates (Ang and McKibbin, 2007; Said et al., 2008).⁸⁶

Figure 4.3 Overview of Malaysian Banking Sector Development, 1997-2011

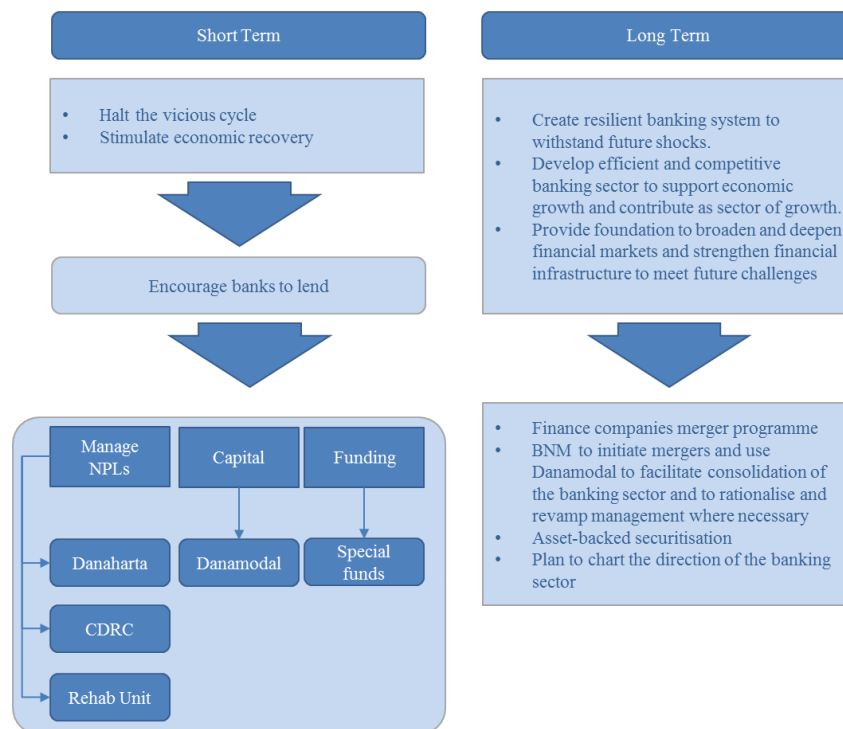


Note: This figure is prepared using various sources from Bank Negara Malaysia Annual Reports (1997-2011), Bank Negara Malaysia Financial Stability and Payment System Reports (2006-2011), Financial Sector Master Plan (2001), Financial Sector Blueprint (2011), and various published literature on Malaysian general banking, banking efficiency, and economic performance.

⁸⁶ The weaknesses of the banking system in Malaysia during the Asian financial crisis are outlined by Sulaiman and Govindan (2000); first, in getting funds for the private sector the economy is dependent on the banking system, which leads to a funding mismatch and undue interest rate risks. Second, collateral is used as a basis for granting loans rather than credit assessment, encouraging banks to take excessive risks particularly when collateral value deteriorates. Third, substantial inflows of foreign funds are invested in share markets and properties, resulting in volatility in asset prices.

Subsequent to the crisis, the government provided a generalised guarantee for bank depositors in early 1998, followed by an easing of monetary policy (i.e. reducing the interest rate) and reducing the reserve requirement of the banks to support the liquidity of the banking system (Detragiache and Gupta, 2006; Goh et al., 2003). In addition to the above, two policy responses were introduced at this stage; first, the restructuring and rationalising of the financial system and second, implementing capital controls, by prohibiting the internalisation of Malaysian Ringgit and adopting a fixed exchange rate regime, by pegging the Malaysian Ringgit to US Dollars (Goh et al., 2003).⁸⁷ Two special purpose vehicle (SPV) entities were established to support the policy measures taken and ensure that banks continued to focus on lending activities to support economic growth: first, Pengurusan Danaharta Nasional Berhad (Danaharta), which was in charge of buying non-performing loans (NPL) at a discount from banks; and second, Danamodal Nasional Berhad (Danamodal), which provided new capital in selected institutions (see Figure 4.4) (Detragiache and Gupta, 2006; Sufian, 2009; Abdul Majid et al., 2011). By 1999, the Malaysian economy rebounded after a sharp downturn in 1998. The consolidation programme was also completed in 2001 except for a couple of banks that merged in 2003 due to unresolved shareholders issues. As a result, the number of domestic banks contracted from 23 to 10 banking groups (Abdul Majid et al., 2011; Hon et al., 2011; Ramlee and Mohd Said, 2009).

⁸⁷ First, structural changes were made to rationalise the financial system (where banks are consolidated to reduce their number, consequently increasing the level of banks' capital). The purpose of this initiative was to prepare the domestic banks to face competition when the banking sector is liberalised under the World Trade Organization's provisions (Goh et al., 2003; Khrishnasamy et al., 2004). The consolidation initiative also broadens the product mix, reduces cost and increases the size of capital and assets of the domestic banks to enable them to compete against foreign banks in the local as well as in the international market (Tahir et al., 2008). Second, the Malaysian government also adopted a fixed exchange rate regime during the Asian financial crisis. This initiative helped Malaysia to insulate the economy from external uncertainties. The step taken to peg the Malaysian Ringgit was intended to allow a decline in interest rates without affecting the value of Malaysian currency (Detragiache and Gupta, 2006).

Figure 4.4 Restructuring Plan for the Malaysian Banking Sector

Adapted from Bank Negara Malaysia Annual Reports (Bank Negara Malaysia, 2000, 2001a)

Based on Figure 4.4, one of the long-term objectives of the banking sector restructuring was to provide a future direction for the Malaysian banking system. Therefore, in March 2001, BNM introduced the Financial Sector Master Plan (FSMP). The FSMP set the medium- and longer-term agenda to build a financial sector that is resilient, efficient and competitive, and responsive to changing requirements (Bank Negara Malaysia, 2001a, 2001b). The FSMP played an important role in the liberalisation process of the financial system covering the banking sector, the insurance sector, the Islamic banking sector, financial development institutions, alternative modes of financing and the Labuan International Offshore Financial Centre. The time-frame of the plan spanned ten years in three major sequenced phases. The FSMP adopted a gradual approach, allowing sufficient time for a sequenced and comprehensive development of the

financial sector (see Figure 4.5) (Bank Negara Malaysia, 2001b).⁸⁸ There were 119 recommendations over six sectors with following objectives:

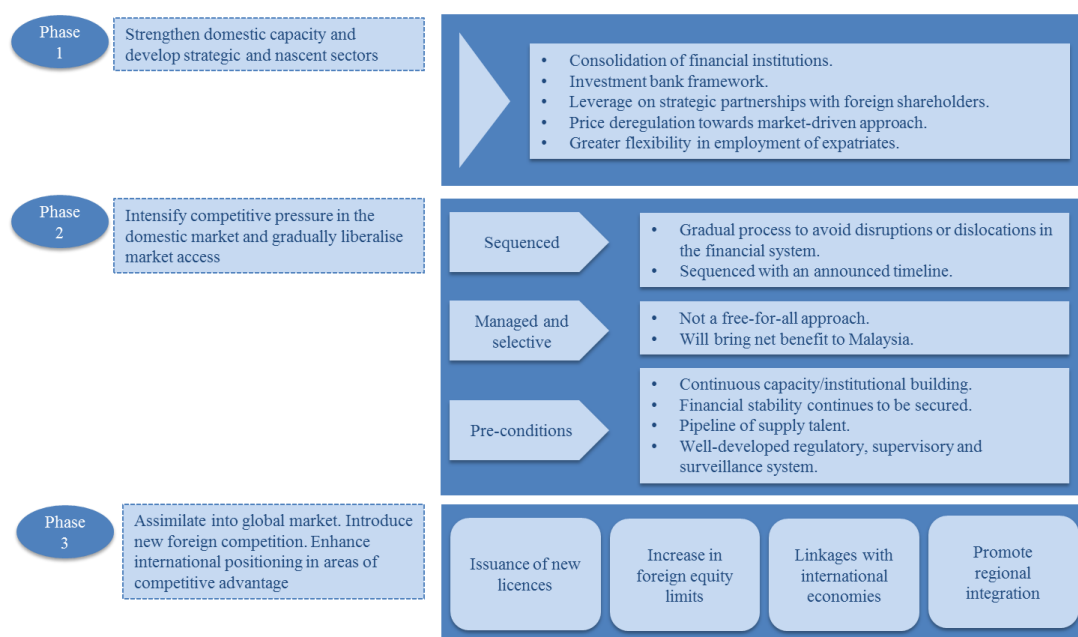
1. First phase (2000–2003)⁸⁹: to enhance the capacity of domestic banking institutions and strengthen financial infrastructure; in this phase, the focus was to strengthen the capability and capacity of domestic banks, and building and enhancing the financial infrastructure of Malaysian banking industry. At the same time, steps were taken to create the necessary infrastructure for a more market-based consumer protection framework to foster further competition among banks. The result from this phase is to expect the emergence of domestic banks that were stronger, more efficient and innovative, and increasingly competitive and resilient (Bank Negara Malaysia, 2001b).
2. Second phase (2004–2007): to intensify competitive pressure in the domestic financial sector; this phase emphasised creating an increasingly levelled playing field for incumbent foreign players. Several restrictions for foreign banks were removed to add further competition to the banking industry, as well as providing wider choices for the consumers. These steps were implemented gradually depending on the overall ability of the financial system to absorb these changes (Bank Negara Malaysia, 2001b); and

⁸⁸ As pointed out by McKinnon (1973) and Shaw (1973), financial liberalisation is hypothesised to benefit the financial system by increasing the efficiency level of players within the industry by exposing them to the competition, which can enhance domestic savings as well as investment. Nonetheless, there are studies indicate that severe crises can occur after the ‘big-bang’ implementation of financial liberalisation (e.g. Demirguc-Kunt and Detragiache, 1999; Weller, 2001). Aware of the adverse experiences of the ‘big-bang’ financial reform approach in other countries, BNM has employed a gradualist approach in the liberalisation process on its financial system to reduce the risks of destabilising implications (Bank Negara Malaysia, 2003; Njie, 2007).

⁸⁹ There were no specific timelines for these three phases of the FSMP. The timelines were subjected to achievements of banking sector in meeting specified objectives of these phases. This flexible approach allows BNM’s policy measures being implemented during the first phase to continue and overlap during second- and third-phase, depending on orderly adjustments made by Malaysian banks to a more competitive and increasingly deregulated and liberalised market (Bank Negara Malaysia, 2001b). For this study, to reflect the policy measures implemented by BNM in these three phases, they are grouped into 2000-2003, 2004-2007 and 2008-2011. The inclusion of years 2000 and 2011 in the first- and third-phase (respectively) is to indicate the implementation of several policy measures of the FSMP being performed outside the planned timeline (i.e. 2001-2010). For example, due to unforeseen circumstances, events such as global credit crisis in 2007 had delayed some of planned liberalisation initiatives of the third-phase FSMP, with them being performed beyond 2010 (e.g. commencement of operation of new foreign banks). This delay was however necessary to ensure that the outcome of policy measures is achieved without destabilising impacts to the banking sector (Bank Negara Malaysia, 2010b, 2011b).

3. Third phase (2008–2011): to introduce new foreign competition and assimilation into the global arena; in third phase of the FSMP, new foreign banks were permitted to provide competitive pressure to incumbent banking institutions, which would also serve as an incentive against complacency and to remain competitive. Additionally, with intensifying degree of global competition and greater assimilation into the global arena, the banking sector was introduced with more liberalisation measures (Bank Negara Malaysia, 2001b).

Figure 4.5 The three Phases of the Financial Sector Master Plan, 2001-2010



Adapted from Financial Sector Master Plan 2001 (Bank Negara Malaysia, 2001b)

During the consolidation of the banking sector, issues relating to branches and staff strength became inevitable in minimising a duplication of resources and maximising consolidation benefits. There were not more than 10,000 personnel affected by redundancy during this consolidation exercise. In view of that, through the staff training fund, reskilling initiatives were introduced to facilitate the smooth transition of the workforce into other industries in the economy (Bank Negara Malaysia, 2000, 2001a). At the same time, several deregulation measures were taken by BNM prior to the implementation of the FSMP. For example, in recognising the manpower capabilities, expertise and a degree of innovation that give rise to the competitive edge of banking

institutions, BNM abolished the guideline on wage moratorium for the banking sector. With this liberalisation measure, banks could retain an experienced and high quality pool of bankers who were in the industry as well as attracting best talents into the industry. Banks were also provided with greater autonomy to determine the salaries and wages of their employees; a remuneration committee on the Board of Directors had to be established to determine the remuneration package for senior management and key personnel (Bank Negara Malaysia, 2000).

In recognising the impact of technology as an enabler to improve efficiency and competition of the banks, an alternative delivery channel was required to reach out to the customers of banking institutions. Therefore, BNM allowed banking institutions to offer a full range of banking products and services over the Internet. The approach taken by BNM for this move was in line with the principles of FSMP's gradual liberalisation programme, and domestic banks were given a head start in June 2000. The foreign banks, on the other hand, were allowed to provide communicative Internet banking services from January 2001 and transactional Internet banking services in January 2002 (Bank Negara Malaysia, 2000, 2001). This technology driven delivery channel was expected to benefit the foreign banks, particularly when branching was restricted by the regulator (Khrishnasamy et al., 2004).

As consolidation and liberalisation of the financial system took place, there was an increase in the blurring segregation of distinctions between activities of different groups of financial institutions. Hence, in removing this artificial barrier, in 2000, BNM allowed banks to freely cross-sell financial products and services of entities within the same group, including those belonging to their related non-bank financial institutions (Tahir et al., 2008, Bank Negara Malaysia, 2000). This provided a more level playing field for smaller banking institutions as they were able to tap a larger pool of customers using broader marketing channels with greater cost savings, and relying on the expertise of larger banking institutions. This would also allow the improvement of banking groups operations via centralisation and prevent the duplication of similar operational functions (Bank Negara Malaysia, 2000). In the same year, to enhance the competition and efficiency of banks, the banking sector was granted blanket approval to outsource

its non-core operational functions, such as payroll and data entry processes, to third-party service providers. With this, banks may not have had to invest a significant amount of time and funds to undertake such functions, which allowed them to concentrate their resources on developing competencies in their core areas of business (Bank Negara Malaysia, 2000).

Learning from the 1997/98 Asian financial crisis, which highlighted that poor credit risk management is a source of vulnerability in banking institutions, BNM introduced in 2001 guidelines on best practices for the Management of Credit Risk, which assisted banks towards upgrading the practice of banking institutions with regards to credit management and administration. The guidelines introduced four pillars: First, they increased the responsibility of the board of directors and management for the credit risk management process. Second, they improved the infrastructure for effective credit risk management. Third, they developed an integrated risk management process, and fourth, they established comprehensive internal controls and audit procedures for credit risk management (Bank Negara Malaysia, 2001a). Additionally, BNM also initiated an improved and up-to-date centralised electronic database of credit information of banking institutions' borrowers known as Central Credit Reference Information System (CCRIS). CCRIS would assist banking institutions in their credit assessment and monitoring process by conducting background checks and the levels of indebtedness of borrowers in the banking system, as well as developing a sound credit culture in the banking sector (Bank Negara Malaysia, 2001a).

In 2002, further deregulation was made on new product approval processes. Before the introduction of this guideline, banking institutions were required to meet a pre-approval requirement prior to introducing new products or activities. This was imposed to protect consumers as well ensure the orderly development of the financial market. Recognising the importance of regulatory environments that are market-driven, the new guideline provided a simpler product notification and approval process, inspired by FSMP's principle of 'what is not prohibited is allowed'. As a consequence, this has incentivised banks to improve research and development for their niche products and services (Bank Negara Malaysia, 2002).

As part of the first phase of FSMP (2000-2003) in preparing domestic banks to face market liberalisation, the newly-established consolidated banking groups were allowed to hold two licences; one, to carry on banking business, and the other, to carry finance company business. This is further improved by amending the Banking and Financial Institution Act 1989 to allow the merger of commercial banks and finance companies within a domestic banking group as a single legal entity. The amendment of the legal framework changed the structure of internal operations of domestic banking groups to face competition as well as improved efficiency (Bank Negara Malaysia, 2003).

As the banking sector moved into the second phase of the FSMP (i.e. 2004-2007), the liberalisation measures continued, in which BNM initiated substantial changes in the conduct of monetary policy. On 23 April 2004, BNM introduced a New Interest Rate Framework (NIRF) to strengthen the monetary policy transmission mechanism and promote better pricing based on demand in the financial system (Bank Negara Malaysia, 2004). The NIRF is designed for the implementation of monetary policy, which consequently facilitates the transmission of a policy rate based on the new Overnight Policy Rate (OPR).⁹⁰ This deregulation regarding pricing is one of the key initiatives implemented under FSMP for better allocation and distribution of resources in the financial system (Bank Negara Malaysia, 2004). The implementation of deregulated interest rates requires banks to compete, as autonomy in pricing can be an important factor for gaining market share based on respective cost structures and economic conditions, which in turn can benefit consumers and the banking sector as a whole.

Another significant milestone to the liberalisation of banking system relates to the development of Islamic banking. The planned liberalisation of the Islamic banking sector was brought forward from 2007 to 2004 with the issuance of three new Islamic banking licences under the Islamic Banking Act 1983 to Islamic financial institutions

⁹⁰ The OPR replaced the three-month intervention rate and corresponding overnight interbank rate, and was used as an operating target and indicator of the monetary policy stance. Another important aspect of the new interest rate framework, it removed both the ceiling on the base lending rates (BLRs) and the prescribed lending spreads. Hence, Malaysian banks are allowed to set their own BLRs using their respective cost structures, risks and business strategies. This change allowed the transmission of policy rates into other market rates and enabled banks to implement more efficient pricing of products and services (Bank Negara Malaysia, 2004).

from the Middle East, namely the Kuwait Finance House, Al Rajhi Banking and Investment Corporation and a consortium of Islamic financial institutions represented by the Qatar Islamic Bank, RUSD Investment Bank Incorporated and Global Investment House (Bank Negara Malaysia, 2004).⁹¹ The introduction of new foreign Islamic banks into the Malaysian banking system provided the required competition within the market and improved the Islamic banking industry in Malaysia and the South East Asia region (Suffian, 2010). These new entrants promoted the exchange of knowledge and expertise as well as improving financial linkages between Malaysia and the Middle East. Additionally, to further enhance the institutional structure and human capital development of the Islamic banking system, there was a review of the Islamic Banking Scheme (IBS)⁹², which uses the ‘window-based’ institutional structure, with the result that both domestic and foreign banks with IBS were encouraged to establish their respective Islamic subsidiaries and license them under the Islamic Banking Act 1983. The primary aim of converting IBS to an Islamic bank subsidiary was to improve the strategic focus of Islamic banks, while at the same time continuing to rely on a synergy of conventional banking operating infrastructures (Bank Negara Malaysia, 2004).

The presence of foreign banks, which is supported by extensive research capability from their international operations network, has contributed to a significant development of the banking sector in Malaysia, particularly through capital investment, the creation of employment opportunities and the transfer of skills and technology (Ang, 2008, 2009). Thus, another gradual liberalisation measure taken under FSMP initiatives is to introduce a more competitive environment for domestic banks. For this, BNM allows foreign banks to open up to four additional branches subject to a predetermined ratio of additional branches in market centres, urban areas and non-urban areas (Bank Negara

⁹¹ Nevertheless, these new Islamic banks commenced operations in 2006. This initiative is in line with the aspiration of FSMP, which is to make Malaysia an international Islamic hub (Bank Negara Malaysia, 2004).

⁹² Under the IBS framework, the conventional banking institutions were allowed to provide Islamic banking products and services. This arrangement was called ‘Islamic windows’: in order to operate, the commercial banks should have a separate and dedicated Islamic Banking Division within its structure. The operation and maintenance of Islamic banking operations as profit and cost entities should be separate from conventional banking (i.e. there should be a firewall), although the physical capital and staff can be shared with conventional banks. Banks with IBS should disclose separate Islamic financial reports with their financial statements (Abdul Majid et al., 2011; Kamaruddin et al., 2008).

Malaysia, 2005). With this policy, foreign banks are expected to provide a greater level of competition in the banking market, while at the same time to improve their socio-economic obligations to serve non-urban areas.

Realising the importance of the capital market within the financial system, BNM and Securities Commission (SC) introduced the investment banking framework in March 2005. The framework aims to broaden the scope of activities of merged entities based on the types of licences held prior to consolidation and to share the regulatory burden between BNM and SC. Investment banks can be formed via various merger permutations. For example, newly-established investment banks can be materialised from either the merger of a merchant bank and a stockbroking company within the same banking group, or a merger of two stand-alone discount houses with a stockbroking company (Bank Negara Malaysia, 2005). Hence, this investment banking framework resulted in a dual regulatory regime, in which the responsibilities of prudential regulations and market conduct are placed under the purview of BNM and SC, respectively (Bank Negara Malaysia, 2005).

The introduction of deposit insurance is identified in the FSMP to provide an adequate safety net for the promotion of financial stability, particularly in protecting depositors. The Malaysian Deposit Insurance Corporation (Perbadanan Insurans Deposit Malaysia, or PIDM) was established in 2005, with the enactment of the Malaysia Deposit Insurance Act 2005, which aimed to provide least-cost resolution options when dealing with problem banks (Rajoo, 2008). With the implementation of PIDM's differential premium system, sound risk management and a greater degree of efficiency in banks is expected from banking institutions (Bank Negara Malaysia, 2008b). In the same year, under the consumer protection framework, BNM established a Credit Counselling and Debt Management Agency (Agensi Kaunseling dan Pegurusan Kredit, or AKPK) to ensure consumers had sufficient avenues to seek assistance and redress when they faced problems with their financial institutions. The agency also helped consumers to seek advice on managing credit and equip them with the necessary skills to manage their finances (Bank Negara Malaysia, 2005).

Following the pegging of the Malaysian Ringgit during the Asian financial crisis in 1998, BNM finally announced the floating of the Malaysian Ringgit (MYR) on 21 July 2005 through a managed floating exchange rate regime, where the value of ringgit was determined by economic fundamentals (Bank Negara Malaysia, 2005). Malaysia also relaxed some of its capital control measures by liberalising a number of foreign exchange administration rules, for example by simplification of the reporting requirements for exporters, increasing the overnight limit of foreign currency accounts for residents, extending loans in Malaysian Ringgit to non-residents and allowing residents to enter into forward foreign exchange contracts (Bank Negara Malaysia, 2004).

In 2006, a transformation exercise within BNM was carried out to streamline both regulatory and supervisory functions. Financial stability mandates became necessary to ensure the smooth progress of financial development and liberalisation. Hence, a dedicated macro-prudential orientation surveillance was further enhanced with the setting up of a Financial Surveillance department. The department was entrusted to undertake integrated identification of vulnerabilities through appropriate assessment via collaborative mechanisms within BNM and with other domestic regulators to emanate risks from domestic and international developments for the overall stability and functioning of the financial system. The department worked closely with the supervision departments to ensure that their risk assessment incorporated both macro-prudential and micro-prudential standpoints (Abdul Ghani, 2010; Bank Negara Malaysia, 2006b).

In order to improve product innovations and the efficiency of the banking sector, in 2007, BNM introduced a 'launch-and-file' system that removed the regulatory notification and approval requirements prior to the introduction of new products by banking institutions. The deregulation initiative can reduce the time needed to market new products, allowing banks to respond faster to the changes in the market. It is however applicable only to conventional banks, but not Islamic banks. However, investment products that could expose customers to losses exceeding the principal sum

invested must undergo the normal process where approvals should be attained from BNM (Bank Negara Malaysia, 2007b).

As the timeline progressed into the end of second phase of the FSMP (i.e. 2004-2007), the Malaysian economy experienced a steady growth (i.e. GDP) averaging close to 6.0% from 2005 to 2007. Following the early onset of the global credit crisis, the Malaysian financial system was insulated from the global turmoil. In 2007, the downside risks from US subprime and credit market events were felt only in increased volatility in equity and bond markets, due to the small direct and indirect exposure of financial institutions from subprime related instruments, which accounted for about 0.3% of the banking system capital base (Bank Negara Malaysia, 2007b). The continued accumulation of a capital buffer during the earlier first two phases of financial liberalisation had resulted in improved capacity of the banking sector to withstand higher risk levels from greater market volatility and financial shocks (Tan, 2011; Bank Negara Malaysia, 2007a, 2007b).⁹³

However, Malaysia is an open economy and relies heavily on external trade. The impact of the global recession was felt strongly and recorded a heavy decline in exports in the second half of 2008. The effect of the decline in the trade flow via the real economy has affected outputs, employment opportunities, and private investment and consumption activities (Tan, 2011). Malaysia also experienced an increase in prices reflected from the higher global inflation rates.⁹⁴ With inflation concerns, the international financial markets were also experiencing disruptions mainly due to the collapse of the US subprime mortgage market in the middle of 2007, which led to disorder in the financial market. The valuation of asset-backed securities (ABS) using subprime markets as underlying assets became more complicated as delinquencies increased. As a result, several major international financial institutions suffered large write-downs and losses.

⁹³ Responding to emerging risks during the early period of the crisis, several Malaysian banks preemptively increased their capital positions through rights issues and Tier-1 capital instruments in anticipation of a worsened global crisis; this resulted in a stronger capital position for absorbing a potential adverse impact (Bank Negara Malaysia, 2008b).

⁹⁴ For example, in 2007, the global oil prices broke USD 100 per barrel for the first time in history, which in return affected other commodity prices. The impact of high food and fuel prices on inflation was also felt in emerging economies, such as Malaysia (Bank Negara Malaysia, 2008a).

The decline of the capital base and inability of US banks to raise sufficient funds indicates that the US banking sector was facing a severe crisis. In September 2008, following the aftermath of the collapse of Lehman Brothers, there was a freeze in the interbank money market, which hindered ‘business as usual’ activities (Bank Negara Malaysia, 2008a). The global financial crisis resulted in widespread effects on an unprecedented scale across the world. Given its trade and investment linkages with the rest of the world, Malaysia was not spared from the impact of the global crisis (Lee, 2011). As a result, the Malaysian economy contracted by 1.7% in the full year of 2009 (Bank Negara Malaysia, 2009a).

Malaysia responded by initiating some policy actions to maintain confidence in the market. On 16 October 2008, the Ministry of Finance and BNM announced a blanket Government guarantee through PIDM on all Ringgit and foreign currency deposits placed with financial institutions regulated by BNM (Bank Negara Malaysia, 2008a). The safety net was also further extended to insurance and the takaful⁹⁵ sector to ensure adequate access to liquidity. The guarantee was in force until December 2010 (Ibrahim, 2011). During this period of uncertainty, BNM also encouraged banks to engage proactively with borrowers and provide debt management solutions and advice, such as restructuring and rescheduling of financial obligations and controlling new delinquencies that could be higher during an economic downturn (Bank Negara Malaysia, 2009b).

In November 2008, to mitigate the impact of the economic and financial crisis between 2008 and 2009 on domestic demand and to accommodate monetary policy, BNM via its Monetary Policy Committee (MPC) reduced OPR by 25 basis points followed by another reduction of 75 and 50 basis points in January 2009 and February 2009, respectively. The overnight interbank rate in the money market decreased from 3.50% in November 2008 to 2.00% in 25 February 2009. In ensuring that policy rates are

⁹⁵ Takaful is a type of joint guarantee insurance mechanism based on ‘the law of large numbers’, in which a group of societal members pool their financial resources together against certain loss exposures. Takaful-branded insurance is based on Shariah (Islamic religious principles/laws, which prohibits riba (interest), al-maisir (gambling), and al-gharar (uncertainty) principles), and explains how it is the responsibility of individuals to cooperate and protect each other (Maysami and Kwon, 1999; Kwon, 2007).

transmitted to retail market rates, the statutory reserve requirements (SRR)⁹⁶ of the banks were also reduced from 4.0% to 1.0% (Bank Negara Malaysia 2009a). As monetary policy eased between 2008 and 2009, there was uncertainty in the debt market due to lower yields; and with the low interest rate offered by the banking sector, it became more attractive for corporates to raise funds from banking institutions. Therefore, the government announced the establishment of Financial Guarantee Institution (Danajamin Nasional Berhad, or Danajamin) that provides credit guarantee to selected issuers. The establishment of Danajamin provided credit enhancements to bonds issuances, which in return, reduced financing costs, addressed potential maturity mismatches and diversified funding sources (Bank Negara Malaysia, 2008b, 2009a). Concurrently, realising that Malaysia was an economy with a high savings rate, BNM issued the Merdeka Savings Bond and Sukuk Simpanan Rakyat (Islamic Bond for citizen savings), totalling approximately MYR 7.0 billion for savers who depended on income from deposits. These household bonds had higher rate of return compared to the market interest rate (Bank Negara Malaysia, 2009a).

Learning from the 1997–1998 Asian Financial crisis, this global credit crisis was managed by a collective and coordinated response by both BNM and the entire government machinery. Therefore, the Malaysian government introduced a couple of stimulus packages between 2008 and 2009 to revive the economy. The first stimulus package was announced in November 2008 amounting to MYR 7.0 billion to support private spending. The second stimulus was announced in March 2009 as the global crisis worsened. A total of MYR 60.0 billion was provided in the second stimulus package to support economic activity. Although the amount was significantly larger than the first stimulus package, only MYR 15.0 billion was used for direct fiscal injection. The balance of the second package was used in the form of government guarantees, as equity investment by government's investment company (Khazanah

⁹⁶ The statutory reserve requirement (SRR) is commonly employed by central banks as a regulation that sets the minimum fraction of customers' deposits and notes that each commercial bank must hold as reserves (not to be lent out). SRR is normally in the form of cash stored physical in a bank vault or deposits made with a central bank (e.g. BNM). SRR could also be used as a tool to manage monetary policy of a country that influences the country's borrowing and interest rates, which is done by adjusting the amount of funds available for commercial banks to lend out to customers (Weiner, 1992; Furfine, 2000).

Nasional Berhad), private finance initiatives and tax incentives. The total overall stimulus packages totalling to 9.6% of GDP were implemented to spur businesses and households during the period of uncertainty (Bank Negara Malaysia, 2009). The accelerated implementation of fiscal stimulus measures, easing of monetary policy, and improved access to financing have stabilised the economy, and the Malaysian economy recovered in the second half of 2009, particularly after a sharp contraction in the first half of 2009 (Bank Negara Malaysia, 2009a).

During the period of unsettled global crisis, BNM continued to implement its liberalisation agenda in the third phase of FSMP (i.e. 2008-2011). Hence, in 2009, BNM announced three broad strategies to further liberalise its financial system by granting new issuance of banking licences, an increase in foreign equity limits and operation flexibilities in the establishment of new branches of foreign institutions as well as expatriates' employment (Ibrahim, 2011; Bank Negara Malaysia, 2009b). After the evaluation process took place, five new commercial banking licences were granted in 2010 and one in 2009 to foreign financial institutions with expertise, global networks and good business capacity that could facilitate the growth and diversity of the Malaysian financial system as well as supporting international trade and investment flows (Bank Negara Malaysia, 2010b).⁹⁷ At the same time, branching liberalisation continued: four foreign banks obtained approval from BNM to operate 12 new branches nationwide (Bank Negara Malaysia, 2010b).

In relation to capital regulation, BNM also introduced Basel II for a standardised approach (Basel II SA) and the Internal Rating Based approach (Basel II IRB); they successfully implemented these new capital regulatory regimes in 2008 and 2010, respectively (Bank Negara Malaysia, 2008b, 2010b).⁹⁸ By implementing this new international risk-based capital framework, the Malaysian banking sectors were

⁹⁷ The licences were granted to Industrial and Commercial Bank of China, BNP Paribas, National Bank of Abu Dhabi, Mizuho Corporate Bank, Sumitomo Mitsui Banking Corporation and PT Bank Mandiri (Bank Negara Malaysia, 2010b).

⁹⁸ Basel II is the second series of capital regulatory measures of the Basel Accords. Basel II introduced operational risk and a new measurement of credit and market risks. It also included embedded governance guidelines in managing these risks, which affected the risk-weighted capital ratio (Lastra, 2004; Basel Committee on Banking Supervision, 2005; Herring, 2005; Kupiec, 2007).

expected to improve their risk management, data gathering, information disclosure and business conduct.

The Malaysian economy expanded by 7.2% in 2010. The growth was driven by strong domestic demand and expansion in private sector activity, as well as strong support by public sector programmes on infrastructure and delivery systems (Bank Negara Malaysia, 2010a). Hence, the OPR was adjusted upwards to manage the potential risk of financial imbalances.⁹⁹ BNM normalised the OPR by gradually raising it by 25 basis points in March, May and July 2010. The SRR also increased by one percentage point at the same time intervals from 1% to 4%. Instruments such as Bank Negara Monetary Notes were also issued to mop excess liquidity, particularly from large short-term capital flows from large deposit inflows from European financial institutions (Bank Negara Malaysia, 2010a).

A rapid increase in household indebtedness took place after the Asian Financial crisis in 1997 (due to a shift of loan portfolios from corporates and businesses to households and retail segments). Supported by a series of low interest rates faced by the domestic banking market, this trend had the potential to become a risk concern. Therefore, a series of hikes in interest rates by BNM could have slowed down the rapid growth of household loans. However, to ensure the effectiveness of this monetary policy, a macro-prudential policy was supplemented in November 2010 to deter excessive investment and speculative activity in the residential property market at certain locations. Using Loan-to-Value (LTV)¹⁰⁰ measures, borrowers were subjected to an LTV ratio of 70% for third and subsequent house financing facilities. To reinforce prudent underwriting

⁹⁹ Financial imbalances can be formed in many ways such as unhealthy build-up of leverage, underpricing of risks, excessive yield-seeking activities, over-investment in certain markets, asset prices that are not fundamentally substance and increased risks to financial stability. Keeping a low interest rate for too long may translate into risks that suspend growth (Assenmacher-Wesche and Gerlach, 2010).

¹⁰⁰ The loan-to-value (LTV) ratio is a financial term that is used to describe the ratio of a loan to the value of the asset being financed (e.g. home mortgage). LTV is used by banks as one of the key risk factors when assessing the underlying assets that the borrowers intend to purchase. A higher LTV ratio may indicate higher risk, in which banks face higher likelihood of absorbing losses due to low amount of equity to cover the potential default for loans (Qi and Yang, 2009). LTV measure is also used as macro-prudential regulation instruments on commercial banks in several countries (e.g. Hong Kong, Korea, Malaysia and China) to reduce the risk of bank credit becoming a source of procyclicality risk and to strengthen banking sector resilience against asset price volatilities (Wong et al., 2011; Igan and Kang, 2011).

practice, BNM revised the risk weight from 75% to 100% for LTV exceeding 90%. In March 2011, another pre-emptive measure was introduced for credit card debt, where BNM tightened the requirement for credit cards for individuals earning monthly gross earnings of MYR 3,000 or less, where the number of cards and their credit limit should be available according to their debt repayment ability (Bank Negara Malaysia, 2010b).

At the end of the FSMP period, a new 10-year plan was introduced in 2011 called Financial Sector Blueprint (FSBP).¹⁰¹ FSBP moved away from the sector-based approach as adopted in the FSMP by introducing a blueprint that exhibits a more integrated financial sector with increased sectoral and cross-border interlinkages (Bank Negara Malaysia, 2011b, 2011c). FSBP specifies the direction to support the transformation of the financial landscape in the next decade with the aspiration to become a country with a high-value added and high-income economy. In addition, as the Competition Act 2012 (CA 2012) came into effect on January 2012 in all sectors in Malaysia, BNM and the Malaysian Competition Commission cooperated to ensure that smooth implementation would take place within the financial sector. The introduction of the Act complemented BNM efforts to improve a consumer protection framework by promoting and safeguarding process of competition (Bank Negara Malaysia, 2011b).

¹⁰¹ The FSBP is worked out based on four broad areas which are (1) evolving a financial sector that best serves the Malaysian economy, (2) enhancing regional and international financial linkages, (3) safeguarding the stability of the financial system, and (4) enhancing key enablers that support the development of the financial system. There are 69 recommendations to support these strategic outcomes (Bank Negara Malaysia, 2011b).

Table 4.2 Liberalisation Measures in Malaysia, 2000-2011

Year	Salient Liberalisation Measures
2000	Internet banking guidelines for internet banking services introduced, beginning with domestic banks in 2000, followed by foreign banks in 2001, to commence transactional services in 2002.
2000	Increase of capital funds for banking institutions, allowing adequate buffers to absorb unexpected shocks.
2000	Banking institutions allowed to outsource their non-core and non-strategic operations to third-party providers.
2000	Wage moratorium for banking sector was lifted and banks allowed to determine the salaries of personnel, to attract talent into the industry.
2000	Banking institutions allowed to cross-sell various products within the financial group.
2001	Benchmarking project by BNM to assess and evaluate banks and their peers in term of health and niche business operations and strategies.
2001	Best practices introduced for management of credit risks for banking institution references.
2001	Anti-Money Laundering Act 2001 introduced to combat illegal funds within banks.
2001	Introduction of Central Credit Reference Information System (CCRIS), which collects data from borrowers on real time basis.
2001	Financial Sector Master Plan introduced to chart the direction of Malaysian financial industry for the next 10 years including various liberalisation initiatives.
2002	Liberalisation of internet payment gateway. Banks were allowed to use other operators.
2002	Introduction of new product approval guidelines, which can improve the process of generic and plain vanilla products without having to follow long list of requirements.
2003	Amendment to Central Bank of Malaysia Act 1958 and Bank and Financial Institutions Act 1989 to broaden authority of BNM.
2003	Special relief provided by banking institutions to borrowers in Severe Acute Respiratory Syndrome-(SARS-)affected Industries.
2003	Financial Mediation Bureau established to support consumers in redress financial issues.
2003	'Banking Info' initiated to ensure consumers understand the common financial products and their features as well as risk embedded in those products.
2004	Implementation of Basel II timeline for Malaysian banks, which was to commence in 2008 (Standardised Approach) and 2010 (Internal Ratings-Based Approach).
2004	Introduction of New Interest Rate Framework. OPR introduced as policy and indicative interest rate for market rate. Banking institutions allowed to set their own interest rate based on cost structure and business strategies.
2004	Liberalisation of Islamic banks pushed forward by granting three new Islamic banking licences to foreign financial institutions
2005	Branching liberalisation for locally incorporated foreign banks subjected to ratio of urban and non-urban branch network.
2005	De-pegging of Malaysian Ringgit with several currency capital controls lifted to support fundamental economic activities.
2006	Transformation of investment banks to which these banks are regulated by both BNM and SC.
2006	Transformation of Islamic window operation of Islamic banking scheme to Islamic bank subsidiaries.
2006	Operations of three new Islamic foreign banks, namely Kuwait Finance House, Al Rajhi Bank and Asian Finance Bank.
2006	Credit counselling and debt management agency (AKPK) set up to support consumers in addressing issues they have with banking institutions as well as providing financial knowledge and awareness.
2007	Liberalisation to the foreign exchange administration rules, providing better conditions for real sector players in performing transactional activities with foreign counterparts.
2007	Introduction of launch and file system for new conventional banking products.
2008	Establishment of Financial Guarantee Institution (Danajamin) to assist corporates in raising funds via debt capital market.
2009	Relaxed monetary policy where OPR reduced by 75 basis points to overcome global crisis and introduction of various stimulus packages by the government to support fiscal growth.
2009	Issuance of new banking licences to foreign bank: Industrial and Commercial Bank of China.
2010	Issuance of five new banking licences to foreign banks: BNP Paribas, National Bank of Abu Dhabi, Mizuho Corporate Bank, Sumitomo Mitsui Banking Corp and PT Bank Mandiri.
2011	Introduction of Financial Sector Blueprint to steer the direction of banking sector for the next 10 years.
2011	BNM work with competition commission for banking sector upon the introduction of Competition Act 2012.

In summary, the Malaysian banking system has undergone a vast transformation after experiencing declines in the economy during the 1997–1998 Asian financial crisis. Structural reforms made since the Asian financial crisis have led to changes in the financial landscape and at the same time, contributed towards strengthening of fundamentals and the resilience of domestic financial systems. A series of systematic and sequenced liberalisation (see Table 4.2) was introduced following the Asian financial crisis, which resulted in a banking system that was able to withstand economic shocks (as observed and evidenced in the impact of the global credit crisis in 2008-2010 and discussed later in this chapter).

4.4 The Structure of Malaysian Banks

The discussions in the previous section focused on the background of Malaysian banking system and the liberalisation measures (e.g. FSMP) implemented by the Malaysian government and BNM. In this section the discussions are consequently continued on the structure of the Malaysian banks according to ownership and specialisation. As mentioned earlier in Chapter 3 ('Empirical Studies of Banking Efficiency'), the structure of banking institutions is useful to explain the variation in efficiency of the banks (e.g by specialisation and by ownership structure). Additionally, this section outlines the financial position (e.g. balance sheet and income statement) of Malaysian banks and the effect of liberalisation measures (e.g. FSMP) between 2000 and 2011.

4.4.1 The Ownership of Malaysian Banks

There are two major types of ownership in Malaysia: foreign banks and domestic banks. Foreign banks in Malaysia are subsidiaries owned by foreign parent banks. On the other hand, domestic banks can be owned by either Malaysian or foreign entities. However, foreign ownership in domestic banks is limited to only 30% of the total equity.

As discussed in the previous section, foreign banks held over 90% of the share of the banking system in 1957, but this share declined due to branching restrictions and the suspension of new banking licences to foreign entities (Bank Negara Malaysia, 1999b, Suffian, 2011). Nevertheless, due to the demands of World Trade Organisation (WTO) and as part of liberalisation initiatives under the FSMP, BNM issued three new Islamic banking licences to foreign banks and six new conventional banking licences between 2004 and 2010 (Bank Negara Malaysia, 2004a, 2009a, 2010a). As mentioned earlier, these new foreign banks were also required to be locally incorporated. This regulation was implemented to ensure that the existence of foreign banks in Malaysia was subjected to permanent capital, facing the same prudential requirements as domestic banks (Loke, 2011).¹⁰² As of December 2011, the share of total assets of foreign banks was at 24.3% (see Table 4.3).

Table 4.3 Total Assets of Malaysian Banks by Ownership

Year	Total Assets	Domestic Banks' Total Assets	Percentage of Domestic Banks	Foreign Banks' Total Assets	Percentage of Foreign Banks
	(MYR millions)	(MYR millions)	(%)	(MYR millions)	(%)
2000	482,454	360,982	74.82	121,472	25.18
2001	522,042	390,591	74.82	131,452	25.18
2002	549,782	412,814	75.09	136,967	24.91
2003	609,325	449,755	73.81	159,570	26.19
2004	715,168	530,974	74.24	184,194	25.76
2005	827,462	627,237	75.80	200,225	24.20
2006	999,342	774,092	77.46	225,250	22.54
2007	1,142,305	868,738	76.05	273,567	23.95
2008	1,239,268	939,359	75.80	299,909	24.20
2009	1,344,062	1,035,941	77.08	308,121	22.92
2010	1,457,707	1,124,875	77.17	332,832	22.83
2011	1,622,197	1,227,503	75.67	394,695	24.33

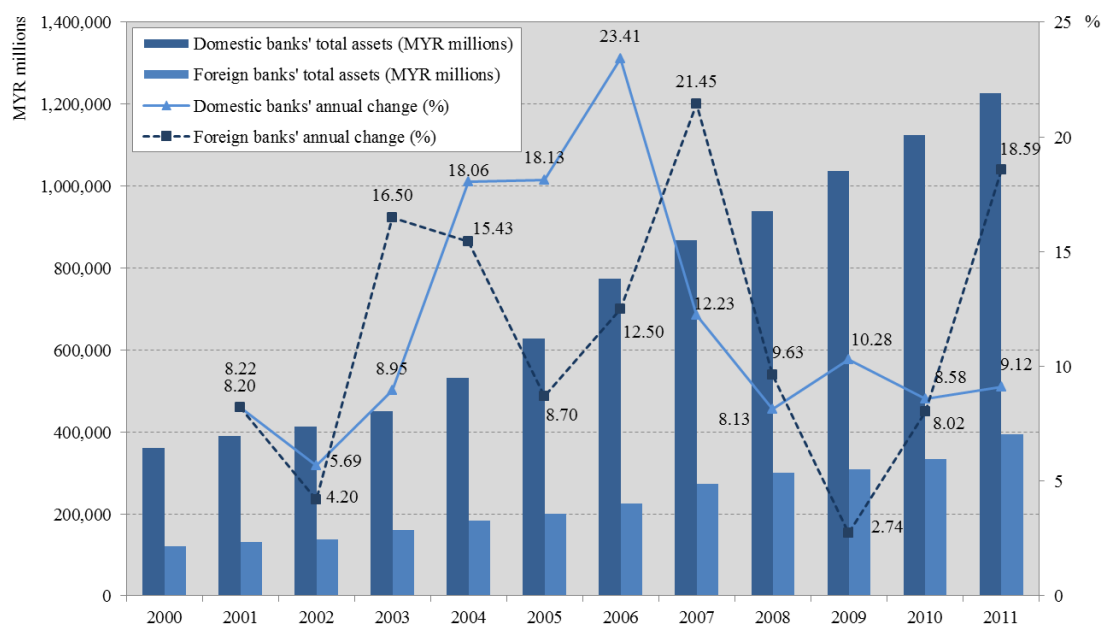
Source: Confidential dataset from a Malaysian financial organisation

Note: As of December 2011, USD 1.0000 was equal to MYR 3.1265

¹⁰² By incorporating a local entity bank, its relationship with the parent bank can be 'ring-fenced' by creating a legal separation between these two entities. By doing this, BNM could ensure that the capital and assets of foreign banks in Malaysia would remain in Malaysia and safeguard local depositors by giving them the first priority should there be any adverse occurrence in the foreign parent company (Loke, 2011).

Based on Table 4.3, the foreign banks' market share (based on total assets) was constantly around 24.0% from 2000 to 2011. Total average annual growth of the total assets for the years under study for foreign banks was 11.5%. Figure 4.6 exhibits the growth of total assets of domestic and foreign banks for the years 2000 to 2011. Between 2000 and 2003, the average annual growth of the total assets was 9.6%. In the second phase of the FSMP (between 2004 and 2007), the average growth was 14.5%. The sharp increase in the second phase of FSMP was attributed to the introduction of the new foreign Islamic banks and deleveraging activities by European Banks. As the financial system was undergoing the third phase of FSMP (between 2008 and 2011), the average annual growth was moderated to 9.6%. The asset growth for foreign banks declined in the third phase of FSMP due to the economic slowdown faced by Malaysia.

Figure 4.6 Total Assets of Domestic and Foreign Banks in Malaysia



Source: Confidential dataset from a Malaysian financial organisation

4.4.2 Malaysian Banks by Specialisation

The main types of banking institutions that make up the banking system in Malaysia are conventional banks, Islamic banks and investment banks. These types of banks can be owned either by domestic or foreign parties (see Table 4.1).

Commercial banks (conventional and Islamic banks) are the largest component in the Malaysian banking system. They provide banking services such as accepting deposits, granting loans, and providing trade-finance facilities (Tahir et al., 2008; Abdul Majid et al., 2011).¹⁰³ Unlike commercial banks, investment banks are more involved in capital market activities and stockbroking as well as providing short-term credit for working capital (Sufian, 2011).¹⁰⁴ However, investment banks are not comparable with conventional commercial banks or Islamic banks due to huge differences in their business activities. Hence, due to non-homogeneous structures and different business activities from commercial and Islamic banks, investment banks are excluded for the purpose of this research.

Islamic banking has grown rapidly all over the world since the establishment of the first Islamic bank in Egypt in 1963 by Nasser Social Bank (Batchelor and Wadud, 2004; Iqbal and Molyneux, 2005).¹⁰⁵ The oil boom in 1975 further intensified the development of Islamic banking due to huge capital inflows to Islamic countries. In Malaysia, the development of Islamic banking by the government in 1963 began with the establishment of the pilgrimage fund institution (Tabung Haji) to mobilise funds for Muslims to perform their pilgrimage to Mecca. Based on this experience, the first Islamic bank was established on 1 July 1983 (called Bank Islam) together with the introduction of the Islamic Bank Act, which came into force on 7 April 1983. In Malaysia, separate Islamic banking regulations exist side by side with those for conventional banking (Abdul Majid et al., 2011). Ten years later, in 1993, the Islamic

¹⁰³ Domestic banking groups in Malaysia undertake the full range of commercial banking, investment banking and Islamic banking activities. Many of these conglomerates also provide insurance protection as well as undertake fund management and management of unit trust funds within their groups (Tahir, Abu Bakar and Haron, 2008; Lee, 2011).

¹⁰⁴ In addition to the above, investment banks also play an active role in the short-term money market, stockbroking and capital raising activities such as financing, syndicating, corporate financing, management advisory services, arranging for the issue and listing of shares as well as managing investment portfolios (Sufian, 2011).

¹⁰⁵ The key features of Islamic banking are the prohibition of interest payment transactions, and the prohibition of undertaking of financing anti-social and immoral activities such as gambling, pornography, and consumption of alcohol and narcotics. Islamic banks are different from conventional banks because Islamic banks have to follow the concept of shariah (an Islamic principle), which employs the principles of justice, fair dealings and harmony, coupled with equitable wealth distribution as a basis of conducting business (Abdul Majid et al., 2011). For example, the products of Islamic banks are normally supported by various underlying assets, whereas the equivalent of loans in conventional banks is termed as 'financing'. It commonly involves the buying and selling of assets, and the profits gained for these transactions are amortised, equivalent to interest income for conventional banks.

Banking Scheme (IBS) was introduced by BNM in Malaysia, the first country to implement a dual banking system (i.e. operating both conventional and Islamic banking (IBS) under one banking institution) (Batchelor and Wadud, 2004). BNM also introduced an Islamic cheque clearing and settlement system, and an interbank money market system as well as Government Islamic Bonds and Notes, which operated alongside but separately from conventional banks (Abdul Majid et al., 2011). In 2005, BNM encouraged commercial banks under IBS to establish their fully-fledged Islamic subsidiaries to ensure a more prominent role in promoting the Islamic banking industry (Bank Negara Malaysia, 2005). As exhibited in Table 4.4, the total assets of Islamic banks in Malaysia as at the end of 2011 was at MYR 310.3 billion, which amounted to 19.1% of the total assets in Malaysian banks.

Table 4.4 Total Assets of Malaysian Banks by Specialisation

Year	Total Assets	Conventional Banks' Total Assets	Percentage of Conventional banks	Islamic Banks' Total Assets	Percentage of Islamic Banks
	(MYR millions)	(MYR millions)	(%)	(MYR millions)	(%)
2000	482,454	469,766	97.37	12,688	2.63
2001	522,042	505,639	96.86	16,403	3.14
2002	549,782	530,854	96.56	18,927	3.44
2003	609,325	587,693	96.45	21,632	3.55
2004	715,168	693,624	96.99	21,544	3.01
2005	827,462	794,760	96.05	32,702	3.95
2006	999,342	941,780	94.24	57,562	5.76
2007	1,142,305	1,050,058	91.92	92,247	8.08
2008	1,239,268	1,064,375	85.89	174,893	14.11
2009	1,344,062	1,127,952	83.92	216,110	16.08
2010	1,457,707	1,205,039	82.67	252,668	17.33
2011	1,622,197	1,311,832	80.87	310,365	19.13

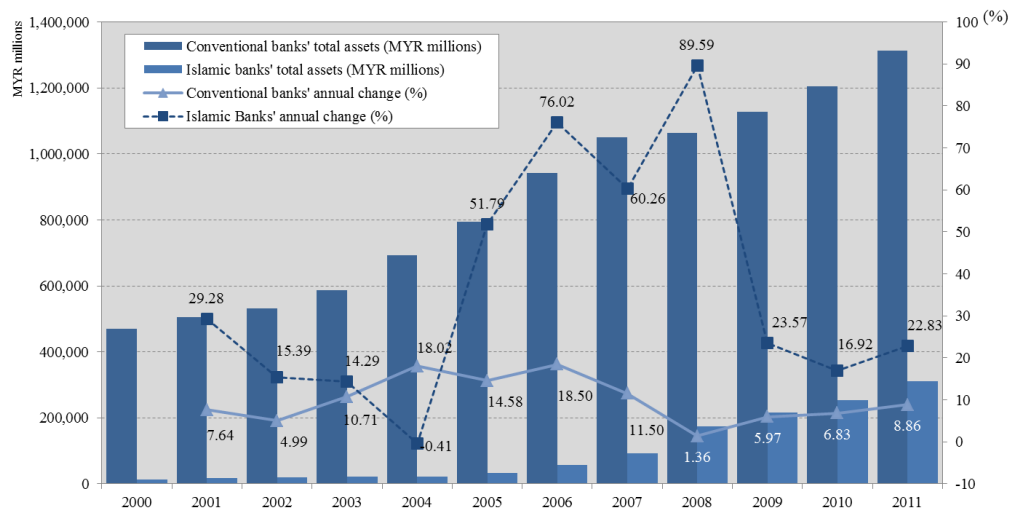
Source: Confidential Dataset from a Malaysian financial organisation

Note: As of December 2011, USD 1.0000 was equal to MYR 3.1265. Prior to 2005, most banks with IBS had Islamic assets within their banking structure

At the end of 2011, there were 16 Islamic banks in Malaysia. Five of them were foreign banks and the rest were owned by domestic banks. There were five standalone fully fledged banks (two domestic- and three foreign-owned), and the balance of 11 banks were Islamic banking subsidiaries within their respective banking groups. As of 2011,

the market share held by the Islamic banks was 19.1%, and was on an increasing trend since 2000. The average annual growth for Islamic banks between 2000 and 2011 was at 36.3%. Following the three different phases of FSMP, the average annual growth for years 2000–2003, 2004–2007 and 2008–2011 were registered at 19.7%, 46.9% and 38.2%, respectively. From Figure 4.7, there were significant increases in the total assets of Islamic banks between 2005 and 2008, mainly because of the formation of new Islamic banking subsidiaries within their respective banking groups and the commencement of operation of new foreign Islamic banks.

Figure 4.7 Total Assets of Conventional and Islamic Banks in Malaysia



Source: Confidential Dataset from a Malaysian financial organisation

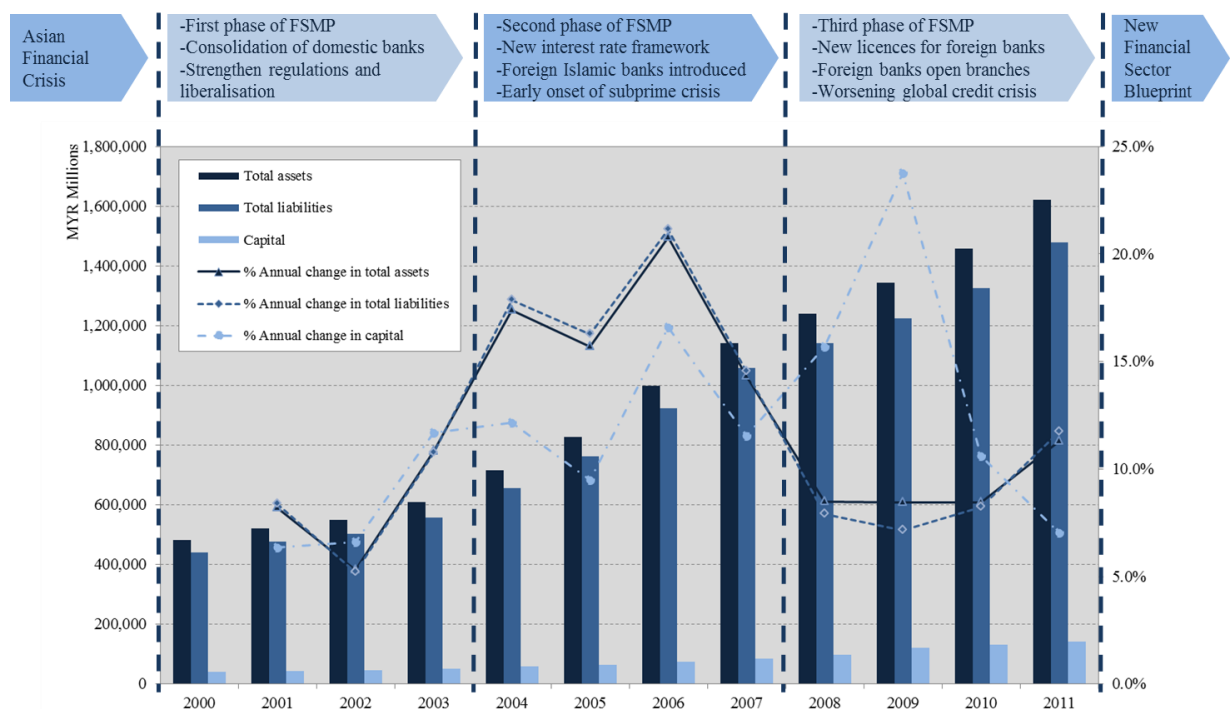
4.4.3 Assets and Liabilities Structure of Malaysian Banks

Figure 4.8 exhibits the growth of assets and liabilities of Malaysian banks between 2000 and 2011. As of 2011, the total assets of Malaysian banks amounted to MYR 1,622.2 billion. The average annual growth of the total assets for the period 2000–2011 was at 11.7%. Based on the three phases of FSMP, the average annual growth of total assets for the periods 2000–2003, 2004–2007 and 2008–2011 were registered at 8.1%, 17.0% and 10.9%, respectively.

During the first phase of FSMP, the main policy thrust focused on enhancing the performance and competitiveness of domestic banks and ensured that they were ready to face a more liberalised operating environment (Bank Negara Malaysian, 2001b). The domestic banks exhibited steady growth in the years 2000–2003. The average annual growth during the first phase of FSMP for total assets, total liabilities and capital were at 8.1%, 8.1% and 8.2%, respectively.

During the initial second phase of FSMP, BNM intensified the implementation of gradual liberalisation and increased competitive pressures in the banking sector. It can be observed that the growth of assets, liabilities and capital of Malaysian banks were on an increasing trend during this period. Growth was driven by the introduction of NIRF and the commencement of new foreign Islamic Banks in 2004 and 2006, respectively. However, at the end of the second phase of FSMP in 2007, the growth of assets, liabilities and capital declined due to slower export activities by the real sectors (weakened global demand) and the emerging subprime crisis, which affected financial markets in US and Europe.

Figure 4.8 Total Assets, Total Liabilities and Capital in Malaysian Banks



Source: Confidential dataset from a Malaysian financial organisation

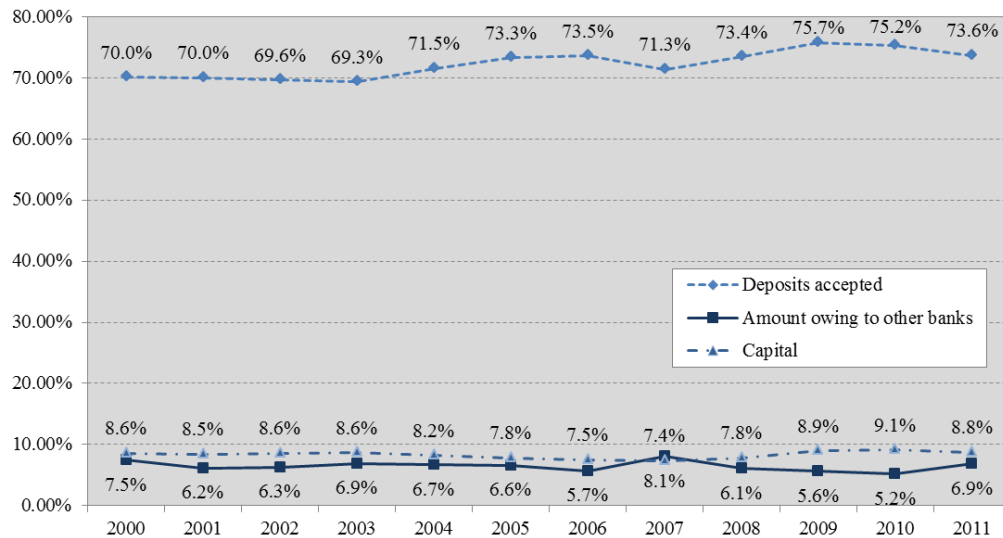
As the global economy worsened in 2008, the Malaysian banks were cautious in their lending strategies. As the timeline moved into the third phase of FSMP (2008–2011), the average annual growth for total assets, total liabilities and capital recorded growths of 9.2%, 8.8% and 14.3% respectively; which were lower compared to the second phase of FSMP. However, the growth of Malaysian banks' capital trended in a different direction. This is because most Malaysian banks undertook pre-emptive measures to increase their capital position in order to absorb potential losses from adverse shocks. Despite the difficulties in raising capital during economic turmoil, the Malaysian banks were successful in raising high quality capital (Tier-1 capital) within a short period of time at a reasonable cost (Bank Negara Malaysia, 2009b). Consequently, in 2011, the assets and liabilities trended upward, which implied a sign of recovery for the banking sector.

As intermediary institutions, banks' assets and liabilities are influenced by the composition of sources and uses of funds (i.e. accumulating and allocating of funds). First, the sources of funds for Malaysian banks are represented by (1) deposits from customers, (2) amounts owing to other banks (i.e. interbank borrowing/deposit) and (3) capital.

Figure 4.9, Figure 4.10 and Table 4.5 display the structure and annual growth of the funding of Malaysian banks. Based on Figure 4.9, the customers' deposit is the largest portion of the sources of funds in Malaysian banks. The average composition of deposit accepted during the period under study (2000–2011) is 72.2%. In terms of growth, the average annual growth for the periods 2000–2003, 2004–2007 and 2008–2011 are registered at 7.7%, 17.9% and 10.0%, respectively. Higher average annual growth in deposits is observed in the second period of FSMP. The main reason for the higher increase during the period 2004–2006 was mainly due to the improved confidence of customers in banking institutions; this was supported by the introduction of new foreign

Islamic banks and initiation of deposit insurance (PIDM) (Hon et al., 2011).¹⁰⁶ Additionally, growth in the second phase of FSMP expanded the deposits' product range and diversity (Bank Negara Malaysia, 2006b).¹⁰⁷

Figure 4.9 Percentage of Capital, Amounts Owing to Other Banks and Deposits Accepted relative to Liabilities and Capital, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

Another important component of the sources of funds is the amounts borrowed from other banks. This interbank item exhibited a sudden upsurge of inflow in 2007, driven by the early stage of the subprime crisis and financial market turbulence in the US and Europe. The banks in US and Europe released their position from these inferior asset-backed structured products and placed the excess liquidity elsewhere in search of a higher yield, such as Malaysia. The increase was more pronounced in foreign banks (Bank Negara Malaysia, 2007b).¹⁰⁸ Consequently, as deleveraging activities took place to strengthen the parent and head office in US and Europe, the interbank deposits of Malaysian banks declined in 2008 (Bank Negara Malaysia, 2008b, 2011b).

¹⁰⁶ With the establishment of PIDM in 2005, depositors were guaranteed up to MYR 60,000 on total deposits placed with a single member bank (Bank Negara Malaysia, 2007b).

¹⁰⁷ For example, the deposit products have broadened from 'plain vanilla' deposits to structured products that offer returns based on the performance of the underlying assets and currency movements (Bank Negara Malaysia, 2007b).

¹⁰⁸ This condition can lead to unsustainable growth in asset prices and financial imbalances. The cumulative liquidity surplus in the banking sector may carry the potential risk of higher liquidity demands and unexpected withdrawals (Bank Negara Malaysia, 2007b)

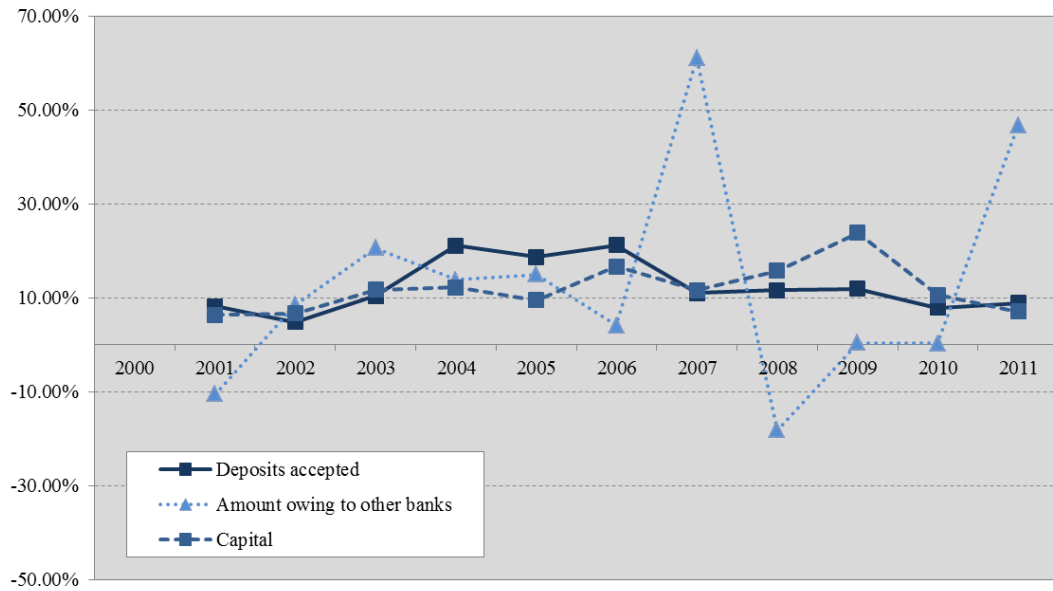
The final component of sources of funds is capital. The average percentage of the capital against the total sources of funds is approximately 8.6%.¹⁰⁹ Based on the three phases of the FSMP, the average capital ratios were 8.5%, 7.7% and 8.7% for the periods 2000–2003, 2004–2007 and 2008–2011 respectively. The reason for the lower capital ratio in the second phase of FSMP was the influence of active capital management performed by the banks. During this period, the fraction of Tier-1 capital declined as higher issuances of Tier-2 subordinated debt capital were observed where banks had undertaken to diversify and improve the overall cost of capital as well as providing attractive long-term returns to shareholders (Bank Negara Malaysia, 2005, 2006b).¹¹⁰ However, in the third phase of the FSMP, the banks pre-emptively responded to the global credit crisis by increasing their capital position. The capital position was strengthened through rights issues and issuance of Tier-1 capital instruments (Bank Negara Malaysia, 2008b). Moreover, the introduction of a regulatory capital framework, namely Basel II (both Standardised Approach and Internal Rating Based Approach) also influenced the need for banks to have stronger capital position.¹¹¹ As such, there were significant increases of 15.7% and 23.8% of capital in Malaysian banks in 2008 and 2009 respectively (see Figure 4.10).

¹⁰⁹ However, this ratio does not indicate the regulatory risk-based capital as measured by Basel's Risk Weighted Capital Ratio (RWCR). The average RWCR had consistently exceeded 12.0% over the years under study. The capital ratio shown is indicative of the capital available without taking into account any risk-weighted factors on assets held by the banks.

¹¹⁰ Tier-1 capital is described as core capital, which usually includes equity capital and retained earnings. Additionally, Tier-2 capital consists of supplementary capital such as revaluation reserves, hybrid instruments and subordinated term debt (Basel Committee on Banking Supervision, 1999, 2005).

¹¹¹ Despite the greater capital savings that can be achieved from lower risk-weights of retail loan portfolios, the new capital charges from operational risks can diminish and limit the capital savings from the implementation of Basel II (Bank Negara Malaysia, 2008b).

Figure 4.10 Annual Growth of Deposits Accepted, Amounts Owing to Other Banks and Capital of Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

Table 4.5 Sources of Funds of Malaysian Banks (Liabilities and Capital), 2000–2011

Description	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Total Liabilities and Capital (MYR millions)	482,454	522,042	549,782	609,325	715,168	827,462	999,342	1,142,305	1,239,268	1,344,062	1,457,707	1,622,197
Annual Growth of Total Liabilities and Capital (%)	0	7.58	5.05	9.77	14.80	13.57	17.20	12.52	7.82	7.80	7.80	10.14
Domestic Banks' Total Liabilities and Capital (%)	74.82	74.82	75.09	73.81	74.24	75.80	77.46	76.05	75.80	77.08	77.17	75.67
Foreign Banks' Total Liabilities and Capital (%)	25.18	25.18	24.91	26.19	25.76	24.20	22.54	23.95	24.20	22.92	22.83	24.33
Conventional Banks' Total Liabilities and Capital (%)	97.37	96.86	96.56	96.45	96.99	96.05	94.24	91.92	85.89	83.92	82.67	80.87
Islamic Banks' Total Liabilities and Capital (%)	2.63	3.14	3.44	3.55	3.01	3.95	5.76	8.08	14.11	16.08	17.33	19.13
Total Deposits from Customers (MYR millions)	337,796	365,314	382,735	422,245	511,116	606,234	734,648	814,947	909,345	1,017,010	1,096,884	1,193,930
Deposits from Customers as % of Total Liabilities and Capital	70.02	69.98	69.62	69.30	71.47	73.26	73.51	71.34	73.38	75.67	75.25	73.60
Annual Growth of Total Deposits from Customers		7.53	4.55	9.36	17.39	15.69	17.48	9.85	10.38	10.59	7.28	8.13
Domestic Banks' Total Deposits from Customers (%)	75.59	76.52	76.64	75.22	76.19	77.14	77.98	77.55	76.95	77.98	78.02	76.95
Foreign Banks' Total Deposits from Customers (%)	24.41	23.48	23.36	24.78	23.81	22.86	22.02	22.45	23.05	22.02	21.98	23.05
Conventional Banks' Total Deposits from Customers (%)	96.87	96.33	95.81	95.65	96.46	95.51	93.61	90.94	84.42	82.81	81.04	79.60
Islamic Banks' Total Liabilities and Capital (%)	3.13	3.67	4.19	4.35	3.54	4.49	6.39	9.06	15.58	17.19	18.96	20.40
Total Amount Due to Banks (MYR millions)	35,951	32,182	34,908	42,085	47,892	55,025	57,281	92,181	75,369	75,639	75,881	111,264
Amount Due to Banks as % of Total Liabilities and Capital (%)	7.45	6.16	6.35	6.91	6.70	6.65	5.73	8.07	6.08	5.63	5.21	6.86
Annual Growth of Total Due to Banks (%)		-11.71	7.81	17.05	12.13	12.96	3.94	37.86	-22.31	0.36	0.32	31.80
Domestic Banks' Total Amount Due to Banks (%)	60.68	54.39	56.50	50.53	43.61	54.38	62.85	55.24	62.78	64.42	64.87	57.27
Foreign Banks' Total Amount Due to Banks (%)	39.32	45.61	43.50	49.47	56.39	45.62	37.15	44.76	37.22	35.58	35.13	42.73
Conventional Banks' Total Amount Due to Banks (%)	99.83	98.07	99.48	99.65	99.25	99.25	96.37	97.09	90.98	83.22	83.31	76.54
Islamic Banks' Total Amount Due to Banks (%)	0.17	1.93	0.52	0.35	0.75	0.75	3.63	2.91	9.02	16.78	16.69	23.46
Total Capital (MYR millions)	41,516	44,159	47,082	52,593	58,992	64,588	75,303	84,006	97,181	120,273	133,033	142,408
Capital as % of Total Liabilities and Capital (%)	8.61	8.46	8.56	8.63	8.25	7.81	7.54	7.35	7.84	8.95	9.13	8.78
Annual Growth of Total Capital (%)		5.98	6.21	10.48	10.85	8.66	14.23	10.36	13.56	19.20	9.59	6.58
Domestic Banks' Total Capital (%)	77.35	75.77	76.92	77.85	78.19	78.87	78.83	77.97	76.15	77.95	78.24	77.30
Foreign Banks' Total Capital (%)	22.65	24.23	23.08	22.15	21.81	21.13	21.17	22.03	23.85	22.05	21.76	22.70
Conventional Banks' Total Capital (%)	96.86	96.98	97.07	97.20	97.18	95.85	94.94	90.86	86.69	85.96	85.03	85.01
Islamic Banks' Total Capital (%)	3.14	3.02	2.93	2.80	2.82	4.15	5.06	9.14	13.31	14.04	14.97	14.99

Source: Confidential dataset from a Malaysian financial organisation

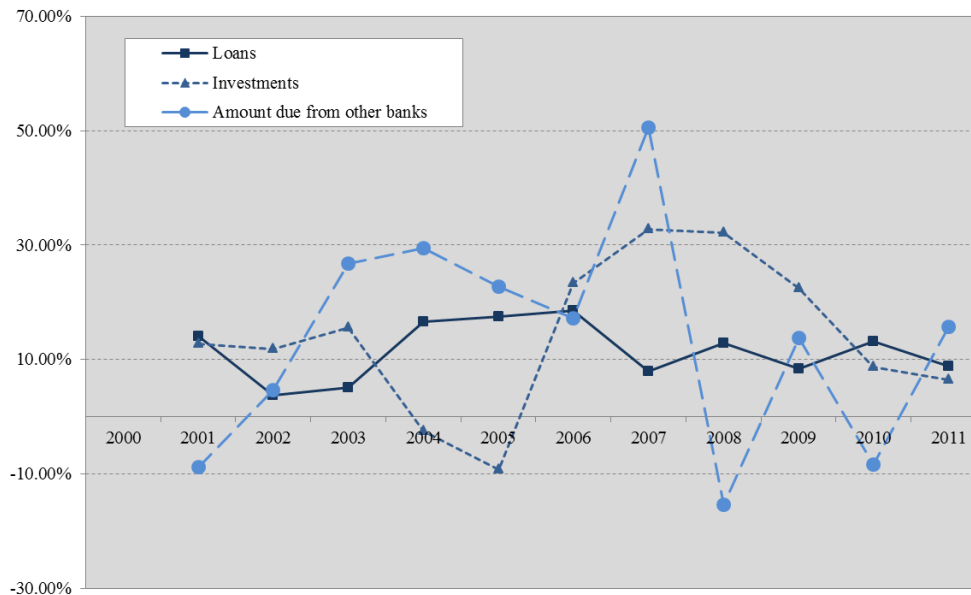
Note: As of December 2011, USD 1.0000 is equal to MYR 3.1265

Second, the uses of funds on the assets side of Malaysian banks are mainly used to (1) provide loans and financing to customers, namely households and businesses as well as the public sector; (2) hold securities, particularly government papers and bonds; 3) place deposits into other banks.

Loans (also termed as ‘financing’ in Islamic banks) remained the biggest and most important component within the assets of banks, with a share of approximately 60.0% of total assets (Bank Negara Malaysia, 1999b). A closer look at Table 4.6 highlights the increasing trend of loans in Malaysian banks during the years under study. The annual average growth of the loans for the period 2000–2003, 2004–2007 and 2008–2011 (based on indicative timeline grouped for three different phases of FSMP) were registered at 7.6%, 15.2% and 10.9% respectively. The significant average growth during the second phase of FSMP was influenced by the liberalisation of interest rates, in which banks were allowed to determine their own BLR as signalled by the BNM’s OPR (see Figure 4.11). Thus, the demand for loans increased when lower interest rates were offered by the banks (as a consequence of an increasingly competitive environment). The increase in loans was also driven by the household sector, particularly for the purchase of residential properties and consumer credit, which accounted for approximately 57.0% of total loans (Bank Negara Malaysia, 2005).¹¹² Consequently, amidst the global subprime crisis and during the third phase of FSMP, the average annual growth of loans decreased to 10.9%, implying that banks were more cautious in lending and busy managing delinquent loans’ accounts.

¹¹² After learning from the 1997–1998 Asian financial crisis and the consolidation of fragmented domestic banks, the Malaysian banks shifted their strategy by rebalancing the loans portfolio, aiming to capture a higher market share in the retail sector and invest in debt instrument from capital markets for exposure to large corporates. The higher growth during the second phase of FSMP was also influenced by consumer confidence and was supported by increased private consumption with continued high income levels, a stable labour market environment and improved business conditions and productivity (Bank Negara Malaysia, 2006a). This was further augmented by initiatives and measures introduced by the government and other regulatory agencies to support the residential property sector such as exemptions from the real property gains tax, a 50% exemption from stamp duty for houses below the value of MYR 250,000, and allowing monthly withdrawals for Employees Provident Funds (EPF) accounts for repayment of housing loans (Bank Negara Malaysia, 2007a).

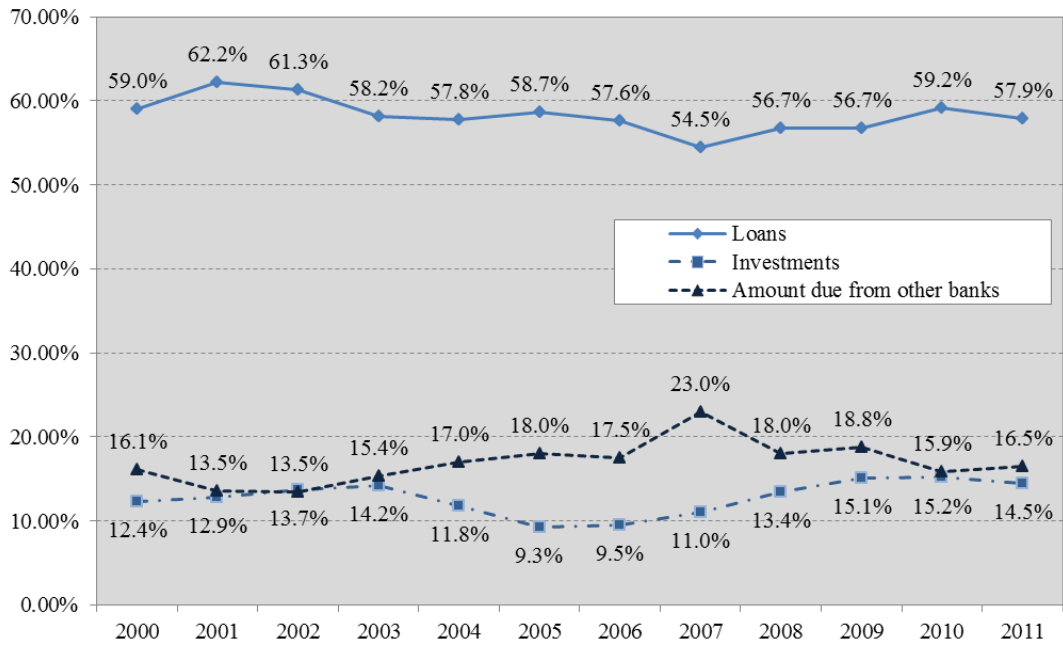
Figure 4.11 Annual Growth of Loans, Investments and Amounts Due from Other Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

The second largest class of assets is the amount due from other banks, accounting for an average of 16.9% of total assets over the period 2000–2011. Based on Figure 4.12, the growth of interbank assets increased significantly in 2007 as there was a persistent inflow of funds into Malaysia which posed liquidity challenges to Malaysian banks in 2007 (Bank Negara Malaysia, 2007b). The sudden influx of foreign funds posed a challenge to the banking system, where excess liquidity and short-term maturity was matched by placing deposits into other banks. Soon after that, as deleveraging of assets took place in US and European banks in 2008, substantial withdrawals from interbank assets were observed. In addition, the investment items also displayed a significant increase between 2006 and 2008. This was primarily driven by the excess liquidity during the onset of the subprime crisis. Between 2007 and 2008, the excess liquidity in the banking industry was mopped up by BNM through its money market operations, which resulted in a higher investment of BNM instruments in Malaysian banks (Bank Negara Malaysia, 2008b).

Figure 4.12 Percentage of Loans, Investments and Amounts Due from Other Banks Relative to Total Assets, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

Table 4.6 Uses of Funds of Malaysian Banks (Total Assets), 2000–2011

Description	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Total Assets (MYR millions)	482,454	522,042	549,782	609,325	715,168	827,462	999,342	1,142,305	1,239,268	1,344,062	1,457,707	1,622,197
Annual Growth of Total Assets (%)		7.58	5.05	9.77	14.80	13.57	17.20	12.52	7.82	7.80	7.80	10.14
Domestic Banks' Total Assets (%)	74.82	74.82	75.09	73.81	74.24	75.80	77.46	76.05	75.80	77.08	77.17	75.67
Foreign Banks' Total Assets (%)	25.18	25.18	24.91	26.19	25.76	24.20	22.54	23.95	24.20	22.92	22.83	24.33
Conventional Banks' Total Assets (%)	97.37	96.86	96.56	96.45	96.99	96.05	94.24	91.92	85.89	83.92	82.67	80.87
Islamic Banks' Total Assets (%)	2.63	3.14	3.44	3.55	3.01	3.95	5.76	8.08	14.11	16.08	17.33	19.13
Total Loans (MYR millions)	284,791	324,861	336,882	354,377	413,247	485,635	576,060	622,008	702,519	761,918	862,777	939,759
Loans as % of Total Assets (%)	59.03	62.23	61.28	58.16	57.78	58.69	57.64	54.45	56.69	56.69	59.19	57.93
Annual Growth of Total Loans		12.33	3.57	4.94	14.25	14.91	15.70	7.39	11.46	7.80	11.69	8.19
Domestic Banks' Total Loans (%)	75.22	75.22	74.56	73.82	75.49	77.66	79.36	78.22	78.05	79.23	78.95	77.49
Foreign Banks' Total Loans (%)	24.78	24.78	25.44	26.18	24.51	22.34	20.64	21.78	21.95	20.77	21.05	22.51
Conventional Banks' Total Loans (%)	98.13	97.76	97.56	97.32	97.52	96.87	95.03	92.93	86.06	83.74	82.37	80.55
Islamic Banks' Total Loans (%)	1.87	2.24	2.44	2.68	2.48	3.13	4.97	7.07	13.94	16.26	17.63	19.45
Total Investments (MYR millions)	59,614	67,185	75,139	86,795	84,567	76,804	94,735	125,786	166,310	203,556	221,154	235,547
Investments as % of Total Assets (%)	12.36	12.87	13.67	14.24	11.82	9.28	9.48	11.01	13.42	15.14	15.17	14.52
Annual Growth of Total Investments (%)		11.27	10.59	13.43	-2.64	-10.11	18.93	24.69	24.37	18.30	7.96	6.11
Domestic Banks' Total Investments (%)	75.14	73.49	79.43	75.34	73.94	75.61	76.19	73.97	73.96	80.66	81.50	79.02
Foreign Banks' Total Investments (%)	24.86	26.51	20.57	24.66	26.06	24.39	23.81	26.03	26.04	19.34	18.50	20.98
Conventional Banks' Total Investments (%)	94.92	94.60	93.06	92.76	93.01	91.55	91.11	89.99	85.97	83.80	79.80	79.66
Islamic Banks' Total Investments (%)	5.08	5.40	6.94	7.24	6.99	8.45	8.89	10.01	14.03	16.20	20.20	20.34
Total Amount Due from Banks (MYR millions)	77,463	70,704	73,981	93,815	121,452	149,032	174,598	262,909	222,512	253,203	231,938	268,412
Amount Due from Banks as % of Total Assets (%)	16.06	13.54	13.46	15.40	16.98	18.01	17.47	23.02	17.96	18.84	15.91	16.55
Annual Growth of Total Due from Banks (%)		-9.56	4.43	21.14	22.76	18.51	14.64	33.59	-18.16	12.12	-9.17	13.59
Domestic Banks' Total Amount Due from Banks (%)	68.15	69.18	69.73	70.52	73.81	70.41	70.89	71.22	74.20	68.63	66.79	64.97
Foreign Banks' Total Amount Due from Banks (%)	31.85	30.82	30.27	29.48	26.19	29.59	29.11	28.78	25.80	31.37	33.21	35.03
Conventional Banks' Total Amount Due from Banks (%)	96.03	94.66	94.91	96.23	97.23	94.10	90.21	88.26	79.36	79.13	78.87	74.71
Islamic Banks' Total Amount Due from Banks (%)	3.97	5.34	5.09	3.77	2.77	5.90	9.79	11.74	20.64	20.87	21.13	25.29

Source: Confidential dataset from a Malaysian financial organisation

Note: As of December 2011, USD 1.0000 was equal to MYR 3.1265

4.4.4 Income Statements Structure of Malaysian Banks

The previous section reviews the balance sheet position of the Malaysian banks. This section, on the other hand, explains the structure of the income statements of the Malaysian banks and discusses the liberalisation measures and regulatory initiatives introduced by BNM that affected the banking institutions' income, costs and profits.

First, total income of the Malaysian banks and its major components are exhibited in Table 4.7, Figure 4.13 and Figure 4.14. The Malaysian banks primarily depended on interest income as a source of income amounting to more than 65.0% of their total income during the period 2000–2011.¹¹³ The average annual growth rates between 2000 and 2011 in total income, interest income and non-interest income were 10.6%, 8.9% and 16.9% respectively. A steady growth of total income is observed particularly during the post-consolidation period from 2000 to 2003. In the second phase of FSMP, with the introduction of new liberalisation measures (e.g. a new interest rate framework (NIRF), new Islamic foreign banks and safety net institutions (PIDM)), demand from the retail sector subsequently increased competition among the banks, as evident from their greater lending activities (Bank Negara Malaysia, 2006a).¹¹⁴ Also, in the second phase of FSMP (2004–2007), the average annual growth of total income was at 19.0%, higher than in the first and third phases of FSMP, which displayed 6.6% and 5.3% respectively.

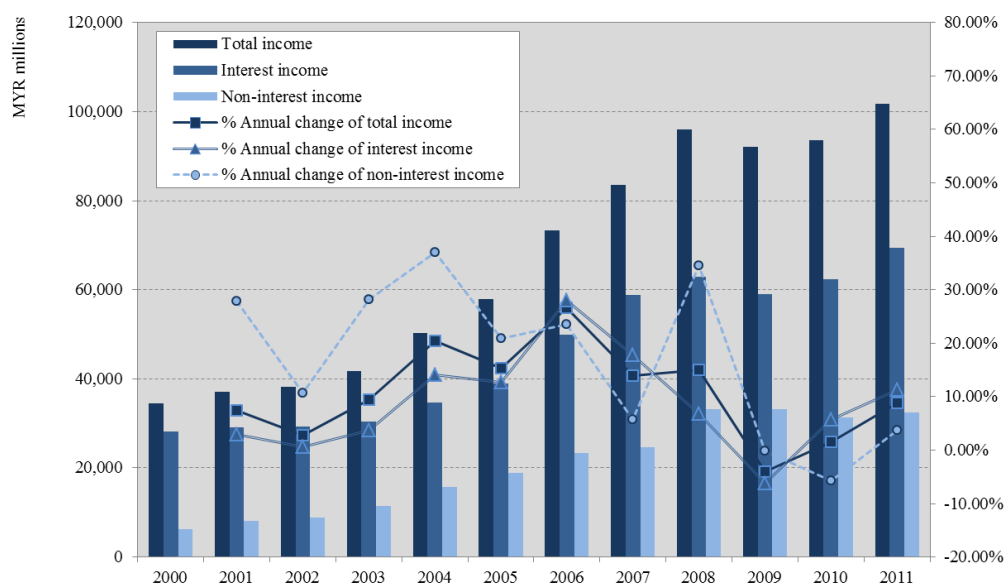
Consequently, in 2007, as the subprime crisis set in, there was higher volatility in the total income of the banks. The annual growth of total income declined slightly due to the anticipation of global economic uncertainty. Subsequently, in 2009, the gross interest rate margin narrowed as BNM reduced the OPR rapidly in three stages from 3.25% to 2.00%. As the global economic condition worsened, the income of Malaysian

¹¹³ Nevertheless, for Islamic banks, the profits received from financing and other assets are equivalent to interest income, which is encapsulated as interest income to banks for the purpose of this research.

¹¹⁴ Partly, the higher interest income in 2006 was also influenced by the BNM's decision on increasing OPR in November 2005, which saw an increase of 40 basis points in the 3-month Kuala Lumpur interbank overnight rate (KLIBOR) from 2.80% to 3.20%. In return, this move improved the interest income of banks during the year 2006. The strong growth in 2004–2007 shows that the income of banking institutions was increasingly diversified, mainly from revenues of financing and related advisory activities, provision of remittance and settlement services, portfolio and wealth management and trading and investment activities (Bank Negara Malaysia, 2007b).

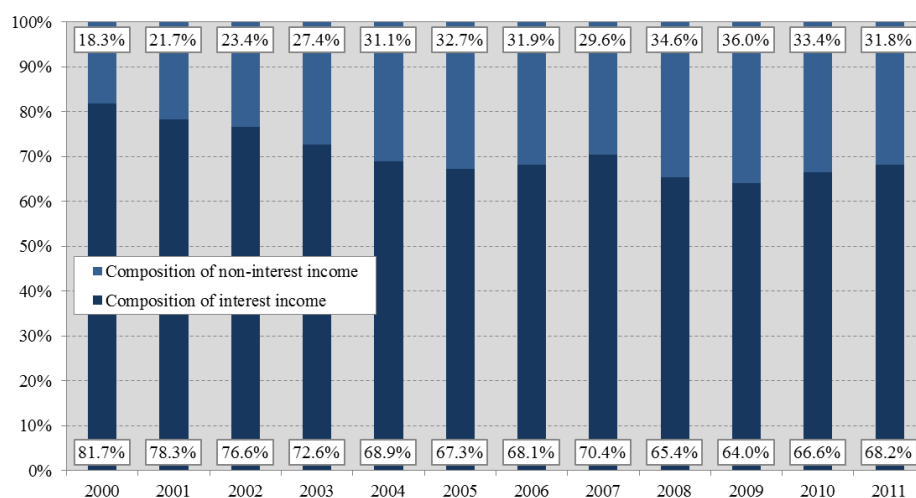
banks declined further; this was attributed to lower interest income, higher valuation losses from trading and investment portfolios, and selective activities of banks in attracting and retaining customers' pool base with strong credit standings (Bank Negara Malaysia, 2008b) (see Figure 4.13). As economic recovery emerges, BNM increased the OPR in late 2010, which in return increased the total interest income of banking sector.

Figure 4.13 Growth and Structure of Income of Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

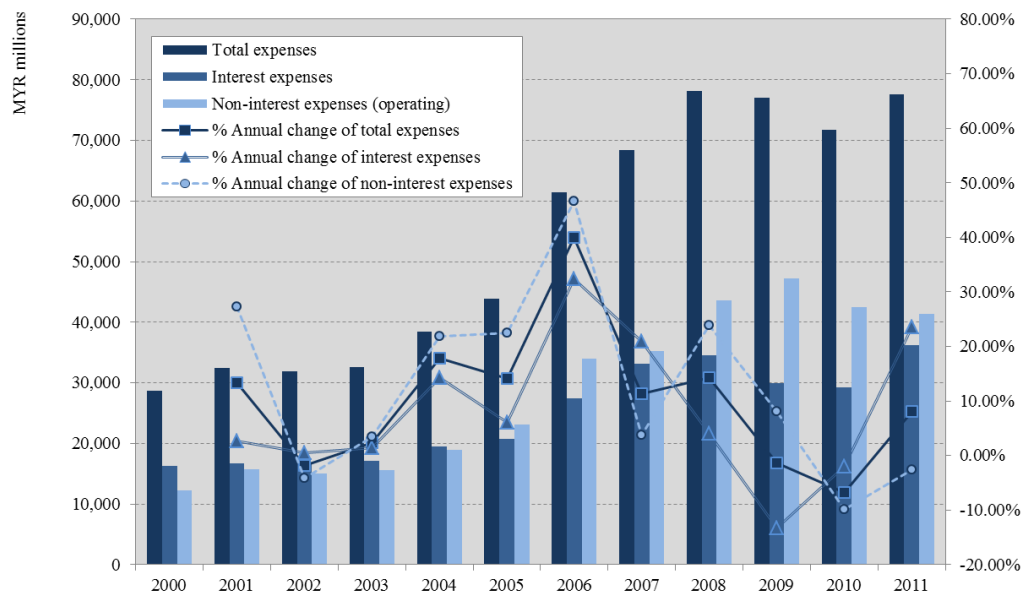
Figure 4.14 Composition of Income of Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

Second, Figure 4.15, Figure 4.16 and also Table 4.7 display the growth and sources of expenses of Malaysian banks between 2000 and 2011. The growth of the total expenses had increased steadily during the first phase of FSMP. The average annual growth for this period (2000–2003) was at 4.6%. Then, during the second phase of FSMP between 2004 and 2007, there was greater volatility in the growth of total expenses. The average annual growth for period 2004–2007 increased by 20.8%. During this period, the composition of non-interest expenses was more than the composition of interest expenses. This can be explained from the increasing operating expenses incurred by banks in an environment of intensified competition due to liberalisation of interest rates, an end to the of wage moratorium on bank personnel, and the introduction of new foreign players.¹¹⁵

Figure 4.15 Growth and Structure of Expenses in Malaysian Banks, 2000–2011

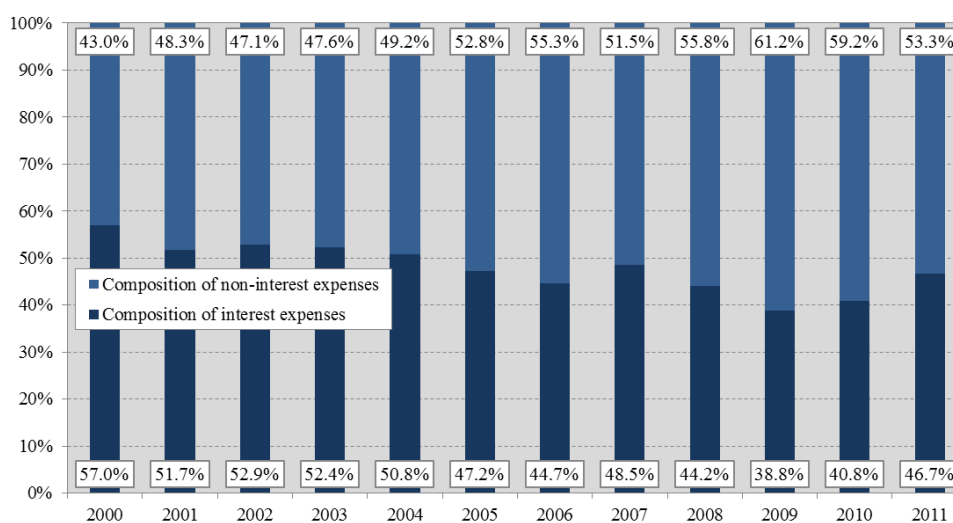


Source: Confidential dataset from a financial organisation

¹¹⁵ With regards to staff expenditure, investment in the acquisition and development of skills is necessary when banks are experiencing business expansion and have to retain existing, and attain new, talents (Bank Negara Malaysia, 2005). The highly competitive labour market in the banking sector has led to higher personnel-related expenditure, particularly with the introduction of new foreign Islamic banks where specific experience and skills were sought for these new establishments (Bank Negara Malaysia, 2006b). Moreover, amidst intensified competition, there is also a significant increase in operating overheads, which implies the aggressive marketing and sales strategies performed by banks in Malaysia (Bank Negara Malaysia, 2005). Concurrently, the policy interest rate via OPR was also increased by BNM, where banks were expected to pay higher interest rates to customers in 2007.

During the third phase of FSMP (i.e. 2008–2011), the total expenses of Malaysian banks were on a decreasing trend. The declining trend of expenses was driven by the reduction of OPR in 2008 and 2009. This can be evident from the proportion between interest and non-interest expenses where non-interest expenses exhibited a higher percentage than interest expenses in 2009 and 2010 (see Figure 4.16). This also suggests that Malaysian banks were affected by higher operating costs such as managing liquidity surpluses and administering potential loan losses, particularly during this period of uncertainty. Additionally, non-performing loans had also increased marginally (2008: 3.7%; 2009: 3.8%), reflecting the slowdown in economic condition (as evidenced from Malaysia's GDP, which contracted by 1.7% in 2009) (Bank Negara Malaysia, 2009a).

Figure 4.16 Composition of Expenses in Malaysian Banks, 2000–2011

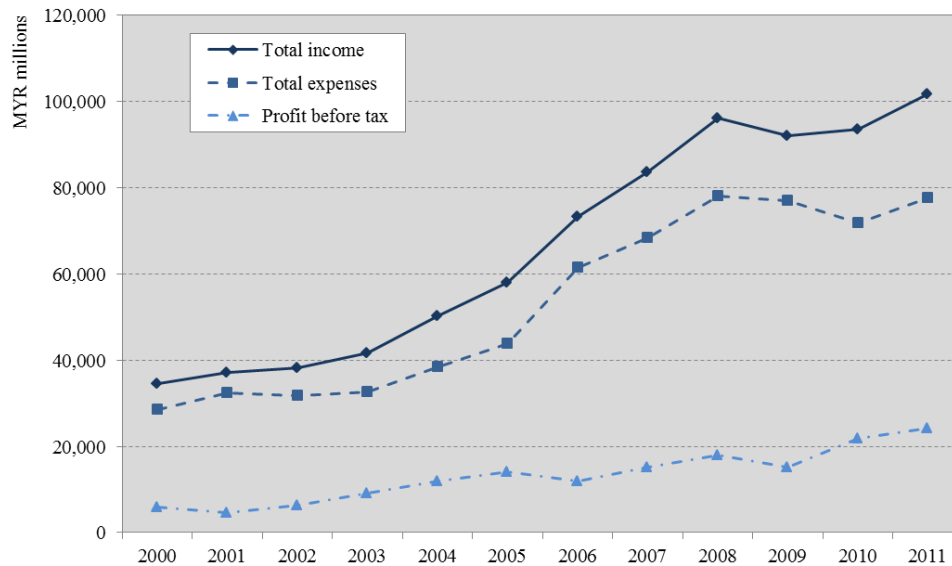


Source: Confidential dataset from a Malaysian financial organisation

Third, the trend of the profit before tax (PBT) for Malaysian banks is also shown in Table 4.7, Figure 4.17 and Figure 4.18. The PBT of Malaysian banks has increased steadily since 2000. During the period of first FSMP (i.e. 2000–2003), the gap between income and expenditure was fairly constant. Most of the products and services offered by banks during this period were ‘plain vanilla’ and traditional. The steady growth of PBT was attributed to the consolidation of domestic banks, which managed to synergise their operations in a short period of time.

Whilst liberalisation initiatives escalated in the second phase of FSMP (2004–2007), the banking system continued to record stronger profit performance at the early phase, underpinned by favourable macroeconomic conditions and financial markets. Competition in the lending market, particularly in the retail segment, resulted in higher returns from lending interest income. Moreover, greater competitive pressure had forced banks to improve their product innovativeness and cross-selling activities. The profit of the banks was also influenced by revenues generated from diverse products and services, and not limited to lending – such as wealth management-related products (e.g. unit trust and bancassurance), remittance services and trade related products (e.g. guarantees and commission form issuance of bankers' acceptance) (Bank Negara Malaysia, 2005). Consequently, the profitability of Malaysian banks declined in 2006 (see Figure 4.17). This was influenced by significant losses faced by Bank Islam (one of the domestic full-fledged Islamic banks), which was experiencing a substantial write-down of large financing (or loans, in conventional banking terms). The write-offs led to higher losses in Bank Islam amounting to almost 30.0% of the financing portfolio, caused by cross-border currency advances with large exposure in Bosnia, Indonesia and the Middle East via its Labuan offshore branch. Nevertheless, Bank Islam had managed to recapitalise its capital position within a year and reduced its NPFs to almost the industry average after four years (Rating Agency Malaysia, 2010).

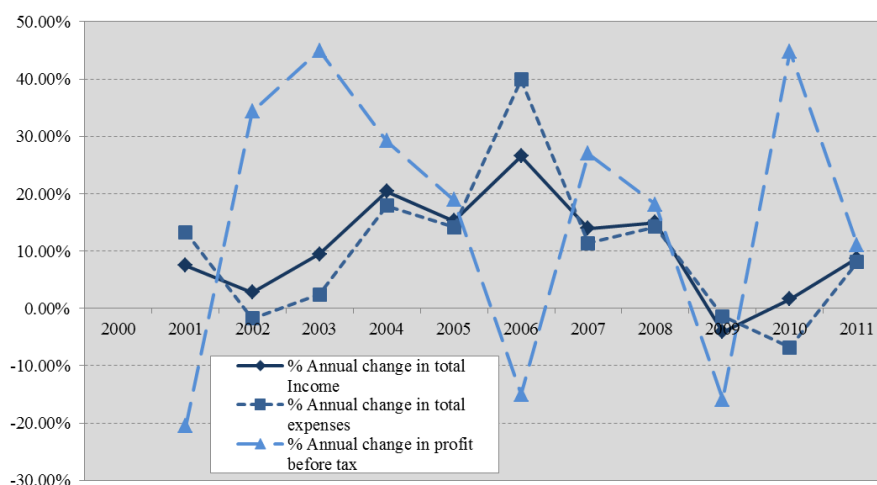
Figure 4.17 Trends of Profit before Tax, Total Income and Total Expenses of Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

In tandem with global economic slowdown, the period 2008–2011 saw greater volatility in the profitability of Malaysian banks. As mentioned earlier, Malaysia suffered contraction of GDP by 1.7% in 2009. The lower profit of Malaysian banks in 2009 was influenced by lower interest income (OPR in 2009 was reduced to 2.0% from 3.5% in 2008) as well as a slight increase in delinquency level of household and business loans (Bank Negara Malaysia, 2009b). Consequently, the profit of Malaysian banks bounced back and expanded by 7.2% in 2010. The growth was mainly driven by improved domestic demand for private sector activities, attributed to fiscal stimulus by the government, improved labour market conditions, a steady increase in household and business income, and continued access to credit (Bank Negara Malaysia, 2010b).

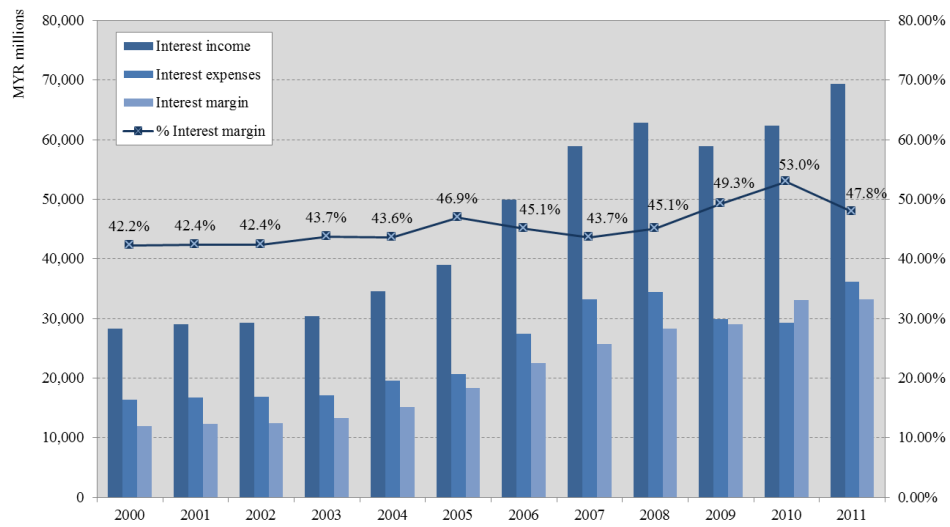
Figure 4.18 Annual Growth in Total Income, Total Expenses and Profit before Tax in Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

In addition, Figure 4.19 exhibits the trend of interest margin in percentage and MYR amount.¹¹⁶ The average interest margin for the period under study was recorded at 45.4%. The interest rate margin was fairly stable during the first and early second phase of FSMP, but greater fluctuations were observed as the timeline enters the third phase of FSMP, which was influenced by the uncertainty of the global economic crisis. The increasing trend of interest margin from 43.7% in 2007 to 53.0% in 2010 was attributed by lower interest expense to interest income. The interest expenses of Malaysian banks were lower in 2009 and 2010 due to lower interest offered to customers for deposits and higher interest rate on loans offered to customers. The lower interest rates on deposits were caused by the movement of overnight policy rate (OPR), which may affect banks' BLR. On the other hand, although interest rates for loans are expected to decline (in tandem with the movement of OPR), it was however observed that Malaysian banks charged higher interest rates that were attributed by higher demand of credit by both household and business sectors. The higher demand may also imply that borrowers during this period of uncertainty require funds to remain afloat during economic downturn and willing to be charged at higher interest rates.

¹¹⁶ The interest rate margin is measured by a ratio of interest margin over total interest income. The ratio indicates the percentage of interest profitability derived from the interest income after deducting the interest expenses incurred by the banks.

Figure 4.19 Trend of Interest Margin in Malaysian Banks, 2000–2011

Source: Confidential dataset from a Malaysian financial organisation

Table 4.7 Income Statement Structure for Malaysian Banks, 2000–2011

Description	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Total Income (MYR millions)	34,545	37,129	38,164	41,767	50,278	57,953	73,357	83,581	96,078	92,147	93,618	101,811
Annual Growth of Total Income (%)		6.96	2.71	8.63	16.93	13.24	21.00	12.23	13.01	-4.27	1.57	8.05
of which Interest Income (MYR millions)	28,241	29,071	29,240	30,331	34,618	39,023	49,965	58,856	62,825	58,965	62,340	69,392
Interest Income as % of Total Income (%) ¹	81.75	78.30	76.62	72.62	68.85	67.34	68.11	70.42	65.39	63.99	66.59	68.16
Annual Growth of Interest Income (%)		2.86	0.58	3.60	12.38	11.29	21.90	15.11	6.32	-6.55	5.41	10.16
of which Non-Interest Income (MYR millions)	6,305	8,058	8,924	11,435	15,660	18,930	23,392	24,725	33,252	33,182	31,277	32,418
Non-Interest Income as % of Total Income (%) ²	18.25	21.70	23.38	27.38	31.15	32.66	31.89	29.58	34.61	36.01	33.41	31.84
Annual Growth of Non-Interest Income (%)		21.76	9.70	21.96	26.98	17.27	19.07	5.39	25.64	-0.21	-6.09	3.52
Total Expenses (MYR millions)	28,618	32,420	31,840	32,599	38,435	43,871	61,389	68,387	78,133	77,063	71,778	77,551
Annual Growth of Total Expenses (%)		11.73	-1.82	2.33	15.18	12.39	28.54	10.23	12.47	-1.39	-7.36	7.44
of which Interest Expenses (MYR millions)	16,310	16,756	16,833	17,076	19,521	20,713	27,416	33,163	34,499	29,890	29,301	36,191
Interest Expenses as % of Total Expenses (%) ³	56.99	51.68	52.87	52.38	50.79	47.21	44.66	48.49	44.15	38.79	40.82	46.67
Annual Growth of Interest Expenses		2.66	0.46	1.42	12.53	5.76	24.45	17.33	3.87	-15.42	-2.01	19.04
of which Non-Interest Expenses (MYR millions)	12,308	15,664	15,007	15,523	18,914	23,157	33,973	35,223	43,634	47,173	42,477	41,360
Non-Interest Expenses as % of Total Expenses (%) ⁴	43.01	48.32	47.13	47.62	49.21	52.79	55.34	51.51	55.85	61.21	59.18	53.33
Annual Growth of Non-Interest Expenses		21.43	-4.38	3.33	17.93	18.32	31.84	3.55	19.28	7.50	-11.06	-2.70
Profit Before Taxation (MYR millions)	5,927	4,709	6,324	9,168	11,843	14,082	11,968	15,194	17,945	15,084	21,840	24,260
Annual Growth of Profit before Taxation		-25.87	25.54	31.02	22.59	15.90	-17.67	21.24	15.33	-18.97	30.93	9.98
Interest Margin	11,931	12,315	12,407	13,256	15,097	18,310	22,549	25,693	28,326	29,075	33,039	33,202
% of Interest Margin (%) ⁵	42.25	42.36	42.43	43.70	43.61	46.92	45.13	43.65	45.09	49.31	53.00	47.85

Source: Confidential dataset from a Malaysian financial organisation.

Note: As of December 2011, USD 1.0000 is equal to MYR 3.1265.

- 1 Interest income includes interest/profits received from loans/financing and amounts due from other banks.
- 2 Non-interest income includes fee-based income, commissions and income from investments.
- 3 Interest expenses include interest/profits sharing paid on deposits from customers and amounts due to other banks.
- 4 Non-interest expenses include staff expenses, maintenance costs, administration costs and depreciation.
- 5 Percentage of interest margin equals interest income minus interest expenses, divided by interest income.

4.5 Financial Ratio Analysis of Malaysian Banks

Financial ratio analysis uses accounting information (rather than production theory using an input and output mix) and is employed by various parties (e.g. regulators, analysts, investors and bank managers) to measure the performance of the banks. Thus, in this section, two different perspectives of financial ratios are analysed, namely profit and cost performance analysis.¹¹⁷

First, Figure 4.20 displays the return of assets (ROA) and return of equity (ROE) of Malaysian banks between 2000 and 2011.¹¹⁸ The ROE declined in 2001 because of the consolidation initiatives and rationalisation in domestic banks. This was influenced by the higher level of equity as well as the lower income generated by the Malaysian banks, where the ROE declined to 10.2% in 2001 from 15.9% in 2000. The greater level of equity in Malaysian banks can be explained by the greater rise in the equity of domestic banks following the completion of the consolidation exercise in the banking sector. At the same time, the consolidation exercise also forced banks to adapt new operations and business strategies, where domestic banks were affected by higher operational expenditure such as the closing of duplicate branches, the preparation of new branding images and a voluntary separation scheme for personnel (redundancy exercise) (Bank Negara Malaysia, 2001a).

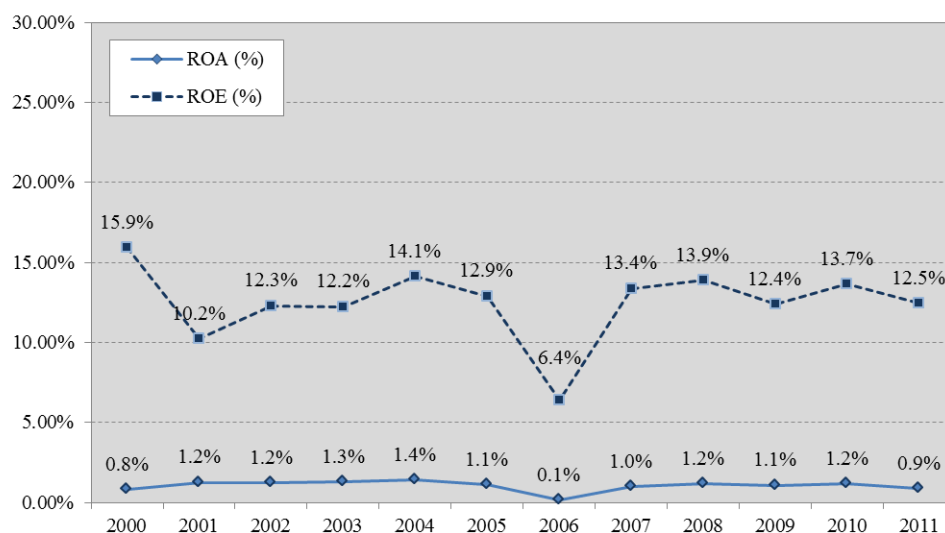
Thereafter, the ROE for 2002 and 2011 remained positive, hovering between 12.2% and 14.1% except in 2006, when ROE was at its lowest point at 6.4%. The lower rate of ROE in 2006 was influenced by substantial losses through write-offs of large loans (financing) experienced by a domestic Islamic bank and the establishment of new Islamic foreign banks. Additionally, the new foreign players exhibited negative ROE's due to losses

¹¹⁷ The financial ratios are 'built-in' ratios that are commonly used in the financial statement analysis. In this section, the ratios that measure the ability of banks to generate income are return on assets (ROA) and return on equity (ROE). On the other hand the ratios that represent the capability of banks in managing costs are total cost–total revenues, cost–income ratio, staff cost–total assets ratio and Non-performing loans (NPL) ratios. However, these financial ratios may not implicate the effect of cost- and profit-efficiency based on the frontier analysis.

¹¹⁸ First, ROA is defined as or calculated using PBT over total assets, and measures the bank profits relative to total assets, implying the ability of available assets in generating profits (Suffian, 2011). Second, ROE is measured by dividing PBT with the equity, implying the rate of return of shareholders from the investment.

incurred in the first year of operations. Hence, the ROA for 2006 also declined to 0.1%, which is lower than other years that maintained ROAs between 0.8% and 1.4%.

Figure 4.20 Trends of Profit Performance Indicators in Malaysian Banks, 2000–2011



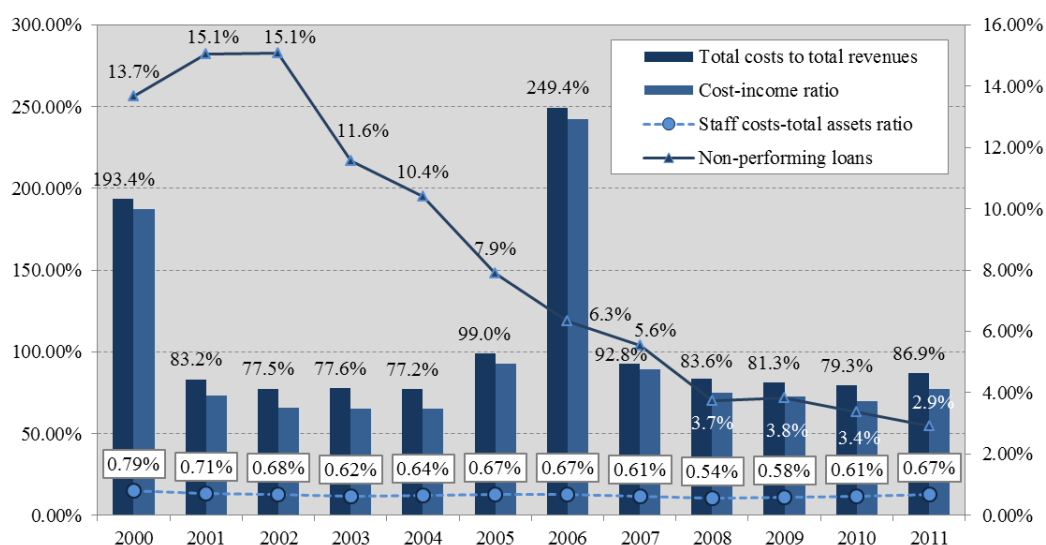
Source: Confidential dataset from a Malaysian financial organisation

Second, Figure 4.21 exhibits the trends of the cost–income ratio, total cost–total income ratio, NPL ratio, and staff cost ratio for the period 2000–2011.¹¹⁹ From Figure 4.21, the trend of the cost–income and total cost–total income ratios is almost similar. These two ratios increased significantly in the years 2000 and 2006. In 2000, the domestic banks were in the process of consolidation following the 1997/98 financial crisis. Income for the Malaysian banks declined in the post-crisis period and at the same time, incurred a substantial amount of expenses relating to the offering voluntary separation scheme (VSS) to reduce the number of staff, the new entity branding, and the repositioning of branches and

¹¹⁹ There are several ratios used to measure the cost performance measures, namely the cost–income ratio, total cost–total income, staff cost to total assets and NPLs over loans. Firstly, the cost–income ratio is measured and defined by the US Federal Financial Institutions Examination Council where it can be composed from non-interest expenses/(non-interest income plus net interest income). It measures the operating expenses used from the non-interest income and net interest income received from the business operations. Secondly, the total cost–total income ratio informs the user how much income is used to cover the total cost of the banks. Thirdly, the NPL–NPF ratio indicates the materialised and potential losses of the banks from loans and financing. It also implies that additional expenses may be incurred by banks, particularly in administering and monitoring loan and financing losses within their credit portfolio. Finally, the ratio of staff cost over total assets is used to assess the manpower expenses incurred by the banks. For all the ratios mentioned above, smaller ratios suggest greater cost savings for the banks.

personnel. As mentioned earlier, in 2006, the cost–income ratio was significantly greater due to significant losses from write-offs of a couple large financing (loans) by a full-fledged domestic Islamic bank. The losses materialised in 2006 was due to a weak financing structure coupled with the ‘evergreen’¹²⁰ effects from vulnerable restructuring and rescheduling arrangements during the last Asian Financial crisis in 1997–1998.

Figure 4.21 Trends of Cost Performance Indicators in Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

From Figure 4.21, the NPL ratio has been on a declining trend since the period of Asian financial crisis, indicating a greater cost saving for Malaysian banks. The declining trend in the first and second phases of the FSMP was due to the higher reclassification of NPLs to performing loans status and complete write-offs (Bank Negara Malaysia, 2005). During this period, the Malaysian banks’ proportion of loans and financing was shifted to the household where credit risks are more widespread and diversified compared to the pre-1997–1998 condition, in which risks were concentrated on large businesses in vulnerable sectors. Furthermore, the introduction of the credit bureau (i.e. CCRIS) enhanced the data collection on the financial position of households and businesses at a more granular level, which contributed to the declining trend in NPLs (Bank Negara Malaysia, 2007b). Concurrently,

¹²⁰ For example, the ‘evergreen’ effect can be seen in loan facilities such as ‘revolving loan’ or ‘credit card’, where the principal amount is required to be paid off within a specified period of time. The loans are routinely renewed after maturity, leaving the principal to be remained and repaid over in a long-term period.

with the establishment of advisory and debt management services (AKPK), the financial literacy of banking customers might have improved, particularly in their management of debt involving banks' credit facilities. In addition, the staff cost ratio remained consistent during the period under study. The ratio was slightly higher in 2000 due to the consolidation exercise where higher expenses were incurred on staff redundancy.

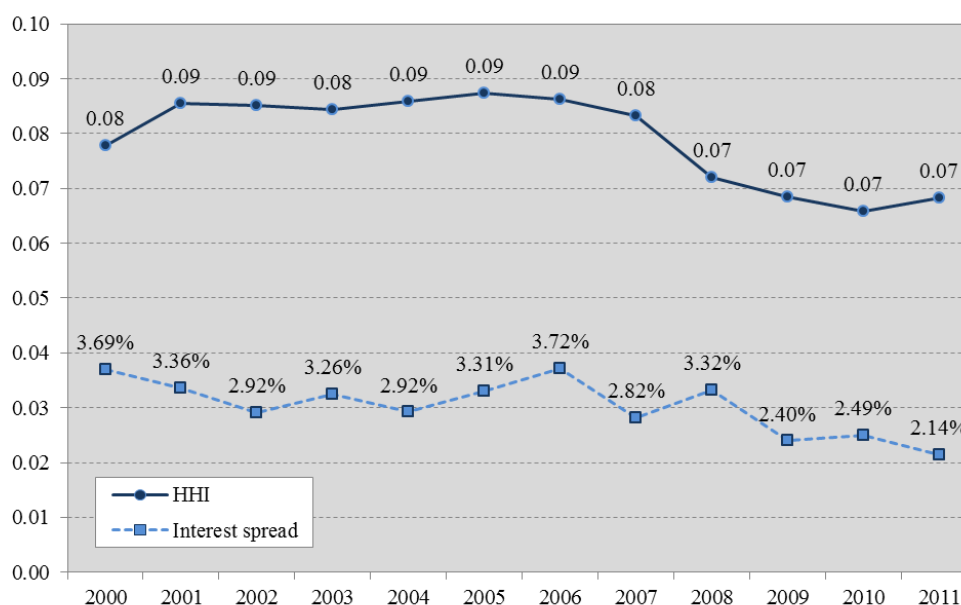
4.6 Competition Indicators of the Malaysian Banking Sector

One of the main objectives of the liberalisation efforts in the banking industry is to promote competition. As mentioned earlier, this can be done through various initiatives such as allowing foreign banks to compete in the domestic market, allowing foreign investors to hold equity in domestic banks and providing autonomy for banks to determine their own interest rate (i.e. BLR) on products and services offered. To infer the level of competition of the banking industry, the common indices used are Herfindahl–Hirschman Index (HHI) and the spread of interest rates between loan interest and deposit interest (Berger and Humphrey, 1997). HHI can be computed using the sum of the squared market share of all the banks in the market and it can be written as:

$$HI = \sum_{i=1}^n (x_i/x)^2 \quad (4.1)$$

where HHI is the Herfindahl–Hirschman Index (HHI), which is used to infer the level of market concentration, x_i indicates the deposits of the i th bank and x represents the total deposits for the banking industry as a whole. HHI can take the value between 0 and 1. The market is highly concentrated and has a low level of competition when HHI is approaching 1 and on the other hand if HHI is approaching 0, it indicates that the market is more competitive. As asserted by Molyneux et al. (1996), the number and dispersion of firms in the market could influence HHI and thus, it can be determined as a better measure of concentration within the market.

Figure 4.22 Trends of Concentration and Competitive Indicators in Malaysian Banks, 2000–2011



Source: Confidential dataset from a Malaysian financial organisation

Based on Figure 4.22, the HHI trend is quite consistent throughout the period under study, but slightly higher during the first and second phases of the FSMP (i.e. 2000–2003 and 2004–2007). The competition level was lower during this period due to restrictions put on foreign banks and the gradual liberalisation initiatives undertaken by BNM. In these phases, domestic banks were strengthened and given a head start to face greater competition at a later stage. Subsequently, it is evidenced that the competitive environment improved in the third phase of FSMP (i.e. 2008–2011) after the creation of three new Islamic banks, the establishment of Islamic banks' subsidiaries and an increase in foreign banks' branches. These liberalisation measures introduced by BNM have generally improved the degree of competition among Malaysian banks, where the number of entities competing in the banking sector has increased (Bank Negara Malaysia, 2007b).

Another way to gauge the competition level of the banking industry is by using the spread of the interest rate of banks, which can be computed based on the difference between the interest rates on loans and the interest rate on deposits. An inverse relationship between interest spreads and the competition level is commonly cited where an increasing

competitive condition exhibits a lower level of interest spreads and conversely, a higher interest spread may indicate a less competitive environment within the banking sector, with banks exercising their market power to impose a higher interest rate (Molyneux et al., 1996). From Figure 4.22, the lower spread shown in 2004 is a result of the introduction of a new interest rate framework, where banking institutions were allowed to price their own products based on their own cost structures and market interest rates (e.g. OPR). OPR was increased by BNM in late 2005 and mid-2006 in order to contain inflation expectations. The increment of OPR influenced banking institutions to adjust their BLR accordingly, which resulted in a lower interest spread.¹²¹ As the subprime crisis surfaced in 2007, there was a large influx of foreign capital inflows into Malaysia by overseas investors searching for higher yield income. Consequently, banks were competing for these cheaper funds, resulting in lower spreads for Malaysian banks. As the global economic crisis worsened and the Malaysian economy contracted in 2009, BNM adjusted the OPR by reducing it from 3.25% to 2.00% between late 2008 and mid-2009. With this implementation, the interest spread of Malaysian banks was also reduced to the floor-rate interest imposed by BNM on fixed deposits and lower interest rates on loans granted to customers. The declining interest spreads during the third phase of FSMP may also indicate that the Malaysian banks were facing a more competitive market structure (Bank Negara Malaysia, 2007b, 2009b).

4.7 The Development of Research Hypotheses

From the earlier discussion, this present research aims to examine the efficiency of Malaysian banks based on the financial liberalisation (via FSMP) that was being implemented in Malaysia in the years 2000 to 2011. One of the objectives of financial liberalisation (i.e. FSMP) is to increase market competitiveness in the banking sector. In Chapter 2, the discussion of the neoclassical market structures, such as perfect competition, monopoly, monopolistic competition and oligopoly, explained that the market structure can affect the level of efficiency of the firm. With insufficient competitive pressure in the banking industry (market imperfections), the senior managers of the banks may pursue goals

¹²¹ The Malaysian banks were speedier in adjusting the interest rate from the asset side (e.g. loans/financing and advances) compared to adjustments made on the liability side (i.e. deposits and capital), which resulted in a lower interest spread (Bank Negara Malaysia, 2006).

other than maximising profits and they often enjoy a ‘quiet life’ (Berger and Hannan, 1998). Thus, a more competitive market structure could provide the required pressure on bank managers to improve the efficiency of the banks and continue to survive, or face the risk of being forced out of the market (Machlup, 1967).

In Chapter 3, it was noted that banks’ operations may be highly influenced by regulations imposed by the policy makers (McKinnon, 1973; Shaw, 1973). Kumbhakar and Sarkar (2003) state that heavy governmental intervention in banking activities is seen as one of the factors that influence lower efficiency and poor performance in banks (particularly for banks in the developing/transition countries). Thus, an increasingly competitive market structure resulting from financial liberalisation may result in the higher efficiency of banks, particularly banks in developing countries. Financial liberalisation in the banking industry could force the management to adopt strategic management choices in banks to reduce operating expenses, improve banking technologies, seek new ways to increase revenues, diversify businesses, and reduce inherent risks (Gart, 1994). With the increasingly competitive market structure from the implementation of financial liberalisation measures, banks may need to optimise the production of output through banking services and products based on the optimal use of resources (Kumbhakar and Sarkar, 2003). Despite having various positive outcomes in implementing financial liberalisation on banks, there are several drawbacks. For instance, with more freedom from regulators, banks have a tendency to invest in higher-risk products (e.g. asset back securities, derivatives products and high-yield notes) (Kraft and Tirtiroglu, 1998). Moreover, they may aspire to increase profitability by competing for cheaper funds. This can be done by increasing the interest rate for deposits, which in turn can narrow the profit margin of the banks (Gant, 1994). Such behaviour by banks could threaten the safety and soundness of the financial industry.

This chapter (Chapter 4) discussed the financial liberalisation (via FSMP) implemented by BNM, which had the objective of improving the competitive structure of the Malaysian banking industry. Various deregulation initiatives were introduced in order to attain a more competitive market structure. Thus, in this chapter, several measurements such as HHI and interest spread trends were used to measure the competitive conditions of the banking industry in Malaysia. Based on the competition indicators’ trends in Figure 4.22, the

financial liberalisation measures initiated in 2001 have shown improvements during the third phase of the FSMP (i.e. 2008-2011). Greater market competition is evidenced from declining market concentration (as shown in Figure 4.22; from 0.077 (2000) to 0.068 (2011)) in the third phase of the FSMP. Moreover, the interest spread was also trending downward (from 3.69% (2000) to 2.14% (2011)). This may indicate an increase in the level of competition in the banking market based on simple ratio indicators. Nonetheless, a superior measurement such as frontier measurement is needed to examine the level of competitiveness and the efficiency performance of the Malaysian banks (see Section 2.3). Therefore, a hypothesis on the impact of financial liberalisation and market competitiveness on Malaysian banking efficiency is developed based on the discussions made on (1) the market and firm theories, and the measurement of production efficiency (Chapter 2), (2) the empirical results and findings of past literature on banking efficiency (Chapter 3), and (3) banking industry conditions and financial liberalisation initiatives to improve the Malaysian banking market structure (Chapter 4).

In deriving the hypothesis on the relationship between financial liberalisation and bank efficiency, past literature suggests that most banking efficiency studies on developing countries exhibited a positive impact from liberalisation measures introduced by the policy makers (see Chapter 3) (e.g. Gilbert and Wilson, 1998; Attaullah et al., 2004; Sensarma, 2006). Nevertheless, there are also banking efficiency studies in the developing countries that conclude that financial liberalisation does not have a significant impact on the efficiency of banks (e.g. Cook et al., 2001; Drake et al., 2006; Kraft et al., 2006). Based on past literature, despite mixed findings on the impact of financial liberalisation on the efficiency of the banks in developing countries (see Chapter 3), Berger and Humphrey (1997) state that the result of financial liberalisation in the banking industry can be influenced by the regulatory conditions of the industry and the type of liberalisation initiatives being introduced. In other words, factors regarding the institutional structure of the banking system could lead to successful financial liberalisation. If the institutional structure of banking industry is weak (e.g. it has weak legal structure and enforcement, ineffective bureaucracies, high free-riding activities and high corruption), greater instability and greater inefficiency could occur immediately after the implementation of financial liberalisation (Demirguc-Kunt and Detragiache, 1998). Therefore, as discussed earlier in

this chapter, deregulation measures introduced by BNM in the Malaysian banking industry and its impact on the efficiency of the banks can be examined from this hypothesis:

Hypothesis 1 Banking efficiency has improved since the implementation of the Financial Sector Master Plan (strengthening and liberalisation measures) in the Malaysian banking system between 2000 and 2011.

This particular hypothesis is examined by looking at the trend of the efficiency scores of Malaysian banks from 2000 to 2011. By comparing the efficiency scores of Malaysian banks over this period, it is expected that the trends will provide an indication of the relationship (either positive or negative) between financial liberalisation and the efficiency of Malaysian banks. To further support this relationship, dummy variables are assigned in the specification of the stochastic frontier analysis (SFA) model to account for different years in the three different phases (i.e. 2000-2003, 2004-2007 and 2008-2011) of financial liberalisation that are initiated through the FSMP by BNM, in which the coefficient and significance testing (e.g. the *t*-test) from the model specification could show the relationship and significance of these two factors, particularly in the three different phases of the FSMP.

Additionally, there are several other hypotheses that can be derived regarding the endogenous factors (internal bank-specific factors, such as ownership structure, specialisation and size) and exogenous factors (external factors, such as market concentration structure), and their impact on the efficiency of Malaysian banks. As outlined by Berger and Humphrey (1997), efficiency measurement could be employed in various contexts. Hence, this present research also examines several contextual areas regarding bank-specific factors and the market structure. For example, in Chapter 2, discussions on the theory of the firm and the production function demonstrate that a firm faces a series of processes of transforming inputs (resources) into maximum outputs (goods and services), subject to technological conditions. Moreover, to transform these inputs into outputs efficiently, the firm requires dedicated managerial ability (as a consequence of the separation between owner and manager). Modern theories such as managerial theories and the behavioural theory of the firm show that managers may have their own utility to satisfy, which may result in lower efficiency of the firms (banks). In addition, the effect of endogenous and exogenous factors may also influence managers' decision-making process

in transforming inputs into outputs efficiently, based on the technological conditions. In view of the discussions in Chapters 2, 3 and 4, this research will also examine the following hypotheses:

- Hypothesis 2* *Foreign banks are more cost (profit) efficient than domestic banks.*
- Hypothesis 3* *Conventional banks are more cost (profit) efficient than Islamic banks.*
- Hypothesis 4* *The impact of the global economic slowdown led to lower cost (profit) efficiency of Malaysian banks.*
- Hypothesis 5* *Large banks are more cost (profit) efficient than small banks.*
- Hypothesis 6a* *A high concentration in the banking sector exhibits lower cost (profit) efficiency of Malaysian banks.*
- Hypothesis 6b* *A high market share exhibits higher cost (profit) efficiency of Malaysian banks.*

For the above, the variables and rationale that are used to consider the differences among Malaysian banks in hypotheses 2 to 6 are discussed in detail in Chapter 5.

4.8 Conclusion

In this chapter, the main attributes of Malaysian banking system between 2000 and 2011 are discussed. In essence, the Malaysian banking system consists of commercial banks (both conventional and Islamic banks) and investment banks.¹²² Commercial banks dominate the banking system in Malaysia, and can be owned by either foreign or domestic parties. The banking and financial system accounted for 11.5% of the Malaysian real GDP, which has had an average annual growth of 7.1% since 2000. This indicates that the financial role is imperative in providing economic growth and development in Malaysia.

Gradual liberalisation measures were introduced by the Malaysian government and BNM in 2001 via its 10-year Financial Sector Master Plan (FSMP). This master plan was initiated after the Asian financial crisis in 1997–1998; the crisis reveals the vulnerability of the Malaysian banking and financial system. The FSMP consists of three major phases. The

¹²² However, as mentioned earlier and for the context of this current research, investment banks are excluded from the analysis due to the differences in operations and objectives in comparison to conventional and Islamic banks.

first phase of FSMP focused on enhancing the capacity of the Malaysian domestic banks and strengthening the financial infrastructure. The second phase of FSMP concentrated on intensifying competitive pressures on the domestic financial sector. The third phase of FSMP introduced new foreign competition and assimilation into the global arena. Within these phases of FSMP, various liberalisation initiatives were implemented such as removing controls on interest rate, staff remuneration and employment of expatriates, issuance of new foreign banking licences (for both conventional and Islamic banks) and establishment of a deposit insurance. Additionally, despite the various liberalisation measures introduced, Malaysian banks also improved their regulatory and supervision structure by introducing a new department within BNM to strengthen financial surveillance at the macro level; this department works closely with the supervision departments. Furthermore, in promoting financial stability and soundness, BNM adopted Basel II to enhance the risk management culture for Malaysian banks, particularly regarding capital management and governance.

The structure of the Malaysian banking scenario changed after the implementation of the FSMP. The implementation of the various liberalisation measures and the macroeconomic environment has affected the financial position of the Malaysian banks. The banks were stronger in terms of capital and customers outreach due to the consolidation, which prepared them to compete with their foreign counterparts. Additionally, with the introduction of New Interest Rate Framework (NIRF), the banks have had to compete and determine pricing based on their cost structure, risks and demands. Hence, the income sources of banks also diversified into fee-based income, rather than concentrating on interest income (composition of non-interest income in has increased from 19.3% in 2000 to 31.8% in 2011). The asset structure of banks also changed; their loans portfolio is more concentrated on the retail market, which reduces the concentration risks of the banks. The banks focus on investment rather than loans when it comes to corporate exposure, particularly in bonds from the debt market. Liberalisation measures such as the introduction of new foreign banks in 2006 and 2010/11 provided competitive pressures on domestic banks; the concentration ratio, measured by HHI, was on a decreasing trend from 2007 onwards (HHI in 2006 and 2010 were at 0.09 and 0.07, respectively). At the same time, the interest rate spreads of Malaysian banks were also on a downward trend (interest spreads declined from 3.7% in 2000 to 2.1%

in 2011), which implies that competitive forces existed within the Malaysian banking sector, despite facing economic uncertainties from the global credit crisis.

The remaining chapters will empirically examine the efficiency of the Malaysian banks using the parametric model and nonparametric approaches as a comparison to measure how the impact of liberalisation measures (via FSMP) affected the performance and efficiency of the banks. Furthermore, conceptual issues relating to the structure of Malaysian banks (e.g. ownership, specialisation, inherent risks and market structure) are discussed in the next chapter to explore the differences that can occur among banks in terms of efficiency levels.

Chapter 5 Methodology and Data

5.1 Introduction

In Chapter 2, theoretical discussions were introduced on how market structure could affect a firm's decision-making and level of efficiency. The transformation of inputs into outputs (via the production function) and frontier measurements were also briefly discussed. The discussion continued in Chapter 3 regarding banks' role in the economic development of a country, and particularly as an intermediary institution that mobilises funds which can be allocated efficiently into areas that could generate growth (Levine, 1997). Realising the importance of banks to the economy and the general public, governments generally place strict regulation on this sector (Kumbhakar and Sarkar, 2003). This however, could result in banks' increased inefficiency, particularly when they face limited options to transform their liabilities (e.g. deposits and capital) into profitable assets (e.g. investment and loans) (Fry, 1995). Thus, liberalisation in the banking sector could help banks to innovate, increase competitive market pressure and improve their level of efficiency (Arestis, 2005). Chapter 4 focused on such financial liberalisation measures namely, the Financial Sector Master Plan (FSMP) which was introduced between 2000 and 2011.¹²³ Subsequently, parametric stochastic frontier analysis (SFA) and nonparametric data envelopment analysis (DEA) models are employed to measure and capture the effect of financial liberalisation mentioned above.¹²⁴ This is in line with the objective of this research, which is to examine the impact of financial liberalisation initiatives introduced by Bank Negara Malaysia (BNM) on the cost- and profit-efficiency of Malaysian banks.

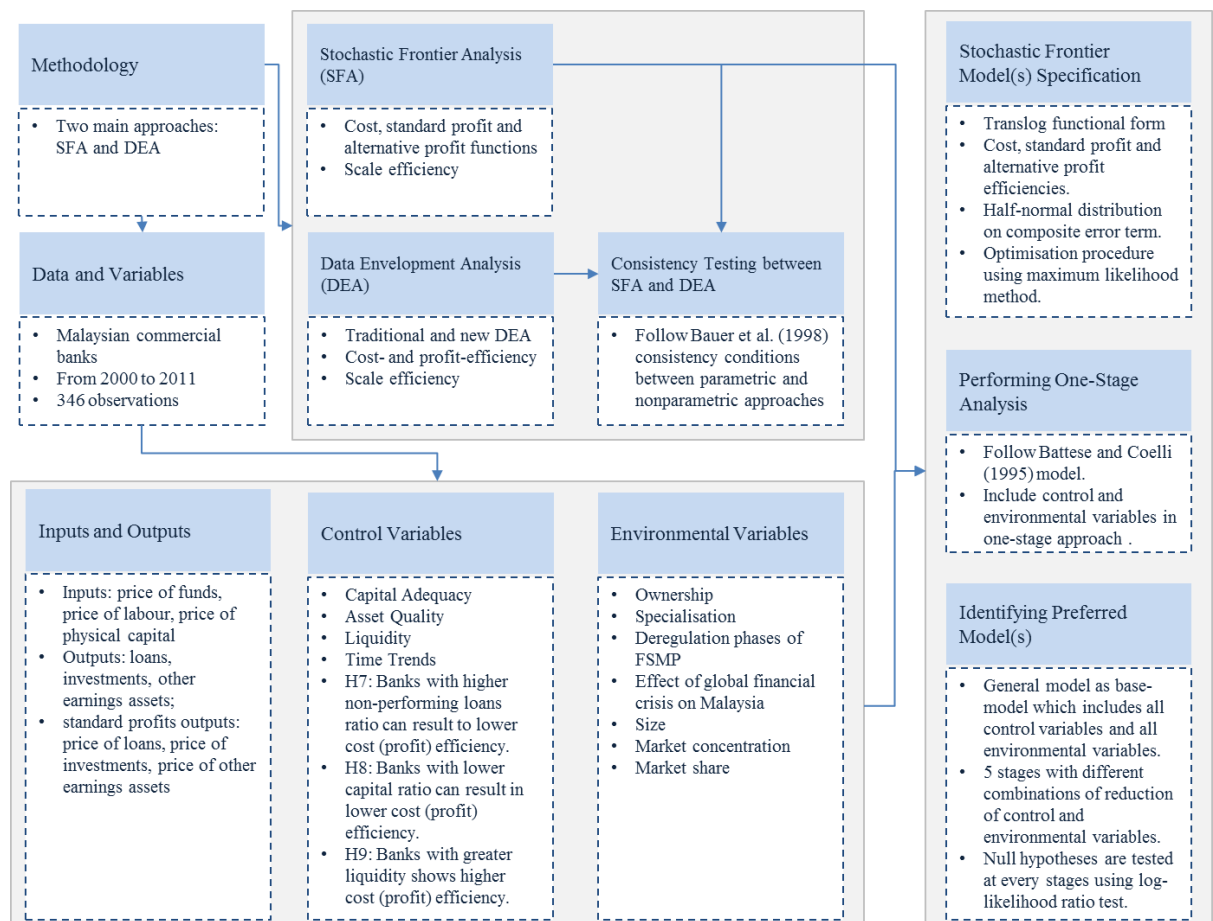
Section 5.2 outlines the SFA and DEA methodologies employed in this study to estimate the cost- and profit-efficiency of Malaysian banks. Additionally, this section discusses a

¹²³ Bank Negara Malaysia (BNM) introduced the Financial Sector Master Plan (FSMP), which sets the medium- and longer-term agenda of building a financial sector that is resilient, efficient and competitive, and responsive to changing requirements (Bank Negara Malaysia, 2001a, 2001b).

¹²⁴ As mentioned earlier in Chapter 2, following Resti (1997), the DEA models are used by way of comparison and to test the consistency of the SFA model (Bauer et al., 1998).

test of the conditions required to measure the consistency between parametric and nonparametric methods. Subsequently, Section 5.3 focuses on the data used in this research, which comprises all banks in Malaysia between 2000 and 2011. The description and rationale of different variables used in this research are also discussed. Section 5.4 outlines the stochastic frontier analysis model's specification, which includes the functional form that is used in the SFA model to measure the cost and profit functions based on the variables employed. This section also compares the different models and structural tests used in finding the preferred model. Once the preferred model has been identified, it is then used to measure the scores of cost- and profit-efficiencies (i.e. standard profit- and alternative profit-efficiency).

Figure 5.1 Information Flow for Chapter 5



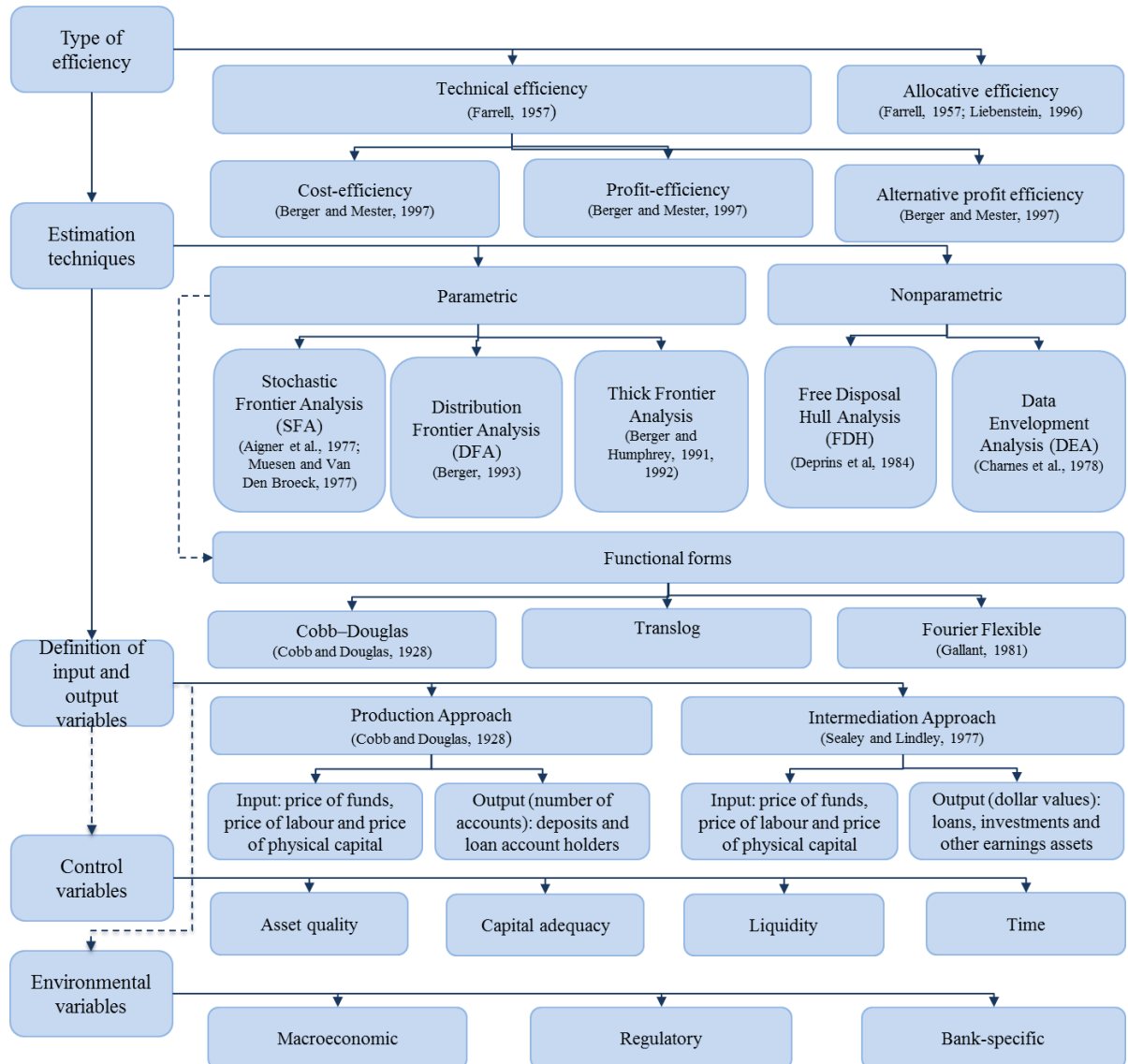
5.2 Methodology

As discussed in Chapters 2 and 3, there are two main approaches to measuring the efficiency of banks based on frontier estimations: the parametric and the nonparametric approaches. Both approaches display similar objectives using relative benchmarks to measure the efficiency of firms but the estimation methods and assumptions are significantly different. Following discussion of empirical studies of Malaysian banking (see Chapter 3), to the author's knowledge, there are no studies of Malaysian banking that perform comprehensive consistency testing (except for rank-order correlations) between parametric and nonparametric approaches (see Bauer et al. (1998)). The results of such consistency comparisons can help policy makers to be more confident in making policy decisions. Following Resti (1997), this study employs a parametric approach (i.e. SFA) to estimate the cost- and profit-efficiency levels of Malaysian banks. At the same time, a nonparametric approach (i.e. DEA) will be used to test the consistency of the SFA estimations.

Figure 5.2 gives an overview of the common approaches of frontier measurement. In this study, for the SFA model, the frontier estimation has focused on cost- and profit-efficiency (standard- and alternative profit-efficiency) using the translog functional form (model specification) and half-normal assumption (random error distribution). In terms of variables, the SFA model will utilise inputs and outputs based on the intermediation approach (see Sealey and Lindley, 1977). Concurrently, this research also considers the effect of heterogeneity on the efficiency of Malaysian banks. Thus, control and environmental variables are employed using the 1-stage approach of the Battese and Coelli (1995) model. Additionally, the DEA model is used in this study to examine the consistency between two different approaches (i.e. parametric versus nonparametric). In view of this, the DEA model employs the same dataset and intermediation approach

variables as the SFA model for both cost- and profit-efficiency estimations.¹²⁵ The methods employed in this study are discussed further in subsequent sections.

Figure 5.2 Overview of Frontier Efficiency Estimations



Adapted from Mokhtar et al. (2006, p. 5)

¹²⁵ However, the effect of heterogeneity is not examined for the DEA model due to differences of estimations. For instance, in this study, SFA uses a 1-stage approach, but for DEA to measure the effect of heterogeneity it requires a 2-stage approach, which is therefore not directly comparable.

5.2.1 Stochastic Frontier Analysis (SFA)

SFA is one of the most widely used methods in banking efficiency studies (e.g. Berger, 1993; Cebenoyan et al., 1993; Allen and Rai, 1996; Berger and Mester, 1997; Bhattacharya et al., 1997; Chang et al., 1998; Bauer et al., 1998; Altunbas et al., 2001; Casu and Girardone, 2002; Bonin et al., 2005; Bos and Schmiedel, 2007). SFA was independently proposed in 1977 by Aigner et al. (1977) and Meeusen and Van den Broeck (1977). The SFA model assumes that in producing a certain level of output, firms face various technical inefficiencies and a given combination of input levels. Additionally, the firm's production is influenced by the sum of a parametric function of known inputs, with unknown parameters, and a random error (associated with the measurement error of the level of production and inefficiency). SFA requires a functional form, such as cost or profit, with a two-component error terms: random error and inefficiency. By way of illustration, the single-equation stochastic cost function model is shown below:

$$\ln Y_{it} = \beta_{x_{it}} \ln X_{it} + \ln \varepsilon_{it} \quad (5.1)$$

where $\ln Y_{it}$ is the natural logarithm of output for the i -th bank at time t , $\ln X_{it}$ is a vector of inputs of the i -th bank at time t , $\beta_{x_{it}}$ is a vector of unknown parameters to be estimated and $\ln \varepsilon_{it}$ is the error term. Following Aigner et al. (1997), the assumption of the composed error term is:

$$\varepsilon_{it} = V_{it} + U_{it} \quad (5.2)$$

where V_{it} and U_{it} are independently distributed; V_{it} represents a random uncontrollable error and is assumed to be normally distributed with zero mean and variance σ_v^2 is drawn from a one-sided distribution that is assumed to capture inefficiency. U_{it} is assumed to be drawn from a half-normal distribution with mean zero and variance σ_u^2 . U_{it} , can be estimated by using the conditional mean of the inefficiency term, given the composed error term, as proposed by Jondrow et al. (1982) and derive the log-likelihood for inefficiency, which is expressed in terms of the two variance parameters, $\sigma^2 = \sigma_v^2 + \sigma_u^2$ capturing the variance of the composed error term $\lambda = \sigma_u^2 / \sigma_v^2$, which measures the fraction of

inefficiency relative to statistical noise. Moreover, U_{it} can be used to measure the level of inefficiency of banks. For instance, if U_{it} is equal to 0, it indicates that there is no inefficiency based on the production function imposed. On the other hand, if U_{it} is more than 0, it indicates that inefficiency is present. In the past, most studies using SFA were directed towards inefficiency prediction and this inefficiency is commonly measured using technical efficiency (TE). Equation 5.3 exhibits the common output-orientated measure of TE using the ratio of observed output to corresponding frontier output, which can be written as (Coelli et al., 2005):

$$TE_{it} = \frac{y_{it}}{\exp(\beta_{x_{it}} + v_{it})} = \frac{\exp(\beta_{x_{it}} + v_{it} - u_{it})}{\exp(\beta_{x_{it}} + v_{it})} = \exp(-u_{it}) \quad (5.3)$$

where TE is the technical efficiency of i -th bank at time t , y_{it} is the observed output and $\exp(\beta_{x_i} + v_i)$ is the corresponding frontier output. As mentioned earlier, TE has a value between 0 and 1, in which TE derives from the output of i -th bank relative to a fully-efficient bank's output, located on the estimated frontier curve that utilises the same input vector (Coelli et al, 2005).

5.2.1.1 Cost- and Profit-Efficiency

Based on earlier discussions in Chapter 3, the bulk of empirical studies of Malaysian banks have focused mainly on cost efficiency, while a limited number looked into the efficiency of the profit function. There are even fewer studies that compare both cost- and profit-efficiency using the same Malaysia banks' data set. Thus, following Berger and Mester (1997), this research employs three main economic functional forms under the parametric approach, in which efficiency estimates can be derived from the: (1) cost function, (2) standard profit function; and (3) alternative profit function. This corresponds with the two paramount objectives: cost minimisation and profit maximisation.

First, cost-efficiency (CE) indicates how close a bank's cost is to that of a best-practice bank, which produces the same outputs using the same technology. The variable costs of the cost function rely on: the prices of variable inputs, the amount of variable outputs,

fixed netputs (or exogenous factors) (if any), random errors, and inefficiency (Berger and Mester 1997). The cost-efficiency of a bank is measured using the observed bank's total cost, relative to the total cost of a bank on the estimated frontier. Hence, the cost function is described as:

$$\ln TC_{it} = f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + \ln v_{it} + \ln u_{it} \quad (5.4)$$

where $\ln TC_{it}$ is the natural logarithm of the observed bank's total cost, $f()$ is the cost frontier's functional form, w_{it} is the vector of input prices, y_{it} is the vector of outputs, c_{it} denotes the vector of control variables (if any) and z_{it} represents the vector of environmental variables (if any). These control c_{it} and environmental z_{it} variables are included in the cost function to capture the heterogeneity effects of cost-efficiency. β is the parameters to be estimated. The term $\ln v_{it} + \ln u_{it}$ is treated as a composite error term, where $\ln u_{it}$ is the inefficiency term: a non-negative and one-sided error component that follows an asymmetric half-normal distribution. $\ln v_{it}$ is the random error term that permits the random variation of the frontier across banks and captures the effects of measurement error, other statistical noise and random shocks outside the bank's control. This error term is assumed to consist of independently and identically distributed normal random variables, with zero mean and variance (Coelli et al., 2005).

The CE of the i -th bank is the estimated cost needed to produce bank i 's output vector if the bank were as efficient as the best-practice bank (on the frontier curve) in the sample facing the same inputs and outputs, control, and environmental variables (w, y, c, z), divided by the actual cost of i -th bank, and adjusted for random error. It can be written as:

$$CE_{it} = \frac{\exp(f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + v_{it})}{\exp(f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + v_{it} + u_{it})} = \frac{1}{\exp(u_{it})} \quad (5.5)$$

where CE_{it} is the cost-efficiency of the i -th bank. The numerator in equation 5.5 indicates the minimum cost that can be incurred by the best practice banks and the denominator in equation 5.5 denotes the actual cost incurred by the i -th bank at time t . Hence, cost-efficiency CE_{it} is measured against the ratio of minimum cost banks (best-practice banks on the frontier) and the actual cost of i -th bank. Cost-efficiency CE_{it} could also be seen as

a proportion of cost that is either being used efficiently or being wasted. For example, if CE_{it} of i -th bank is 0.60, it indicates that i -th bank is 60.0% efficient and 40.0% of its cost is being wasted when compared to the best-practice bank. Cost-efficiency ranges between 0 and 1. Banks with a cost-efficiency of 1 are considered to be best-practice banks within the observed data.

Second, the standard profit-efficiency (SPE) tests measure how close a bank is to achieving the maximum possible profit with given input and output prices (control and environmental variables, if any). The standard profit functions differ from cost functions by specifying the profit variable in place of the cost variable, and takes variable output prices as a given. On the other hand, the cost functions hold all output quantities to be statistically fixed at their inefficient levels. In standard profit-efficiency, the dependent variable (i.e. total profit) takes revenue into account by varying output and input prices. Output prices are employed to allow for inefficiencies in output choice when reacting to the prices of the profit function (Berger and Mester, 1997). In a log form, the stochastic standard profit function is:

$$\ln TP_{it} + \theta = f(w_{it}, p_{it}, c_{it}, z_{it}, \beta) + \ln v_{it} + \ln u_{it} \quad (5.6)$$

where TP_{it} is the profit before tax of i -th bank at time t , which includes all the interest and fee income earned from outputs minus total costs; θ is a constant added to every firm's profit so that the natural log is taken of a positive number; p_{it} is the vector of output prices; w_{it} is the vector of input prices, c_{it} is the vector of control variables (if any) and z_{it} is the vector of environmental variables (if any). $\ln u_{it}$ represents inefficiency that reduces profits and $\ln v_{it}$ is the random error.

SPE can be viewed as the ratio of the actual profits of i -th bank to the maximum profits that could be earned if the i -th bank was as efficient as the best bank in the sample, net of random error. The measurement for standard profit-efficiency of i -th bank, based on the ratio of actual profit to maximum profit, can be written as:

$$SPE_{it} = \frac{\exp(f(w_{it}, p_{it}, c_{it}, z_{it}, \beta) + v_{it} + u_{it})}{\exp(f(w_{it}, p_{it}, c_{it}, z_{it}, \beta) + v_{it})} = \exp(u_{it}) \quad (5.7)$$

where SPE_{it} is standard profit-efficiency, which exhibits the proportion of maximum profits that are earned compared to the best-practice banks. If i -th bank's standard profit-efficiency ratio is at 0.60, it indicates that the bank is losing 40% of the profits it could be earning due to excessive costs, deficient revenues or even both. Similar to cost-efficiency, for best-practice banks that maximise profits in given conditions within the observed data, the standard profit-efficiency would be equal to 1. However, by contrast, standard profit-efficiency can be negative if firms throw away more than 100% of their potential profits (Berger and Mester, 1997).

Third, alternative profit-efficiency (APE) employs the same dependent variables as the standard profit function and the same independent variables as the cost function. Berger and Mester (1997) further suggest that alternative profit-efficiency could provide useful information if one or more of the following conditions are met: (1) there are substantial unmeasured differences in the quality of banking services, (2) outputs are unable to achieve output scale and product mix, (3) output markets are not perfectly competitive, (4) output prices are not accurately measured. Alternative profit-efficiency tests measure how close a bank comes to earning maximum profit, given its output levels compared to its output prices (as in standard profit-efficiency). Hence, rather than compute deviations from optimal output as inefficiency, as in the standard profit function, the outputs are held constant as in the cost function, while allowing output prices to vary and affect profits. The log form of the alternative profit function can be written as:

$$\ln TP_{it} + \theta = f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + \ln v_{it} + \ln u_{it} \quad (5.8)$$

Equation 5.8 is identical to the standard profit function except that p_{it} is replaced by y_{it} (i.e. the vector of outputs) in the function, $f()$, yielding different values for $\ln u_{it}$ and $\ln v_{it}$, which represents inefficiency and random error respectively. Alternative profit-efficiency is a ratio of predicted actual profits to the predicted maximum profits for a best-practice bank and can be written as:

$$APE_{it} = \frac{\exp(f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + v_{it} + u_{it})}{\exp(f(w_{it}, y_{it}, c_{it}, z_{it}, \beta) + v_{it})} = \exp(u_{it}) \quad (5.9)$$

Here, the output prices are allowed to vary efficiency values in choosing output quantities and do not affect alternative profit-efficiency except through the point of evaluation $f(w_{it}, y_{it}, c_{it}, z_{it})$, to the extent that the best-practice bank is not operating at the same level $f(w_{it}, y_{it}, c_{it}, z_{it})$ as the i -th bank, owing to the effects mentioned earlier: market power and the quality of banking services, for example. Similarly to standard profit-efficiency, an alternative profit efficiency of 1, indicates best-practice banks within the sample are maximising profits (Berger and Mester, 1997).

There are several important steps involved in estimating the efficiency score of cost, standard profit- and alternative profit-efficiency when using the SFA model. The first step in computing the efficiency score is to estimate the parameters (β) within the stochastic equation. Since stochastic frontier models involve a composite error of $\varepsilon_i = v_i + u_i$ on the right side of the equation, it makes the estimation of parameters (β) complex. If there is no composite error term ($\varepsilon_i = v_i + u_i$) in the equation, such as $y = f(w, x, \beta)$, it will be relatively easy to estimate the parameters (β) by using the ordinary least squares (OLS) method. Even with the composite error term (ε) the estimation can still be made, as it is unbiased, consistent and efficient among linear estimators.

However, Greene (2010) found that the estimator of the intercept in the OLS method is commonly not consistent. To overcome this limitation, a common method used to estimate the parameters (β) in the stochastic frontier model is to use the maximum-likelihood (ML) function, particularly when dealing with non-linear equations. The concept of ML entails a particular sample of observations that is likely to be generated by the distribution of other samples. In other words, estimation via ML is derived by observing given data and provides a means of choosing an asymptotically efficient estimator for a set of parameters (Greene, 2002). ML estimators are preferred due to their large sample properties, which are consistent and asymptotically normally distributed (Coelli et al., 2005).

When using the ML to estimate parameters (β), a strong assumption on the composite error (ε_i) is needed. As mentioned above, the composite error (ε_i) consists of random error (v_i) and inefficiency (u_i), and it is assumed that v_i is independent of u_i , just as other input and output variables are independent. It is also assumed that v_i is normally distributed with zero mean and variance σ_v^2 ($v \sim iid N(0, \sigma_v^2)$) and u_i is non-negative half-normally distributed with zero mean and variance σ_u^2 ($u \sim iid N^+(0, \sigma_u^2)$). This assumption of a half-normal model on u_i is commonly used in stochastic frontier models. Other assumptions such as truncated normal, exponential and gamma models are also available to facilitate the ML process (Coelli et al., 2005, Kumbhakar and Lovell, 2000).

Apart from the assumptions on the distribution and independent density functions of v_i and u_i , the second part of the ML process involves a joint probability function (as a result of integrating individual density functions). The joint probability density function (pdf) is also known as the likelihood function and expresses the likelihood of observing a sample from the unknown parameters of (β) and σ^2 . The likelihood function is maximised via an iterative process to estimate the parameters (β) and is normally in logarithm form (for detailed discussion, see Coelli et al 2005, pg. 217, Kumbhakar and Lovell, 2000, pg.75).

Maximising the log-likelihood function is often involves taking the first derivatives of the unknown parameters and setting them to zero. Since the first derivative of this log-likelihood function is highly non-linear and normally the parameters cannot be solved, the log-likelihood function has to employ iterative optimisation processes, which begin with placing the starting values of unknown parameters and are updated iteratively until the log-likelihood function is converged at maximum value (see Coelli et al., 2005, p. 246 for details). In this procedure, the joint pdf of v_i and u_i is considered to be one, within the ML technique in order to estimate the parameters in the cost, standard and alternative profit functions. The log-likelihood function can be written:

$$\ln l(TC|\beta, \sigma, \lambda) = \frac{N}{2} \ln \frac{2}{\pi} - N \ln \sigma + \sum_{i=1}^N \ln \left[\Phi \left(\frac{\varepsilon_i \lambda}{\sigma} \right) \right] - \frac{1}{2\sigma^2} \sum_{i=1}^N \varepsilon_i^2 \quad (5.10)$$

where TC is the vector of log-total costs, β, σ, λ is a set of parameters to be estimated, N is the number of banks. $\sigma^2 = \sigma_v^2 + \sigma_u^2$, $\varepsilon_i = v_i + u_i = \ln TC_i - f(w_{it}, y_{it}, c_{it}, z_{it}, \beta)$ and $\Phi(x)$ are the standard normal cumulative density function (cdf) of the standard normal random variable evaluated at x (Allen and Rai, 1996; Coelli et al., 2005).

Once the parameters (β, σ, λ) have been estimated in equation (5.10), the next step is to compute the composite error term $\varepsilon_i = v_i + u_i$. The existence of inefficiency (u_i) within the composite error should be first determined before computation of inefficiency is carried out. In order to test for the presence of inefficiency, the skewness of the joint distribution of the composite error ($\varepsilon_i = v_i + u_i$) is observed. The domination random error (v_i) and inefficiency (u_i) depend on the skewness of distribution of the composite error (ε_i). For cost-efficiency frontier, if the distribution is positively skewed (or negatively skewed for the profit-efficiency frontier), it indicates that the component-inefficiency (u_i) dominates the random error (v_i), which sums up the composite random error (ε_i). Alternatively, Battese and Corra (1977) introduced parameter γ which measures the deviations of the composite random error (ε_i). The Parameter γ is measured as:

$$\gamma = \frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2} \quad (5.11)$$

where σ_u^2 is the variance of inefficiency and, σ_v^2 is the variance of random error. This γ -parameter lies between 0 and 1. The γ -parameter indicates the cause of deviations either by inefficiency u or random error v relative to ε . If γ is approaching 1, it indicates that the composite error (ε) is dominated by inefficiency u ; on the other hand, if γ is approaching 0, then it is noise or random error v dominating the composite error (ε).

The main objective in the next step is to estimate the mean inefficiency of all banks. The mean inefficiency will be used as a basis to compute individual banks' inefficiency at a

later stage. Based on Aigner, Lovell and Schmidt (1977), the mean inefficiency of a half-normal model can be estimated via the equation:

$$E(u) = \sigma_u \sqrt{\frac{2}{\pi}} \quad (5.12)$$

where $E(u)$ is the expected mean for inefficiency (u_i). To further calculate the average cost-efficiency of all banks within the sample, Aigner, Lovell and Schmidt (1977) suggest that $[1 - E(u)]$ should be used. Once an estimation of average inefficiency within the sample is obtained, the next step is to compute the individual banks' level of inefficiency. To do so, this study follows Jondrow, Lovell, Moterov and Schmidt's (1982) model, which estimates an individual bank's efficiency using a half-normal case. Their method can separate inefficiency u and random error v from the composite error ε for individual banks. This can be done from the distribution of ε , as it contains necessary information on inefficiency u . That is, individual banks' inefficiency can be obtained from the conditional distribution of u given ε . Based on the half-normal case,¹²⁶ Jondrow et al. (1982) proposed a point estimator for cost-inefficiency that involves the conditional expectation value of u from the composed error term ε_i , which can be written as:

$$E(u_i|\varepsilon_i) = \frac{\sigma_\lambda}{(1 + \lambda^2)} \left[\frac{\varphi(\varepsilon_i \lambda / \sigma)}{\Phi(\varepsilon_i \lambda / \sigma)} + \frac{\varepsilon_i \lambda}{\sigma} \right] \quad (5.13)$$

where $E[.]$ is the expectation operator for point estimator, $\varphi[.]$ is the joint pdf of standard normal distribution, $\Phi[.]$ is the cdf, $\sigma = \sqrt{\sigma_u^2 + \sigma_v^2}$ and $\lambda = \frac{\sigma_u}{\sigma_v + \sigma_u}$. The value of λ must lie between 0 and 1. A value closer to 1 indicates that deviation from the frontier is influenced more by cost inefficiency. On the other hand, a value closer to 0 denotes that the deviation can be explained by noise.¹²⁷ The estimates of efficiency are unbiased but inconsistent

¹²⁶ For this, the half normal distribution assumption is employed on the composite error term ($\varepsilon_i = v_i + u_i$). For other distribution assumptions (e.g. exponential, gamma and truncated distributions), the log-likelihood function differs based on its respective density functions.

¹²⁷ Kumbhakar and Lovell (2000) explain the effect of λ in more details. They mention that when λ approaches 0, it can be that either σ_v^2 approaches ∞ or σ_u^2 approaches 0 and result in a random error component that dominates the one-sided error component in determining ε_i . On the other hand, as λ

because the variations associated with the distribution of estimator $(u_i|\varepsilon_i)$, which is independent of i and remain non-zero (Greene, 2010).¹²⁸ Thus, the cost-efficiency measure from equation 5.13 takes values over the interval $(1,\infty)$ and a value equal to 1 means fully efficient. Given this, the cost-efficiency scores can be obtained from $1/E(u_i|\varepsilon_i)$ (Coelli et al., 2005).

The next step after estimating the inefficiency levels of individual banks is to test the hypotheses concerning parameters (β) . Commonly, if OLS regression is used to measure the coefficient, or a small sample is employed in the regression process, t - and F - tests are sufficient to test the hypothesis concerning the parameters (β) . However, as most SFA operates by using ML estimation, testing the hypothesis of the coefficients is commonly done with the z -test due to a normally asymptotic distribution of unconstrained ML estimators (Coelli et al., 2005). The t -test ratios can also be used to test the coefficient hypothesis by dividing the coefficients and the standard errors from the model generated (Qasrawi, 2009).

In addition, Coelli et al. (2005) mention that testing the absence of inefficiency effects is one of the things that stochastic frontier researchers are interested in. Thus, the null hypothesis is to reject if there are no inefficiency effects. Common methods used for this procedure are the likelihood ratio, the Wald test and the Lagrange multiplier tests (Greene, 2010). The most frequently employed statistical method to test a hypothesis using ML estimation is the likelihood ratio (LR). LR compares the closeness of the of unrestricted estimates to the log-likelihood function of restricted estimates (Coelli et al., 2005). That is, the test is used to evaluate a simple null hypothesis against a simple alternative hypothesis (Greene, 2002). This index is also particularly useful when researchers use several different models, with a number of restricted variables, to estimate firms' efficiency (where the LR index is often used to find the preferred model). The LR statistic function can be described as (Coelli et al., 2005):

approaches ∞ , it can be either that σ_u^2 approaches ∞ or that σ_v^2 approaches 0, and the one-sided error component dominates the random error in the determination of ε_i .

¹²⁸ Greene (2010) iterates that an estimator is consistent if its values approach the true parameter value and if its variances become smaller when the sample size increases indefinitely.

$$LR = -2[\ln L_r - \ln L_u] \quad (5.14)$$

where $\ln L_r$ is the value of maximum log-likelihood in the restricted model (OLS model) and $\ln L_u$ is the value of maximum log-likelihood in the unrestricted model (ML estimates model).

The LR statistics are asymptotically equivalent to chi-squared (χ^2) distributions (Coelli, 1995). Based on the null hypothesis of the restricted (i.e. OLS) model, it has a limiting chi-squared distribution with degrees of freedom equal to the number of restrictions being tested (Greene, 2010). The degrees of freedom of the chi-squared statistic are equal to the reduction in the number of dimensions in the parameter space due to the restrictions being imposed (Greene, 2010). For example, if the restrictions in the restricted model is 1, which is setting $\sigma_u^2 = 0$, the null hypothesis is often tested by observing the result of LR statistics. If the LR statistic exceeds the chosen (1%, 5% or 10%) critical value of chi-squared distributions, the null hypothesis is rejected, which implies the existence of inefficiency in the unrestricted model. However, if the null hypothesis is accepted, it indicates that the model does not generate inefficiency via its stochastic functions. This study employs a 5% critical value with 1 degree of freedom, which gives a value of 3.84.

5.2.1.2 Scale-Efficiency

As discussed in Chapter 2, scale economies can be described as a ratio of the proportionate increase in cost in relation to the proportionate increase in output. Theoretically, given the total cost function defined by $TC = f(Q)$ where Q is output, the average cost can be derived as $ATC = f(Q)/Q$ and marginal cost (MC) is $\partial TC/\partial Q$. The average cost will decline as long as marginal cost lies below average cost. Hence, the scale economies (SE) are given as ATC/MC , which is the elasticity of cost with respect to output (Altunbas et al., 1996). In this research, following Mester (1996) and Altunbas et al. (2001), the SE can be estimated as the mean output, input prices and other control variables for different size quartiles of all banks. Therefore the SE can be mathematically computed by differentiating the cost function in equation (5.4) with respect to output. This is given as:

$$SE = \sum_{i=1}^N \frac{\partial \ln TC}{\partial \ln y_i} \quad (5.15)$$

where $\partial \ln TC / \partial \ln y_i$ is the partial derivative of the logarithm of the cost function with respect to logarithm of output i . The sum of individual cost-elasticity measures the proportional change in TC due to the proportional change in output when all other factors are held constant. The change in output could change the scale of output but not the composition of output bundles. A bank is said to have economies of scale if $SE < 1$, where an equal proportionate rise in all outputs results in a less than proportionate rise in total costs. If $SE > 1$, the total costs increase is more than proportionate with the increase in outputs and the bank is said to face diseconomies of scale. On the other hand, if $SE = 1$, the bank is operating at an optimal level of production and demonstrating constant returns to scale.

5.2.2 Data Envelopment Analysis (DEA)

In order to provide more accurate and useful information for regulatory purposes, it is imperative to have reliable methods for measuring efficiency. For this purpose, DEA will be employed to cross-check the robustness of the SFA's results. As seen in the Chapter 2, SFA is a useful and powerful tool in determining the efficiency frontier as well as in estimating the efficiency levels of respective banks. However, there is a limitation on estimation specifications, which may influence the error term and may not be reflective of industry's conditions. Thus, DEA provides a more flexible approach in measuring efficiency, as it does not involve an explicit estimation of the cost/profit functions and it avoids misspecification risks of the error term associated with the estimation function. Nevertheless, DEA is a deterministic model and does not take into account random errors in the data. Therefore, although both methods are not fully comparable due to differences in estimation of the random error, DEA will still be employed in this study for the purpose of policy analysis and decision-making, given that different techniques or methods contain different information and may generate different efficiency scores (Eisenbeis et al., 1999; Bauer et al., 1998). DEA is a nonparametric approach that employs a linear programming method by constructing a piece-wise combination yielding a convex production possibility

set that envelope all decision-making units (DMUs) in the sample. Consequently, based on the frontier created by efficient DMUs (i.e. banks), the efficiency of other banks is calculated relative to the frontier (Coelli et al., 2005). For methodological purposes, the mathematical expression is explained in the next section.

Charnes, Cooper and Rhodes (1978) present a model with an input-orientation that is able to deal with many inputs and outputs and assumed constant returns to scale (CRS). The input-orientation DEA measures efficiency by maximising the ratio of all outputs to all inputs involved within the sample. DEA used observations of the sample to measure efficiency. Following Coelli et al. (2005), DEA can be explained by assuming the data of N inputs and M outputs for each of I firms. The column vectors x_i and y_i represent the input and output data for i -th firm where $(N \times I) = X$ and $(M \times I) = Y$ for both input and output matrices, respectively, which represent the data for all I firms.

Coelli et al. (2005) further iterate that DEA could be explained through ratio form where maximisation of outputs can be made via computation of the ratio of all outputs over all inputs (input-orientated). To illustrate this, the ratio can be shown as $(u'y_i/v'x_i)$, where u is an $(M \times I)$ vector of output weights, and v is a $(N \times I)$ vector of input weights. The optimal weights are attained by solving the following mathematical programming problem:

$$\begin{aligned} \text{Max}_{u,v} \quad & (u'y_i/v'x_i) & (5.16) \\ \text{subject to} \quad & u'y_j/v'x_j \leq 1, \quad j = 1, 2, \dots, I, \\ & u, v \geq 0. \end{aligned}$$

The above encompasses finding values for u and v that maximise the efficiency of i -th firm, subject to the constraint that all the efficiency measures should be less than or equal to one for respective firms. However, the fractional formulation has a problem with an infinite number of solutions which can be overcome by imposing the optimal constraint $v'x_j = 1$, and then written in the form of (Coelli et al., 2005):

$$\begin{aligned}
& \text{Max}_{u,v} && (u'y_i) && (5.17) \\
& \text{subject to} && v'x_i = 1 \\
& && \mu'y_j - v'x_j \leq 0, \quad j = 1, 2, \dots, I, \\
& && u, v \geq 0
\end{aligned}$$

where the notation of u and v is changed to u and v and is used to emphasise that this is a different linear programming problem. This linear programming is known as a multiplier form. Using the duality in linear programming,¹²⁹ the dual of a maximisation problem can derive an envelopment form and be written as (Coelli et al., 2005):

$$\begin{aligned}
& \text{Min}_{\theta,\lambda} && \theta && (5.18) \\
& \text{subject to} && -y_i + Y\lambda \geq 0, \\
& && \theta x_i - X\lambda \geq 0, \\
& && \lambda \geq 0
\end{aligned}$$

where θ , a scalar or efficiency score of the i -th firm, and λ is a $I \times 1$ vector of the constant, which denotes the weights employed to evaluate the performance of i -th firm relative to the most efficient firms that sit on the efficiency frontier. When a firm is inefficient, the weight (λ) implicates the percentage of input reference firms that are required to become efficient. According to Farrell (1957), a technically efficient firm would achieve a value of 1, which is located on the frontier and be satisfied from the value $\theta \leq 1$, which indicates the efficiency score of i -th firm. Concurrently, the linear programming should be solved I times, once for each firm within the sample for a value of θ to be obtained for each firm. The envelopment form shown above has fewer constraints than the multiplier form, which makes it preferable to solve the envelopment form (Coelli et al., 2005). The above linear programming has a fine intuitive interpretation where DEA seeks to minimise the efficiency score θ of the i -th firm and radially contract the input vector x_i , as much as

¹²⁹ The formulation of linear programming model can either be primal or dual. The former normally treats the rows and dual treats column as representing the model. Given this, duality describes the changing objective from maximising output/input ratio to minimising input usage. The dual for equation 5.17 is minimising $v'x_i = 1$, represented by θ , and dual for $\mu'y_j - v'x_j \leq 0$ is λ . Since the results from the two formulations (i.e. primal and dual) are equal, though expressed differently, the choice between them is based on computational efficiency, or perhaps ease of interpretation. The dual form is more efficient in computation if the number of DMUs is large compared to the number of input and output variables. Primal form entails constraints represented by the number of DMUs, while in dual form, these constraints are replaced by the number of input and output variables (Cooper et al., 2007).

possible, whilst remaining within the feasible input set. The observed data points represented by all DMUs will determine the piece-wise linear isoquant, which is the inner boundary of the data set. In this process, a projection point $(X\lambda, Y\lambda)$, which is a linear combination of observed data points (not an explicit functional form) located on the efficient frontier, is being produced from a radial contraction of input vector x_i that also limits the projection points belonging to a feasible production set (Coelli et al., 2005). Hence, inefficiency is a measure of the difference between vector x_i and projection point $(X\lambda, Y\lambda)$ on the efficient frontier.

As discussed earlier in Chapter 2, CRS assumes all firms are operating at an optimal scale where the CRS linear programming is imposed on all observations. However, Coelli et al. (2005) argue that imperfect competition, financial constraints and government regulation have an impact on firms, causing them not to operate at an optimal scale.¹³⁰ Hence, it is at times reasonable to adjust the CRS model to account for variable returns to scale (VRS). VRS allows the computation of technical efficiency without confounding the effects of scale-efficiencies. Banker, Charnes and Cooper (1984) who introduced their BCC model, state that VRS ensures that a similar size of firms is compared, devoid of the impact of scale-efficiencies. VRS adds the convexity constraint $\sum \lambda = 1$ from the CRS linear programming equation (5.18) shown above. The linear programming of VRS can be written as (Coelli et al., 2005):

$$\begin{aligned} \text{Min}_{\theta, \lambda} \quad & \theta & (5.19) \\ \text{subject to} \quad & -y_i + Y\lambda \geq 0, \\ & \theta x_i - X\lambda \geq 0, \\ & \sum \lambda = 1 \\ & \lambda \geq 0 \end{aligned}$$

where $\sum \lambda$ is an $I \times 1$ vector of ones, and by placing the convexity constraint $\sum \lambda = 1$, it ensures different scales of firm are recognised as efficient; consequently, envelopment is generated from multiple convex linear combinations of best practice (Coelli et al., 2003). A

¹³⁰ In other words, the CRS model relates to the ability of managers to utilise a firm's given resources, while the VRS model refers to exploiting scale economies by operating at a point where the production frontier exhibits CRS (Fethi and Pasiouras, 2010).

convex hull of intersection planes will be formed by using this approach, where the envelopment of data points is tighter than the CRS conical hull; this consequently, can generate technical efficiency scores that are greater than or equal to the scores obtained from the CRS model (Coelli et al., 2005). Additionally, unlike the CRS model, where firms can be benchmarked against firms that are substantially larger or vice versa, the VRS model would only allow an inefficient firm to be benchmarked against a similar sized firm by imposing convexity constraints (Coelli et al., 2005).

5.2.2.1 Cost- and Profit-Efficiency

The previous section discussed two variants of the DEA model, which focus on technical efficiency. But such measurement does not consider price input and output. Input and output prices are required for economic efficiency computation (e.g. cost- and profit-efficiency). By including input and output prices, the optimal proportionate allocation of firms on their input and output can be analysed according to prices. Hence, the DEA model can also measure cost- and profit-efficiency to support the firm's objectives, driven primarily by behavioural assumptions such as cost minimisation or profit maximisation (Cooper et al., 2007).

Cost-efficiency involves a minimisation of cost behaviour. Thus, the DEA cost efficiency model for this study adopts an input-orientated measure with VRS. The model of cost-efficiency can be written as (Coelli et al., 2005):

$$\begin{aligned}
 \text{Min}_{\lambda, x_i^*} \quad & w'x_i^* & (5.20) \\
 \text{subject to} \quad & Y\lambda \geq y_i^* \\
 & x_i \geq X\lambda \\
 & \Sigma\lambda = 1 \\
 & \lambda \geq 0
 \end{aligned}$$

where w_i is a $(N \times 1)$ vector of input prices for the i -th firm. x_i^* (computed by linear programming) is the cost minimisation vector of input quantities for the i -th firm, given the

input prices w_i and output level of y_i . Therefore, the total cost of efficiency of the i -th firm is computed as:

$$CE = w'_i x'_i^* / w'_i x'_i \quad (5.21)$$

where CE is cost-efficiency, $w'_i x'_i^*$ is the minimum cost and $w'_i x'_i$ is the observed cost. From the mathematical fractional formula, it can be shown that CE is the ratio of minimum cost to observed cost, for i -th firm. In addition, CE is interactive with allocative efficiency (AE) and TE where $AE = CE/TE$. These three efficiencies (CE , AE and TE) can take values ranging from 0 to 1, where 1 implies full efficiency.

Profit-efficiency involves the maximisation of profits for i -th firm. Thus, in this study, output-orientated efficiency is employed and the profit model can be written as (Coelli et al., 2005):

$$\begin{aligned} \text{Max}_{\lambda, y_i^*, x_i^*} \quad & (p'_i y'_i - w'_i x_i^*) & (5.22) \\ \text{subject to} \quad & Y\lambda \geq y_i^* \\ & x_i^* \geq X\lambda \\ & \Sigma\lambda = 1 \\ & \lambda \geq 0 \end{aligned}$$

Where p_i is a vector of output prices for the i -th firm. y_i^* is the revenue maximising vector of output quantities for the i -th bank, given the output prices p_i and the input level of x_i , w_i is a $(N \times 1)$ vector of input prices for the i -th firm, and x_i^* is the cost minimisation vector of input quantities for the i -th firm. To compute the profit- efficiency, the fractional formula is shown as (Coelli et al., 2005):

$$PE = (p'_i y_i - w'_i x'_i) / (p'_i y_i^* - w'_i x'_i^*) \quad (5.23)$$

where PE is the profit-efficiency, $(p'_i y_i - w'_i x'_i)$ is the observed profit and $(p'_i y_i^* - w'_i x'_i^*)$ is the maximum profit. Therefore, PE is a ratio of observed profit to maximum profit for the i -th firm. The value of PE ranges between 0 and 1, where a value of 1 indicates full efficiency.

With the exception of the widely used Berger and Mester's (1997) parametric alternative profit-efficiency, which considers the impact of market power and regulatory effects, only two studies were found that estimated alternative profit-efficiency using DEA models (see Maudos and Pastor (2003) and Ariff and Can (2008)).¹³¹ Fethi and Pasiouras (2010) found that estimation of profit efficiency using the DEA model is rather limited, based on a review of past literature. This may be attributed to the difficulty in collecting reliable information for output prices. Due to the lack of studies using alternative profit-efficiency for the DEA model, this research focuses on the new cost- and profit-efficiency DEA model devised by Tone (2002) which is different to the traditional DEA model. The new DEA models consider the impact of market power by relaxing the assumption of homogeneous of degree 1 in the input prices (as exhibited by traditional DEA models) (Tone, 2002).

5.2.2.2 New DEA Cost- and Profit-Efficiency

The traditional cost- and profit-efficiency in equations (5.20) and (5.22) assume that input prices are the same across all DMUs. In this scenario, a traditional DEA model might assume that the market is performing under perfect competition. However, this may not always be the case; input prices might not be identical across all DMUs. Therefore, Tone (2002) points out that the cost- and profit-efficiency in equations 5.20 and 5.22 do not take into consideration the fact that costs (profits) can be reduced by reducing the inputs (outputs) prices. For example, two banks have similar inputs and outputs, but one of them faces greater input prices than the other; under the traditional DEA model, the cost function is homogeneous of degree 1 in input price and the scaling factor is cancelled in the cost-efficiency ratio, which could result in both banks having the same cost-efficiency scores, even though they have different levels of input price. This is because traditional DEA focuses on the technical aspect (i.e. technical efficiency) of these two banks, and does not take into account variations in input prices. Therefore, in overcoming this

¹³¹ The alternative profit-efficiency of the DEA model introduced by Maudos and Pastor (2003) employs a solution that corresponds to the revenue R^* and input demand $x = (x_1, x_2, \dots, x_N)$, which maximises profits given the prices of inputs w . This solution uses a linear combination of firms that produce at least as much of each of the outputs using a smaller or equal quantity of inputs, and obtain at least as much revenue as i -th firm (Maudos and Pastor, 2003; Arif and Can, 2008).

drawback, Tone (2002) introduces a new DEA model that is distinct from being homogeneous of degree 1 in input price (as in the traditional DEA model). The new DEA model deals with different input prices by taking the TC of each unit, which is equal to the price of inputs multiplied by the quantity of inputs. Thus under Tone's (2002), the new DEA cost-efficiency model, the cost-efficiency of i -th bank, is determined to be the following linear programming problem:

$$\begin{aligned}
 \text{Min}_{\lambda, \bar{x}} \quad & e\bar{x} & (5.24) \\
 \text{subject to} \quad & \bar{x} \geq \bar{X}\lambda \\
 & y_i \geq Y\lambda \\
 & \lambda = 1 \\
 & \lambda \geq 0
 \end{aligned}$$

where e is a row vector of all elements equal to 1, $\bar{x} = w'_i \times x_i^*$. w'_i is the price of inputs and x_i^* is the quantity of inputs. Tone (2002) assumes that the elements of \bar{x} are denominated in homogeneous units in monetary terms. Unlike traditional DEA which retains the unit cost of i -th bank fixed at w'_i and searches for the optimal input mix x_i^* for producing y_i , the New DEA (NDEA) cost efficiency (CE) searches for the optimal input mix x_i^* for producing y_i based on the independently current unit price w'_i of i -th bank.

The NDEA CE for the i -th bank can be measured as:

$$\text{NDEA CE} = \frac{e\bar{x}_i^*}{e\bar{x}_i} \quad (5.25)$$

where $e\bar{x}_i^*$ is the minimum cost and $e\bar{x}_i$ is the observed cost. For the NDEA profit-efficiency, the NDEA profit efficiency (PE) of i -th bank can be mathematically described as:

$$\begin{aligned}
& \text{Max}_{\bar{x}, \bar{y}, \lambda} \quad (e\bar{y} - e\bar{x}) & (5.26) \\
& \text{subject to} \quad \bar{x} \geq \bar{X}\lambda \\
& \quad \quad \quad y_i \geq Y\lambda \\
& \quad \quad \quad \lambda = 1 \\
& \quad \quad \quad \lambda \geq 0
\end{aligned}$$

where e is a row vector with elements equal to 1, $\bar{y} = p'_i \times y_i^*$, where p'_i is the price of outputs and y_i^* is the quantity of outputs and $\bar{x} = w'_i \times x_i^*$, where w'_i is the price of inputs and x_i^* is the quantity of inputs. The NDEA PE of the i -th bank can be written as:

$$\text{NDEA PE} = \frac{e\bar{y}_i - e\bar{x}_i}{e\bar{y}_i^* - e\bar{x}_i^*} \quad (5.27)$$

where $e\bar{y}_i - e\bar{x}_i$ is the observed profit and $e\bar{y}_i^* - e\bar{x}_i^*$ is the maximum profit.

5.2.2.3 Scale-Efficiency

Scale-efficiency can be attained by using a ratio of CRS efficiency to the VRS efficiency. When both CRS and VRS efficiencies are available, the scale-efficiency for the i -th firm can be computed. The scale-inefficiency of firms occurs when the CE obtained under CRS differs from the CE under VRS (Coelli et al., 2005). The scale-efficiency (SCALE) can be defined and written as:

$$\text{SCALE} = \frac{CE_{\text{CRS}}}{CE_{\text{VRS}}} \quad (5.28)$$

where SCALE is the scale-efficiency, CE_{CRS} is the cost-efficiency under CRS and CE_{VRS} is cost-efficiency under VRS.

5.2.3 Consistency Tests between SFA and DEA Models

From the discussion in Chapter 2, both SFA and DEA have their own advantages and disadvantages. The result of efficiency scores derived from these different frontier approaches contains different information (Dong et al., 2014). Until now, there is no consensus on a single best frontier approach to measure efficiency. Hence, the discussion

about the frontier-efficiency approach that would result in the best estimation should not be sought. In view of that, Bauer et al. (1998) introduces a set of consistency condition tests in which two or more methods are used, such as parametric and nonparametric models. The empirical analysis of this research follows the consistency tests proposed by Bauer et al. (1998), which are:

1. Compare and analyse means, standard deviations and other distributional properties of SFA and DEA models.
2. Rank-order banks' efficiency scores using both SFA and DEA models (using the pairwise Spearman correlation coefficients) and analyse the sequence of banks in the ranking.
3. Cluster efficiency scores of banks into best-practice and worst-practice categories for both SFA and DEA models and examine overlapped proportion of banks that appear in the top 25% and lowest 25% of banks by efficiency score for each methods.
4. Test stability across different time intervals and compare stability of efficiency scores for both SFA and DEA models.
5. Compare and examine the correlations of efficiency scores from SFA and DEA models with non-frontier performance indicators (e.g. financial ratios) for consistency.

Bauer et al. (1998) suggest that the first three consistency conditions are introduced to measure the degree to which the different approaches are identical. Subsequently, the next two conditions are the degree to which the efficiencies estimated by the different techniques are consistent with reality. However, neither the effect of heterogeneity nor the determinant of cost- and profit-efficiencies for DEA are analysed in this research, owing to differences in their approach to accounting for environmental and bank-specific factors. For example, the SFA model in this study includes control and environmental variables in its 1-stage approach to capture heterogeneity effects (discussed later in this chapter) (Battese and Coelli, 1995), whereas the DEA model generally requires a 2-stage approach, in which efficiency scores derived from this model are then regressed with control and environmental variables to test their significance and relationship (Pit and Lee, 1981).

5.3 Data and Variables

This research employs data from a confidential database of a Malaysian financial organisation on the Malaysian banking industry from 2000 to 2011 (please note that investment banks are not included in this analysis due to distinct differences in operational structures and business models), which also covers the period of implementation of the ten-year Financial Sector Master Plan (FSMP) introduced in 2001. The sample includes 32 banks in 2000 followed by 39 banks in 2011, operating in Malaysia.¹³² The final sample consists of 354 observations after a number of missing data and outliers were taken out prior to the analysis. These excluded observations are due to the following conditions: (1) banks operational for less than 2 years that ceased operations after the merger and acquisition exercise in 2000, (2) banks with zero non-performing loans (NPLs),¹³³ (3) several newly formed Islamic banks' subsidiaries without a fixed asset value,¹³⁴ and (4) several outliers that exhibited high volatility in independent and dependent variables during the cleaning process of panel data.

In Chapter 3, it was noted that Hon et al. (2011) attempted to examine the impact of the FSMP. However, the study did not span the FSMP's entire timeline (i.e. between 2001 and 2005) and focused only on domestic banks. Generally, the data timeline used for most Malaysian banking efficiency studies are short, on average, spanning three to eight years. Additionally, to the author's knowledge, this is the first study that will cover the period of the global credit crisis. Therefore, to overcome these gaps, the period 2000–2011, is imperative to this analysis, when measuring the efficiency of Malaysian banks; as it covers the events of various liberalisation initiatives under the FSMP, including the interest rate deregulation in 2004, the introduction of new foreign banks and the impact of the global credit crisis into Malaysia between 2008 and 2010.

¹³² This includes all domestic conventional, foreign conventional, domestic Islamic and foreign Islamic banks in Malaysia.

¹³³ In particular, newly established banks may not have NPL in their books for a certain period of time. The NPL ratio is needed in this empirical analysis as control variables represent asset quality.

¹³⁴ These newly formed Islamic bank subsidiaries rely on banking groups' fixed assets during their early years of operation. For this study, fixed asset values are required to compute physical capital variables.

5.3.1 Input and Output Variables

Once the frontier measurement methods (cost- and profit-efficiency using SFA and DEA) are decided, the variables for such measurement have to be determined and defined. Dyson et al. (2001) state that inputs and outputs should cover the full range of resources used and outputs generated, particularly if one wants to evaluate the banking institutions' efficiency in transforming inputs into outputs. To evaluate banking efficiency, using input and output variables that do not contribute to the identification of bank efficiency should be avoided. Past literature suggests that, based on the definition of a bank approach (e.g. intermediation, production, user-cost and asset approaches), common sense and expert judgement can play a role in determining the input and output variables. Thus, it is imperative to include resources as input and output objectives that are regarded as relating to a bank's success performance as outputs (Avkiran, 1999). Therefore, for the SFA and DEA models, variables such as total costs and profits, inputs, outputs, input prices and output prices need to be specified. Unlike manufacturing firms, which use physical quantities as their output measurement, banks' outputs are rather complex: their inputs and outputs have to be based on the definition of the banks' operations. As discussed earlier, banks are seen as intermediary institutions that link savers and borrowers, and offer other services (such as fee transaction-based products and services, payment services, and investments) (Molyneux et al., 1996) (see Chapter 3 on the roles of banks). Thus, a proper definition that reflects the products of banks and how they are measured could affect the efficiency estimations of Malaysian banks (Goddard et al., 2001).

In the literature, there have been many long-standing arguments on the functions of banks and the output of banking institutions (Berger and Humphrey, 1992; Grigorian and Manole, 2002). The inputs and outputs to be employed have to be defined according to banking functions and services (Grigorian and Manole, 2002). Two main approaches have received the most attention in banking efficiency: the intermediation approach and the production approach (Berger and Humphrey, 1997; Goddard et al., 2001).¹³⁵

¹³⁵ In addition, there are other models as well, such as the asset approach, the value-added approach and the user-cost approach (Berger and Humphrey, 1992). First, the asset approach refers to banks as financial

The production approach considers a bank to be a production unit, which provides financial services to its customers. Under the production approach, banks are normally treated as firms that employ capital and labour to produce services for both deposit and loan account holders. Therefore, the deposit accounts provided to customers should be included as one of the outputs of banks. Outputs are measured by the number of account services, as opposed to dollar values, and are commonly treated as stock (Aly et al., 1990; Fu and Heffernan, 2007). One of the disadvantages of this production approach is that it does not count interest expenses as an input: interest generally forms approximately 60.0% of the banks' total costs (Ferrier and Lovell, 1990; Molyneux et al., 1996; Goddard et al., 2001). Moreover, data for the number of deposits and loan accounts are difficult to acquire, which makes the production approach less favourable for analysis (Aly et al., 1990).

On the other hand, under the intermediation approach, banks are treated as financial intermediaries between borrowers and depositors rather than producers of loan deposit services. Sealey and Lindley (1977) state that an individual banks' decision-making process focuses on the production of earning assets, in which 'loanable funds' borrowed from depositors and serviced by the bank. They further suggest and assume that banks collect funds (deposits and purchased funds with the assistance of labour and capital), and convert these into loans and other assets.

Nevertheless, Berger and Humphrey (1997) and Fu and Heffernan (2007) find that both approaches are imperfect because they do not fully capture the dual role of financial institutions, which includes both the provision of transactions and document processing services and the transfer of funds from savers to borrowers. The multi-product nature of banking firms is widely recognised, but until now there has been no consensus as to which approach is the 'best', particularly on how the 'production flow' characterising a bank's

intermediaries between liability holders (i.e. savers) and the beneficiaries of funds. The inputs are deposits, and other liabilities and the outputs of the banks are defined as assets and loans (Favero and Pappi, 1995). Second, the value-added approach considers that assets and liabilities possess some output characteristics and the inputs and outputs of a bank are defined based on their share of value-added, such as loans, demand deposits and savings deposits. Other outputs are not utilised and considered unimportant (Bhattacharya et al., 1997). Third, the user-cost approach can be described as the contribution of inputs or outputs to bank revenue. For example, a transaction is defined as an output if the financial return (e.g. ROE and ROA) exceeds the opportunity cost of the funds, or is defined as cost (liability) if the financial cost is less than the opportunity cost of those funds (Aly et al., 1990).

outputs and inputs is defined. The appropriateness of each method (i.e. production and intermediation) varies according to the circumstances (Tortosa-Ausina, 2002). As suggested by Berger and Humphrey (1997), each approach has its advantages, particularly in view of differences in the areas within the financial institutions. Of studies evaluating the efficiency of branches of banks, Berger and Humphrey (1997) suggest that the production approach should be employed since branch managers typically have little influence over bank funding and investment decisions. On the other hand, for evaluating entire banks, the intermediation approach is more appropriate because it includes interest expenses, which account for between one-half and two-thirds of total costs.

The intermediation approach has the advantage of being more inclusive and capturing the essence of financial intermediaries (Berger et al., 1997). The intermediation approach utilises deposits as an input, and is more realistic as an approach, as banks use deposits and other funds to generate loans and investments (see Aly et al., 1990; Zaim, 1995; Berger et al., 1997; Casu and Girardone, 2002; Fu and Heffernan, 2007; Girardone et al., 2004). Using deposits as an input is more convincing than the production approach (which treats deposits as an output), since they are paid in part by interest payments, and the funds raised provide the bank with its basic 'raw material': that is, investable funds. In accordance with this view, Elyasiani and Mehdian (1990) support the idea that banks buy rather than sell deposits.

The unit of bank inputs and outputs, under the intermediation approach, are measured in terms of monetary values that might also determine the market share of individual banks. Moreover, some services cannot be measured in terms of numbers, such as investment in securities. Hence the intermediation approach is very suitable, because it addresses questions related to cost minimisation and emphasises the overall costs of banks (Ferrier and Lovell, 1990), which include interest expenses on deposits and other purchased funds.

For the purpose of analysing the efficiency of Malaysian banks, this research employs the intermediation approach introduced by Sealey and Lindley (1977), specifying banks as intermediary institutions, which collect deposits to produce loans and investments with the support of labour and capital. From Table 5.1, the input variables used in this analysis are

funds, labour and physical capital; and the outputs are loans, investments and other earning assets.¹³⁶ In addition, the control variables used in this research are asset quality, capital adequacy, liquidity, and time trends to capture the effect of heterogeneity. Moreover, a number of environmental variables are introduced: ownership, specialisation, liberalisation phases (based on the FSMP three-stage liberalisation process), global credit crisis period, bank size, market concentration and market share. These variables are employed for the cost, standard profit- and alternative profit-efficiencies estimations (discussed in detail later in this chapter).

For the estimation of cost-efficiency of SFA, total cost (TC) is employed as the dependent variable, which includes interest expenses on deposits and amounts owing to other banks, cash and cash equivalents, staff expenses, depreciation, provisions for expenses and other expenses related to the operations of banks. On the other hand, for estimating profit-efficiencies of SFA, the dependent variable total profit (TP) consists of total profit before tax, which is equivalent to income and revenues from loans, investments, amounts owed by other banks, trading income and other revenues from banking operations, minus total costs.¹³⁷

5.3.1.1 Input Prices, Outputs and Output Prices

In view of the discussion above, the most common inputs employed from previous bank efficiency studies using intermediation the approach are the price of labour, the price of physical capital and the price of funds. Hence this research employs three input prices: price of funds (W1), price of labour (W2) and price of physical capital (W3).

From earlier discussions, the role of deposits has been controversial as deposits have both input and output characteristics. Deposits could be considered as inputs because interest

¹³⁶ For standard profit-efficiency, the outputs to be employed in this study are the price of loans, the price of investments and the price of other earning assets.

¹³⁷ As shown in equations 5.6 and 5.8, the dependent variable TP is added with a constant θ to every bank's profit so that the natural log is taken as a positive number. This is required due to some banks showing negative profits or facing losses. Since the log does not accept negative values, the dependent variable TP is transformed using $\ln TP + 1 + TP^{min}$, where $TP^{min} = \theta$ is the minimum value of TP or maximum loss from the sample of panel data (Berger and Mester, 1997).

expenses should be paid for them (Elyasiani and Mehdian, 1990). Hence, the price of funds (W1) is used in this research, where banks are deemed to pay interest expenses on borrowed funds as well as on deposits placed by bank customers. In banking efficiency studies, the price of funds is often expressed as interest expenses incurred by the bank on borrowed or purchased funds and deposits over total funds and deposits (see Berger, 1995, Kwan and Eisenbeis, 1996, Bauer et al., 1998, Casu and Girardone, 2002, Huang and Wang, 2002).

The price of labour (W2) is commonly used as an input for the intermediation approach (see Berger, 1995; Berger and Mester, 1997; Humphrey and Pulley, 1997; Weill, 2004; Fu and Haffernan, 2007). The variable most often used for the price of labour is the average expenditure on each employee. The ratio of labour expenses to the number of employees is used for most of the previous studies (e.g. Mester, 1993; Kaparakis et al.; 1994; Resti, 1997; Berger and di Patti, 2003). Some studies also adopt the ratio of labour expenses to total assets (e.g. Weill, 2004; Altunbas et al., 2001). Additionally, there are also studies that employ only the total amount of labour costs, due to unavailable data on the number of employees or other suitable variables to normalise the amount of expenditure (Noulas, 1997). The drawback of using labour expenses without implying employee numbers is that it could result in a bias, particularly with banks that hire quality workers at a higher cost (Drake et al., 2006). In cases where data on the number of employees are unavailable or incomplete, several variables and assumptions are used to proxy these data. Fu and Haffernan (2007) employed growth assumptions based on total asset growth information to be applied to the growth in employee numbers. This method was also applied in Rezvanian and Mehdian (2002) and Vander and Vennet (2002). The assumption is useful, particularly when the data are incomplete or where missing data are replaced with total asset growth information. On the other hand, when the number of employee data is significantly missing or completely unavailable, total assets are commonly used as a proxy (e.g. Altunbas et al., 2000, 2001; Fiordelisi et al., 2011; Weill, 2004; Hasan and Marton, 2003; Maudos et al., 2002). Total assets can be a better proxy compared to other financial statement items such as loans or deposits because, ideally, they cover the entire banking business as financial intermediation. Fries and Taci (2005) asserted from their study that if proxy variables such as total assets are not used, the intended study will not be complete due to substantial

missing data. It is also worth noting that when comparing a ratio to the total assets with a ratio to the number of employees, the result is unlikely to be constant over the period under study, which can yield different parameter estimates (Altunbas et al., 2001). Based on the discussion above, total assets are found to be a reasonable proxy when employee number data is unobtainable (Altunbas et al., 2001).

For the price of physical capital (W3), the basic indication is based on expenses made on physical assets over physical assets such as buildings and equipment. Physical capital is considered as an input in the intermediation approach because banks are required to incur costs in assets other than labour and funds in producing loans, deposit and securities (Mester, 1996; Mahajan et al., 1996).

For this study, three output quantities and output prices are employed: loans (Y1), investments (Y2), other earnings assets (Y3), and price of loans (P1), price of investments (P2), and price of other earnings assets (P3). Unlike cost function efficiency, when standard profit-efficiency is used, the output variables differ from the cost function outputs. The standard profit function specifies variable profits in place of variable costs, and takes variable output prices as part of its equation. Profit dependent variables allow for consideration of revenues that can be earned by varying outputs as well as inputs (Berger and Mester, 1997). Instead of taking the amount or magnitude of the loans and deposits, the prices of loans and deposits were considered. The prices of loans and deposits in general are based on income or interest received from loans and deposits over loans and deposits (see Berger and Mester, 1997; Isik and Hassan, 2002).

In the intermediation approach, banks are assumed to produce loans from the input variables. Thus, output loans (Y1) is included, consisting of items such as loans and advances, credit cards, hire purchase and overdrafts.¹³⁸ On the other hand, in estimating

¹³⁸ Some of the drawbacks can be seen in Resti (1997), who argues that loans could be a possible source of bias due to the fact that they are taken from the financial statements, which conceals heterogeneity in credit quality (different levels of credit management and administration). Additionally, loans have different time horizons and structures, which require different types of monitoring and cost embedded to it. Nevertheless, loans have been used heavily in previous studies, as they are one of the main business activities of banks (e.g. Berger and Humphrey, 1992; Akhavein et al., 1997, Resti, 1997; Al-Jarrah and Molyneux, 2005; Bos and Schmeidel, 2007).

the standard profit-efficiency, the price of loans (P1) is used and measured as a percentage of interest income derived from loans (Berger and di Patti, 2003).

The second output in this study is investments (Y2). Investments (Y2) include dealing securities, investment securities, government and private debt securities. The price of investment (P2) can be measured using the ratio of income from dealing and investment securities to investment (Y2) (Drake et al., 2006; Berger et al., 2009).¹³⁹

The third output is other earnings assets (Y3), which comprises interbank deposits placed in other banks. The interbank deposit placed in other banks generates interest income, which can be received either from customers or deposits from other banks. The price of other earnings assets (P3) is measured as a percentage of interest income from deposits placed in other banks over other earnings assets (Casu and Girardone, 2002).

¹³⁹ Investments are considered to be intermediation activities and included as an output, because they naturally allocate resources into productive activities, which generates more income for the banks. They could represent assets that are drawn when demand for loans increases, or added when such demand declines (Allen and Rai, 1996).

Table 5.1 Descriptive Statistics of Banks' Inputs and Outputs, 2000–2011

Variables		Descriptions	Mean	Std Dev.	Min.	Max.
Dependent variables						
TC	Total Cost (MYR million)	Variable operating costs of deposits, personnel and investments	1,438.86	97.97	17.05	12,413.74
TP	Total Profit (MYR million)	Variable profits, revenues from loans and investments minus total cost	805.57	52.51	1.00	7,063.54
Independent variables/output prices						
Y1	Loans (MYR million)	Total loans, net of interest in suspense and NPL	15,273.20	19,496.82	19.85	99,506.07
Y2	Investments (MYR million)	Dealing securities, investment securities, government and private debt securities	3,393.27	4,349.12	0.05	27,420.31
Y3	Other earning assets (MYR million)	Interbank deposits in other banks	4,473.50	5,374.63	25.04	30,884.64
Independent variables/input prices						
W1	Price of funds (%)	Interest expenses on deposits/deposits	2.41	0.81	0.25	10.06
W2	Price of labour (%)	Personnel expenses/total assets	0.64	0.25	0.03	1.85
W3	Price of physical capital (%)	Other operating expenses/physical assets	11.50	13.89	0.00	121.02
Independent variables/output prices						
P1	Price of loans (%)	Interest income from loans/average loans	5.80	1.67	0.57	25.66
P2	Price of investments (%)	Income from dealing and investment securities/investment (dealing, investment, government debt, private debt, shares securities)	3.65	2.03	0.07	15.13
P3	Price of other earnings assets (%)	Interest income form interbank deposits/deposits in other banks	5.21	13.5	0.29	175.64

Notes: All financial values are inflation-adjusted to the base year 2000.

As of December 2011, USD 1.0000 is equal to MYR 3.1265

5.3.2 Control and Environmental Variables

Based on the earlier discussions, banks are deemed to allocate resources, such as deposits, labour and capital to produce loans, investments and other earning assets. By employing only inputs and outputs for the estimation of cost- and profit-efficiency, banks are assumed to face similar conditions with no differences found among them (homogeneous). However, as discussed in Chapter 2, a bank faces various challenges in transforming inputs into outputs, which could be attributed to bank-specific characteristics and/or external environmental factors.¹⁴⁰ Omitting control and environmental variables may lead to some efficient elements of banking operations being incorrectly measured in cost- and profit-inefficiency. As discussed in Chapter 3, studies of explanatory factors or determinants of inefficiency in Malaysian banks are found lacking, which may result in a failure to reveal those factors. To overcome these limitations, control and environmental variables are included in the efficiency analysis of Malaysian banks in this research.

Figure 5.3 exhibits the common methods of testing heterogeneity factors using 2-stage and 1-stage approaches. A 2-stage procedure is commonly employed by estimating cost and profit frontiers. This is performed by deriving efficiency scores at the first stage and regressing them against a set of possible determinants at the second stage (Pitt and Lee, 1981; Cebenoyan et al., 1993; Bhattacharya et al., 1997; Berger and Hannan, 1997; Akhigbe and McNulty, 2003; Bonin et al., 2005). This procedure, however, has limitations and suffers from serious econometric problems. Inefficiencies are assumed to be identically distributed in the first stage, and to have a functional relationship with a set of variables in the second stage (Kumbhakar and Lovell, 2000). Berger and Mester (1997) argue that this analysis is suggestive but not necessarily conclusive because the dependent variable in the regression (i.e. efficiency) is an estimate: the standard error of this estimate is not accounted

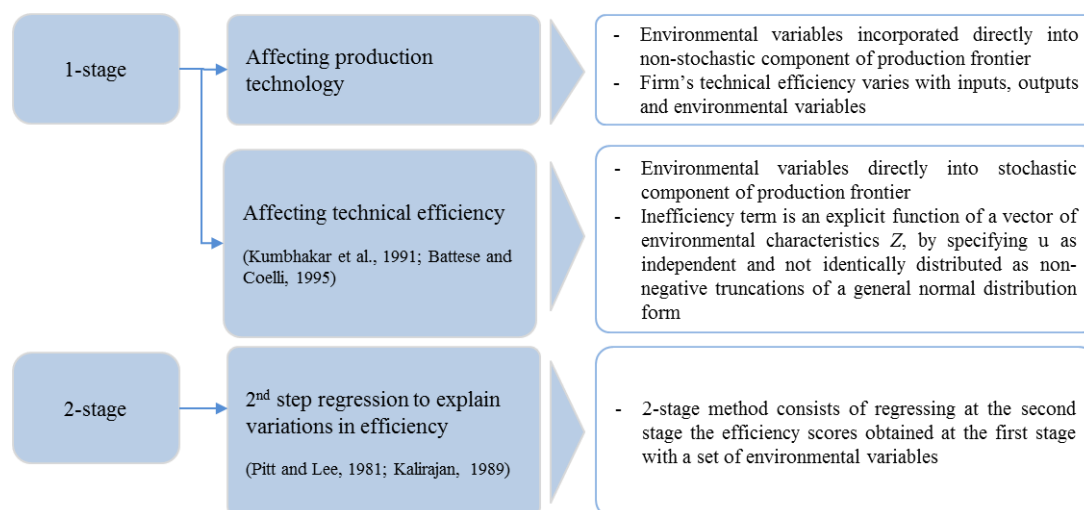
¹⁴⁰ This study incorporates several control and environmental variables that may influence a bank's cost and profit that are useful to policy makers, particularly if bank efficiency results are to be used to inform policy analysis (Drake et al., 2006). In addition, these environmental factors that could influence a bank's performance are generally not in managerial control. The objective of control and environmental variables is to identify possible sources of inefficiency. This can be done by investigating the relationship between the variation in bank efficiencies and the variation in the exogenous environmental variables (Kumbhakar and Lovell, 2000).

for in the subsequent regression or correlation analysis. Thus it is possible to determine inferences only about correlation, and not about causality.

An alternative 1-stage estimation procedure may overcome this problem. This is done by incorporating environmental factors into the non-stochastic function of the production frontier. However, these external factors are assumed to affect either the structure of the frontier, through which conventional inputs are processed into outputs, or directly to affect technical efficiency. Assuming that external environmental variables have a direct effect on the production structure, they are included in the frontier function as control variables (Drake et al., 2006). Another way of incorporating environmental variables into a 1-stage procedure is to assume that exogenous variables influence technical efficiency, rather than the structure of production technology (see also Berg and Kim, 1998; Allen and Rai, 1996; Drake et al., 2006; Bos and Schmeidel, 2007). In other words, this 1-stage procedure provides an explanation for variations in efficiency, and characterises the production environment. Battese and Coelli (1995) propose a model that assumes non-negative cost-inefficiencies as a function of firm-specific variables, which are independently distributed as truncations of normal distributions with constant variance but with means that are a linear function of observable variables of a general normal distribution form. Battese and Coelli (1995) demonstrated a model that can include both control and environmental variables into the stochastic functions. This model derives from the weaknesses found in Battese and Coelli's (1992) time-invariant and varying model, which includes environmental variables directly into the model. Battese and Coelli (1995) also assumes the time-varying model is helpful in overcoming statistical problems such as the multicollinearity problems associated with the translog functional form (Battese and Coelli, 1995).^{141 142}

¹⁴¹ The translog functional form is more flexible than other functional forms (e.g. the Cobb–Douglas function), and hence, multicollinearity might exist among variables, which can result in parameter estimates. However, multicollinearity may not be a severe problem, particularly when the scores from efficiency analysis are utilised for forecasting purposes (Coelli et al., 2005; Kumbhakar and Lovell, 2000).

¹⁴² If the multicollinearity problem is mainly created by a positive correlation between the second-order terms in the translog form of the cost and profit function, maximum likelihood estimates are still unbiased and efficient. However, multicollinearity problems can result in the estimated standard error of the coefficients to be large and small values of *t*-ratios, which in turn could lead to results that are biased towards accepting a null hypothesis with coefficients that are equal to 0 (for details, see Gujarati, 2003).

Figure 5.3 Accounting for Environmental Variables in Stochastic Frontier Analysis

Adapted from Coelli et al. (2005)

In this empirical study, as well as input and output variables, control and environmental variables are also introduced and considered in the SFA specification to capture the heterogeneity effect. The control variables are assumed to have a direct influence on banks' performances that is incorporated directly into the non-stochastic component of the production frontier. These control variables are to be fully interactive with outputs and input price variables, which consequently could affect the production technology (Coelli et al., 2005). On the other hand, the environmental variables are treated by incorporating them directly into the stochastic component, which influences inefficiency (u) and can affect the technical efficiency of the banks (Battese and Coelli, 1995). Based on past banking efficiency studies, this present study employs bank-specific factors, such as capital adequacy, liquidity and asset-quality as control variables, and exogenous factors, such as ownership, specialisation, financial liberalisation phases, the global crisis effect on Malaysian banking, bank size, market concentration, and market share as environmental variables.

5.3.2.1 Control Variables

Control variables, which have the ability to influence a bank's costs, are often included into the SFA model. These variables are treated similarly to input and output variables where

they interact with other parameters in the model. The control variables are included in the measurement of the cost- and profit-efficiency of the banks to account for their inherent risks. Moreover, risk-taking choices by the banks are an important part of the banking technology, and can affect the banks' efficiency (Hughes and Mester, 2008). Failure to consider heterogeneity effects may result in biased efficiency scores. For example, some banks may be indicated as inefficient although their prudent operations exhibit risk-averse behaviour, and on the other hand, some banks could also be efficient, even though their quality of outputs is poorer than others (Mester, 1997). Thus, by including heterogeneity effects, using control variables in the specification of cost- and profit-efficiency functions could result in greater accuracy in efficiency scores. These control variables are incorporated into the non-stochastic component of the production functions and fully interact with the inputs and outputs to consider the effect of risks in the estimated cost- and profit-efficiency (Coelli et al., 2005). In this study, there are four control variables (see Table 5.2) that are included in this present research: (1) capital adequacy (C1), (2) asset quality (C2), (3) liquidity (C3); and (4) time effects (T).

First, capital has been one of the common control variables used in measuring banking efficiency (e.g. Clark, 1996; Mester 1996; Berger and Mester, 1997; Girardone et al., 2004). A bank's insolvency or capital adequacy ratio depends on the financial capital availability to absorb potential loan and investment losses. Thus, the capital ratio (defined as capital adequacy for this control variable) is a ratio of capital to total assets. Capital ratio provides a proxy to regulatory conditions for measuring capital adequacy. In the absence of complete data on regulatory capital adequacy ratios (i.e. Basel's risk-weighted capital ratio (RWCR)), this analysis uses a capital ratio (i.e. capital to total assets ratio) that reflects the proportion of total assets (not risk-weighted) financed by financial capital. While a capital to total assets ratio is less risk-sensitive than the regulatory capital adequacy ratio, the changes in the capital ratio could still reveal shifts in a bank's balance sheet's structure and shifts in in the bank's risk-taking (Fiordelisi et al., 2011).

In general, the degree of insolvency risk influences the risk premium of a bank's borrowings. The higher the financial capital available to absorb potential losses, the lower the borrowing risk premium expected for a bank. Hence, capital has a direct influence on a bank's

borrowing cost and should be considered as an input in its production process, in which capital is also one of many sources of loanable funds other than deposits (Berger and Mester, 1997). Therefore, failure to control for capital could yield a scale bias (Mester 1997). A risk-averse bank normally holds higher capital, and if capital is ignored, the efficiency of these banks will be mis-measured, even though they behave at an optimal level (Mester, 1996).¹⁴³ Furthermore, Berger and Mester (1997) state that a well-capitalised firm is more efficient than its undercapitalised counterpart. They suggest that the relationship between capital and efficiency is positive, and undercapitalised banks with a higher appetite for risk normally exhibit lower levels of efficiency.

Hughes and Mester (1993) argue that the level of a bank's financial capital (i.e. equity) provides an alternative to deposits and other borrowed funds as a source of loanable funds, which may have a direct impact on a bank's borrowing costs and should be considered as an input to the bank's production process. However, owing to the small proportion of financial capital in Malaysian banks (averaging between 8.0% and 12.0%) compared to other sources of funds (e.g. deposits and borrowed funds), capital in this study is viewed as a measure of risk preferences, rather than as part of the banks' production process. Furthermore, the introduction of Basel I in 1988 on international banks' capital standards emphasised the importance of capital in banks' risk management (Saurina, 2004). This is also in line with Fiordelisi et al. (2011), where the capital ratio was employed as a measure of risk, using a ratio of capital to total assets.¹⁴⁴ Santos (1999) and Diamond and Rajan (2000) also argued that using the capital ratio as a measure of risk is a better concept of capital adequacy than using the book value of equity.¹⁴⁵

¹⁴³ Raising capital generally results in higher costs relative to deposits. Thus, risk-averse banks prefer equity financing and for that, they would appear inefficient if financial capital were omitted from the efficiency analysis. On the other hand, unlike expenses incurred on depositors, dividends paid on financial capital are not considered as cost (based on an intermediation approach), thus, if capital is not controlled for banks holding a greater level of financial capital, they would appear more efficient (Mester, 1996; Berger and Mester, 1997).

¹⁴⁴ For this research, the capital ratio variable (C1) affects a bank's cost structure due to its direct interactions with inputs and outputs in the cost and profit efficiencies' specifications.

¹⁴⁵ The capital ratio can be used to compare banks of different sizes. This ratio may also imply the level of equity to acquire assets with the aim of increasing the return on equity (Diamond and Rajan, 2000). Moreover, regulators have used the capital adequacy ratio as an instrument to protect depositors and promote the stability and efficiency of banks: banks are required to hold a certain ratio of capital to assets, as specified in Basel's capital regulatory standards (Santos, 1999, 2001; Diamond and Rajan, 2000).

As shown in Chapter 4, the capital ratio of Malaysian banks was steadily constant, averaging around 8.0% over the years 2000–2011. The capital ratio of Malaysian banks declined slightly in the second phase of the FSMP (2004–2007) owing to active capital management by the banks. Instead of issuing Tier-1 capital, Malaysian banks were active in issuing Tier-2 subordinated debt capital to diversify the overall cost of capital (Bank Negara Malaysia, 2005, 2006b). However, the degree of capital over total assets increased slightly from 7.7% in the second phase of FSMP (2004–2007) to 8.7% in the third phase of FSMP (2008–2011). During this period of uncertainty (the global credit crisis), Malaysian banks managed to increase their capital position via rights issues and Tier-1 capital at an accepted price to buffer for potential losses, something generally unachievable in a time of economic turmoil (Bank Negara Malaysia, 2008b). Hence, the seventh hypothesis for this study is:

Hypothesis 7 Banks with a lower capital ratio exhibit higher costs and lower profits.

Second, as mentioned in Chapter 3, one of the functions of a financial intermediary is to transform credit risk. Credit risk is commonly measured as an asset quality (C2) in the CAMEL framework.¹⁴⁶ It is also known as the largest risk that banks face as intermediating institutions and, owing to its high importance, this type of risk is explicitly measured in the computation of RWCR in Basel I and Basel II regulatory standards (Basel Committee of Banking Supervision, 1989, 2005). In view of that, non-performing loans (NPLs) (as proxies of banks' asset quality) are added to the model to control for the bank's risk structure (see Clark, 1996; Mester, 1996; Berger and Mester 1997).¹⁴⁷ It is paramount to control for NPLs because, when comparing efficiency, banks must have a homogeneous output quality; otherwise, unmeasured differences in loan quality may be mistakenly measured as inefficiency (Berger and Mester, 1997). By omitting NPLs as a control variable, banks with a good quality loan portfolio may appear inefficient because they use more labour and physical capital in managing and monitoring loan quality (Mester, 1996; Berger and Mester, 1997).

¹⁴⁶ This is termed according to CAMEL bank rating methodology of the Federal Reserve. C stands for capital adequacy, A is asset quality, M is management, E is earnings and L stands for liquidity. One of the key measures for asset quality is the NPL ratio.

¹⁴⁷ In the banking sector, NPLs are an indication of asset quality and the management's capability in monitoring the credit portfolio.

A high proportion of NPLs against total loans suggests distress in a bank, exposing the bank's capital to risk and potential failure, reflecting the bank management's quality. Kwan and Eisenbeis (1994) established that problem loans are negatively related to efficiency, and Berger and De Young (1997) consequently found a link between management quality and problem loans, and report that an increase in management quality reduces banks' problem loans, supporting Cebenoyan et al.'s (1993) and Wheelock and Wilson's (1995) whose findings suggest that the relationship between operating inefficiency and failure rates is positive.

Whether or not to include the NPL ratio variable in the bank's cost, standard and alternative profit functions depends on the extent to which these variables are exogenous. Following Berger and Mester (1997) and Berger and De Young (1997), NPLs should be exogenous to the bank if they are generally caused by negative economic shocks ('bad luck hypothesis'). A 'bad management hypothesis' indicates that a bank may incur extra expenses by administering bad loans, reflecting bad bank management, and a 'skipping hypothesis' argues that banks may save costs now by not investing in loan-monitoring expenses but face high default loans later. Banks are required to incur extra administrative expenses and managerial efforts in overcoming the impact of problem loans in their operations, resulting in a potential reduction in cost-efficiency. In this regard, adding NPLs as a control variable helps to remove (by statistical means) the costs of dealing with problem loans. Nevertheless, in the context of this study, which measures efficiency in an environment characterised by a series of events such as economic crisis and deregulation, the NPL variable should remain controlled in the bank's cost, standard and alternative profit functions.

As discussed in Chapter 4, the NPL ratio has been on a declining trend since the Asian financial crisis: 13.7% in 2000 declined to 2.9% in 2011. During the global economic crisis between 2008 and 2010 (Bank Negara Malaysia, 2010b), the NPL share increased slightly from 3.7% in 2008 to 3.8% in 2009. Various economic stimuli were introduced during this period and, supported by a lower interest rate environment, the Malaysian banks were able to withstand the external crisis. Thus, the NPL rate has remained below 4.0% since 2008. Accordingly, for this study, the control variable for the NPL ratio (defined as asset quality)

is a measure of non-performing loans over total loans. Thus, the eighth hypothesis for this empirical study is:

Hypothesis 8 Banks with a higher non-performing loans' ratio exhibit higher costs and lower profits.

Third, another key control variable to be included in the SFA is the liquidity risk (C3). The introduction of the liquidity coverage ratio in the Basel III international banks regulatory standards amplifies the importance of liquidity risk in banking operations (Basel Committee on Banking Supervision, 2010). Liquidity risk in this study is measured using the loans over deposit ratio. Deposit taking and lending by banks are closely related. Both activities (i.e. deposit taking and granting loans) exhibit the liquidity transformation function of banks and incur similar costs for converting deposits into loans (Kashyap et al., 2002). Thus, deposits and loans can be analysed in tandem through the loan-to-deposit (LTD) ratio. This ratio is a proxy for the degree to which banks are active in traditional forms of financial intermediation (i.e. lending). In other words, the LTD ratio captures differences among banks in terms of their ability to convert deposits into loans (Claessens and Horen, 2011). This ratio measures the ability of the bank to meet its liability obligations, maintain adequate liquid assets and collect receivables (e.g. loans and other earning assets). In a general sense, lower liquidity ratios suggest a bank has a larger margin of safety and ability to cover its short-term obligations. As liquid assets are controlled in outputs, one would expect banks with higher liquid assets (all other things being equal) to be more cost efficient.

Since saving accounts and transaction deposits normally can be withdrawn at any time, there is a high liquidity risk for banks. Banks can experience liquidity problems, especially when withdrawals exceed available deposits significantly over a short period of time (Samad and Hassan 2000). At the same time, when loans exceed the deposit base, banks face a funding gap for which they have to seek funds from the financial market. A high funding gap implies a high dependence on market funding, which could be more expensive than retail funding (de Haan and End, 2013). Goodhart et al. (2013) suggests that capital alone is not sufficient to contain the problems during a crisis, in which liquidity adds value to the funding measures to mitigate systemic risks. Due to the importance of a bank's

liquidity structure as evidenced from the 2007–2008 global credit crisis, liquidity issues are reflected in the development of a regulatory framework (i.e. Basel III), which intends to contain the risks during the financial intermediation process of the banking industry. Therefore, the liquidity ratio (C3) is included as a control variable for analysing cost- and profit-efficiency.

The average liquidity ratio for Malaysian banks between 2000 and 2011 was at 75.1%. Based on three different phases of the FSMP, the liquidity ratio for the first, second and third phases of FSMP was at 83.0%, 68.4% and 73.9% respectively. Malaysian banks were more liquid in the second phase of the FSMP, which implies that during good times, liquidity tends to rise, particularly when market funding is abundantly available to finance credit growth (de Haan and End, 2013). This can be evidenced from Chapter 4 where it was observed that the deposit from 2003 to 2006, from 69.3% to 73.5% of total capital and liabilities (see Figure 4.9). Nevertheless, instead of transforming the deposits into loans, Malaysian banks' deposits were channelled into the interbank market and investments due to short-term maturity of deposits received. In the third phase, the deposits of Malaysian banks continued to strengthen and recorded 75.7% of total capital and liabilities in 2009. One of the reasons found to support strong deposits in Malaysian banks was due to creation of the deposit insurance corporation (PIDM) and explicit government guarantee on deposits during the stressed conditions of the global credit crisis (see Chapter 4)(Bank Negara Malaysia, 2008b). In addition, in response to various economic stimuli implemented by the Malaysian government, banks continued to play their part as intermediary institutions and provide loans to deserving customers, as reflected from higher LTD ratio of 73.9% in the third phase of the FSMP. Hence, the ninth hypothesis is:

Hypothesis 9 Banks with greater liquidity exhibit lower costs and higher profits.

Fourth, a time trend variable (T) is also employed to control changes in technology over the time period under study. Time trend is estimated using $T_1 = 2000$, $T_2 = 2001, \dots$, $T_{12} = 2011$. Isik and Hassan (2002) point out that it is important to include the time variable particularly in the continuously changing business environment where banking technology may vary from time to time. A bank may be efficient one year but not the next; this change

may be influenced by technology. Altunbas et al. (1999) state that the time trend (T) variable is a ‘catch-all’ variable that captures technological factors, which exhibits that banks learn by performing organisational changes such as more efficient utilisation of inputs and other factors (e.g. adjustments made resulting from changes to environmental regulations). Thus, technical change can be estimated from a ratio of changes in total cost to changes in technology, as shown from the measurement of partial derivative of the estimated cost function, with respect to the time trend. This can be mathematically written as $\partial \ln TC / \partial T1$ (Altunbas et al., 1999).

Table 5.2 Descriptive Statistics for Control Variables, 2000–2011

Variables	Descriptions	Mean	Std Dev.	Min.	Max.	
C1	Capital adequacy (%)	Capital/total assets	10.43	6.38	1.19	36.85
C2	Asset quality (%)	Non-performing loans/loans	8.00	7.72	0.07	54.23
C3	Liquidity (%)	Loans/deposits	66.79	22.55	0.63	120.34
T	Time trend	T1 = 2000, T2 = 2001, ..., T12 = 2011	6.95	3.51	1.00	12.00

Table 5.2 shows that the Malaysian banks are heterogeneous in terms of capital adequacy asset quality and liquidity. This can be seen from the greater dispersion in the standard deviation, in comparison to the sample means. Thus, it is important to include all the control variables mentioned above into the measurement of cost- and profit-efficiency and test the hypotheses developed.

5.3.2.2 Environmental Variables

The inclusion of environmental variables in the cost and profit functions of the SFA models could influence the efficiency of banks (Battese and Coelli, 1995). Kasman and Yildirim (2006) iterate that market structure, geographical areas and other financial depth variables (e.g. banks size, market concentration and market share) are often employed for certain attributes of the banking sector when measuring efficiency estimates. By having these variables in the study, the different environments in which the banks operate may demonstrate significant variations, although they can be relatively homogeneous (Kasman and Yildirim, 2006).

Environmental variables are used to observe the banking and market conditions of the period under study. When accounting for environmental variables that are exogenous, one can theoretically explain that the scores generated are a measure of managerial performance (Coelli et al., 1999, Abdul Majid et al., 2011). That is, the efficiency scores resulting from SFA, using the 1-stage approach is net of the impact of environmental variables on input and output variables (Coelli et al., 2005). From Table 5.3, the environmental variables employed for this research are ownership structure (Z1), specialisation (Z2), FSMP liberalisation phase 2 (Z3), FSMP liberalisation phase 3 (Z4), effect of global financial crisis in Malaysia (Z4), bank size (Z6), market concentration (Z7) and market share (Z8).

First, the ownership structure (Z1) is introduced into the stochastic frontier specification to consider the effects of bank-specific factors. As mentioned in Chapter 4, there are two types of bank ownership in Malaysia: domestic and foreign. To measure the effect of ownership, dummy variables are assigned: 0 for domestic banks and 1 for foreign banks. Foreign banks in Malaysia are locally incorporated subsidiaries that are fully owned by international foreign banks. On the other hand, domestic banks can either be fully owned by local parties or partly owned by foreigners as minority shareholders. As discussed in Chapter 3, banking institutions have heterogeneous ownership, corporate controls, market and risk characteristics. Relying on different hypotheses and theories, such as the theory of the firm, agency theory, managerial theories, the behavioural theory of the firm and market structure hypotheses (see Chapter 2), many past studies have investigated whether ownership structure or organisational form are related to the differences in frontier-efficiency. These theories are implicit in terms of different types of ownership or/and organisational forms that could provide stronger incentives to control costs and/or increase profits for better efficiency. Therefore there might be differences associated with foreign versus domestic ownership (Isik and Hassan, 2003).

Berger and De Young (2001) state that greater control by parent banks may imply that the efficiency of their subsidiaries will be similar to that of their parents in the respective home countries. The controlling parent banks can export their quality of managerial skills, policies and procedures. A greater level of competition and improved corporate governance practices are expected from the knowledge and technology transfers of foreign banks. In

other words, foreign banks can improve the quality, pricing and availability of financial services, both directly as providers of such enhanced services and indirectly through competition with domestic financial institutions (Levine, 1996). The existence of foreign banks in developing countries can improve the infrastructure of the financial system (e.g. accounting standards, transparency and financial regulation), and at the same time stimulate the increased presence of supporting agents such as ratings agencies, auditors and credit bureaus (Glaessner and Oks, 1998).

From Chapter 3, past literature on the comparative efficiency of foreign banks and domestic banks shows conflicting conclusions. It has been found that foreign banks are more efficient than domestic banks in developing countries, while it is the other way round in developed countries (Claessens et al., 2001). Many foreign-owned banks in many developing/transition economies now strongly dominate the banking markets (Kasman and Yildirim, 2006).¹⁴⁸ As a result, it has often been argued that majority foreign-owned banks are likely to be more efficient than their domestic counterparts (Hasan and Marton, 2003; Bonin et al., 2005a; Berger et al., 2009). In addition, the alternative global advantage hypothesis argues that foreign institutions can be more efficient for a number of reasons (see Berger et al., 2005; Berger and Mester, 2003; Buch, 2003; Weill, 2004): (1) Foreign shareholders may contribute their superior managerial skills, high-quality human capital, best-practice policies and procedures, and sophisticated investment and risk management skills; (2) Foreign banks, generally being part of a large banking organisation, face the same scale economies and diseconomies as domestic banks; (3) Foreign banks normally serve profitable multinational customers; (4) Foreign banks have better access to capital markets, superior ability to diversify risks, and the ability to offer some services to multinational clients not easily provided by domestically owned banks; (5) Foreign-owned institutions from developed countries have access to superior information technologies for collecting and assessing 'hard' quantitative information; and (6) Foreign-owned banks may benefit from better control by private shareholders since these banks are mostly privately owned, resulting in more incentive for managers to operate efficiently (Berger et al., 2005; Berger and Mester, 2003; Buch, 2003; Weill, 2004).

¹⁴⁸ Foreign banks' potential advantages include superior managerial skills, high-quality human capital, lower-cost funds and adequate capital supply. (McCauley and Seth, 1992; Terrell, 1993; Berger et al., 2005).

The second environmental variable introduced in this research is the specialisation of banks (Z2). For this environmental variable, a dummy variable is assigned for two types of banks: 0 for conventional banks and 1 for Islamic banks.¹⁴⁹ Conventional and Islamic banks differ in terms of legal principles (traditional principles versus shariah principles), although in practice, the operations of the banking models are similar. Islamic banks have been actively imitating conventional banks in terms of products and services, but still following shariah principles (Beck et al., 2013, Chong and Liu, 2009; Khan, 2010).

Conventional banks undertake the practice of traditional commercial banking activities, such as accepting deposits and providing loans to customers. Thus, conventional banks incur interest expenditure and earned income from deposits and loans respectively. On the other hand, the key features of Islamic banking are the prohibition of interest payment transactions, and of financing anti-social and immoral activities such as gambling, pornography, and consumption of alcohol and narcotics. Islamic banks are also different from conventional banks, as the former have to follow the concept of shariah (Islamic principle), which employs the principles of justice, fair dealing and harmony, coupled with equitable wealth distribution as the basis of conducting business (Abdul Majid et al., 2011). In this case, Islamic banks are considered to be specialised banks that require specialised personnel to produce distinct services, different product mixes, and intense use of inputs. Hence, Islamic banks generally face substantially higher costs compared to conventional banks (Abdul Majid et al., 2011). Hassan (2005), Kamaruddin et al. (2008), Hamilton et al. (2010) and Abdul Majid et al. (2011) found that Islamic banks are less cost efficient than conventional banks due to higher costs incurred, particularly by the specialised personnel and complex business operations in a regulatory environment that does not support Islamic banking. Moreover, Islamic banking products and services are required to be packaged in accordance with Shariah principles and laws; hence, the costs associated with these products (e.g. personnel, legal, and administration costs) are higher than products and services offered by conventional banks. Therefore, hypothesis 3 discussed in Chapter 4 is valid to be tested.

¹⁴⁹ As mentioned earlier in Chapter 4, another type of banking that is not considered in this research is investment banks. This is due to the non-homogeneous nature of the sector's business models, operations and structures. Moreover, the financial services offered to customers are significantly different from those available in commercial banks.

The third and fourth environmental variables relate to the impact of the financial liberalisation phases implemented in the FSMP. To include these environmental factors, dummy variables are assigned to Z3 and Z4 to control for the liberalisation measures introduced in three phases of FSMP for the years 2000–2003, 2004–2007 and 2008–2011, respectively. As discussed in Chapter 4, there are three phases of FSMP in which different measures of financial liberalisations were introduced to Malaysian banks. In the first phase of the FSMP (2000–2003), the focus was more on building the capacity of the domestic banks, enhancing the financial infrastructure and introducing the consumer protection framework. These initial steps are taken to prepare domestic banks for a greater level of competition in the second phase of FSMP. There were several initiatives taken by BNM in this phase, such as (1) implementation of an electronic credit bureau database, (2) internet banking platform activities for foreign banks, (3) permission for banks to cross-sell products within the banking group, (4) simpler product approval processes, (5) introduction of best-practice management for credit risks, and (6) introduction of ten-year consumer education programme (Bank Negara Malaysia, 2001a, 2002, 2003; Khrishnasamy et al., 2004; Tahir et al., 2008).

The FSMP moved into the second phase (2004–2007) with a focus on increasing the degree of competition in the Malaysian banking industry. A number of liberalisation measures were introduced during this phase (see Chapter 4), for example (1) the introduction of a new interest rate framework, in which banks were allowed to set their own lending and deposit rates, (2) the approval of banking licences for three new foreign Islamic banks to operate in Malaysia, (3) authorisation for existing foreign banks to open additional branches, subject to several conditions (e.g. they had to include non-urban areas), (5) the introduction of a deposit insurance corporation (PIDM), and (6) the implementation of a ‘launch and file’ system for the product approvals process (Bank Negara Malaysia, 2004, 2005, 2006a, 2006b, 2007b). To capture the effect of liberalisation during the second phase of the FSMP, a dummy variable value of 1 is assigned for years 2004–2007 and 0 for other years.

Further financial liberalisation was introduced in the third phase of the FSMP (2008–2011) to increase the competitive conditions of the Malaysian banking industry. In this phase, Malaysia was affected by the global credit crisis and faced various challenges to further

liberalise the banking market. Nevertheless, BNM continued its financial liberalisation agenda by introducing several initiatives, such as (1) the Basel II regulatory framework, (2) six new banking licences for foreign banks, (3) permission for foreign banks to open new branches subject to approval, and (4) a financial sector blueprint for another ten-year development plan for Malaysian banks (Bank Negara Malaysia, 2008b, 2009b, 2010b, 2011b, 2011c). Similar to the second phase of the FSMP, a dummy variable equal to 1 for years 2008–2011 is assigned, otherwise 0.

The fifth environmental variable is the effect of the global credit crisis in Malaysia (Z5). As discussed in Chapter 3, to the author's knowledge, there are no studies examining the impact of the global financial crisis on the efficiency of Malaysian banks. Hence, a dummy variable is assigned for the years 2008–2010 with a value of 1 and 0 for the other years. From Chapter 4, although the credit crisis began in the USA in 2007 (International Monetary Fund, 2007), the crisis really affected Malaysia in the second half of 2008. The impact of the global recession was felt most strongly and recorded a heavy decline in exports in the second half of 2008. The effect of the decline in the trade flow via the real economy affected output, employment and private investment and consumption activities (Tan, 2011). As a result, the Malaysian economy contracted by 1.7% in the full year of 2009 (Bank Negara Malaysia, 2009a). Between 2008 and 2009, to mitigate the impact of the economic and financial crisis on domestic demand and accommodate monetary policy, BNM reduced OPR from 3.50% in November 2008 to 2.00% on 25 February 2009 (Bank Negara Malaysia 2009a). The Malaysian government also introduced a couple of stimulus packages between 2008 and 2009 to revive the economy. The accelerated implementation of fiscal stimulus measures, the easing of monetary policy and improved access to financing stabilised the economy, allowing Malaysia to recover in the second half of 2009, particularly after a sharp contraction in the first half of 2009 (Bank Negara Malaysia, 2009a). The Malaysian economy expanded by 7.2% in 2010. This growth was driven by strong domestic demand and expansion in private sector activity, as well as strong support from

public sector programmes on infrastructure and delivery systems (Bank Negara Malaysia, 2010a).¹⁵⁰

From the discussion above, the Malaysian economy felt the effect of global crisis in 2008; in 2010 however, the Malaysian economy recovered as evidenced by the raising of OPR (policy interest rate) by BNM in the second half of 2010. Therefore, for this analysis, the years 2008–2010 are considered to be the time that Malaysia was most affected by the worldwide credit crisis.

The sixth environmental variable included in the analysis of efficiency of Malaysian banks is the size of banks (Z6) (measured via total assets). From past literature, the total assets variable is used to control for bank size (Evanoff and Israilevich, 1991; Casu and Girardone, 2002). Bank size is employed in this study to control for potential scale biases in the estimating process. It is also noted that bank size is an important determinant of net interest margins and spreads if there are economies of scale in the Malaysian banking sector. In other words, one bank may be more efficient than another as a result of the economies of scale that arise from size rather than because of better management (Casu and Girardone, 2002). This can be evidenced from Beck and Hesse (2006), who argue that small banks are less able to diversify risks, which could result in a higher risk premium in the interest rates at which they can borrow. On the other hand, larger banks may have a greater number of professional management teams that could be more cost conscious due to greater pressure from owners concerned with profit maximisation (Evanoff and Israilevich, 1991).¹⁵¹ Therefore, this research analyses the impact of size on banking efficiency in Malaysian banks, which in turn provides useful information for regulators and allows bank managers to evaluate the optimal scale at which to conduct their operations. In this analysis, the logarithm of total assets is used as a proxy for bank size (instead of using a dummy for different size categories: i.e. small, medium and big). The advantage of employing this method is to capture the effects of scale on cost (profit) efficiency while avoiding

¹⁵⁰ In line with the positive growth in 2010, the OPR was adjusted upwards to manage the potential risk of financial imbalances. BNM normalised the OPR by raising it three times from 2.00% in March to 2.75% in July 2010 (Bank Negara Malaysia, 2010a).

¹⁵¹ Berger et al. (1993) state that larger banks can reduce their costs from the output side, in order to reach their optimal mix and scale of outputs. That is, larger banks may reap efficiency benefits from economies of scale or/and scope (Casu and Girardone, 2006).

misspecification from inappropriate break points when artificially dividing the range of banks into different group sizes (Beck and Hesse, 2006).

The final two environmental variables that are employed in the present research relate to the market structure. To capture the effect of the market structure along the lines of the structure–conduct–performance paradigm on the efficiency of Malaysian banks, the market concentration (i.e. HHI) (Z7) and market share (Z8) variables are used to capture market concentration and market share (Molyneux et al., 1996). Based on the discussion in Chapter 2, an oligopoly market structure suggests that there is a positive relationship between concentration and profitability. The degree of concentration, as an indicator of market structure, may influence a bank's profitability and efficiency. The level of concentration in the banking sector is usually measured using the Herfindahl–Hirschman index (HHI), which proxies the bank's market power or the intensity of competition among banks. A higher value of HHI would indicate a more concentrated banking market (a lower degree of competition).

As discussed in Chapter 2, there are two hypotheses regarding the relationship between market structure and the efficiency of banks: the market power hypothesis and the efficient structure hypothesis. The market power hypothesis asserts that only banks with large market share and well differentiated products are able to exercise market power and earn supernormal profits (Shepherd, 1982; Berger, 1995). In this case, banks may find less pressure to control their costs and might thus enjoy the 'quiet life' (Berger and Mester, 1997). Banks may utilise their own market power through size (Berger, 1995); hence a market share variable (Z8) is included in this analysis to control for 'relative market power' hypothesis (Berger, 1995). On the other hand, Demsetz (1973) iterates that the relationship between market concentration and efficiency could be positive. A greater market concentration may result from superior production efficiency rather than from the 'quiet life' (Berger, 1995). In other words, relatively efficient banks with lower costs can compete aggressively, maximise profits and consequently, gain a bigger market share, which supports the efficient structure hypothesis (see Chapter 2). Therefore, from the discussions above, two environmental variables (Z7 and Z8) are employed in this research in order to examine the relationship between the market structure and cost (profit) efficiency of

Malaysian banks: market concentration and market share. Market concentration (HHI) is defined as the sum of squared asset market shares of all Malaysian banks. On the other hand, market share is calculated as a share of bank deposits in relation to the total deposits of all banks (Berger and Mester, 1997). Both the HHI and market share variables are included in the specification of cost- and profit-efficiency because the HHI is an aggregate measure that only changes over time; while the market share variable differs from bank to bank and over time.

Table 5.3 Descriptive Statistics for Environmental Variables, 2000–2011

Variables	Descriptions	Mean	Std Dev.	Min.	Max.	
Z1	Ownership	Foreign = 1, domestic = 0 (dummy)	0.47	0.50	0.00	1.00
Z2	Specialisation	Islamic banks = 1, conventional banks = 0 (dummy)	0.25	0.43	0.00	1.00
Z3	Deregulation FSMP phase 2	Dummy for 2004 – 2007 = 1, otherwise = 0	0.31	0.46	0.00	1.00
Z4	Deregulation FSMP phase 3	Dummy for 2008 – 2011 = 1, otherwise = 0	0.40	0.49	0.00	1.00
Z5	Effect of global financial crisis on Malaysia	Dummy for 2008 – 2010 = 1, otherwise = 0	0.31	0.46	0.00	1.00
Z6	Bank size (MYR million)	Total assets	32,320.00	41,679.13	721.17	229,504.17
Z7	Market concentration (%)	Herfindahl–Hirschman index (concentration of banks' assets)	7.79	0.81	6.58	8.73
Z8	Market share (%)	Bank deposits/total banking deposits	3.39	3.98	0.08	19.69

5.4 Stochastic Frontier Model(s) Specification

To compute the cost- and profit-efficiency of Malaysian banks, coupled with the identification of dependent and independent variables for SFA models, discussed in the earlier section, a functional form has to be specified. For this study, a widely used transcendental logarithm or translog functional form¹⁵² is employed for the cost function.¹⁵³ The estimation of cost-efficiency can be explicitly specified as:

¹⁵² As discussed in Chapter 2, the translog function may not fit well with data far from the mean of output size. Thus, the FF functional form overcomes this problem with better approximation across a broad range of

$$\begin{aligned}
\ln\left(\frac{TC_{it}}{w_3}\right) &= \alpha_0 + \sum_{n=1}^2 \beta_n \ln\left(\frac{w_{nit}}{w_3}\right) + \sum_{k=1}^3 \delta_k \ln y_{kit} + \frac{1}{2} \left[\sum_{n=1}^2 \sum_{j=1}^2 \beta_{nj} \ln\left(\frac{w_{nit}}{w_3}\right) \ln\left(\frac{w_{jit}}{w_3}\right) \right] \\
&+ \frac{1}{2} \left[\sum_{k=1}^3 \sum_{m=1}^3 \delta_{km} \ln y_{kit} \ln y_{mit} \right] + \sum_{n=1}^2 \sum_{k=1}^3 \eta_{nk} \ln\left(\frac{w_{nit}}{w_3}\right) \ln(y_{kit}) \\
&+ \sum_{r=1}^3 \chi_r \ln c_{rit} + \frac{1}{2} \left[\sum_{r=1}^3 \sum_{h=1}^3 \chi_{rh} \ln c_{rit} \ln c_{hit} \right] + \sum_{n=1}^2 \sum_{r=1}^3 \rho_{nr} \ln\left(\frac{w_{nit}}{w_3}\right) \ln(c_{rit}) \quad (5.29) \\
&+ \sum_{k=1}^3 \sum_{r=1}^3 \omega_{kr} \ln y_{kit} \ln c_{rit} + \tau_1 \ln T + \frac{1}{2} \tau_{11} \ln T^2 + \sum_{k=1}^2 \varphi_{1n} \ln T \ln\left(\frac{w_{nit}}{w_3}\right) \\
&+ \sum_{k=1}^3 \kappa_{1k} \ln T \ln y_{kit} + \sum_{r=1}^3 \xi_{1r} \ln T \ln c_{rit} + u_{it} + v_{it}
\end{aligned}$$

where $\ln TC_{it}$ is the natural logarithm of total cost of i -th bank at time t ; $\ln y_{kit}$ is the natural logarithm of the outputs for the i -th bank at time t where; y_1, y_2 and y_3 are loans, investments and other earning assets respectively; $\ln w_{nit}$ is the natural logarithm for the input price vectors for i -th bank at time t where; w_1 is the price of funds, w_2 is the price of labour and w_3 is the price of physical capital; $\ln T$ is the natural logarithm of time trend ($T=1$ for 2000, $T=2$ for 2001, ..., $T=12$ for 2011); $\ln c_{rit}$ are the natural logarithm of control variables imposed for the i -th bank at time t where the control variables consist of capital (c_1), asset quality (c_2) liquidity (c_3) and time (T); u_{it} is inefficiency and is assumed to be one-sided and independently distributed from v_{it} ; v_{it} is a random error and assumed to be two-sided with normally distributed $v_{it} \sim N(0, \sigma_v^2)$ and independent from u_{it} ; $\alpha, \beta, \delta, \eta, \chi, \rho, \omega, \tau, \varphi, \kappa$ and ξ are parameters to be estimated.

outputs by incorporating additional parameters for the Fourier trigonometric terms (Girardone et al., 2004). However, additional parameters and trigonometric terms can give rise to a couple of problems. First, with trigonometric terms, for the FF function to hold it requires a larger sample of observations (i.e. a higher degree of freedom) and, second, the additional parameters can give rise to multicollinearity issues in estimating the parameters of the model specification (Chambers, 1988; Coelli et al., 2005, Fu and Haffernan, 2007). Given that 346 observations were employed for this study, should FF functional form be employed, it could lead the loss of a degree of freedom and would result in biased efficiency scores.

¹⁵³ For standard profit- and alternative profit-efficiency, the functional form is similar to the cost function using the translog functional form. The difference in the standard profit and the alternative profit model specification is that the dependent variable becomes total profit (TP) and composite error of $\varepsilon_{it} = v_{it} - u_{it}$. Additionally, the output variables for standard profit-efficiency are output prices (p_i) rather than output quantities (y_i).

In Coelli et al. (2005), it is mentioned that cost has to be linearly homogeneous in input prices (i.e. an increase in input prices must result in an increase in the total cost variable). Therefore, the TC and input price variables are normalised by input price, w_3 . The normalisation imposes a theoretical condition that the model is linearly homogeneous (Lang and Welzel, 1996). Hence, the following restrictions should be applied to the parameters of cost function in equation (5.29); $\sum_{n=1}^3 \beta_n = 1$; $\sum_{n=1}^3 \beta_{nj} = 0$; $\sum_{n=1}^3 \theta_{ln} = 0$; $\sum_{n=1}^3 \eta_{nk} = 0$; $\sum_{n=1}^3 \rho_{nr} = 0$. Furthermore, the second order parameters of the model of standard symmetry restrictions $\delta_{km} = \delta_{mk}$ and $\beta_{jn} = \beta_{nj}$ are applied to the above cost function. Doing so is consistent with the duality theorem when normalisation is imposed on input prices and total costs, coupled with restrictions on second order parameters (Lang and Welzel, 1996).

For this study, to estimate the cost and profit frontiers, the control variables c_{rit} are to be included and specified directly into the model, similar to the approach taken with the input and output variables. By implementing this approach, the control variables will interact with both input prices and output and influence the shape of the frontier, as displayed in equation (5.29) (Kumbhakar and Lovell, 2000).

An additional way to account for exogenous factors on banks' performance is to include environmental variables. Unlike control variables, environmental variables do not affect the cost and profit frontiers but influence the degree of cost-inefficiency. Following Battese and Coelli (1995), the environmental variables will be included in the stochastic component of the cost and profit functions.

This approach where the environmental variables will influence only the distance of each bank from the best-practice cost function, is also called 1-stage analysis. In other words, this approach of handling environmental variables will define the inefficiency (u_i), which includes non-negative random variables and independently distributed as truncations at 0 of $N(m_{it}, \sigma_u^2)$ distribution, and can be written as (Battese and Coelli, 1995):

$$m_{it} = z_{it}\psi \quad (5.30)$$

where m_{it} is the mean inefficiency of cost and profit function that was obtained by truncation (at 0) of normal distribution with mean ψZ_{it} . The vector z_{it} is for environmental variables containing numerical values that could influence the inefficiency of i -th bank at time t ; and ψ is the vector comprising coefficients that need to be estimated. In measuring the cost- and profit-efficiency, the inefficiency effects are estimated simultaneously with the cost- and profit-function, which can be specified as:

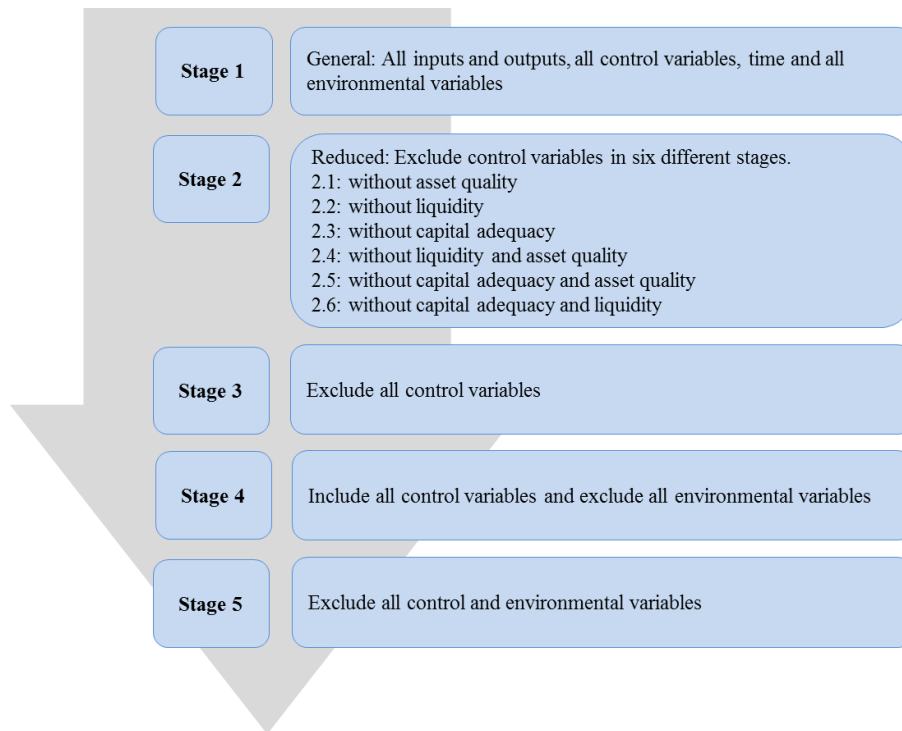
$$m_{it} = \psi_0 + \psi_1 z_1 + \psi_1 z_1 + \psi_2 z_2 + \psi_3 z_3 + \psi_4 z_4 + \psi_5 z_5 + \psi_6 z_6 + \psi_7 z_7 + \psi_8 z_8 \quad (5.31)$$

where z_1 is ownership, z_2 is bank specialisation, z_3 is dummy variables for the second phase of FSMP (years 2004-2007), z_4 is dummy variables for the third phase FSMP (years 2008-2011), z_5 is dummy variables for global crisis impact on Malaysia, z_6 is natural log asset size, z_7 is natural log market concentration and z_8 is natural log market share.¹⁵⁴

5.4.1 Performing 1-Stage Analysis

In measuring the cost- and profit-efficiency of Malaysian banks, three models are estimated: the cost-efficiency model, the standard profit-efficiency model and the alternative profit model. A 1-stage approach will be employed to derive each cost- and profit-efficiency model. In deriving a preferred model, a number of models will be estimated for the Malaysian banks and from of these models, only one preferred model will be chosen for each cost, standard profit and alternative profit model.

¹⁵⁴ Asset size, market concentration and market share are taking logarithm form and there were no significant differences between the results of logged value of Z_6 , Z_7 and Z_8 and raw value of Z_6 , Z_7 and Z_8 .

Figure 5.4 Stages Performed in Deriving Preferred Model

As shown in Figure 5.4 and Table 5.4, there are basically five main stages to be tested in choosing the best of each of the cost- and profit-efficiency models. Each stage involving a reduced model has its own null hypothesis, which is to test whether the reduced model is better than the general model estimated at Stage 1. Stage 1 is a general model, which includes all inputs and outputs and all control variables (C1–C3) and all environmental variables (Z1–Z8), which will be compared at different stages with different models, consisting various combinations of the control and environmental variables.

In essence, using the null hypothesis at different stages, the reduced models are compared against the general model to determine whether the reduced model performs better than the general model. The preferred model is selected based on the log-likelihood ratio test (discussed later in this chapter).

Table 5.4 Stages Involved in Searching for Preferred Model

Stage	Model	Inputs	Outputs	Time	Control variables	Environmental variables
A1.0	General	W1, W2, W3	Y1, Y2, Y3	T	C1, C2, C3	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.1	Reduced – without C3	W1, W2, W3	Y1, Y2, Y3	T	C1, C2	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.2	Reduced – without C2	W1, W2, W3	Y1, Y2, Y3	T	C1, C3	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.3	Reduced – without C1	W1, W2, W3	Y1, Y2, Y3	T	C2, C3	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.4	Reduced – without C2, C3	W1, W2, W3	Y1, Y2, Y3	T	C1	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.5	Reduced – without C1, C3	W1, W2, W3	Y1, Y2, Y3	T	C2	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A2.6	Reduced – without C1, C2	W1, W2, W3	Y1, Y2, Y3	T	C3	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A3.0	Without C1, C2, C3	W1, W2, W3	Y1, Y2, Y3	T	–	Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8
A4.0	Without Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8	W1, W2, W3	Y1, Y2, Y3	T	C1, C2, C3	–
A5.0	Without C1, C2, C3 and Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8	W1, W2, W3	Y1, Y2, Y3	T	–	–

Notes: W1: Price of funds, W2: Price of labour, W3: Price of physical capital, Y1: Loans, Y2: Investments, Y3: Other earnings assets, T: time trend, C1: capital ratio, C2: liquidity, C3: asset quality, Z1: ownership structure, Z2: specialisation, Z3: deregulation period, phase 2 of FSMP (2004–2007), Z4: deregulation period, phase 3 of FSMP (2008–2011), Z5: bank size, Z6: effect of global credit crisis on Malaysia (2008–2010), Z7: market concentration, Z8: market share

Table 5.4 exhibits in detail a series of combinations of control and environmental variables that are reduced and added to the models in five different stages. In determining the preferred model, several validity tests need to be made on these different estimations at different stages. These tests are shown in Tables 5.5, 5.6 and 5.7.¹⁵⁵

¹⁵⁵ Tables 5.6 and 5.7 display the results of estimation of different models for standard profit-efficiency (B1.0–B5.0) and alternative profit-efficiency (C1.0–C5.0), respectively. These models follow the same procedure as the cost-efficiency model (A1.0–A5.0) to derive the efficiency estimations. Hence, only summarised estimation results of standard profit- and alternative profit-efficiency are shown in Tables 5.6 and 5.7.

5.4.2 Identifying the Preferred Model

Stage A1.0: The General Cost Model

This stage estimates the general cost model using the stochastic frontier approach. The specification of this model includes input prices (w_n), outputs (y_k) and control variables (c_r). In addition, for the general model, using the Battese and Coelli (1995) 1-stage approach in the estimation of cost-efficiency, a set of environmental variables (Z_i) is included in the stochastic component of the cost function. In other words, the cost inefficiency (U_i) in this model is the function of a set of environmental variables mentioned earlier. Hence, to estimate the cost-efficiency of this general model, the specification in the equation 5.29 and 5.30 are used.

After the general model is estimated, the first thing to do is to check for the existence of inefficiency.¹⁵⁶ A systematic way of checking for the existence of inefficiency within the random error (ε) is exhibited in Table 5.5, which exhibits an estimated general cost model for Malaysian banks. From Table 5.5, the first six columns ((1) to (6)) are used to test the existence of inefficiency in the models estimated. Additionally, columns (7) to (10) are used to choose the preferred model. Column (1) is the maximum log-likelihood estimate for the model. Column (2) is the log-likelihood ratio test or LR test of maximum log-likelihood estimate for OLS model and maximum log-likelihood value for stochastic model. The LR ratio is the key measure that determines the existence of inefficiency in the composite error. For the purpose of testing this hypothesis, the LR test is derived by employing a three-step method performed in Coelli (1996). In the first step, the process includes the estimation of parameters and log-likelihood value using the OLS estimates. Accordingly, in the second step, the skewness of the joint distribution of u and v is examined. This test determines the level of domination of inefficiency (u) or random error (v) in the total random error (ε). In the third step, the log-likelihood ratio of the stochastic model is computed should there be inefficiency (u) in the random error term (ε). The LR test can be computed as shown in equation 5.14, which indicates the differences in the OLS model against the stochastic

¹⁵⁶ As discussed earlier in this chapter, the existence of inefficiency can be checked from composite error term (ε), which can be disentangled into random error (v) and inefficiency (u).

model. If the LR test exceeds the critical value of the chi-squared distribution, which in this case is at 5% at 1 degree of freedom, which is at 3.84, the null hypothesis is rejected, which implies the presence of inefficiency in the stochastic model (unrestricted model). On the other hand, if the LR test is lower than 3.84, the null hypothesis is not rejected, indicating that no inefficiency exists and the model could not follow the stochastic path.

Column (3) is another test to observe the existence of inefficiency. Using the Battese and Corra (1977) parameter γ to measure the deviations of composite random error (ε), one is able to test the presence of inefficiency in the stochastic model. The parameter γ can be measured as $\frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2}$ as shown in equation (5.11), where σ_u^2 is the variance of inefficiency and, σ_v^2 is the variance of the random error. γ indicates the deviations of random error whether its influence by inefficiency (u) or noise (v). If the γ parameter is approaching 1, the composite error (ε) is dominated by inefficiency (u). On the other hand if γ is approaching 0, the composite error (ε) is influenced by noise or random error (v). For the validity test of the existence of inefficiency, the null hypothesis to be tested is, $H_0: \gamma = 0$, which indicates there is no inefficiency (u) in the composite error (ε). The alternative hypothesis is $H_1: \gamma \neq 0$, which implies the existence of inefficiency (u) in the composite error (ε) (see Figure 5.7).

Column (4) indicates the number of restrictions made in the estimation. The number of restriction is one (i.e. $\sigma_u = 0$) where there is no inefficiency (u) and the estimates only consist of random error (v). With this one restriction, column (5) shows the value of 1 degree of freedom at 5%. Column (6) on the other hand determines the presence of inefficiency using the parameter γ and the LR test. For instance, Table 5.5 displays that the LR of the general model exceeds the 5% critical value at 1 degree of freedom, which indicates that inefficiency is present in the general model (A1.0)

Next, reduced models are estimated in several stages (see Tables 5.4 and 5.5), including the test of existence of inefficiency in each reduced model as shown in columns (1) to (6). In these stages, a number of null and alternative hypotheses are examined regarding the estimation of the various reduced models against the general model.

Stage A2.0: General Cost Model with Reduced Control Variables (C1, C2, and C3) and with Environmental Variables (Z1–Z8)

At this stage, based on the three control variables (i.e. capital adequacy (C1), asset quality (C2) and liquidity (C3)) being introduced in this study, the following models were estimated:

- Model A2.1 General cost model with reduced control variables – without liquidity (C3) but with environmental variables (Z1–Z8)
- Model A2.2 General cost model with reduced control variables – without asset quality (C2) but with environmental variables (Z1–Z8)
- Model A2.3 General cost model with reduced control variables – without capital adequacy (C1) but with environmental variables (Z1–Z8)
- Model A2.4 General cost model with capital adequacy (C1) – without asset quality (C2) and liquidity (C3) – and with environmental variables (Z1–Z8)
- Model A2.5 General cost model with asset quality (C2) – without capital adequacy (C1) and liquidity (C3) – and with environmental variables (Z1–Z8)
- Model A2.6 General cost model with liquidity (C3) – without capital adequacy (C1) and asset quality (C2) – and with environmental variables (Z1–Z8)

For the models shown above, six null hypotheses are to be tested using LR test:

- H1₀ The general cost model specification without liquidity (C3) is better statistical fit according to the LR test than the general cost model.
- H2₀ The general cost model specification without asset quality (C2) is better statistical fit according to the LR test than the general cost model.
- H3₀ The general cost model specification without capital adequacy (C1) is better statistical fit according to the LR test than the general cost model.
- H4₀ The general cost model specification without asset quality (C2) and liquidity (C3) is better statistical fit according to the LR test than the general cost model.
- H5₀ The general cost model specification without capital adequacy (C1) and liquidity (C3) is better statistical fit according to the LR test than the general

cost model.

- H6₀ The general cost model specification without capital adequacy (C1) and asset quality (C2) is better statistical fit according to the LR test than the general cost model.

The alternative hypotheses on the other hand can be specified as follows:

- H1₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without liquidity (C3) (A2.1).
- H2₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without asset quality (C2) (A2.2).
- H3₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without capital adequacy (C3) (A2.3).
- H4₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without asset quality (C2) and liquidity (C3) (A2.4).
- H5₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without capital adequacy (C1) and liquidity (C3) (A2.5).
- H6₁ The general cost model is better statistical fit according to the LR test than the general cost model specification without capital adequacy (C1) and asset quality (C2) (A2.6).

Stage A3.0: General Cost Model without Control Variables (C1–C3) but with Environmental Variables (Z1–Z8)

In this model specification (A3.0), the control variables are excluded (C1–C3) but the environmental variables (Z1–Z8) remain to influence the inefficiency (u_i) of the banks.

- Model A3.0 General cost model without the control variables – without capital adequacy (C1), asset quality (C2) and liquidity (C3) – but with the environmental variables (Z1–Z8)

For this model, the null and alternative hypotheses are:

- H7₀ The general cost model specification without the control variables (A3.0) is better statistical fit according to the LR test than the general cost model (A1.0).
- H7₁ The general cost model (A1.0) is better statistical fit according to the LR test than the general cost model specification without control variables (A3.0).

Stage A4.0: General Cost Model with Control Variables (C1–C3) but without Environmental Variables (Z1–Z8)

For stage A4.0, the environmental variables are excluded from the specification of the general cost-efficiency model. In this case, the Malaysian banks are assumed not to face any exogenous factors that could affect the efficiency of Malaysian banks, but variations in the banks' production are controlled using capital adequacy (C1), asset quality (C2) and liquidity (C3). Hence, the specification of the model at stage 4.0 is shown as:

Model A4.0 General cost model with all the control variables (C1–C3) but without the environmental variables (Z1–Z8)

Thus, the null and alternative hypotheses to be examined are:

- H8₀ The general cost model specification without the environmental variables (A4.0) is better statistical fit according to the LR test than the general cost model (A1.0).
- H8₁ The general cost model (A1.0) is better statistical fit according to the LR test than the general cost model specification without the environmental variables (A 4.0).

Stage A5.0: General Cost Model without Control Variables (C1–C3) and without Environmental Variables (Z1–Z8)

Stage A5.0 has fewer variables in its specification of cost frontiers, and does not account for the effect of heterogeneity. The general cost model without control variables (C1–C3) and

environmental variables ($Z1-Z8$) and is left with only outputs (y_{kit}) and input prices (w_{nit}). For this stage, it is assumed that all Malaysian banks possess the same production technology and face the same environmental conditions. Thus the model specification for this stage can be shown as:

Model A5.0 General cost model without any control variables (C1–C3) and without any environmental variables (Z1–Z8)

The null and alternative hypotheses at this stage are:

- H9₀ The general cost model specification without the control variables or the environmental variables (A5.0) is better statistical fit according to the LR test than the general cost model (A1.0).
- H9₁ The general cost model (A1.0) is better statistical fit according to the LR test than the general cost model specification without the control variables or the environmental variables (A 5.0).

After testing all the models for the presence of inefficiency, the next stage is to find the preferred model based on the null hypotheses developed at each stage. In doing so, the process starts from column (7), which shows the LR test value for the reduced model in comparison to the general model. Coelli (1996) suggests using the log-likelihood ratio test to determine the best model among various competing models estimates. Coelli et al. (2005) and Kumbhakar and Lovell (2000) further suggest that the log-likelihood ratio test provides a convenient process for testing the hypothesis of different competing models. The LR test is employed by providing a test of whether reduced models offer a better fit than the general model. It also provides an indication of whether the parameters of the reduced model are significantly different from the parameters of the general model. The LR test for the reduced models can be written as:

$$LR = -2[\ln L_{reduced} - \ln L_{general}] \quad (5.32)$$

where $\ln L_{reduced}$ is the log-likelihood for the reduced model and $\ln L_{general}$ is the log-likelihood for the general model. LR statistic is asymptotically equivalent to the chi-squared

(χ^2) distributions (Coelli, 1995). Based on null hypothesis of the restricted model (i.e. the OLS model), it has a limiting chi-squared distribution with degrees of freedom equal to the number of restrictions being tested (Greene, 2002).

The LR test for a reduced model is employed to test the hypothesis at every stage of model estimations. The hypothesis is used to determine whether the reduced model is a better fit compared to the general model (Coelli, 1996). At every stage of the null hypothesis H_0 ;, the model estimation with a lower number of control and environmental variables provides a better estimation than the general model. To test this hypothesis, the LR test of a reduced model is used. The result of the LR test is then compared against the 5% critical value of χ^2 distribution at the given degrees of freedom. The 5% critical value of χ^2 distribution at the given degrees of freedom is indicated in column (9) based on the degrees of freedom shown in column (8).

If the value of the LR test exceeds 5%, the critical value of χ^2 distribution at the given degrees of freedom, the null hypothesis that the reduced model provides a better estimation is rejected. Conversely, if the LR test's value is less than the 5% critical value of χ^2 distribution at the given degrees of freedom, the null hypothesis is not rejected, implying that the reduced model provides a better estimation than the general model. Consequently, the selected preferred models are those estimated models that fail to reject the null hypothesis (shown in column (10)). Should there be several models that fail to reject the null hypotheses, the most preferred model would be based on the highest maximum log-likelihood value (Coelli et al., 2005).

Table 5.5 lists the stages and processes taken to determine the preferred model. From the test results, it was found that the effect of heterogeneity could significantly influence the stochastic cost frontier estimation. Therefore, heterogeneity effects should be considered by including control and environmental variables. Without control and environmental variables, the estimation of cost-efficiency could be mis-specified, which could result in an inappropriate estimation of the parameter and cost-efficiency (Coelli et al., 1999). Hence, this research does not only focus on the efficiency level but also examines the sources or determinants of cost-inefficiencies. From Table 5.5, the preferred cost model is the general

model without capital adequacy (C3) (A2.3). The reduced model is found to be a better-fitting model due to an LR value of less than 5% critical value of χ^2 distribution at seven degrees of freedom. The result fails to reject the null hypothesis at stage A1.3, which implies that the reduced model without capital adequacy (C3) is a better estimation than the general model. Tables 5.6 and 5.7 report the same procedures at different stages applied to the standard profit- and the alternative profit-efficiency models. For the standard profit-efficiency model, the preferred model is the general model (B1.0): a model consisting of all the control variables (C1–C3) and all the environmental variables (Z1–Z8). For the alternative profit-efficiency model, the preferred model is C1.0, which is also a general model that includes all the control variables (C1–C3) and all the environmental variables (Z1–Z8). Following Coelli (1996), reduced models at different stages were tested based on null hypotheses, although all the null hypotheses at stages 2.0 to 5.0 (B2.0–A50 and C2.0–C5.0) were rejected. Greene (2008) argued that when more variables are introduced into the functional form (i.e. control variables), less variation is left to be included in the error term. Additionally, when exogenous variables enter the mean of u (i.e. environmental variables), the estimation of efficiency changes significantly, not only in the score but also in the rank. Hence the introduction of heterogeneity factors into the general model (B1.0 and C1.0), which affect the profit specifications with more regressors than other models might result in a different model fit (Greene, 2008).

In addition to the above procedures, several model variants were also tested. For example, a model without a time trend variable (T) and the Battese and Coelli (1992) time-invariant model were also tested in this analysis. A null hypothesis was developed for the model without a time trend variable (T) against the general model. From the analysis, the null hypothesis was rejected, which demonstrated that the general model is better than a model without a time trend variable (T). Additionally, the Battese and Coelli (1992) time-invariant model was also analysed to see the impact of assumption on cost- and profit-efficiency being constant through time, allowing some degree of flexibility in the distribution of inefficiency (u), including using a truncated or half-normal distribution on the stochastic component of the production efficiency function (Coelli et al., 2005). Battese and Coelli (1995) state that, without the time-varying effect, the Battese and Coelli (1992) model does not allow change in the rank ordering of firms over time; for instance, a firm that is ranked

i-th at the first period is ranked *i*-th in all periods in the sample (Coelli et al., 2005). Battese and Coelli (1995) argue that the Battese and Coelli (1992) model could not consider firms' efficiency change, and also fails to account for any environmental effects in their model. Therefore, the Battese and Coelli (1995) model was introduced to overcome the issues arising from the Battese and Coelli (1992) model by explicitly expressing technical inefficiency effects in terms of appropriate explanatory variables, where the parameters of the stochastic frontier and the inefficiency model are then estimated simultaneously, using the ML method from panel data (Coelli et al., 2005). However, Battese and Coelli (1995) and Battese and Coelli (1992) are not comparable as the methods and statistical tests needed to compare them are not available.¹⁵⁷ Hence, the efficiency scores resulting from a Battese and Coelli (1992) analysis of Malaysian banking are not shown in this research.

¹⁵⁷ These two model specifications are non-nested and hence, no set of restrictions can be defined to allow a test of one specification versus the other (Coelli et al., 2005).

Table 5.5 Estimated Cost Frontier Models

Models	(1) Log-likelihood	(2) Log-likelihood ratio against OLS model	(3) γ	(4) Number of restrictions	(5) χ^2 critical value at 5%	(6) Presence of inefficiency $\gamma \neq 0$	(7) Log-likelihood ratio against general model	(8) Degree of freedom	(9) χ^2 critical value at 5%	(10) $H_0 =$ reduced model is better than general model
A1.0: General – W1, W2, W3, Y1, Y2, Y3, T, C1, C2, C3, Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8	54.62	133.98	0.87	1	3.84	yes				
A2.1: Reduced without liquidity (C3)	28.24	133.54	0.88	1	3.84	Yes	52.75	10	18.31	Reject
A2.2: Reduced without asset quality (C2)	25.28	161.01	0.99	1	3.84	Yes	58.68	10	18.31	Reject
A2.3: Reduced without capital (C1)*	46.02	150.65	0.89	1	3.84	Yes	17.19	10	18.31	Fail to Reject
A2.4: Reduced without asset quality or liquidity (C2 and C3)	-19.92	116.50	0.90	1	3.84	Yes	149.08	19	30.14	Reject
A2.5: Reduced without capital and liquidity (C1 and C3)	16.57	142.26	0.99	1	3.84	Yes	76.09	19	30.14	Reject
A2.5: Reduced without capital and asset quality (C1 and C2)	13.39	163.38	0.99	1	3.84	Yes	82.44	19	30.14	Reject
A3.0: Without control variables (C1, C2, C3)	-16.95	160.90	0.99	1	3.84	Yes	143.14	27	40.11	Reject
A4.0: Without environmental variables (Z1–Z8)	7.19	194.01	0.99	1	3.84	Yes	94.86	1	3.84	Reject
A5.0: Without control variables (C1, C2, C3) or environmental variables (Z1–Z8)	-88.90	208.50	0.99	1	3.84	Yes	287.03	25	37.65	Reject

Notes: (1) Log-likelihood: The maximum log-likelihood value from the unknown parameters being estimated based on given data (2) LR against OLS model: LR test based on ratio of log-likelihood of stochastic model subtracted from log-likelihood of OLS model and multiplied by (-2). (3) γ : Lies between 0 and 1. Approaching 0, composite error (ε) dominated by random error (v). Approaching 1, composite error (ε) influenced by inefficiency (u). (4) Number of restrictions: Restrictions on the composite error where it consists only of random error (v) and no inefficiency (u). One restriction is made on $\sigma_u = 0$. (5) χ^2 critical value at 5%: 5% critical value of chi-squared distribution at 1 degree of freedom is 3.84. (6) Presence of inefficiency: If $\gamma \neq 0$; inefficiency exists. If $\gamma = 0$; there is no inefficiency. (7) LR against general model: LR test based on the ratio of log-likelihood of reduced model subtracted from log-likelihood of general model and multiplied by (-2). (8) Degree of freedom: The difference between the numbers of parameters of the general model subtracted against the number of parameters of the reduced model. (9) χ^2 critical value at 5% , the chi-squared distribution value based on the degree of freedom derived from column (8). (10) $H_0 =$ reduced model better than general model: H_0 is failed to be rejected if the LR value is less than the chi-squared distribution at 5% critical value. Reject H_0 if the LR value exceeds the chi-squared distribution at 5% critical value. Columns (1) to (6) are to test for inefficiency and columns (7) to (10) are for choosing preferred model. * - Preferred cost frontier model (A2.3).

Table 5.6 Estimated Standard Profit Frontier Models

Models	(1) Log-likelihood	(2) Log-likelihood ratio against OLS model	(3) γ	(4) Number of restrictions	(5) χ^2 critical value at 5%	(6) Presence of inefficiency $\gamma \neq 0$	(7) Log-likelihood ratio against general model	(8) Degree of freedom	(9) χ^2 critical value at 5%	(10) $H_0 =$ reduced model is better than general model
B1.0: General – W1, W2, W3, P1, P2, P3, T, C1, C2, C3, Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8*	-332.48	120.68	0.00	1	3.84	yes				
B2.1: Reduced without liquidity (C3)	-352.85	132.82	0.00	1	3.84	Yes	40.74	9	16.92	Reject
B2.2: Reduced without asset quality (C2)	-354.37	101.03	0.00	1	3.84	Yes	43.78	9	16.92	Reject
B2.3: Reduced without capital (C1)	-355.68	191.36	0.44	1	3.84	Yes	46.41	9	16.92	Reject
B2.4: Reduced without asset quality or liquidity (C2 and C3)	-373.25	124.10	0.66	1	3.84	Yes	81.53	17	27.59	Reject
B2.5: Reduced without capital or liquidity (C1 and C3)	-372.41	218.01	0.48	1	3.84	Yes	79.87	17	27.59	Reject
B2.5: Reduced without capital or asset quality (C1 and C2)	-372.29	189.15	0.05	1	3.84	Yes	79.62	17	27.59	Reject
B3.0: Without control variables (C1, C2, C2)	-387.21	215.34	0.47	1	3.84	Yes	109.46	24	36.42	Reject
B4.0: Without environmental variables (Z1–Z8)	-408.20	5.23	0.04	1	3.84	Yes	151.45	1	3.84	Reject
B5.0: Without control variables (C1, C2, C3) or environmental variables (Z1–Z8)	-504.71	12.23	0.52	1	3.84	Yes	344.46	24	36.42	Reject

Notes: (1) Log-likelihood: The maximum log-likelihood value from the unknown parameters being estimated based on given data (2) LR against OLS model: LR test based on ratio of log-likelihood of stochastic model subtracted from log-likelihood of OLS model and multiplied by (-2). (3) γ : Lies between 0 and 1. Approaching 0, composite error (ϵ) dominated by random error (v). Approaching 1, composite error (ϵ) influenced by inefficiency (u). (4) Number of restrictions: Restrictions on the composite error where it consists only of random error (v) and no inefficiency (u). One restriction is made on $\sigma_u = 0$. (5) χ^2 critical value at 5%: 5% critical value of chi-squared distribution at 1 degree of freedom is 3.84. (6) Presence of inefficiency: If $\gamma \neq 0$; inefficiency exists. If $\gamma = 0$; there is no inefficiency. (7) LR against general model: LR test based on the ratio of log-likelihood of reduced model subtracted from log-likelihood of general model and multiplied by (-2). (8) Degree of freedom: The difference between the numbers of parameters of the general model subtracted against the number of parameters of the reduced model. (9) χ^2 critical value at 5%, the chi-squared distribution value based on the degree of freedom derived from column (8). (10) $H_0 =$ reduced model better than general model: H_0 is failed to be rejected if the LR value is less than the chi-squared distribution at 5% critical value. Reject H_0 if the LR value exceeds the chi-squared distribution at 5% critical value. Columns (1) to (6) are to test for inefficiency and columns (7) to (10) are for choosing preferred model. * - Preferred standard profit frontier model (B1.0).

Table 5.7 Estimated Alternative Profit Frontier Models

Models	(1) Log-likelihood	(2) Log-likelihood ratio against OLS model	(3) γ	(4) Number of restrictions	(5) χ^2 critical value at 5%	(6) Presence of inefficiency $\gamma \neq 0$	(7) Log-likelihood ratio against general model	(8) Degree of freedom	(9) χ^2 critical value at 5%	(10) $H_0 =$ reduced model is better than general model
C1.0: General – W1, W2, W3, Y1, Y2, Y3, T, C1, C2, C3, Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8*	-139.55	7.65	0.49	1	3.84	yes				
C2.1: Reduced without liquidity (C3)	-207.81	50.47	0.45	1	3.84	Yes	136.51	10	18.31	Reject
C2.2: Reduced without asset quality (C2)	-148.88	10.56	0.48	1	3.84	Yes	18.66	10	18.31	Reject
C2.3: Reduced without capital (C1)	-163.96	78.38	0.52	1	3.84	Yes	48.82	10	18.31	Reject
C2.4: Reduced without asset quality or liquidity (C2 and C3)	-220.31	23.12	0.43	1	3.84	Yes	161.51	19	30.14	Reject
C2.5: Reduced without capital or liquidity (C1 and C3)	-220.09	103.54	0.51	1	3.84	Yes	161.08	19	30.14	Reject
C2.5: Reduced without capital or asset quality (C1 and C2)	-168.09	25.08	0.50	1	3.84	Yes	57.08	19	30.14	Reject
C3.0: Without control variables (C1, C2, C2)	-230.77	93.77	0.48	1	3.84	Yes	182.43	27	40.11	Reject
C4.0: Without environmental variables (Z1–Z8)	-151.63	12.08	0.01	1	3.84	Yes	24.15	1	3.84	Reject
C5.0: Without control variables (C1, C2, C3) or environmental variables (Z1–Z8)	-244.69	65.92	0.01	1	3.84	Yes	210.28	27	40.11	Reject

Notes: (1) Log-likelihood: The maximum log-likelihood value from the unknown parameters being estimated based on given data (2) LR against OLS model: LR test based on ratio of log-likelihood of stochastic model subtracted from log-likelihood of OLS model and multiplied by (-2). (3) γ : Lies between 0 and 1. Approaching 0, composite error (ε) dominated by random error (v). Approaching 1, composite error (ε) influenced by inefficiency (u). (4) Number of restrictions: Restrictions on the composite error where it consists only of random error (v) and no inefficiency (u). One restriction is made on $\sigma_u = 0$. (5) χ^2 critical value at 5%: 5% critical value of chi-squared distribution at 1 degree of freedom is 3.84. (6) Presence of inefficiency: If $\gamma \neq 0$; inefficiency exists. If $\gamma = 0$; there is no inefficiency. (7) LR against general model: LR test based on the ratio of log-likelihood of reduced model subtracted from log-likelihood of general model and multiplied by (-2). (8) Degree of freedom: The difference between the numbers of parameters of the general model subtracted against the number of parameters of the reduced model. (9) χ^2 critical value at 5%, the chi-squared distribution value based on the degree of freedom derived from column (8). (10) $H_0 =$ reduced model better than general model: H_0 is failed to be rejected if the LR value is less than the chi-squared distribution at 5% critical value. Reject H_0 if the LR value exceeds the chi-squared distribution at 5% critical value. Columns (1) to (6) are to test for inefficiency and columns (7) to (10) are for choosing preferred model. * - Preferred alternative profit frontier model (C1.0).

5.5 Conclusion

This chapter presents the data and variables that were employed for SFA and DEA models. There are three types of variables used in the analysis of Malaysian banking efficiency: namely, input and output, control, and environmental variables. First, the inputs and outputs used in this research are based on commonly used inputs and outputs from past literature using the intermediation approach on cost- and profit-efficiency functions; for example, price of funds, price of labour, price of physical capital, loans, investments, other earning assets, price of loans, price of investments and price of other earning assets. Second, the control variables are employed in this analysis to capture the effect of inherent bank-specific risks that could have an impact on the cost- and profit-frontier estimations (e.g. asset quality, capital adequacy and liquidity risk). Third, environmental variables are introduced in this analysis to account for other factors that may affect the levels of inefficiency of the Malaysian banks, such as ownership, specialisation, FSMP deregulation phases, global credit crisis, size, market concentration, and market share. The empirical analysis of this research employs unbalanced panel observations of 32 banks in 2000 followed by 39 banks in 2011 (including domestic, foreign, domestic Islamic and foreign Islamic banks),¹⁵⁸ which represent all commercial banks in Malaysia.

The variables discussed above are employed using the SFA and DEA.¹⁵⁹ Three common economic concepts are utilised using the translog functional form (for SFA) to estimate Malaysian banks' efficiencies: namely, cost-, standard profit-, and alternative profit-efficiency. Following the Battese and Coelli's (1995) model (which is a 1-stage analysis approach), estimations of cost- and profit-efficiency use a number of model variants with different combinations of the control and environmental variables, involving several stages to test several null hypotheses for a preferred model. The baseline model is the A1.0, B1.0 and C1.0, which includes the heterogeneity effects of all control (C1–C3) and environmental (Z1–Z8) variables. Consequently, a number of models with reduced combinations of control and environmental variables are tested on sample banks on which

¹⁵⁸ Investment banks are excluded from this empirical analysis due to a non-homogeneous business model.

¹⁵⁹ As mentioned earlier in Chapter 2, this research follows Resti (1997), which utilises SFA as the main estimation tool and DEA as the consistency testing tool (similar to Bauer et al., 1998).

the empirical analysis is based. As recommended by Coelli et al. (2005), the preferred model for cost, standard profit and alternative profit-efficiencies are chosen using the ML estimates and log-likelihood ratio tests. From the log-likelihood tests of different models at different stages, for the cost efficiency frontier model, the reduced model (A1.3) without capital adequacy (C3), but with all environmental variables (Z1–Z8) was preferred. For standard profit efficiency, the general model (B1.0) with all control variables (C1–C3) and environmental variables (Z1–Z8) was found to be the preferred model. With regards to the alternative profit efficiency, the preferred model is the general model (C1.0) with all control variables (C1–C3) but with all environmental variables (Z1–Z8). Additionally, for SFA, the SE will also be estimated by differentiating the estimated cost function with respect to output. In terms of DEA methodology, both traditional and new DEA (Tone, 2002) models are used as cross-checking method for consistency testing. For DEA, scale-efficiency will also be estimated. In addition, following Bauer et al. (1998), this chapter proposes five conditions for evaluating the potential consistency of the SFA and DEA models. From this consistency testing, should all five consistency conditions be met, the results from different efficiency frontier approaches could be more useful and reliable for policy makers in making decisions.

In the next chapter, the SFA preferred models will be tested using the properties of cost and profit functions in economic theory that were discussed in Chapter 2. Moreover, the results reviewed in Chapter 3 from various empirical studies of banking efficiency will also be compared with the results from this study. The hypotheses developed in Chapters 4 and 5 will also be tested using the SFA preferred models. Using the cost- and profit-efficiency scores of SFA and DEA, the trends from 2000 to 2011 will be analysed to examine to what extent the financial liberalisation of the FSMP has affected the degree of efficiency of Malaysian banks. Additionally, the next chapter will also examine the cost- and profit-efficiency scores in term of profiles (e.g. by ownership, by specialisation, by size, and CAMEL ratings system), and consistency testing of parametric and nonparametric models (Bauer et al., 1998). Scale-efficiency and technological progress in the Malaysian banking industry will also be analysed in the next chapter.

Chapter 6 Empirical Results: Analysis and Discussion

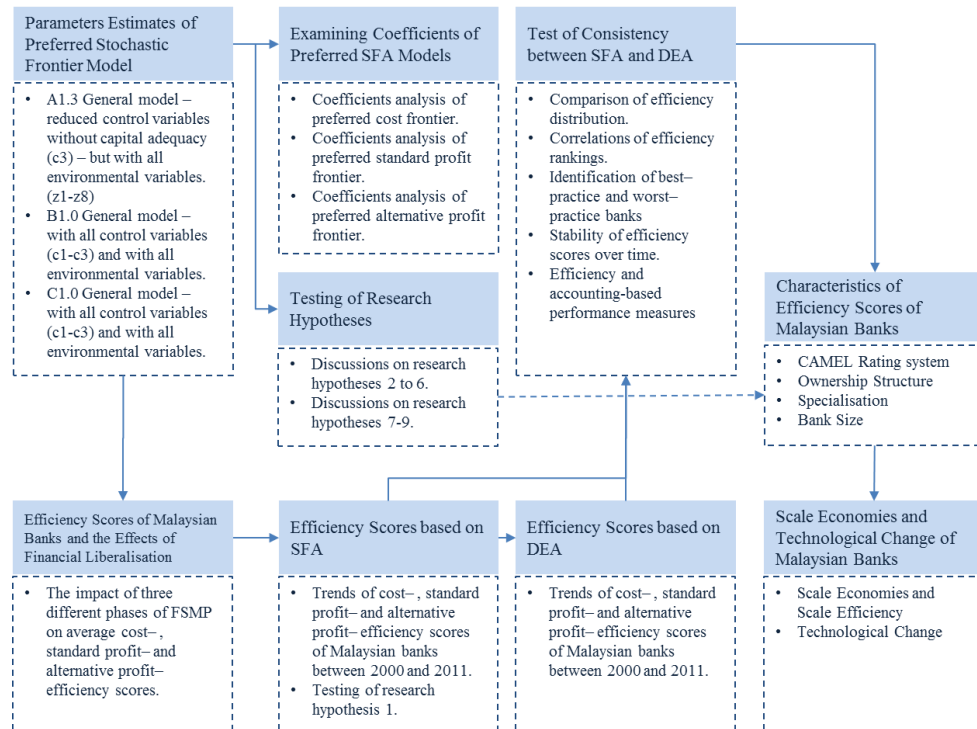
6.1 Introduction

In the preceding chapter (Chapter 5), the methodology employed to measure the efficiency of Malaysian banks between 2000 and 2011 was discussed. The preferred cost and profit efficiency models for SFA were derived, and these will be used to calculate the efficiency scores for the cost and profit frontiers discussed in this chapter. The properties of these scores are also analysed here, to ensure that the theoretical conditions discussed in Chapter 2 are satisfied. The results of past empirical studies, reviewed in Chapter 3, are also compared with the results of this study. In addition, the cost- and profit-efficiency results are compared to corresponding events (e.g. the deregulation of interest rates, the establishment of foreign banks, the global credit crisis) that affected the Malaysian banking industry between 2000 and 2011 (see Chapter 4). The hypotheses developed in Chapters 4 and 5 are also tested in this chapter. Consequently, this chapter analyses the efficiency scores derived using both SFA and DEA, and the trends of the cost and profit efficiency scores that have changed over the period under study. From the efficiency scores of SFA and DEA, the consistency conditions introduced by Bauer et al. (1998) are also tested in this chapter. The scores are then compared, using several main characteristic criteria – such as, the CAMEL rating system, ownership, specialisation and size – in which, finally, the scale efficiency and technological change of Malaysian banks are also examined.

This chapter begins with Section 6.2, which examines the coefficients of the SFA preferred models for cost, standard profit and alternative profit frontiers. The maximum likelihood parameters estimated by the preferred models are also reported in this section. In addition, the properties of cost and profit efficiency are analysed using the parameters estimated. In Section 6.3, research hypotheses 2 to 9, developed in Chapters 4 and 5, are tested using the cost and profit model parameters estimated in Section 6.2. This test indicates the impact of heterogeneity factors (i.e. the control and environmental variables) on the cost and profit

efficiency scores of Malaysian banks using the one-stage approach (see Chapter 5). Section 6.4 provides the cost and profit efficiency scores derived from SFA and DEA. The trends of Malaysian banks' efficiency scores between 2000 and 2011 are analysed and examined by comparing cost and profit efficiency scores, with corresponding financial liberalisation measures introduced by BNM. The analysis looks into how the cost and profit efficiency scores respond to the different policy measures implemented by BNM. From this, the analysis could provide signals or indications as to whether the policies implemented by BNM have resulted in either positive or negative impacts on the level of efficiency of the Malaysian banks. In addition, the main research hypothesis 1, discussed in Chapter 4 – which hypothesised that the implementation of financial liberalisation could result in greater efficiency – is tested. Section 6.5 provides the results of the consistency conditions test, of the SFA and DEA's efficiency scores, using Bauer et al. (1998)'s conditions. Section 6.6 looks into the different characteristics of the efficiency scores of Malaysian banks, for instance based on the CAMEL rating system, the ownership structure, specialisation and size. For the CAMEL rating system, the best and worst practice banks – according to the CAMEL ratios – are compared with best and worst practice efficiency scores. In terms of ownership, the efficiency scores for foreign banks are compared with domestic banks; and for specialisation, the efficiency scores for Islamic banks are compared with conventional banks. Additionally, banks are clustered into different asset size groups and their efficiency scores are compared. Finally, Section 6.7 presents the scale efficiency and technological change of Malaysian banks over the period 2000 to 2011.

Figure 6.1 Information Flow of Chapter 6



6.2 Examining the Coefficients of the Preferred SFA Models

As discussed in Chapter 5, this study employs SFA as the main methodology to measure efficiency, and DEA as a comparative measure of consistency, in Malaysian banks between 2000 and 2011 (for further details, see Resti (1997) and Bauer et al. (1998)). Using the preferred SFA cost, standard profit and alternative profit frontier models (obtained from Tables 5.5, 5.6 and 5.7), the maximum likelihood parameters of the above models are exhibited in Tables 6.1, 6.2 and 6.3. The coefficients for the maximum likelihood estimates for these models are examined to measure the interactions between the dependent variable and inputs, outputs, and control variables. In this study, the total cost, profits, and all the continuous explanatory variables, are divided by their respective sample means before logarithms are applied.¹⁶⁰ This normalisation of variables allows the first-order coefficients of the translog function to be directly interpreted as estimates of cost and profit elasticities at the point of approximation. Therefore, the estimated individual coefficients of the stochastic

¹⁶⁰ Since the mean values of variables are considered as the Taylor series expansion point for the translog function, all variables should be divided by their mean in order to locate a correct evaluation point before estimating the translog function, which can assist in the estimation of scale elasticities.

frontier models, given by the translog functional form, are attributed by many interactions between inputs, outputs and control variables, in which they are not directly interpretable (unlike the Cobb-Douglas functional form, in which the coefficients have specified meaning) (see Chapter 5). In addition, prior to analysing the coefficients of these models, it is worth noting that the flexibility of the translog functional form is likely to result in multicollinearity among variables involved in the SFA specifications, which may lead to inconsistent coefficient estimates in the maximum likelihood model (Coelli, 1996).¹⁶¹ However, multicollinearity may not be a severe problem if the efficiency scores are used for forecasting purposes.¹⁶²

6.2.1 Coefficients of the Preferred Cost Frontier

Based on Table 5.5 in the previous chapter (Chapter 5), the preferred cost frontier model is A2.3, which is the reduced model without capital adequacy (C3) but includes all the environmental variables. The t-ratios of the preferred cost model are used to interpret the level of significance of the estimated coefficients, based on 1%, 5% and 10% critical values. The t-ratio is derived by dividing the estimated coefficients to their corresponding standard errors to indicate the level of significance of the coefficient of banks' inputs and outputs (Hamilton et al, 2010). A statistically insignificant coefficient suggests that the coefficient may not directly affect the banking cost.

Table 6.1 shows that the coefficients of the input prices (i.e. price of funds and price of labour) are positive and significant at 1% significance level. The estimated coefficients for the price of funds β_1 and the price of labour β_2 shows the share of costs attributed from funds and employees.¹⁶³ Based on the cost model estimated, the coefficients for prices of

¹⁶¹ If the multicollinearity problem is mainly created by a strong positive correlation between the second order terms in the translog form of the cost and profit functions, maximum likelihood estimates are still unbiased and efficient. However, this multicollinearity problem could cause the estimated standard error of the coefficients to be large leading to small values for the t-ratios. This, in turn, biases the result towards failure to reject the null hypothesis where the coefficients are equal to zero (see Gujarati, 2003 for more details).

¹⁶² The empirical estimates of the translog cost and profit functions in Tables 6.1, 6.2 and 6.3 demonstrate expected signs (e.g. positive and negative) of the coefficients. A test of multicollinearity using variance inflation factor (VIF) also indicate that there is no serious linear relation among explanatory variables.

¹⁶³ As mentioned earlier, the specification of cost efficiency is imposed with homogeneity of degree 1, in which the inputs price of funds (w_1) and price of labour (w_2) are normalised by the price of physical capital (w_3). Hence, the coefficient of the physical capital is equal to 1 minus the sum of coefficients of price of funds

funds (β_1), price of labour (β_2) and price of physical capital are 0.6720, 0.3711 and -0.0431 respectively. The coefficient values indicate that an increase of 1% in the price of funds (W1), price of labour (W2) and price of physical capital (W3) would result in an increase of 0.67% and 0.37%, and a reduction of -0.04% in cost respectively (this satisfies property number 2 of the cost function (e.g. $TC=f(W1,W2,W3)$), in which overall cost will not decrease when the combination of inputs increases (see Chapter 2)). Given the restriction of homogeneity of degree 1 in input prices imposed in the cost function, the main contributor to the total cost is interest expenses. The coefficient for the price of funds (β_1) reflects the actual data of the banks, where interest expenses appear to be averaging more than 50% of total expenses (see Chapter 4).

On the other hand, the coefficients of the total combination of outputs are also positive. A 1% increase in outputs – loans (Y1), investments (Y2) and other earnings assets (Y3) – could increase costs by 1.04% and 0.02%, and reduce costs by 0.02%, respectively (satisfying property number 3 of the cost function, where overall costs will increase to produce additional combinations of outputs (see Chapter 2)). Since the variables used in the cost function are expressed in log forms, the estimates of the coefficients can be interpreted as output elasticities or input price elasticities of total costs. The estimation of the cost function shows that all three outputs have positive effects on the total costs of the banks, indicating that the total cost variable increases with an expansion in production. The sum of the three output coefficients is 1.0382, implying that if the production is expanded systematically by 1%, the total cost of production would increase by a little more than 1%. This may indicate that the Malaysian banking industry's production could be characterised as facing diseconomies of scale: further expansion in the industry could result in a more than proportional increase in costs. In terms of time (T), the coefficient is negative (-0.04) and significant at 5% significance level. This indicates that a technological change had taken place in the industry over the period under study from technical progress, where banks were able to produce a given level of output at lower levels of cost.

and price of labour. From this formulation, the coefficient for price of capital is equal to 0.0119 ($1-(0.6409+0.3472)$). The effect of imposing restriction using homogeneity of degree 1 on the parameters to being equal to unity before estimation (i.e. $\beta_1 + \beta_2 + \beta_3 = 1$) also satisfies the property number 4 of the cost function, which implies that an increase in input prices will result in proportionate increase in costs.

Table 6.1 The Maximum Likelihood Parameters for the Preferred Cost Frontier

Parameters	Variables	Coefficient		Standard Error	t-ratio
α_0	Constant	-0.1917	***	0.0415	-4.6175
β_1	$\ln \frac{w1}{w3} \left(\frac{\text{price of funds}}{\text{price of physical capital}} \right)$	0.6720	***	0.0426	8.7068
β_2	$\ln \frac{w2}{w3} \left(\frac{\text{price of labour}}{\text{price of physical capital}} \right)$	0.3711	***	0.0731	14.2274
δ_1	$\ln y1$ (loans)	1.0407		0.0388	0.5344
δ_2	$\ln y2$ (investments)	0.0207		0.0429	-0.5406
δ_3	$\ln y3$ (Other earning assets)	-0.0232	***	0.0644	4.7954
β_{11}	$0.5 \left(\ln \frac{w1}{w3} \right)^2$	0.3087	***	0.0544	-6.0310
β_{12}	$\left(\ln \frac{w1}{w3} \right) \left(\ln \frac{w2}{w3} \right)$	-0.3280	***	0.0450	7.7097
β_{22}	$0.5 \left(\ln \frac{w2}{w3} \right)^2$	0.3471	***	0.1771	-2.9988
δ_{11}	$0.5 (\ln y1)^2$	-0.5311	***	0.0999	2.3452
δ_{12}	$(\ln y1)(\ln y2)$	0.2342	***	0.0819	3.1686
δ_{13}	$(\ln y1)(\ln y3)$	0.2596		0.0636	-1.0474
δ_{22}	$0.5 (\ln y2)^2$	-0.0666	***	0.0455	-3.2149
δ_{23}	$(\ln y2)(\ln y3)$	-0.1462	*	0.0541	-1.5049
δ_{33}	$0.5 (\ln y3)^2$	-0.0814	**	0.0756	2.1947
η_{11}	$\left(\ln \frac{w1}{w3} \right) (\ln y1)$	0.1659		0.0324	-0.2921
η_{12}	$\left(\ln \frac{w1}{w3} \right) (\ln y2)$	-0.0095	**	0.0533	-1.8848
η_{13}	$\left(\ln \frac{w1}{w3} \right) (\ln y3)$	-0.1005	***	0.0753	-2.8617
η_{21}	$\left(\ln \frac{w2}{w3} \right) (\ln y1)$	-0.2156	*	0.0263	1.4340
η_{22}	$\left(\ln \frac{w2}{w3} \right) (\ln y2)$	0.0377	***	0.0565	2.4002
η_{23}	$\left(\ln \frac{w2}{w3} \right) (\ln y3)$	0.1357	***	0.0257	5.4262
χ_2	$\ln c2$ (asset quality)	0.1396		0.0134	-0.6207
χ_{22}	$0.5 (\ln c2)^2$	-0.0083		0.0291	-0.1966
ρ_{12}	$\left(\ln \frac{w1}{w3} \right) (\ln c2)$	-0.0057		0.0304	0.9533
ρ_{22}	$\left(\ln \frac{w2}{w3} \right) (\ln c2)$	0.0290	***	0.0384	2.5026
ω_{12}	$(\ln y1)(\ln c2)$	0.0962	**	0.0283	-1.9038
ω_{22}	$(\ln y2)(\ln c2)$	-0.0539		0.0175	-1.2068

(continued)

**Table 6.1 The Maximum Likelihood Parameters for the Preferred Cost Frontier
(Continued)**

Parameters	Variables	Coefficient		Standard Error	t-ratio
ω_{32}	$(\ln y_3)(\ln c_2)$	-0.0211	***	0.1776	-4.6591
χ_3	$\ln c_3$ (liquidity)	-0.8275	***	0.3906	-3.0853
χ_{33}	$0.5(\ln c_3)^2$	-1.2051	***	0.1503	-2.6300
ρ_{13}	$(\ln \frac{w_1}{w_3})(\ln c_3)$	-0.3953	***	0.1535	2.7350
ρ_{23}	$(\ln \frac{w_2}{w_3})(\ln c_3)$	0.4197	***	0.2624	3.2351
ω_{13}	$(\ln y_1)(\ln c_3)$	0.8490	***	0.1650	-2.4045
ω_{23}	$(\ln y_2)(\ln c_3)$	-0.3967	***	0.1079	-3.2602
ω_{33}	$(\ln y_3)(\ln c_3)$	-0.3517		0.0622	-0.4721
χ_{23}	$(\ln c_2)(\ln c_3)$	-0.0294	**	0.0363	-1.9125
τ_1	$\ln t$ (time trend)	-0.0695		0.0405	-1.0833
τ_{11}	$0.5(\ln t)^2$	-0.0439	***	0.0392	3.2971
φ_{11}	$(\ln \frac{w_1}{w_3})(\ln t)$	0.1294	**	0.0382	-2.2944
φ_{21}	$(\ln \frac{w_2}{w_3})(\ln t)$	-0.0876		0.0601	0.9515
κ_{11}	$(\ln y_1)(\ln t)$	0.0572		0.0369	0.0992
κ_{21}	$(\ln y_2)(\ln t)$	0.0037		0.0322	-1.2363
κ_{31}	$(\ln y_3)(\ln t)$	-0.0398	**	0.0240	-1.7810
ξ_{21}	$(\ln c_2)(\ln t)$	-0.0427		0.0982	-1.2751
ξ_{31}	$(\ln c_3)(\ln t)$	-0.1252	*	0.0082	-1.4144
ψ	constant	-46.3125	***	15.7471	-2.9410
ψ_1	Ownership structure (z1)	1.0956	***	0.2654	4.1286
ψ_2	Specialisation (z2)	0.3837	*	0.2636	-1.4556
ψ_3	FSMP phase 2 (2004 – 2007)(z3)	-0.4013		0.4076	-0.9847
ψ_4	FSMP phase 3 (2008 – 2011)(z4)	-0.5224		0.9023	-0.5790
ψ_5	Global crisis(2008 – 2010)(z5)	0.6225	**	0.3289	1.8924
ψ_6	Asset size (z6)	2.4994	***	0.7790	3.2087
ψ_7	Market concentration (z7)	3.6821		3.0993	1.1880
ψ_8	Market share (z8)	-2.9020	***	0.7898	-3.6742
	Log-likelihood			46.0213	
	Log-likelihood ratio against OLS model			150.6543	
	Number of observations			354	
	Degrees of freedom			310	

t-ratio is defined as ratio of estimated coefficient over standard errors. Degree of freedom is defined by netting off number of observations in the sample against the number of estimated coefficients in the maximum likelihood model. ***, **, * is the significance level of estimated coefficients in given degrees of freedom. Using t-ratio critical value at significant levels of 1% (***), 5% (**) and 10% (*) which are at 2.326, 1.645 and 1.282 respectively.

6.2.2 Coefficients of the Preferred Standard Profit and Alternative Profit Frontiers

Tables 6.2 and 6.3 display the preferred standard profit and alternative profit parameter estimations. For such estimations, the general model (B1.0) (see Table 5.6) is preferred. This model includes all the control and environmental variables. The input price of funds (W1) standard profit efficiency shows a negative, but insignificant, coefficient (β_1). An increase in price of fund by 1% could reduce the banks' profits by 0.05%. The negative coefficient implies that an increase in the price of funds could result in a reduction in profitability. This result is influenced by the deregulation of the interest rate introduced by BNM through the New Interest Rate Framework (NIRF) in 2004. The deregulation of interest rates allows banks to price their loans and deposits according to market demand and cost structure. Similar to Dacanay III (2007) findings on the Philippines banking system, the negative elasticities of the price of funds (W1) show that deposits were a cheap source of funds. Thus, the banks had difficulty cutting deposit interest rates, particularly when they were partly determined by the market.

Additionally, from the preferred standard profit frontier estimation, the coefficient of the price of loans (δ_1) was also negative and significant. The elasticity for the price of loans (P1) (-0.3668) was greater than the elasticity of the price of funds (W1) (-0.054). The changes in the price of loans are more sensitive to the standard profit efficiency function compared to the price of funds. Similar to findings by Humphrey and Pulley (1997), Malaysian banks responded to deregulation of interest rates by adjusting their prices for loans in greater proportion than the adjustments made in the price of deposits to gain higher profits.

In addition, the coefficient for the price of labour (β_2) is positive and significant at 1% significance level. This implies that higher expenses incurred in relation to labour input could result in higher profits. The effect on the price of labour (W2) is greater than on the price of funds (W1), indicating that total profits are very sensitive to the price of labour and that, on average, an increase of 1% of the price of labour can increase profits by 1.16%. With the increasing growth in Islamic banking and increasing complexity of financial

products in Malaysia, a highly skilled and specialised workforce is necessary to support the banks. From the restriction of homogeneity of degree 1, the input coefficient for physical capital (β_3) is -0.11. This indicates that an increase of 1% in the input price of physical capital would reduce profits by 0.11%. In this case, it may imply that any expenses incurred on infrastructure and capital expenditure could result in lower profitability. The coefficients for the price of loans (δ_1), price of investments (δ_2) and price of other earning assets (δ_3) are -0.3668, 0.4975 and 0.8693 respectively. The output price of other earning assets is used to normalise other price variables in the frontier regression process, the value for this output is equal to 0.8693 $(1 - (-0.3668) - 0.4975)$.¹⁶⁴ This implies that an increase of 1% of the price of other earning assets could result in an increase of profits by 0.87%, reflecting the excess liquidity that the Malaysian banking sector faced during the height of the subprime crisis in which a large influx of foreign capital occurred in mid-2007 (Bank Negara Malaysia, 2007b). With the flight to liquidity in the US and Europe, the volatility of the interbank market became significant in terms of pricing, where banks generate greater profit from other earning assets, particularly for foreign banks (see Chapter 4).

The coefficient for the price of loans (δ_1) is -0.3668, indicating that an increase of 1% in price of loans (p1) would reduce the profit by 0.37%. Malaysian banks are sensitive to the pricing of loans, where an increase of price might result in lower demand for loans and reduce profitability. For the time trend variable (T), the standard profit frontier model exhibited a positive coefficient (0.21) and 1% significance. In general, the coefficients suggest that over the period under study, the profit efficiency of Malaysian banks has improved and the banks have achieved a higher level of profit for any given level of outputs.

The preferred model for alternative profit efficiency, from 2000 to 2011, is the general model, which includes all controls (C1–C3) and environmental variables (Z1–Z8) (see table 5.7). From Table 6.3, the coefficient for price of funds, price of labour and price of physical capital were estimated at 0.2070, 0.7531 and 0.0399 $(1.00 - 0.2070 - 0.7531)$ respectively. The coefficient for price of funds (β_1) is positive and significant at 5% significant level. The

¹⁶⁴ The price of loans (p1) and price of investments (p2) are normalised by price of other earning assets (p3) to impose homogeneity of degree 1. The coefficient of other earning assets (p3) is equal to 1 minus sum of coefficients of price of loans and the price of investments. Therefore, coefficient of price of other earning assets equals $0.21(1 - 0.05 + 0.74)$.

next input variable is the price of labour (W2) which is also positive and significant at 1%. For the alternative profit model, a 1% increase in the price of labour (W2) would increase profits by 0.75%. For the price of physical capital (W3), an increase of 1% will result in 0.01% increase in profit efficiency. In terms of outputs, an increase in loans (Y1) by 1% could increase profits by 1.63%, while a 1% increase of investment (Y2) and other earnings assets (Y3) could result in decreased profits by 0.45% and 0.25%. These are indicated by coefficient δ_1 , δ_2 and δ_3 at 1.6343, -0.4512 and -0.2545, respectively.

For alternative profit efficiency, the output level of loans is highly sensitive to the Malaysian banks' profitability, suggesting that banks with market power may be able to exercise a different pricing strategy by charging a higher interest rate on loans (and/or a low interest rate for deposits) to customers to generate higher profits (see Table 6.3, positive coefficient for loans (Y1) (δ_1) (1.63%), implying that an increase by 1% loans can result in 1.63% increase in profits). This can also be potentially explained by the homogenous sample used for this study that consists only of commercial banks which mainly perform an intermediary function by transforming customers' deposits into loans and advances.¹⁶⁵ For the time trend variable (T), both models exhibited a positive coefficient (0.21 for standard profit and 0.07 for alternative profit models) but only significant (at 1%) for the standard profit model. In general, the coefficients suggest that, over the period under study, the profit efficiency of Malaysian banks improved and achieved a higher level of profit for any given level of outputs.

¹⁶⁵ Financial institutions such as investment banks are excluded from this study due to different business objectives, customers and operational structure.

Table 6.2 The Maximum Likelihood Parameters Estimation for the Preferred Standard Profit

Parameters	Variables	Coefficient		Standard Error	t-ratio
α_0	Constant	0.2265	**	0.1330	1.7030
β_1	$\ln \frac{w1}{w3} \left(\frac{\text{price of funds}}{\text{price of physical capital}} \right)$	-0.0554		0.1832	-0.3025
β_2	$\ln \frac{w2}{w3} \left(\frac{\text{price of labour}}{\text{price of physical capital}} \right)$	1.1626	***	0.1771	6.5654
δ_1	$\ln \frac{p1}{p3} \left(\frac{\text{price of loans}}{\text{price of other earning assets}} \right)$	-0.3668	**	0.1591	-2.3054
δ_2	$\ln \frac{p2}{p3} \left(\frac{\text{price of investments}}{\text{price of other earning assets}} \right)$	0.4975	***	0.1272	3.9103
β_{11}	$0.5 \left(\ln \frac{w1}{w3} \right)^2$	-0.1246		0.1872	-0.6656
β_{12}	$\left(\ln \frac{w1}{w3} \right) \left(\ln \frac{w2}{w3} \right)$	-0.0578		0.1432	-0.4034
β_{22}	$0.5 \left(\ln \frac{w2}{w3} \right)^2$	0.2130	**	0.1233	1.7281
δ_{11}	$0.5 \left(\ln \frac{p1}{p3} \right)^2$	0.2405	*	0.1753	1.3721
δ_{12}	$\left(\ln \frac{p1}{p3} \right) \left(\ln \frac{p2}{p3} \right)$	-0.1686	*	0.1099	-1.5341
δ_{22}	$0.5 \left(\ln \frac{p2}{p3} \right)^2$	0.0474		0.1035	0.4583
η_{11}	$\left(\ln \frac{w1}{w3} \right) \left(\ln \frac{p1}{p3} \right)$	0.3385	**	0.1866	1.8142
η_{12}	$\left(\ln \frac{w1}{w3} \right) \left(\ln \frac{p2}{p3} \right)$	-0.0363		0.1332	-0.2722
η_{21}	$\left(\ln \frac{w2}{w3} \right) \left(\ln \frac{p1}{p3} \right)$	-0.4147	***	0.1748	-2.3727
η_{22}	$\left(\ln \frac{w2}{w3} \right) \left(\ln \frac{p2}{p3} \right)$	0.0518		0.1198	0.4324
χ_1	$\ln c1$ (financial capital)	-0.4082	**	0.2347	-1.7391
χ_{11}	$0.5 (\ln c1)^2$	-0.8480	***	0.2994	-2.8320
ρ_{11}	$\left(\ln \frac{w1}{w3} \right) (\ln c1)$	-0.3590	**	0.2015	-1.7816
ρ_{21}	$\left(\ln \frac{w2}{w3} \right) (\ln c1)$	0.2025		0.2019	1.0028
ω_{11}	$\left(\ln \frac{p1}{p3} \right) (\ln c1)$	0.1021		0.2944	0.3467
ω_{21}	$\left(\ln \frac{p2}{p3} \right) (\ln c1)$	0.3766	**	0.2033	1.8527
χ_2	$\ln c2$ (asset quality)	-0.1781	**	0.0941	-1.8928

(Continued)

Table 6.2 The Maximum Likelihood Parameters Estimation for the Preferred Standard Profit (Continued)

Parameters	Variables	Coefficient		Standard Error	t-ratio
χ_{22}	$0.5(lnc2)^2$	-0.0759	**	0.0454	-1.6743
ρ_{12}	$(\ln \frac{w1}{w3})(lnc2)$	-0.1050		0.1022	-1.0270
ρ_{22}	$(\ln \frac{w2}{w3})(lnc2)$	0.0685		0.1036	0.6610
ω_{12}	$(lnp1)(lnc2)$	0.0103		0.1052	0.0979
ω_{22}	$(lnp2)(lnc2)$	0.0838		0.0825	1.0161
χ_3	$lnc3$ (liquidity risk)	0.1016		0.1665	0.6106
χ_{33}	$0.5(lnc3)^2$	-0.0364		0.1225	-0.2970
ρ_{13}	$(\ln \frac{w1}{w3})(lnc3)$	0.3294	**	0.1762	1.8696
ρ_{23}	$(\ln \frac{w2}{w3})(lnc3)$	-0.2758	**	0.1641	-1.6803
ω_{13}	$(\ln \frac{p1}{p3})(lnc3)$	0.3509	***	0.1384	2.5364
ω_{23}	$(\ln \frac{p2}{p3})(lnc3)$	0.0539		0.1218	0.4424
χ_{12}	$(lnc1)(lnc2)$	0.4029	***	0.1032	3.9026
χ_{13}	$(lnc1)(lnc3)$	-0.0819		0.2183	-0.3753
χ_{23}	$(lnc2)(lnc3)$	0.0794		0.0932	0.8523
τ_1	lnt (time trend)	0.1005		0.2004	0.5013
τ_{11}	$0.5(t)^2$	-0.1904		0.2038	-0.9346
φ_{11}	$(\ln \frac{w1}{w3})(lnt)$	-0.4172	**	0.1939	-2.1514
φ_{21}	$(\ln \frac{w2}{w3})(lnt)$	0.4227	**	0.1830	2.3091
κ_{11}	$(\ln \frac{p1}{p3})(lnt)$	0.1793		0.1668	1.0746
κ_{21}	$(\ln \frac{p2}{p3})(lnt)$	-0.0229		0.1409	-0.1628
ξ_{11}	$(lnt)(lnc1)$	0.2584	**	0.1486	1.7391
ξ_{12}	$(lnt)(lnc2)$	0.0548		0.0919	0.5959
ξ_{23}	$(lnt)(lnc3)$	-0.3789	**	0.2085	-1.8176
ψ	constant	20.5879		17.6689	1.1652
ψ_1	Ownership structure (z1)	1.1344	***	0.4615	2.4580
ψ_2	Specialisation (z2)	1.4274	***	0.4117	3.4668
ψ_{3a}	FSMP phase 2 (2004 – 2007)(z3a)	0.2102		0.5985	0.3512

(Continued)

Table 6.2 The Maximum Likelihood Parameters Estimation for the Preferred Standard Profit (Continued)

Parameters	Variables	Coefficient	Standard Error	t-ratio
ψ_{3b}	<i>FSMP phase 3 (2008 – 2011)(z3b)</i>	1.4436	1.1874	1.2158
ψ_4	<i>Global crisis(2008 – 2010)(z4)</i>	-0.3631	0.4346	-0.8355
ψ_5	<i>Asset size (z5)</i>	-1.0321	0.8209	-1.2574
ψ_{6a}	<i>Market concentration (z6a)</i>	2.6529	4.6694	0.5681
ψ_{6b}	<i>Market share (z6b)</i>	-0.1046	0.8043	-0.1301
	Log-likelihood		-332.47	
	Log-likelihood ratio against OLS model		-392.82	
	Number of observations		354	
	Degrees of freedom		340	

Notes: t-ratio is defined as ratio of estimated coefficient over standard errors. Degree of freedom is defined by netting off number of observations in the sample against the number of estimated coefficients in the maximum likelihood model. ***, **, * is the significance level of estimated coefficients in given degrees of freedom. Using t-ratio critical value at significant levels of 1% (***), 5% (**) and 10% (*) which are at 2.326, 1.645 and 1.282 respectively.

Table 6.3 The Maximum Likelihood Parameters Estimation for the Preferred Alternative Profit

Parameters	Variables	Coefficient		Standard Error	t-ratio
α_0	Constant	0.3034	***	0.0769	3.9447
β_1	$\ln \frac{w_1}{w_3} \left(\frac{\text{price of funds}}{\text{price of physical capital}} \right)$	0.2070	**	0.0958	2.1609
β_2	$\ln \frac{w_2}{w_3} \left(\frac{\text{price of labour}}{\text{price of physical capital}} \right)$	0.7531	***	0.0986	7.6403
δ_1	$\ln y_1$ (loans)	1.6343	***	0.1738	9.4049
δ_2	$\ln y_2$ (investments)	-0.4512	***	0.0933	-4.8367
δ_3	$\ln y_3$ (Other earning assets)	-0.2545	***	0.0900	-2.8289
β_{11}	$0.5 \left(\ln \frac{w_1}{w_3} \right)^2$	0.1063		0.1135	0.9361
β_{12}	$\left(\ln \frac{w_1}{w_3} \right) \left(\ln \frac{w_2}{w_3} \right)$	-0.1137		0.0963	-1.1804
β_{22}	$0.5 \left(\ln \frac{w_2}{w_3} \right)^2$	0.1544	**	0.0891	1.7337
δ_{11}	$0.5 (\ln y_1)^2$	0.2682		0.2884	0.9297
δ_{12}	$(\ln y_1)(\ln y_2)$	0.1595		0.1495	1.0665
δ_{13}	$(\ln y_1)(\ln y_3)$	-0.0528		0.1459	-0.3617
δ_{22}	$0.5 (\ln y_2)^2$	-0.3179	***	0.1089	-2.9187
δ_{23}	$(\ln y_2)(\ln y_3)$	-0.0529		0.0803	-0.6589
δ_{33}	$0.5 (\ln y_3)^2$	0.0445		0.1015	0.4379
η_{11}	$\left(\ln \frac{w_1}{w_3} \right) (\ln y_1)$	-0.1578		0.1316	-1.1987
η_{12}	$\left(\ln \frac{w_1}{w_3} \right) (\ln y_2)$	-0.0609		0.0578	-1.0535
η_{13}	$\left(\ln \frac{w_1}{w_3} \right) (\ln y_3)$	0.0049		0.0983	0.0500
η_{21}	$\left(\ln \frac{w_2}{w_3} \right) (\ln y_1)$	-0.0313		0.1303	-0.2406
η_{22}	$\left(\ln \frac{w_2}{w_3} \right) (\ln y_2)$	0.1781	***	0.0532	3.3465
η_{23}	$\left(\ln \frac{w_2}{w_3} \right) (\ln y_3)$	0.0777		0.1063	0.7306
χ_1	$\ln c_1$ (Capital ratio)	0.2898	**	0.1573	1.8423
χ_{11}	$0.5 (\ln c_1)^2$	-0.1616		0.2462	-0.6564
ρ_{11}	$\left(\ln \frac{w_1}{w_3} \right) (\ln c_1)$	-0.6483	***	0.1332	-4.8655
ρ_{21}	$\left(\ln \frac{w_2}{w_3} \right) (\ln c_1)$	0.5486	***	0.1294	4.2391

(Continued)

Table 6.3: The Maximum Likelihood Parameters Estimation for the Preferred Alternative Profit (Continued)

Parameters	Variables	Coefficient		Standard Error	t-ratio
ω_{11}	$(lny1)(lnc1)$	0.4877	**	0.2446	1.9935
ω_{21}	$(lny2)(lnc1)$	-0.4303	***	0.1430	-3.0103
ω_{31}	$(lny3)(lnc1)$	-0.1324		0.1442	-0.9179
χ_2	$lnc2$ (asset quality)	0.0031		0.0549	0.0566
χ_{22}	$0.5(lnc2)^2$	-0.0244		0.0279	-0.8755
ρ_{12}	$(ln \frac{w1}{w3})(lnc2)$	0.0231		0.0438	0.5279
ρ_{22}	$(ln \frac{w2}{w3})(lnc2)$	-0.0050		0.0445	-0.1119
ω_{12}	$(lny1)(lnc2)$	-0.0519		0.0620	-0.8371
ω_{22}	$(lny2)(lnc2)$	0.0192		0.0434	0.4425
ω_{32}	$(lny3)(lnc2)$	0.1385	***	0.0439	3.1566
χ_3	$lnc3$ (liquidity)	-3.0914	***	0.3344	-9.2447
χ_{33}	$0.5(lnc3)^2$	1.7053	***	0.5602	3.0444
ρ_{13}	$(ln \frac{w1}{w3})(lnc3)$	0.2653		0.2151	1.2333
ρ_{23}	$(ln \frac{w2}{w3})(lnc3)$	-0.0034		0.2218	-0.0151
ω_{13}	$(lny1)(lnc3)$	-0.9347	***	0.3856	-2.4240
ω_{23}	$(lny2)(lnc3)$	-0.1344		0.2156	-0.6236
ω_{33}	$(lny3)(lnc3)$	0.0951		0.1719	0.5533
χ_{12}	$(lnc1)(lnc2)$	0.1982	***	0.0814	2.4348
χ_{13}	$(lnc1)(lnc3)$	-1.4429	***	0.4352	-3.3156
χ_{23}	$(lnc2)(lnc3)$	0.0741		0.0979	0.7561
τ_1	lnt (time trend)	-0.1549	*	0.0978	-1.5834
τ_{11}	$0.5(t)^2$	-0.1469	*	0.1088	-1.3507
φ_{11}	$(ln \frac{w1}{w3})(lnt)$	0.1787	**	0.0807	2.2152
φ_{21}	$(ln \frac{w2}{w3})(lnt)$	-0.1629	**	0.0801	-2.0339
κ_{11}	$(lny1)(lnt)$	-0.0368		0.1047	-0.3519
κ_{21}	$(lny2)(lnt)$	0.1189	**	0.0692	1.7197
κ_{31}	$(lny3)(lnt)$	-0.0521		0.0656	-0.7939

(Continued)

Table 6.3: The Maximum Likelihood Parameters Estimation for the Preferred Alternative Profit (Continued)

Parameters	Variables	Coefficient	Standard Error	t-ratio
ξ_{11}	<i>(lnt)(lnc1)</i>	0.0453	0.0947	0.4780
ξ_{12}	<i>(lnt)(lnc2)</i>	-0.0491	0.0497	-0.9879
ξ_{23}	<i>(lnt)(lnc3)</i>	-0.1215	0.1609	-0.7551
ψ	<i>constant</i>	200.3159	300.0623	0.6676
ψ_1	<i>Ownership structure (z1)</i>	-0.2366	1.1341	-0.2086
ψ_2	<i>Specialisation (z2)</i>	2.4849	3.2430	0.7663
ψ_{3a}	<i>FSMP phase 2 (2004 – 2007)(z3a)</i>	2.7111	3.4151	0.7939
ψ_{3b}	<i>FSMP phase 3 (2008 – 2011)(z3b)</i>	11.6769	18.3610	0.6360
ψ_4	<i>Global crisis(2008 – 2010)(z4)</i>	-4.0642	3.9827	-1.0205
ψ_5	<i>Asset size (z5)</i>	-9.9378	13.9321	-0.7133
ψ_{6a}	<i>Market concentration (z6a)</i>	4.7069	16.4860	0.2855
ψ_{6b}	<i>Market share (z6b)</i>	8.5324	13.1460	0.6490
	Log-likelihood		-139.55	
	Log-likelihood ratio against OLS model		-143.37266	
	Number of observations		354	
	Degrees of freedom		300	

Notes: t-ratio is defined as ratio of estimated coefficient over standard errors. Degrees of freedom is defined by netting off number of observations in the sample against the number of estimated coefficients in the maximum likelihood model. ***, **, * is the significance level of estimated coefficients in given degrees of freedom. Using t-ratio critical value at significant levels of 1% (***) , 5% (**) and 10% (*) which are at 2.326, 1.645 and 1.282 respectively.

6.3 Testing of Research Hypotheses

Following the analysis and testing of SFA preferred models and the parameters of the estimations, the hypotheses developed in chapters 4 and 5 are discussed in this section, using the results derived from the three respective models. The hypotheses are as follows:

- Hypothesis 1* *Banking efficiency has improved since the implementation of the Financial Sector Master Plan (strengthening and liberalisation measures) in the Malaysian banking system between 2000 and 2011.*
- Hypothesis 2* *Foreign banks are more cost (profit) efficient than domestic banks.*
- Hypothesis 3* *Conventional banks are more cost (profit) efficient than Islamic banks.*
- Hypothesis 4* *The impact of the global economic slowdown led to lower cost (profit) efficiency of Malaysian banks.*
- Hypothesis 5* *Large banks are more cost (profit) efficient than small banks.*
- Hypothesis 6a* *High concentration in banking sector exhibits lower cost (profit) efficiency of Malaysian banks.*
- Hypothesis 6b* *High market share exhibits higher cost (profit) efficiency of Malaysian banks.*
- Hypothesis 7* *Banks with a lower capital ratio exhibit higher costs and lower profits.*
- Hypothesis 8* *Banks with a higher non-performing loans' ratio exhibit higher costs and lower profits.*
- Hypothesis 9* *Banks with greater liquidity exhibit lower costs and higher profits.*

From the discussions in chapter 5, and as mentioned in Coelli et al. (2005), exogenous variables are used to measure the heterogeneity effect on banks' costs and profits. These variables are known as control and environmental variables. Control variables are assigned directly into the frontier regression and interact entirely with the inputs and outputs, which influence the cost and profit frontier structure. Based on hypotheses 6, 7 and 8 the control variables related to this study are: financial capital (C1), asset quality (C2) and liquidity (C3). Since the control variables are assigned directly into the regression, the most appropriate method in testing these hypotheses is via the t-ratio test of the estimated coefficient of the control variables (Coelli et al., 2005).

With regards to hypotheses 2 to 5, similarly to the control variables, inputs and outputs, a t-ratio test is used to measure the significant level of the respective environmental variables. If the coefficient associated with the environmental variable displays a positive relationship and is significant, it denotes that the environmental variables' inefficiency (u_i) has increased. On the other hand, if the coefficient of the environmental variables is negative and significant, it implies that the inefficiency (u_i) has decreased. Unlike with the control variables, the environmental variables do not interact with the inputs and outputs of the frontier estimation and do not affect the structure of the frontier but impact the measurement of inefficiency (u_i). Thus, the direction of inefficiency (u_i) can be determined from the coefficient and significance level of the environmental variables involved in this study, which at the same time are used to test hypotheses 2 to 5 (Battese and Coelli, 1995).

6.3.1 Testing of Hypotheses 2 to 5

In testing hypotheses 2 to 5, the environmental variables are used to determine whether the hypotheses can be accepted or rejected. Furthermore, the t-ratio of the respective coefficients of the environmental variables is assessed for significance tests. To determine the direction of the relationship between inefficiency and the environmental variables, the signs of the coefficients are analysed (i.e. either positive or negative).

Hypothesis 2 Foreign banks are more cost (profit) efficient than domestic banks.

The main environmental variable used to test this hypothesis is ownership ($Z1$). For cost efficiency, the coefficient of (ψ_1) in table 6.1 is positive 1.09 and significant at the 1% significance level. Hence, the coefficient and t-ratio of this variable suggests that the foreign banks are less cost efficient (i.e. more inefficient) than domestic banks. This finding is surprising because the past literature on ownership structure, has in many cases suggested that foreign banks are more efficient when compared to domestic banks, particularly in emerging / developing countries (Claessen, 2001). This however could be explained by the various restrictions established by BNM on locally incorporated foreign banks since the independence of Malaysia such as, restrictions in entering the banking sector, limited branch

activities and restrictions to operate in several areas of banking (e.g. hire purchase) (see Chapter 4).

From the profit efficiency perspective, the standard profit efficiency and alternative profit efficiency exhibit positive and negative coefficients for ownership variables, respectively. For standard profit efficiency (Table 6.2), Z1 is significant at 1% significance level; however, it is insignificant for the alternative profit efficiency. This significant positive coefficient of ownership (ψ_1) (1.13) may indicate that foreign banks are more standard profit inefficient compared to domestic banks. Hence, this hypothesis is rejected on the basis that foreign banks are expected to gain greater profit efficiency compared to the domestic banks. In terms of the alternative profit efficiency model, the coefficient of ownership (ψ_1) displays a negative relationship (-0.23) but is insignificant. This could indicate that the foreign banks are more profit efficient than domestic banks. For instance, in developing countries, foreign banks are known to ‘cherry-pick’ good and quality customers as well as having profit oriented behaviour, which could give better returns to the banks; whereas, domestic banks can be describe as ‘mass’ financial institutions which serve the public and in turn this could leave domestic banks with lower profits (Hassan, 2002). However, the hypothesis for alternative profit efficiency is rejected as there is no difference (insignificant t-ratio) between foreign and domestic banks (the efficiency scores regarding ownership (i.e. domestic and foreign banks) are discussed in detail later in this chapter).

Hypothesis 3 Conventional banks are more cost (profit) efficient than Islamic banks

With the dual banking environment in Malaysia arising from the growth of Islamic banks, this hypothesis intends to test whether the conventional banks are more cost and profit efficient than the Islamic banks. The environmental variable Z2 is used for this purpose and a dummy variable is assigned to the Islamic banks. From Table 6.1, the coefficient of specialisation (ψ_2) for the cost efficiency model revealed a positive relationship with inefficiency (significant at 10%). The Islamic banks are more cost inefficient when compared to their conventional counterparts. This could be attributed to the introduction of new foreign Islamic bank subsidiaries to enhance the prominence of Islamic banking in

Malaysia which might have increased competitive pressure on the industry in the second phase of the FSMP (2004–2007).

The standard and alternative profit efficiency (ψ_2) displays positive coefficients; but only standard profit efficiency is significant at 1% significance level. These results suggest that Islamic banks are more profit inefficient than conventional banks. The lower profit efficiency of the Islamic banks is probably influenced by their briefer period of establishment compared to the conventional banks. Except for Bank Islam, which has been in operation for more than 20 years, the Islamic bank subsidiaries and foreign Islamic banks are at a relatively nascent stage, and have yet to optimally manage their production of outputs and inputs (Kamaruddin et al., 2008; Abdul Majid et al., 2011). Additionally, with the introduction of foreign Islamic banks in the second phase of the FSMP, the Islamic banking concepts from the Middle East were new to Malaysian consumers; at this early stage, products with specific shariah ‘jargon’ are not easy for consumers to comprehend. Furthermore, Islamic banks require specialised knowledge and skills, which could result in lower profit efficiency (Abdul Majid., 2001) (the efficiency scores between Islamic and conventional banks are discussed in detail later in this chapter).

Hypothesis 4 The impact of the global economic slowdown led to lower cost (profit) efficiency in Malaysian banks

This hypothesis relates to the impact of the global economic slowdown due to the credit crisis which occurred in the US and Europe namely, the subprime and sovereign debt crises. The impact was felt in Malaysia from end of 2008 until early 2010 (see Chapter 4). To capture the effect of the global credit crisis on Malaysia, a dummy variable is assigned to years 2008, 2009 and 2010. From the cost efficiency parameters estimation, the coefficient of variable global crisis (ψ_4) is found to be positive 0.62 and significant at 5% significance level. This may well indicate that the global credit crisis resulted in lower cost efficiency in Malaysian banks. As discussed in Chapter 4, the global economy deteriorated further in 2008 and the demand for exports declined, which affected the real sector in Malaysia. Malaysian GDP dropped and contracted in 2009 (Bank Negara Malaysia, 2009). During this period of uncertainty, higher costs were incurred by Malaysian banks in managing excess

liquidity from the inflow of foreign funds, and preventing loan portfolios from going delinquent. For standard profit and alternative profit efficiencies, the coefficients for the financial crisis variable (ψ_4) were insignificant and negative -0.36 and -4.06, respectively. Thus, the hypothesis is rejected for both standard profit and alternative profit efficiency models. The negative coefficients indicate that Malaysian banks show higher profit efficiency during the period of the global credit crisis. This could be likely explained from the effect of lower interest environment, with a couple of economic stimulus introduced by the government of Malaysia.

Hypothesis 5 Large banks are more cost (profit) efficient than small banks

This hypothesis relates to asset size variable (Z6) (see Chapter 5). It is hypothesised that bigger banks have a positive relationship with cost and profit efficiency. The result stemming from the cost frontier estimation indicates that the coefficient of bank size (ψ_6) in table 6.1 is positive 2.45 and significant at 1% significance level. Thus, hypothesis 5 is rejected due to larger banks being less cost efficient than smaller banks. According to Akhigbe and McNulty (2003), in situations where the banking industry faced financial liberalisation, large banks were found to be less efficient than small banks. This may be because large banks needed to invest in major restructuring and cost-cutting initiatives, more so than the small banks. During the first phase of the FSMP (2000–2003), domestic banks needed to consolidate and these recently expanded entities suffered higher costs than the smaller banks (see Chapter 4) (Suffian, 2004; Bonin et al., 2005). These new consolidated banks were also publicly traded and with diverse ownership comes with greater agency costs which could result in lower cost efficiency scores (Kwan, 2006). In addition, large banks are less efficient than smaller banks because they normally enjoy competitive market power, in which their managers do not work as hard to keep costs under control and enjoying ‘quiet life’ (Berger and Hannan, 1998).

The standard profit and alternative profit efficiencies displayed negative (-1.03 and -9.93, respectively) and insignificant coefficients for asset size (ψ_6). The directions indicated that larger banks faced a lower level of profit efficiency compared to smaller banks; however the t-ratios for these coefficients are insignificant. This leads us to reject this hypothesis on

standard and alternative profit efficiencies (the efficiency scores regarding bank size are discussed in detail later in this chapter).

- Hypothesis 6a*** ***High concentration in banking sector exhibits lower cost (profit) efficiency of Malaysian banks***
Hypothesis 6b ***High market share exhibits higher cost (profit) efficiency of Malaysian banks***

Hypotheses 6a and 6b relate to market structure. The variables representing market structure are market concentration (Z7) and market share (Z8). Based on the SCP hypothesis, market structure can be measured via market concentration, market share, product differentiation, barriers to entry and exit, vertical integration and diversification. In view of the hypotheses regarding the market structure, two variables are employed to test the impact to cost- and profit-efficiency. Berger (1995) argued that there is a positive statistical relationship between measures of market structure (concentration and market share) and profitability. In this regard, the positive relationship between profitability and market structure elements can be explained by two market power hypotheses: namely, the traditional SCP hypothesis and the relative-market power (RMP) hypothesis. In the context of the banking industry, the SCP hypothesis says that banks can set prices that are less favourable to consumers in a concentrated market, due to imperfect competition within the market. Banks act in a monopolistic manner in this imperfect competitive market structure by lowering the deposit rates and increasing the loan rates. On the other hand, the RMP hypothesis asserts that only firms with a large market share and well-differentiated products can exercise market power in pricing these products and earn supernormal profits (Shepherd, 1982).

For hypothesis 6a, a high market concentration exhibits lower cost/profit efficiency of banks. For the cost efficiency estimation, the result has shown a positive coefficient (3.68) but insignificant, implying that the higher the concentration of the market, the lower the cost efficiency of Malaysian banks. Hence, the null hypothesis 6a for cost efficiency is rejected due to insignificant t-ratio value. The standard profit- and alternative profit- efficiency were also insignificant and exhibited positive coefficients. Hence, hypothesis 6a, for standard profit- and alternative profit-efficiency, is also rejected.

Hypothesis 6b relates to the market share of the banks. RMP hypothesis is employed to test this hypothesis, in which larger banks are expected to exercise their competitive market power through pricing to earn abnormal profits (Shepherd, 1982). The result from Table 6.1 for cost efficiency shows a negative coefficient (-2.9) of variable market share (ψ_8) and a notable 1% significance level. This evidence seems to support the RMP hypothesis under which banks with large market share and well-differentiated products exert more market power to maximise profits and performance (Shepherd, 1982). On the other hand, the standard profit and alternative profit efficiencies display negative and positive coefficients, respectively, but are insignificant (see tables 6.2 and 6.3). Hence, hypothesis 6b for the standard profit- and alternative profit-efficiency are rejected.

6.3.2 Testing of Hypotheses 7 to 9

From the earlier discussions, three control variables (capital adequacy (C1), asset quality (C2) and liquidity (C3)) are introduced in this study, which relate to hypotheses 7 to 9. These hypotheses are related to different types of risks that could affect Malaysian banks' levels of efficiency. Following Battese and Coelli's (1995) 1-stage model, the control variables in this study have a direct influence on the frontier structure (i.e. control variables interact directly with input and output variables). Therefore, with an increasing number of variables being placed into the model specification, multicollinearity issue may arise (Gujarati, 2003). However, as mentioned earlier, multicollinearity problems might not be severe if used for forecasting purposes, particularly when the estimation results produce expected signs on the coefficients which may also suggest that multicollinearity issues may not be a critical issue (Belsley, 1984). Similarly to the other variables, the t-ratio is used to test the significance of the variables and the direction of the relationship is determined from the coefficient signs.

Hypothesis 7 Banks with lower financial capital ratio (C1) (low capital adequacy) exhibit higher cost and lower profits

Hypothesis 8 Banks with greater percentage of non-performing loans (C2) (low asset quality) exhibit higher cost and lower profits

Hypothesis 9 Banks with lower loans to deposit ratio (C3) (high liquidity) exhibit lower cost and higher profits

First, based on the cost efficiency model, the preferred model is the general model without capital adequacy (C1) (stage A2.3) (see Table 5.5). Since the preferred cost efficiency model (A2.3) has a better fit than the general model (A1.0), hypothesis 7 is rejected. Hence, from the preferred cost model (A2.3), only two control variables are to be tested; asset quality (C2) and liquidity (C3). For asset quality (C2), the coefficient for variable asset quality (χ_2) was positive (0.1396) but insignificant. The direction of this coefficient is aligned to past literature where banks with low asset quality or high NPLs are likely to exhibit higher costs (Fu and Haffernan, 2007; Mester, 1997; Pastor, 2010). As shown in Chapter 4, the NPL levels of Malaysian banks were low but they faced higher operating costs during the global credit crisis. This may suggest that the banks were assigning more attention and resources to loan origination, monitoring and other credit judgement activities, learning from their experience of economic turmoil in 1997–98 (Girardone et al., 2004). Moreover, during the global credit crisis, in the third phase of FSMP, banks were actively monitoring and negotiating potential loan delinquencies through reclassification and restructuring of loans (Abd. Karim et al., 2010). However, the coefficient (χ_2) is insignificant and may not directly affect the cost structure of banks, which leads to hypothesis 8 being rejected. Additionally, the coefficient for liquidity risk (χ_3) (measured using LTD ratio, see Chapter 5) showed a negative relationship (-0.83) and positive at the 1% significance level. This suggests that a 1% reduction of liquidity risk would result in an increase of 0.83% in cost. From the discussions made in Chapter 5, a higher level of liquidity (i.e. lower liquidity risk) in banks could reduce costs. However, conflicting findings are observed where a high level of liquidity could result in higher operational costs of banks. This can be likely caused by higher price of funds and increasing utilisation of resources in managing and matching maturities of assets and liabilities from the excess of liquid assets from substantial influx of foreign funds into Malaysian banks in 2007 (see Chapter 4). Despite being significant, at 1% significance level, hypothesis 9 is rejected due to the opposite sign of coefficient (χ_3) expected for the cost efficiency model.

Second, the preferred model for standard profit efficiency is the general model, which includes all the control variables and environmental variables. From the preferred model, the coefficients for capital adequacy (χ_1) and asset quality (χ_2) are found to be negative (-0.41) and (-0.18) respectively, and significant at 5% significance level (hypothesis 7 and 8). However, the coefficient for liquidity (χ_3) is insignificant (hypothesis 9). The negative coefficient (0.41) of capital adequacy (χ_1) suggests that an increase in capital would reduce profits. The result however, differs from the expectation derived in hypothesis 7; and is hence rejected. This could be potentially explained from banks being over-capitalised during the consolidation of domestic banks and during the global credit crisis, which took place in the first and third phases of the FSMP (see Chapter 4). A higher capital level implies that banks face low inherent capital risk. But if banks are over-capitalised, they might face higher capital cost with declining profits, due to higher agency costs, as well as increased costs in the form of dividends to be paid to shareholders (Kashyap et al., 2010; Baker and Wurgler, 2013). For asset quality (C2), the coefficient (χ_2) shows negative (-0.18) and significance at 5%. Hence hypothesis 8 cannot be accepted, because it supports the notion that an increase in the NPL ratio could result in declining profits. The coefficient (χ_2) indicates that a 1% increase in NPLs (low asset quality) will result in a 0.18% decline in profit efficiency. In terms of liquidity (C3), the coefficient for liquidity is positive and insignificant. Therefore, hypothesis 9 for standard profit efficiency is rejected.¹⁶⁶

On the third economic function, the preferred model for alternative profit efficiency is also the general model, which includes all control and environmental variables (see Table 5.7). The coefficient for capital adequacy (χ_1) is positive (0.28) and significant at 5% significance level. Hence, the coefficient and t-ratio of the capital adequacy variable fails to reject hypothesis 7. For the case of the alternative profit efficiency function, a lower level of capital adequacy could result in lower profitability. According to Mester (1993, 1996), efficiency is positively correlated with financial capital. Generally, a greater capital adequacy describes a bank which has a larger margin of safety and a greater ability to absorb potential risks. Hence, a bank with a strong capital level normally enjoys reliable

¹⁶⁶ Similar to cost efficiency model, banks could have incurred higher cost in managing excess liquidity, which consequently reduces profits.

access to sufficient sources of funds on favourable terms (Berger, 1995). Moreover, highly efficient banks normally have more profits and are able (by holding dividends constant) to retain earnings as capital (Girardone et al., 2004). The coefficient for asset quality (χ_2) displays the negative sign and is insignificant. The asset quality for the preferred alternative profit efficiency is unable to fully influence the effect on profitability of Malaysian banks. Thus, hypothesis 8 is rejected. On the other hand, the coefficient for liquidity (χ_3) is negative and significant at a 1% significance level. Similar to preferred cost efficiency model, a high level of liquidity (low LTD ratio) may result in higher profit. Therefore, hypothesis 9 is accepted, as high liquidity can lead to higher profits. As evidenced, in 2007 the liquidity of Malaysian banks increased due to large inflows of foreign funds. Malaysian banks probably earn greater income from conversion of excess funds into short-term assets.

Table 6.4 Testing of Research Hypotheses 2–9

	Null Hypothesis	Cost Efficiency	Standard Profit Efficiency	Alternative Profit Efficiency
Hypothesis 2	Foreign banks are more cost (profit) efficient than domestic banks.	Reject	Reject	Reject
Hypothesis 3	Conventional banks are more cost (profit) efficient than Islamic banks.	Accept	Accept	Reject
Hypothesis 4	The impact of the global economic slowdown led to lower cost (profit) efficiency of Malaysian banks.	Accept	Reject	Reject
Hypothesis 5	Big banks are more cost (profit) efficient than small banks.	Reject	Reject	Reject
Hypothesis 6a	High concentration in banking sector exhibits lower cost (profit) efficiency of Malaysian banks.	Reject	Reject	Reject
Hypothesis 6b	High market share exhibits higher cost (profit) efficiency of Malaysian banks.	Accept	Reject	Reject
Hypothesis 7	Banks with a lower capital ratio exhibit higher costs and lower profits.	Reject	Reject	Accept
Hypothesis 8	Banks with a higher non-performing loans' ratio exhibit higher costs and lower profits.	Reject	Accept	Reject
Hypothesis 9	Banks with greater liquidity exhibit lower costs and higher profits.	Reject	Reject	Accept

6.4 Impact of Liberalisation via the FSMP on the Efficiency of Malaysian Banks

The key research hypothesis for this study relates to the impact of gradual liberalisation via the FSMP (see chapter 4) where the trend of the cost- and profit-efficiency scores of Malaysian banks are evaluated for the period 2000 to 2011. This section shows the cost- and profit-efficiency scores and average trends of SFA for the preferred models, from 2000 to 2011. Additionally, the DEA scores for cost and profit efficiency are also analysed for the years 2000–2011.

6.4.1 SFA Efficiency Scores

6.4.1.1 SFA Cost Efficiency Scores

The average cost efficiency scores for the Malaysian banks for the years 2000–2011 are shown in Table 6.5. The average cost efficiency score for Malaysian banks between 2000

and 2011 is 82.7%. The average cost efficiency scores suggests that Malaysian banks wasted around 20.0% of their inputs to produce the same level of outputs of the best performing banks. This finding (approximately 20% inefficiency) is consistent with past findings in the literature, where SFA was used to derive the cost efficiency function (e.g De Young, 1997; Berger and De Young, 2001; Akhigbe and McNulty, 2003; Bonin et al., 2005).

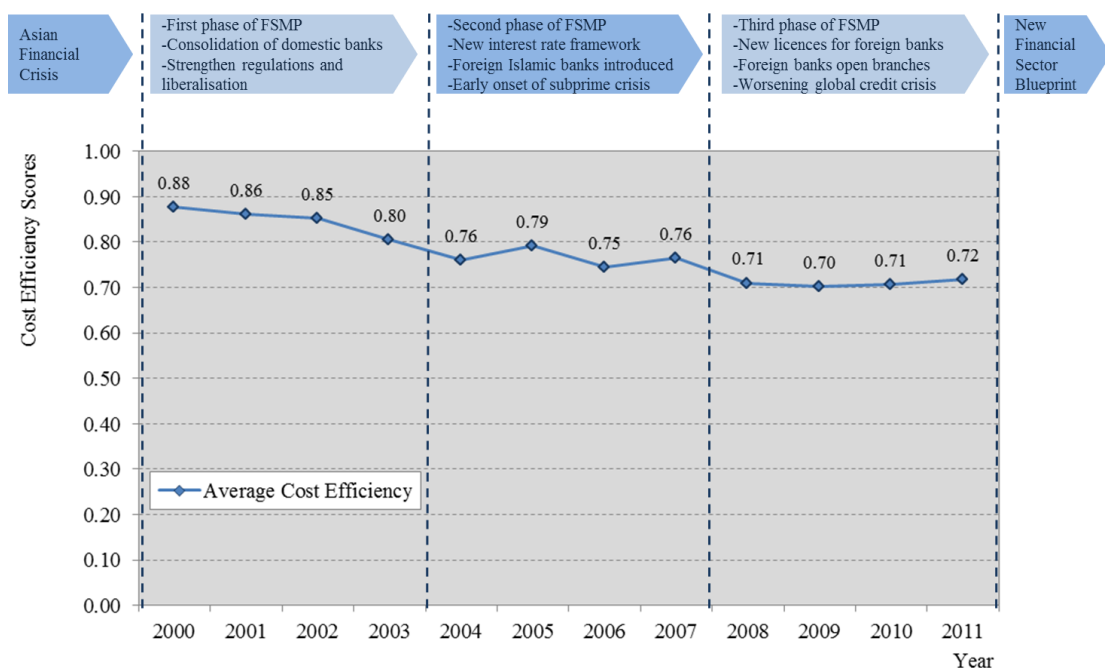
Table 6.5 Average SFA Cost Efficiency Scores, 2000–2011

Year	Count	Mean	Standard Deviation	Minimum	Maximum
2000	24	0.8760	0.1519	0.2600	0.9721
2001	26	0.8614	0.1537	0.2698	0.9787
2002	24	0.8523	0.1524	0.3253	0.9725
2003	24	0.8103	0.1528	0.3731	0.9582
2004	25	0.7596	0.1734	0.2404	0.9270
2005	26	0.7913	0.1554	0.2088	0.9418
2006	29	0.7450	0.2095	0.2033	0.9317
2007	31	0.7648	0.1440	0.2704	0.9402
2008	36	0.7098	0.1783	0.0509	0.9324
2009	37	0.7025	0.1877	0.0296	0.9271
2010	37	0.7065	0.1884	0.0738	0.9478
2011	35	0.7177	0.2157	0.0848	0.9702
2000-2003	98	0.8502	0.1523	0.2600	0.9787
2004-2007	111	0.7647	0.1707	0.2033	0.9418
2008-2011	145	0.7090	0.1909	0.0296	0.9702
2000-2011	354	0.7655	0.1834	0.0296	0.9787

From Figure 6.2, based on the three different phases of the FSMP for the years 2000–2003, 2004–2007 and 2008–2011, the cost efficiency average scores were 85.0%, 76.5% and 70.9% respectively, indicating the overall trend of the cost efficiency for 2000 - 2011 was on a downward trend. From previous literature, many studies found that deregulation of the banking sector resulted in greater efficiency (e.g Tortosa-Austina, 2003; Cuesta and Orea, 2002; Leightner and Lovell, 1998). On the other hand, deregulation of the banking sector could also result in lower efficiency as evidenced from past studies. For instance, Berger and Humphrey (1991) found that financial liberalisation forces banks to cut costs substantially in a very short period of time. However, this has not been the case as the banks

displayed a lower level of cost efficiency, and a slower response in adjusting to changes to minimise their costs (Berger and Humphrey, 1992); and, based on observations by Humphrey and Pulley (1997), they suggested that banks' adjustments following deregulation can take up to four years to complete. Additionally, Girardone et al. (2004) also state that banks may not be efficient when deregulatory initiatives take place at the same time as a macroeconomic downturn.

Figure 6.2 Average SFA Cost Efficiency Scores of Malaysian Banks, 2000–2011



The first phase of the FSMP (2000–2003) is regarded as the period of initial reform in the Malaysian banking industry. After the financial crisis in 1997–98, the Malaysian government took various drastic measures to improve the banking sector (Abdul Majid et al., 2011). Against the backdrop of this crisis, there were significant structural changes in the Malaysian banking sector (Sufian, 2004). For instance, Malaysia did not rely on assistance from the International Monetary Fund (IMF) after the financial crisis, unlike some other Association of South East Asia Nations (ASEAN) member countries. Under the IMF programme, insolvent banks were forced to close down, but Malaysia did not take this path as the social cost involved - in terms of a dislocation of resources - would have been high. Malaysia took a different approach by introducing a consolidation of the banking institutions, in which BNM played an intermediary role, solving issues of fairness to all

parties involved in the merger (Abd. Kadir et al., 2010).¹⁶⁷ During this period (i.e. first phase of FSMP (2000-2003), cost efficiency scores were trending downward. This could be potentially explained by the effect of the consolidation of domestic banks, which resulted in declining cost efficiency scores. These domestic banks had to implement various rationalisation programmes including; restructuring of duplicated branch networks, managing staff redundancy, synchronising technology with the acquiring partner and implementation of internet banking services (Sufian, 2004).

During the FSMP's second phase (2004–2007), BNM introduced the new interest rate framework (NIRF) that aimed to facilitate more efficient pricing of financial products. Following the removal of BNM's intervention rate, Malaysian banks were able to price their funding costs and revenues based on their own cost structure and compete for customers using their own interest pricing structure.¹⁶⁸ This deregulation of interest rates was intended to increase efficiency, productivity, innovation and profitability in the banking system (Leightner and Lovell, 1998; Sensarma, 2006, Berger and Mester, 2003). As a result of the deregulation of interest rates, banks have to adjust their inputs and outputs to remain competitive.¹⁶⁹ During this period, there was a slight increase in the cost efficiency scores in 2005, following the liberalisation of BLR, indicating some increasing level of competition in terms of pricing among Malaysian banks. Towards the end of the second phase of the FSMP, the cost efficiency scores fell marginally due to a significant loss faced by an Islamic bank and the inception of new foreign Islamic banks (see Chapter 4). The overall average cost efficiency scores worsened because these new or de-novo foreign Islamic banks inherently faced higher operational costs during their early phase of operations (Suffian, 2010).

¹⁶⁷ This consolidation programme was also in line with the requirement in having stronger domestic banks to compete regionally when opening its financial industry to the international players in 2003 under the World Trade Organisation (WTO) rules (Abd. Kadir et al., 2010; Sufian and Habibullah, 2009). As a result, in 2001, the consolidation had successfully merged 54 Malaysian banks and financial institutions into 10 anchor banking groups (Dogan and Fausten, 2002; Radam et al., 2009; Ismail and Abdul Rahim, 2009).

¹⁶⁸ The interest rate deregulation program generally increases competitive pressure in the market and forced banks to reduce their cost (Mester, 1993).

¹⁶⁹ In terms of adjustments made to inputs and outputs, Humphrey and Pulley (1997) found that during the deregulation of interest rate, banks tend to respond in three ways. First, to offset higher deposit interest cost with higher explicit and implicit for small deposits. Second, to transfer the higher funding cost to borrowers. And third, to invest in risky assets to obtain higher yield.

In the third phase of the FSMP (2008–2011), Malaysian banks were not significantly affected during the initial stages of the subprime crisis in the US. Malaysian banks were prudent in their investments, particularly relating to derivatives products originated in the US and Europe (only a small portion of these instruments were held by them) (Bank Negara Malaysia, 2007b). However, in mid-2008, as the global economy deteriorated, the demand for Malaysian exports declined, which affected the real sector. Malaysian GDP had consequently contracted by 1.7% in 2009 (Bank Negara Malaysia, 2009b). Consequently, the NPLs of banks increased slightly in 2009, reflecting the contraction experienced by the economy (see Chapter 4). The downward trend of the cost efficiency scores during the third phase of FSMP was also driven by greater operating costs by Malaysian banks in managing potentially delinquent loans (Bank Negara Malaysia, 2009b).

At the same time (i.e. during the third phase of FSMP (2008-2011)), BNM reduced its policy interest rate (overnight policy rate or OPR) from 3.5% to 2.0% in November 2008 to February 2009 (Bank Negara Malaysia, 2009). The policy reduction in interest rates required banks to drastically adjust their input prices and outputs according to the indicative market interest rate (i.e. the OPR). Similar to Berger and Humphrey (1992), a slower response in adjusting to market changes, had resulted in Malaysian banks experiencing lower cost efficiency during the third phase of the FSMP (2008-2011). In addition, Basel II was also implemented between 2008 and 2010 (both Standard Approach and Internal Ratings Based Approach) and Malaysian banks invested heavily in technology, physical assets, external consultants and specialised labour to comply with the new capital regulation (Bank Negara Malaysia, 2008b, 2010b).

From Table 6.1, the variables (Z3) and (Z4) represent the impact of liberalisation made on two different phases of the FSMP. Based on the years 2004 to 2008, the variable Z3 showed a negative coefficient (-0.40), but insignificant. Consequently, the dummy variable Z4 (second phase FSMP, 2008–2011) displayed the similar result to Z3 which has negative coefficient (-0.52), and insignificant. This leads us to reject hypothesis 1 (i.e. t-ratios being statistically insignificant), in which the cost efficiency of Malaysian banks have not improved following the implementation of financial liberalisation.

6.4.1.2 SFA Standard Profit- and Alternative Profit-Efficiency Scores

The average standard profit and alternative profit efficiency scores of the Malaysian banks from 2000 to 2011 were 62.2% and 93.3%, respectively (see Table 6.6). Based on the average of standard profit efficiency scores, banks' profit levels could be increased by approximately 40.0% with the same level of output prices. The average standard profit efficiency scores for the three different phases of FSMP are 66.3% (2000-2003), 65.9% (2004-2007) and 56.7% (2008-2011). On the other hand, the alternative profit efficiency scores for the first (2000–2003), second (2004–2007) and third (2008–2011) phases of the FSMP are higher than average standard profit efficiency scores at 93.2%, 96.3% and 91.2%, respectively.

Figure 6.3 Average SFA Standard Profit and Alternative Profit Efficiencies Scores of Malaysian Banks, 2000–2011

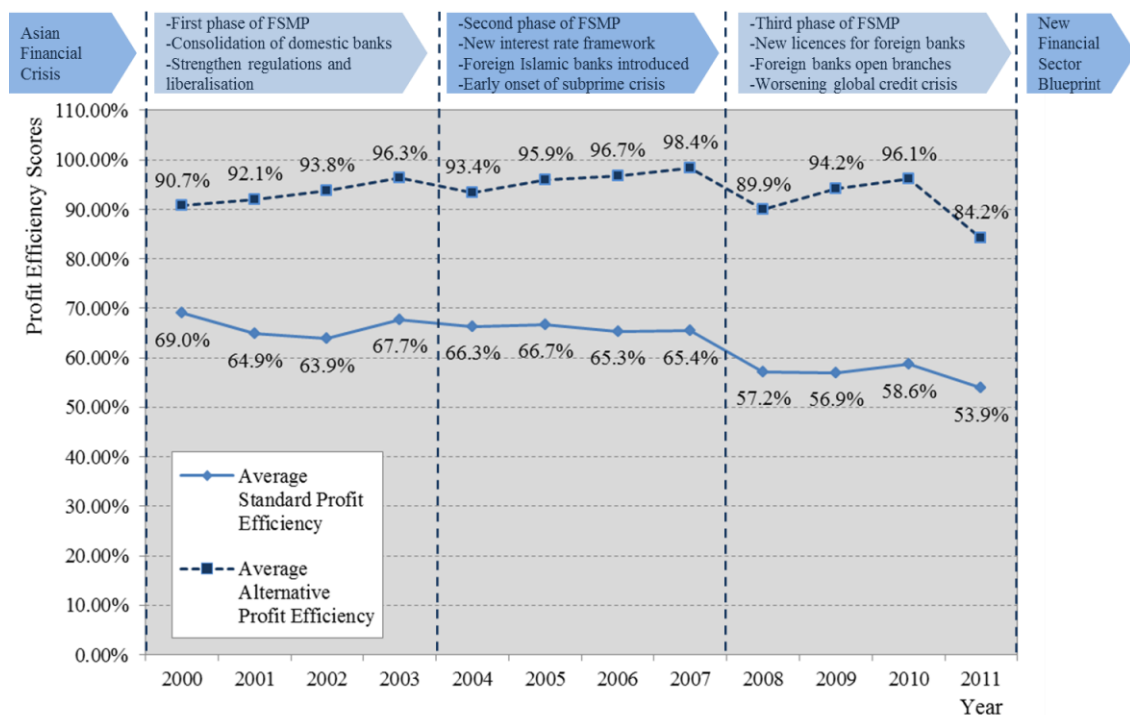


Figure 6.3 shows an increasing trend in the alternative profit efficiency scores but decreasing standard profit efficiency in the first phase of the FSMP (2000-2003), following consolidation of the domestic banks. This could be explained due to the effect of imperfect market. During this period, significant efforts were directed towards the ongoing

transformation of the domestic banking system into one with the capacity and agility to withstand shocks and survive increased financial market volatility (Bank Negara Malaysia, 2005). When domestic banks merged, the fragmented market shares held independently by each institution were consolidated, which resulted in a larger market share for these banks, which then enjoyed the advantages of economies of scale.¹⁷⁰ In terms of profit efficiency models, Akhavein et al. (1997) assert that bank mergers increase profits in three major ways. First, they improve cost efficiency by reducing costs per unit in the production mix of inputs and outputs. Second, bank mergers increase profits through profit efficiency improvements: profits come to be seen as a superior combination of inputs and outputs;¹⁷¹ finally, bank mergers improve profits due to market power – the power to set prices.¹⁷²

During the second phase of the FSMP (2004–2007), the deregulation of interest rates (i.e. NIRF), introduced by BNM in 2004, resulted in a slight increase in standard profit- and alternative profit-efficiency in the following year (i.e. 2005). Banks responded to NIRF by adjusting their loan and deposit prices. This can be evidenced by the interest margin trend after deregulation of the interest rates, in which interest profit margins increased from 43.6% to 46.9% between 2004 and 2005 (see Table 4.19), indicating some level of market power was exercised by some banks in setting prices. From Figure 6.3, standard profit- and alternative profit-efficiency scores demonstrate different trends when responding to deregulation initiatives introduced in the second phase of the FSMP. The average score for alternative profit efficiency scores increased from 93.2% in the first phase to 96.3% in the second. On the other hand, standard profit efficiency's average score worsened from 66.3% to 65.9% in the first and second phases of FSMP, respectively. The differences between these two profit efficiency concepts can be explained by Berger and Mester (1997) (see Chapter 5), who argued that the differences could be viewed from the underlying

¹⁷⁰ For instance, bank mergers can bring new technology and management skills for greater efficiency. Piloff (1996) suggests that greater performance after mergers could be obtained in several ways. First, there will be a transfer of management skills coming from the superior firm and can contribute or provide complimentary skills in the acquired firms. Second, the mergers can improve the financial performance by eliminating the redundant facilities and human resources; and third, there will be a consolidation of technology, skills and resources when firms merged.

¹⁷¹ Profit efficiency is a more inclusive compared to cost efficiency as it involves both revenue and cost. A new merged entity could improve their profit efficiency without improving their cost efficiency, particularly when the outputs associated with the merger increases revenue more than the cost (Akhavein et al., 1997).

¹⁷² By increasing the market share, the new merged institution can charge higher prices for its goods and services which, in return, could increase profitability from the customer surplus without even improving the efficiency (Piloff, 1996; Akhavein et al., 1997).

assumption of each model with respect to market power and unmeasured product quality.¹⁷³ This may also suggest that the competition among Malaysian banks was imperfect (i.e. oligopoly market structure). For example, the lower average alternative profit efficiency score in 2004 (during the introduction of new interest rate framework (NIRF)) implies that some banks with higher market share experienced higher profits by exercising market power in setting input and/or outputs prices to maximise profits. However, as competitors have good knowledge of market demand; and learning from their mistakes, they soon reacted by changing their input or/and outputs prices (i.e. price interdependence), which led to an increasing trend of the average alternative profit efficiency from 2005 to 2007.

Towards the end of the second phase, Malaysian banks progressively faced greater competition from foreign banks and new Islamic bank subsidiaries. However, these banks were unable to provide the necessary competitive pressure during their early stage of operations. In late 2007, Malaysian banks experienced large inflows of foreign funds due to the subprime and sovereign debt crises that occurred in the US and Europe. This led to an abundance of liquidity flowing into the Malaysian banking system, which resulted in higher funding costs. With difficulties in searching for appropriate assets to match these liquidities, banks faced lower revenues and, as a result, lower profit efficiency scores. Hence, a sharp decline in standard profit- and alternative profit-efficiency was observed in 2008. As Malaysian banks began to experience the effect of the global economic crisis of late 2008, the OPR was reduced from 3.5% to 2%, as part of a series of measures taken by BNM to lessen the cost of borrowing and assist consumers in weathering the challenging business and economic environment (Bank Negara Malaysia, 2009a). In this environment, banks were required to adjust their input prices, outputs and output prices according to the changes in policy interest rate (i.e. the OPR), resulting in a gradual increase of both standard profit- and alternative profit-efficiency in 2009 and 2010. As the timeline moved towards the end of the third phase of the FSMP, the Malaysian GDP improved and both profit efficiencies

¹⁷³ According to Berger and Mester (1997) and Maudos and Pastor (2003), the differences between alternative profit- and standard profit-efficiency can be explained with reference to market power, scale and scope economies, unmeasured quality of banking services, and inaccurate measurement of output prices (see Chapter 5). Alternative profit efficiency assumes that banks have some degree of market power in determining the output prices; therefore these output prices are not explicitly specified and are assumed to be endogenous in output variables. However, for standard profit efficiency, perfect competition is assumed and banks have no impact in determining the output prices. Hence, output prices are explicitly specified for in the profit function (Berger and Mester, 1997).

displayed a declining trend in 2011. These results were influenced by the rapid increase in the OPR over three consecutive quarters, which caused complexities for banks in managing their input prices, output prices and outputs. The pass through monetary policy transmission of interest rate may also indicate that Malaysian banks were more responsive in an environment of declining interest rates (i.e. changes in the OPR in 2008) rather than increasing interest rate (i.e. changes in the OPR in 2011) (see Figure 6.3).¹⁷⁴

Based on two variables, (Z3) and (Z4), that reflected the gradual liberalisation in Malaysia (i.e. the three phases), the coefficients (ψ_3) and (ψ_4) of standard profit efficiency exhibit a positive relationship of 0.2102 and 1.4436, respectively, but were insignificant. For the alternative profit efficiency model, the result showed a similar result to standard profit efficiency, which had positive coefficients on (ψ_3) and (ψ_4) of 2.7111 and 11.6769 respectively, but were insignificant. The results stemming from these coefficients may suggest that the liberalisation measures introduced in Malaysia by BNM were unable to improve the scores of standard profit- and alternative profit-efficiency. This occurrence is attributed to two factors: de-novo (new) foreign Islamic banks and the global credit crisis. First, the introduction of new foreign Islamic banks during the second phase of the FSMP (2004–2007) did not result in an improvement in profit efficiency as these banks were unable to generate adequate profits due to their significantly higher operational and capital expenditures, incurred during early operation. Second, the impact of the global economic turmoil on Malaysia was stronger than the liberalisation initiatives introduced by BNM, causing the efficiency scores to decline during the third phase of FSMP (2008–2011).

¹⁷⁴ Lim (2001) asserts that banks generally value their borrowing customers and pass on decreases in loan rates faster than they pass on increases. In term of deposits banks possess greater rigidity in the passing on of increases to deposit rates compared to the passing on of decreases.

Table 6.6 Average SFA Standard Profit- and Alternative Profit-Efficiency

Year	Count	Standard Profit Efficiency				Alternative Profit Efficiency				
		Mean	Std. Dev.	Min.	Max.	Count	Mean	Std. Dev.	Min.	Max.
2000	24	0.6903	0.2175	0.2304	0.9443	24	0.9069	0.0948	0.6504	0.9882
2001	26	0.6485	0.2268	0.1871	0.9438	26	0.9206	0.0796	0.7141	0.9894
2002	24	0.6386	0.2536	0.1243	0.9472	24	0.9381	0.0755	0.7115	0.9923
2003	24	0.6770	0.2415	0.1841	0.9481	24	0.9626	0.0466	0.7870	0.9956
2004	25	0.6630	0.2523	0.0351	0.9437	25	0.9344	0.0750	0.6358	0.9924
2005	26	0.6669	0.2286	0.2207	0.9510	26	0.9590	0.0479	0.7883	0.9960
2006	29	0.6526	0.2368	0.1164	0.9475	29	0.9673	0.0451	0.8379	0.9983
2007	31	0.6544	0.2615	0.0659	0.9513	31	0.9841	0.0202	0.9152	0.9991
2008	36	0.5720	0.2942	0.0179	0.9342	36	0.8991	0.1307	0.4856	0.9953
2009	37	0.5691	0.2947	0.0814	0.9431	37	0.9423	0.0645	0.7683	0.9978
2010	37	0.5865	0.2893	0.0555	0.9480	37	0.9605	0.0445	0.8440	0.9994
2011	35	0.5386	0.2997	0.0624	0.9361	35	0.8421	0.1786	0.2756	0.9945
2000–2003	98	0.6633	0.2323	0.1243	0.9481	98	0.9318	0.0779	0.6504	0.9956
2004–2007	111	0.6588	0.2423	0.0351	0.9513	111	0.9626	0.0519	0.6358	0.9991
2008–2011	145	0.5669	0.2919	0.0179	0.9480	145	0.9121	0.1236	0.2756	0.9994
2000–2011	354	0.6224	0.2646	0.0179	0.9513	354	0.9334	0.0959	0.2756	0.9994

6.4.2 DEA Efficiency Scores

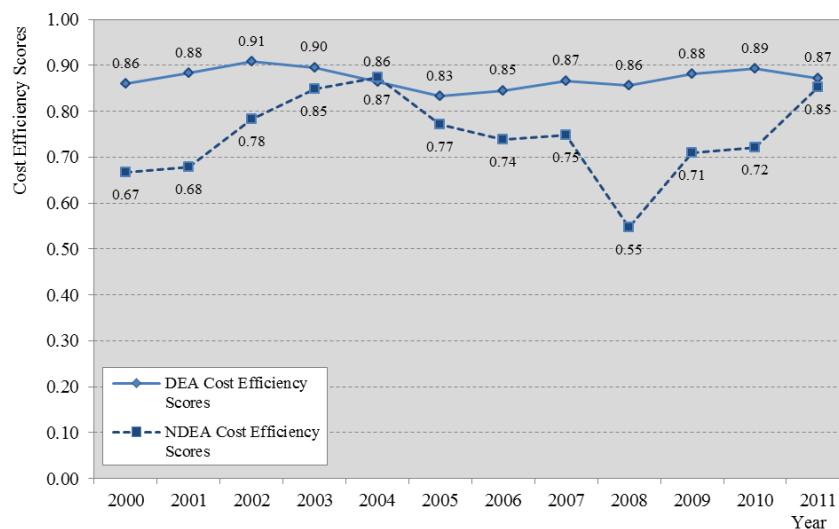
In this study, DEA is used to compare against the results generated by the SFA (Resti, 1997). Bauer et al. (1998) argues that there is no agreement between the parametric (i.e. Stochastic Frontier analysis, Thick Frontier analysis and Distribution Free analysis) and non-parametric methods (i.e. Data Envelopment analysis and Free Disposal Hull analysis) of what constitute the best estimates for efficiency scores. They further argued that these two approaches (parametric versus non-parametric) can be utilised to test each other's consistency. However, it should be noted that the scores produced by SFA may be different to those obtained from DEA due to variations in methods, the variables employed and assumptions regarding the random error. The next section discusses trends of DEA and

NDEA efficiency scores, from both the cost and profit efficiency perspectives. The findings obtained from DEA and NDEA could be utilised to support the findings of SFA in relation to the implementation of financial liberalisation on Malaysian banking institutions.

6.4.2.1 DEA and NDEA Cost Efficiency Scores

The average cost efficiency scores under DEA cost efficiency (CE) for 2000 - 2011 were found to be 87.1%. Based on the three different periods of gradual liberalisation introduced by BNM via its FSMP, the average CE scores were at 88.7%, 85.2% and 87.6% for the phases one (2000–2003), two (2004–2007) and three (2008–2011) of the FSMP respectively. For NDEA, the average CE for 2000 - 2011 was 73.9%. The first, second and third phases of the FSMP averaged CE were 74.3%, 77.9% and 70.7%, respectively.

Figure 6.4 Average DEA and NDEA Cost Efficiency Scores of Malaysian Banks, 2000–2011



From figure 6.4 and Table 6.7, due to differences in methodology, the trend of DEA CE and NDEA CE are different from what is shown by the CE scores of SFA. The differences in trends between DEA CE and NDEA CE indicate that there are differences in input prices which created lower efficiency scores in NDEA, although the inputs and outputs of the two different DMUs are similar. This may suggest that there is opportunity for banks to improve in terms of input prices (which may also reflect market power by several big banks) when compared to the best practice DMUs (Tone, 2002).

Table 6.7 Average DEA and NDEA Cost Efficiency Scores of Malaysian Banks, 2000–2011

Year	Co unt	DEA Cost Efficiency				Count	NDEA Cost Efficiency			
		Mean	Std. Dev.	Min.	Max.		Mean	Std. Dev.	Min.	Max.
2000	24	0.8600	0.0841	0.6924	1.0000	24	0.6661	0.1439	0.4184	1.0000
2001	26	0.8832	0.0858	0.6992	1.0000	26	0.6791	0.1640	0.3854	1.0000
2002	24	0.9075	0.0728	0.7591	1.0000	24	0.7831	0.1350	0.4970	1.0000
2003	24	0.8959	0.1052	0.5036	1.0000	24	0.8489	0.1754	0.3954	1.0000
2004	25	0.8631	0.1029	0.5614	1.0000	25	0.8737	0.1351	0.6367	1.0000
2005	26	0.8330	0.1504	0.3210	1.0000	26	0.7716	0.1841	0.2956	1.0000
2006	29	0.8453	0.0938	0.6392	1.0000	29	0.7386	0.1540	0.4129	1.0000
2007	31	0.8667	0.1202	0.4610	1.0000	31	0.7476	0.2051	0.4065	1.0000
2008	36	0.8563	0.1747	0.2353	1.0000	36	0.5475	0.3490	0.0277	1.0000
2009	37	0.8806	0.1016	0.6062	1.0000	37	0.7104	0.2224	0.1524	1.0000
2010	37	0.8933	0.0917	0.6309	1.0000	37	0.7208	0.2166	0.4016	1.0000
2011	35	0.8724	0.1090	0.5047	1.0000	35	0.8530	0.1689	0.4448	1.0000
2000–2003	98	0.8866	0.0881	0.5036	1.0000	98	0.7430	0.1709	0.3854	1.0000
2004–2007	111	0.8524	0.1175	0.3210	1.0000	111	0.7793	0.1788	0.2956	1.0000
2008–2011	145	0.8758	0.1228	0.2353	1.0000	145	0.7070	0.2684	0.0277	1.0000
2000–2011	354	0.8715	0.1130	0.2353	1.0000	354	0.7396	0.2198	0.0277	1.0000

6.4.2.2 DEA and NDEA Profit Efficiency Scores

The DEA results of profit efficiency scores of the DEA and NDEA models are presented in Figure 6.5 and Table 6.8 for the year 2000 to 2011. For DEA profit efficiency (PE) model, the average PE score for the period under study is 76.8%. On a more granular level, the average profit efficiency for the sub periods of the FSMP were 77.9%, 73.2% and 78.8%, respectively (i.e. first phase (2000–2003), second phase (2004–2007) and third phase (2008–2011) of the FSMP). The DEA average PE scores for these three periods show that the profit efficiency of Malaysian banks was on a decreasing trend in the second phase of FSMP (2004–2007), but rebounded in the third phase of FSMP (2008–2011). This can be explained by the impact of NIRF and new competition of foreign banks. Additionally, the NDEA average PE score for the years 2000–2011 registered at 67.7%. For the first-, second-, and third-phase FSMP, the average NDEA PE scores are 73.8%, 69.6% and 62.2%

respectively. A closer look at the trends of both SFA standard profit efficiency (SPE) and alternative profit efficiency (APE), and DEA- and NDEA PE suggest that the results from respective approaches are not identical but somewhat exhibit an overall similar trend.

Figure 6.5 Average DEA and NDEA Profit Efficiency Scores of Malaysian Banks, 2000–2011

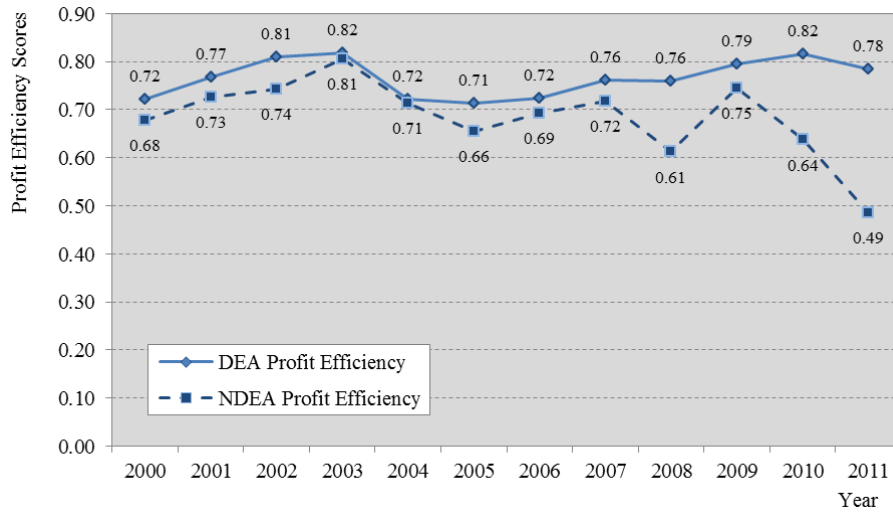


Table 6.8 Average DEA and NDEA Profit Efficiency Scores of Malaysian Banks, 2000–2011

Year	Count	DEA Profit Efficiency				NDEA Profit Efficiency				
		Mean	Std. Dev.	Min.	Max.	Count	Mean	Std. Dev.	Min.	Max.
2000	24	0.7218	0.2221	0.0392	1.0000	24	0.6778	0.2418	0.1480	1.0000
2001	26	0.7679	0.2178	0.0478	1.0000	26	0.7264	0.2110	0.2990	1.0000
2002	24	0.8100	0.2030	0.1151	1.0000	24	0.7433	0.2002	0.2643	1.0000
2003	24	0.8190	0.2079	0.0309	1.0000	24	0.8059	0.2199	0.2773	1.0000
2004	25	0.7222	0.2295	0.0630	1.0000	25	0.7132	0.2167	0.2036	1.0000
2005	26	0.7133	0.2162	0.0171	1.0000	26	0.6554	0.2180	0.2636	1.0000
2006	29	0.7242	0.2488	0.0252	1.0000	29	0.6932	0.2541	0.1100	1.0000
2007	31	0.7625	0.2603	0.0592	1.0000	31	0.7182	0.2762	0.1434	1.0000
2008	36	0.7588	0.2699	0.0173	1.0000	36	0.6126	0.3092	0.0529	1.0000
2009	37	0.7944	0.2028	0.1720	1.0000	37	0.7454	0.2757	0.2289	1.0000
2010	37	0.8166	0.1811	0.2915	1.0000	37	0.6373	0.3047	0.0243	1.0000
2011	35	0.7837	0.2183	0.0184	1.0000	35	0.4858	0.5070	0.0000	1.0000
2000–2003	98	0.7794	0.2131	0.0309	1.0000	98	0.7381	0.2200	0.1480	1.0000
2004–2007	111	0.7319	0.2382	0.0171	1.0000	111	0.6958	0.2427	0.1100	1.0000
2008–2011	145	0.7886	0.2188	0.0173	1.0000	145	0.6222	0.3669	0.0000	1.0000
2000–2011	354	0.7683	0.2243	0.0171	1.0000	354	0.6774	0.2983	0.0000	1.0000

In summary, different models introduced in this section produced non-identical efficiency scores (consistency among different models is tested and discussed later in this chapter). Nevertheless, the trends of these efficiency scores reflect the economic as well as banking industry events through the years 2000–2011. For instance, in the first phase of the FSMP (2000–2003), the efficiency scores of SFA CE were generally on a decreasing trend, implying that consolidation of banks may have adverse effects, particularly in the rationalisation of newly merged entity regarding operations, human resources and business strategies (Napier, 1989; Formbrun and Shanley, 1990). On the other hand, the SFA alternative profit efficiency shows an increasing trend, implying that the effect of market power from consolidation of domestic banks leads to higher profit efficiency. In this case, banks were found to be more innovative and have offered a variety of products and services

to customers to maximise profit. By doing so, banks faced higher costs but at the same time increased revenue in higher proportion (see Berger and Mester, 2003).

During the second phase of the FSMP (2004–2007), BNM deregulated its interest rate in 2004 and new foreign banks commenced operations in 2006–07. Most of these models shows decreasing trends during the early phase of the implementation of the deregulation of interest rate (i.e. NIRF) but bounced back in the later stage of the second FSMP. However, a model such as SFA APE that considers the impact of market power exhibits an increasing trend after the liberalisation of interest rates, which implies that larger banks may exercise their market power through interest rate pricing during the implementation of NIRF. In the third phase of the FSMP (2008–2011), parametric and nonparametric models generally displayed a downward trend of efficiency scores, which explains the impact of the global credit crisis and introduction of new foreign banks.

In general, it is observed that there were differences in the distributional properties of the efficiency scores of different models employed in this study. The differences may result from the differences in the various assumptions on which the methods are based (Weill, 2006). However, these differences are not necessarily a problem for the use of the efficiency scores: if different frontier methods can generate a similar rank order of efficiency, then the regulator can draw some level of reasonable policy conclusion based on the analysis of the efficiency scores.¹⁷⁵ Thus, it is of interest to know whether these approaches generate a consistent ranking of banking efficiency over the period of study.

6.5 Consistency Condition Tests of SFA and DEA

As discussed earlier, SFA and DEA have their own advantages and disadvantages. For instance, the advantage of SFA is that it takes into account the effect of noise in the data. To do this, SFA requires specification of functional forms on inputs and outputs for cost and

¹⁷⁵ In order to understand the relationship of SFA and DEA, this study follows Bauer et al. (1998) in comparing and testing the robustness of the efficiency scores from SFA and DEA. Five consistency checks were implemented which relate to comparable means, ranking of banks, identification of best and worst performance banks, stability over the period under study, and comparison of the financial ratios or non-frontier measures (Dong et al., 2014).

profit efficiencies, and relies on distributional assumptions to disentangle the random errors and inefficiency terms. However, should these specifications and assumptions be misspecified, they could result in biased estimates. Therefore, DEA could avoid these specification errors by enveloping the observed data points rather than rely on *a priori* assumptions for the efficiency frontier structure. However, DEA does not allow for random errors (e.g. all errors are considered to be inefficiency) and sensitive to outliers. Since both methodologies have different techniques and their own advantages and disadvantages, the nonparametric methods (i.e. DEA and NDEA) are applied as complementary models to the preferred SFA model (Resti, 1997).

In this study, as mentioned in Chapter 5, two DEA models (i.e. DEA and NDEA) are used to derive the cost- and profit-efficiency estimates using the variable returns to scale (VRS). VRS are chosen instead of constant returns to scale (CRS) to consider the effect of imperfect competition and constraints on finance that may prevent Malaysian banks from operating at optimal scale.¹⁷⁶ Following Bauer et al. (1998), consistency condition tests can be conducted to seek for evidence on the consistency of two different frontier estimations. However, past literature failed to reconcile what constitutes the best method for frontier estimation, in particular between frontier and non-frontier methodology. Hence, consistency testing can be used to evaluate the impact of the regulatory efforts introduced. For this, Bauer et al. (1998) proposed a set of consistency tests for both parametric and non-parametric measures (see chapter 2).¹⁷⁷

¹⁷⁶ In addition, rather than performing one common frontier for all years, the efficiency frontiers for DEA are estimated based on the three different phases of FSMP (i.e. 2000–2003, 2003–2007 and 2008–2011). This method was introduced in the past literature (e.g. De Young and Hasan, 1998; Isik et al., 2004). Common frontier assumes similar technologies in all years, while clustering frontiers based on FSMP phases do not assume the same technology in different liberalisation phases. This reflects the series of liberalisation measures implemented by BNM, which may have significant impact on banking technology. Moreover, the differences between these two approaches (i.e. common frontier for all years and clustered frontiers based on FSMP phases) are not significant in terms of efficiency scores.

¹⁷⁷ Although the consistency conditions were introduced to reconcile differences between these two approaches, nonetheless, Bauer et al. (1998) found that different approaches within the group of respective methods (i.e. SFA versus TFA or DEA versus DFH) could generate high consistency, but, for different approaches of different groups of methods (i.e. SFA versus DEA), the tests had exhibited low consistency.

6.5.1 Comparison of Efficiency Distributions

Table 6.9 provides a number of important descriptive statistics stemming from the cost and profit efficiency scores of DEA, NDEA and SFA. Based on Table 6.9, the results indicate that the means of SFA and DEA are moderately close. For cost efficiency, in general, SFA and DEA models produce comparable means (between approximately 73.9% and 87.2%). However, the higher cost efficiency scores for DEA (87.2%) compared to SFA (76.5%) is unusual. It is unusual because DEA generally assumes that the deviation from the best practice is due to managerial inefficiency and does not consider random noise. Hence, DEA inefficiency models are expected to exhibit a higher level of inefficiency than SFA. Nevertheless, this finding is similar to Ferrier and Lovell (1990), who suggest that contradictory results can be explained from the ability of the DEA frontier to envelop the data more closely than the translog cost frontier. The evidence showing that DEA envelope data more closely than SFA can be observed from the standard deviation and kurtosis, which displays lower standard deviation and higher kurtosis of 0.11 and 6.66 for DEA CE, compared to 0.18 and 2.27 for SFA CE, respectively. On the other hand, NDEA CE (74.0%) scores are lower than DEA CE (87.2%), perhaps due to the impact of imperfect competition and regulatory changes, in which banks still have room to improve in terms of input prices (Tone, 2002). Nevertheless, the cost efficiency of all the models shows negative skewness, which indicates that the distribution of the cost efficiency scores is skewed to the left, with cost efficiency scores concentrated on the right of means, which is consistent to Bauer et al., (1998) and Delis et al., (2009).

For profit efficiency, it is observed that DEA PE (76.8%) exhibit a greater mean compared to the SFA SPE (62.6%). Nonetheless, a higher mean is exhibited by SFA APE (93.3%) due to imperfect competition of the market, where larger banks exercised market power to maximise profits (Berger and Mester, 1997).¹⁷⁸ The SFA APE also shows higher average profit efficiency scores compared to DEA PE and NDEA PE, which suggests that the DEA does not consider functional form for its estimation of inefficiency and results in lower

¹⁷⁸ Due to price interdependence of oligopoly market of Malaysian banking industry, the competitors had adjusted their pricing of inputs at a given level of outputs nearer to the best practice banks to maximise profits. Hence, the kurtosis (11.80) and standard deviation (0.09) of SFA APE exhibits a greater concentration of profit efficiency scores at the peak of the negatively skewed distribution.

kurtosis and standard deviation that is more sensitive to outliers. In addition, the NDEA PE (67.7%) is lower than DEA PE (76.8%), which suggests that inefficient banks that deviate from best practice have more opportunity to improve in terms of input and output prices (Tone, 2002).

Table 6.9 Descriptive Statistics of Efficiency Scores of SFA, DEA and NDEA

	Mean	Std. Dev.	Min.	Max.	Skewness	Kurtosis
SFA Cost Efficiency	0.7655	0.1834	0.0296	0.9787	-1.4647	2.2718
DEA Cost Efficiency	0.8715	0.1130	0.2353	1.0000	-1.9435	6.6603
NDEA Cost Efficiency	0.7396	0.2198	0.0277	1.0000	-0.6592	0.0730
SFA Standard Profit Efficiency	0.6224	0.2646	0.0179	0.9513	-0.6148	-0.9044
SFA Alternative Profit Efficiency	0.9334	0.0959	0.2756	0.9994	-3.0174	11.8003
DEA Profit Efficiency	0.7683	0.2243	0.0171	1.0000	-1.3735	1.9732
NDEA Profit Efficiency	0.6774	0.2983	0.0000	1.0000	-0.5869	-0.5359

Notes: Skewness refers to the extent to which a distribution is not symmetrical. For a normal distribution, the sample skewness score is asymptotically distributed with a mean of zero and a variance of $6/n$ where n is the sample size. The standardised skewness score reported in the table is (sample skewness score $\times \sqrt{n}$)/ $\sqrt{6}$. Similarly, the standardised kurtosis score shown in the table is (sample skewness score $\times 3 \times \sqrt{n}$)/ $\sqrt{24}$.

6.5.2 Correlations of Efficiency Rankings

The next test of consistency is the rank order correlations of the efficiency scores from both the SFA and DEA models. For this test, if different approaches show similar results in terms of the ranking of the banks, it could be easier for the regulator or policymakers in drawing conclusions based on the efficiency scores (Bauer et al., 1998). Although the distributional characteristics of the efficiency scores differ between different approaches, it is still possible that these methods produce similar rankings of banks in term of efficiency scores. In ranking the banks, the Spearman rank-order correlation coefficient was used and the result derived from this statistical method is then tested for both SFA and DEA to measure the closeness of rankings among different models.¹⁷⁹ For cost efficiency models (see Table 6.10), the rank-order correlations between SFA CE and DEA CE, and NDEA CE are low at 0.0027 and 0.0630, respectively. The above result was expected and could be explained by

¹⁷⁹ The Spearman rank-order correlation provides a measure of association between two or more variables based on the rank order. The rank-order correlation does not measure the linear association between variables (Anderson et al., 1987).

different techniques of estimations between SFA and DEA. On the other hand, the rank-order correlation between DEA CE and NDEA CE is at 0.2688 and significant at 5% significance level. The higher rank-order correlation between DEA CE and NDEA CE was also expected, as many studies in the past have found that efficiency ranking tends to be correlated higher within the group of benchmarking techniques (e.g. Bauer et al, 1998; Huang and Wang, 2002).

For the Profit Efficiency models, the rank-order correlation between SFA SPE and SFA APE is 0.70 and significant at 5% significance level. This may suggest that in most cases, SFA PE and SFA APE rank banks in the same order.¹⁸⁰ For DEA models, the rank order correlation between DEA PE and NDEA PE is around 0.18 and significant at 5% significance level. The lower rank-order correlations may indicate that, although both the methods employ the same linear programming techniques, the differences in the treatment of inputs, outputs, and input prices and output prices may generate significantly different results, particularly when the market is imperfect (Tone, 2002). Between SFA and DEA profit efficiencies, the rank order correlations between SFA and DEA methods ranges between 0.24 and 0.28 and significant at 5% significance level.

Table 6.10 indicates another view regarding the optimisation behaviour towards cost and profit discussed in Chapter 2. The low correlations between cost and profit efficiencies suggest that banks with high cost efficiency may not achieve the same for profits due to differences in behaviour and management approaches towards costs and profits. For example, the rank order correlation between SFA CE and SFA PE and SFA APE is at 0.02484 and 0.0012 respectively, suggesting that a cost efficient bank may not be profit efficient. Delis (2009) and Berger and Mester (1997) argued that a cost efficient bank may not be profit efficient, because of the attitudes taken by the management towards cost and profit optimisation, in which banks may concentrate on either controlling cost or generating greater profits.

¹⁸⁰ These two models were performed using the same technique but having different variables. Moreover, many previous studies in the past found that the efficiency ranking tended to be highly correlated within the group of benchmarking techniques (e.g. Bauer et al., 1998 and Huang and Wang, 2002).

Table 6.10 Spearman Rank-Order Correlation between SFA, DEA and NDEA efficiency Scores

	SFA CE	DEA CE	NDEA CE	SFA SPE	SFA APE	DEA PE	NDEA PE
SFA CE	1.0000						
DEA CE	0.0027	1.0000					
NDEA CE	0.0630	0.2688*	1.0000				
SFA SPE	0.2484*	0.2721*	0.3572*	1.0000			
SFA APE	0.0012	0.2838*	0.2520*	0.6961*	1.0000		
DEA PE	0.0166	0.8470*	0.2679*	0.2787*	0.2396*	1.0000	
NDEA PE	0.0033	0.1271*	0.3955*	0.2419*	0.1838*	0.1867*	1.0000

Notes: * Correlation coefficient statistically significantly different from zero at 5% significance level. SFA CE refers to stochastic frontier analysis cost efficiency, DEA CE refers to data envelopment cost efficiency and NDE CE refers to new data envelopment cost efficiency. SFA SPE refers to stochastic frontier analysis standard profit efficiency, SFA APE refers to stochastic frontier analysis alternative profit efficiency, DEA PE refers to data envelopment profit efficiency and NDE PE refers to new data envelopment profit efficiency.

6.5.3 Identification of Best-Practice and Worst-Practice Banks

Although the consistency in the rank-order correlation provides some information, additional methods for assessing cost efficiency may still be useful (e.g. for regulatory purposes). Thus, another consistency condition introduced by Bauer et al., (1998) is the consistency of best- and worst-efficiency banks' groupings. For this test, banks' efficiency scores are ranked and divided into four quartiles. The first- and fourth-quartile represent best- and worst-performing banks respectively. Banks in these quartiles are identified for different models (SFA, DEA and NDEA) and the overlapped banks are then proportioned in both the top 25% and the lowest 25% efficiency scores (Spong et al., 1995).

Tables 6.11 and 6.12 display the corresponding best- and worst-performing banks of cost- and profit-efficiency respectively. The upper triangular matrix in both tables exhibit the proportion of overlapping banks for each pair of the frontier efficiency approaches of the top 25%. On the other hand, the lower triangular matrix shows the proportion of overlapping banks of pairs of different approaches of frontier efficiency of the 25% lowest efficiency banks. In general, the efficiency scores of overlapping banks' ratios are better for the same family frontier efficiency techniques. For example, based on the 25% top

performers for SFA CE, 21.4% were identified as being in the top 25% of both DEA CE and NDEA CE. Similarly, of the banks identified in the 25% worst performers of SFA CE, 30.3% and 19.1% were also correspondingly identified as 25% worst practice banks for DEA CE and NDEA CE. For DEA CE and NDEA CE, the proportion of overlapping banks between best- and worst-practice banks is 42.7% and 21.4%, respectively.

On the other hand, for profit efficiency, both SFA and DEA methods show higher proportion of overlapping banks within the same techniques (SFA SPE vs. SFA APE (i.e. best-practice (58.4%) and worst-practice (55.1%)), DEA PE vs. NDEA PE (i.e. best-practice (39.3%) and worst practice (43.8%)). However, lower proportions (ranging between 31.5% and 46.0%) are observed when cross techniques are paired (parametric vs. nonparametric).

The results from this test could provide valuable insights for regulators when gauging their policy implications, particularly when identifying problematic banks (Bauer et al., 1997). From the test of identification of best- and worst-practice banks, the profit efficiency models are found to be more consistent in identifying best and worst practices, than cost efficiency models. Generally, this consistency test is able to moderately identify banks based on determined 'best' and 'worst' quartiles. In other words, regulatory policies targeted at efficient and/or inefficient banks may hit different targets, depending upon which frontier techniques are used to determine the policy.

Table 6.11 Correspondence of Best Practice and Worst Practice of Cost Efficiency across Techniques

	SFA CE	DEA CE	NDEA CE
SFA CE		0.2135	0.2135
DEA CE	0.3034		0.4270
NDEA CE	0.1910	0.2135	

Notes: The upper right triangle displays 25% of best-practice efficiency scores. The lower left triangle exhibits 25% of worst-practice efficiency scores. SFA CE refers to stochastic frontier analysis cost efficiency, DEA CE refers to data envelopment cost efficiency and NDE CE refers to new data envelopment cost efficiency.

Table 6.12 Correspondence of Best Practice and Worst Practice of Profit Efficiency across Techniques

	SFA SPE	SFA APE	DEA PE	NDEA PE
SFA SPE		0.5843	0.3146	0.3483
SFA APE	0.5506		0.3483	0.3596
DEA PE	0.4270	0.3708		0.3933
NDEA PE	0.4607	0.3933	0.4382	

Notes: The upper right triangle displays 25% of best efficiency scores. The lower left triangle exhibits 25% of worst efficiency scores. SFA SPE refers to stochastic frontier analysis standard profit efficiency, SFA APE refers to stochastic frontier analysis alternative profit efficiency, DEA PE refers to data envelopment profit efficiency and NDE PE refers to new data envelopment profit efficiency.

6.5.4 Stability of Efficiency Scores over Time

Another consistency measure that is useful, from the regulatory perspective relates to stability over time. The stability over time measure enables policymakers to be informed regarding the reasonableness of the efficiency scores generated over a period of time, where the ranking of efficiency scores should not vary greatly from one year to the next.¹⁸¹ In order to test the stability of efficiency scores over time, the Spearman rank-order correlations were computed for each SFA and DEA efficiency models between each pair of years to examine the year-to-year stability of the efficiency measures over a period of time. For the full sample, there were 66 correlations of k-year-apart efficiencies, where $k = 1, 2, 3, \dots, 11$, are computed in each case. The columns in Table 6.13 display the Spearman rank

¹⁸¹ It is possible for some banks to experiencing improvements or deterioration in performance in following year. Nonetheless, it is unlikely that within a year an efficient bank will experience significant deterioration or improvement (Bauer et al., 1998).

order correlation coefficient for each pair k-year-apart which is used to examine the stability of efficiency scores over the period under study (2000–2011).

From Table 6.13, it is observed that the average correlations of the efficiency scores are relatively less volatile for 1- and 2-years. This result suggests that a bank's efficiency ranking does not show high volatility within a one- or two-year period. After this period (i.e. 1-year and 2-year apart), most of the models exhibit a decline in the correlation coefficients as the number of years between the efficiency scores grows. At some point, several methods – such as, SFA CE, DEA CE, DEA PE and NDEA PE – display insignificant correlations between eight to 11 years apart, resulting in some of average correlations nearing to zero (e.g. DEA CE on 9-year and 10-year apart at 2.6% and 0.51%, respectively). Nevertheless, it could be noted that most of the models are relatively stable in the short term, but apart from SFA SPE, the efficiency scores exhibit greater instability over a longer period of time.

Table 6.13 Stability of Efficiency Scores – The Correlations of k-Year-Apart Efficiencies

	1-Year Apart	2-Year Apart	3-Year Apart	4-Year Apart	5-Year Apart	6-Year Apart	7-Year Apart	8-Year Apart	9-Year Apart	10-Year Apart	11-Year Apart
SFA CE	0.6531*	0.5223*	0.4628*	0.3735*	0.3059*	0.2828	0.2334*	0.1891	0.1527	0.1557	-0.0316
DEA CE	0.6320*	0.3955*	0.3740*	0.2646*	0.2124*	0.1907*	0.0870	-0.0475	0.0260	0.0051	-0.3096
NDEA CE	0.4249*	0.2898*	0.1491*	0.1286	0.1651*	0.3179*	0.4112*	0.2921*	0.1725	0.2612	0.4132
SFA SPE	0.8226*	0.9873*	0.8887*	0.8610*	0.8625*	0.8591*	0.7805*	0.7454*	0.6917*	0.6425*	0.6594*
SFA APE	0.5628*	0.6941*	0.6856*	0.5749*	0.4597*	0.4062*	0.5173*	0.4466*	0.4507*	0.4333*	0.6586*
DEA PE	0.4923*	0.4335*	0.3973*	0.2769*	0.1678*	0.1950*	0.1939*	0.1214	0.2138	0.1741	0.0838
NDEA PE	0.3929*	0.3252*	0.3596*	0.3236*	0.4038*	0.2751*	0.2247*	0.2771*	0.1691	0.2313	0.3873

Notes: Following Bauer et al. (1998) each cell represents mean of correlations of k-year-apart efficiencies for one single efficiency technique within 12-year time period, where for each k, figure reported in each cell is average of correlations between efficiencies. For example, for 4-year apart correlations, there are 8 different correlations to be averaged, that are 2000 with 2004, 2001 with 2005, 2002 with 2006, ..., 2007 with 2011. * denotes that all correlations in that were averaged in each cell of k-year-apart were significant at 5% significance level.

6.5.5 Efficiency and Accounting-based Performance Measures

Regulators, bankers and financial analysts utilise non-frontier measures as part of evaluations and assessments on banks' performances. Therefore, Bauer et al. (1998) state that, if the frontier efficiency scores are correlated with some standard financial ratio of performance, the policymakers could be more confident that the efficiencies are in line with accounting-based performance measures and not simply artificial measures derived from specific assumptions on which the efficiency measures are based. The final consistency test employed to examine the consistency conditions between SFA and DEA is by comparing the efficiency scores with financial performance indicators.

As mentioned earlier, regulators, bank managers and analysts commonly draw on financial performance indicators for their decision-making (Fiorentino et al., 2006). These financial indicators are usually extracted from accounting information and are fairly standard. Therefore, this examination is expected to give confidence to policymakers in providing

industry conclusions (Bauer et al., 1998). Return on assets (ROA), return on equity (ROE), total cost over total revenue ratio (TC/TR), cost-income-ratio (CIR) and staff expenditure over total asset ratio (S/TA) are widely used in evaluating firms' performance. These financial performance indicators are normally used to compare firms with their peers or industry averages (Bauer et al., 1998). For this consistency testing, results from frontier efficiency and financial ratios are not expected to be perfectly correlated, because financial performance indicators do not take into account input prices or output mix in their assumptions (Bauer et al., 1998) (see Chapter 2). The first two ratios (i.e. ROA and ROE) are generally used to evaluate the profitability of banks; commonly, a higher value of these ratios indicates efficiency in utilising assets and equity. These two ratios are expected to exhibit a positive relationship with cost- and profit-efficiency. The second group of ratios relates to expenditure and measures costs according to total revenue and non-interest income plus net interest income. Lower ratios of TC/TR and CIR denote sound management of expenditure based on income received. Hence, these ratios are expected to be negatively correlated to cost- and profit-efficiency. Finally, the ratio of staff cost over total assets is used to assess manpower expenses, in which a small ratio suggests cost savings (Bauer et al., 1998).

Table 6.14 reports the correlations of cost efficiency of DEA and SFA models with the financial performance indicators selected for this consistency examination. From the results, the correlations between frontier efficiency estimates and financial indicators are relatively low. This finding is consistent with the findings of Bauer et al. (1998) and Fiorentino et al. (2006): the correlation is generally greater between cost efficiency scores and expenses related financial ratios (TC/TR and CIR) compared with the correlations of cost efficiency scores and profit related financial ratios (ROA and ROE). The direction is consistent for cost efficiency measures against financial ratios, except for S/TA. The relationship between cost efficiency scores and S/TA is positive, suggesting that greater expenses incurred on staff could generate higher cost efficiency scores.

Table 6.14 Correlations between Frontier Cost Efficiencies and Non-Frontier Performance Measures

	SFA CE	DEA CE	NDEA CE
ROA	0.0148	0.0868	0.2017*
ROE	0.1068*	0.0839	0.2015*
TC/TR	-0.1620*	-0.0856	-0.2725*
CIR	-0.2548*	-0.0782	-0.1982*
S/TA	0.0429	0.0652	0.0149

Notes: * denotes the rank-order correlation is significant at 5% significance level. ROA refers to return on assets, ROE refers to return on equities, TC/TR refers to total cost over total revenue ratio, CIR refers to cost-to-income ratio and S/TA refers to staff expenses over total assets ratio.

For profit efficiency scores, Table 6.15 shows the correlations between financial performance indicators and the frontier PE scores of SFA, DEA and NDEA models. Similar to the observation made to the cost efficiency consistency examination, the correlation of the profit efficiency is greater with the profit related financial ratios (ROA and ROE) when compared to expenses related financial ratios (CIR and TC/TR). The direction of the correlations between profit efficiency scores and financial performance measures is consistent, as expected. For example, the correlations of profit efficiency scores display positive correlations with profit related financial ratios (ROA and ROE) and negative correlations with expenditure related financial ratios (TC/TR and CIR). However, SFA model demonstrates a higher correlation compared to DEA models, indicating that taking the effect of heterogeneity (i.e. control and environmental variables) into the specification of profit efficiency produces greater consistency with financial performance indicators.

Table 6.15 Correlations between Frontier Profit Efficiencies and Non-Frontier Performance Measures

	SFA SPE	SFA APE	DEA PE	NDEA PE
ROA	0.1476*	0.2545*	0.0822	0.2683*
ROE	0.3969*	0.3286*	0.1588*	0.2107*
TC/TR	-0.0031	-0.0938	-0.0868	-0.2085*
CIR	-0.0210	-0.0450	-0.0872	-0.1733*
S/TA	0.1502*	0.1707*	-0.0212	-0.1092*

Notes: * denotes the rank-order correlation is significant at 5% significance level. ROA refers to return on assets, ROE refers to return on equities, TC/TR refers to total cost over total revenue ratio, CIR refers to cost-to-income ratio and S/TA refers to staff expenses over total assets ratio.

As a summary, despite low correlations between frontier efficiency scores and the accounting-based performance ratios, the direction of the relationship has been moderately consistent. Thus, the use of frontier efficiency and financial performance indicators gives some confidence that the frontier models are useful measures of actual accomplishment and not simply artificial products of the assumptions of the efficiency approach and can be of use for interested parties in making decisions (e.g. policymakers, bank managers and analysts).

As mentioned previously, efficiency assessment is one of many important actions for financial regulators and management alike when making decisions. Great care must be taken when choosing an assessment technique(s) from parametric (i.e. SFA) and nonparametric (i.e. DEA and NDEA) frontier efficiency methods. In this section, two nonparametric DEA methods (i.e. DEA and NDEA) were employed for consistency testing with the preferred parametric SFA models (Resti, 1997). Following Bauer et al. (1998), a set of consistency tests were conducted to compare the outcome of different methods. Consistent with the previous empirical literature, in most cases, the findings indicate low to moderate consistency across different methods. The results generated from this examination were in accordance with Bauer et al. (1998) where not all consistency conditions were relatively high. The differences between efficiency scores generated from different methods can be explained mainly from the inherited advantages and disadvantages of the methodologies mentioned earlier. From the five consistency conditions, low consistencies are shown in rank-order correlations of different methods and identification of best- and worst-performing banks (consistency condition test 2 and 3); and moderate consistencies in the

comparable means, stability over time and comparison against financial performance indicators (consistency condition test 1, 4 and 5). These results are similar to previous studies (Bauer et al., 1998). However, it is important for regulators and other interested parties to treat the results for Malaysian banks cautiously; they seem to be sensitive to methods selected for frontier estimation, in which the use of multiple frontier methods for consistency checking is recommended.

6.6 Characteristics of Efficiency Scores of Malaysian Banks

In this section, the characteristics and profile of the banks are analysed from several additional perspectives. The regulators may be interested in looking into the characteristics of banks in relation to the common rating system used by supervisory authorities. The CAMEL rating system is widely used by various financial supervisory authorities and was introduced by the Federal Reserve. There are many new bank rating systems such as, ARROWS (Advanced, Risk-Responsive Operating Framework), PAIRS and SOARS (Probability and Impact Rating System and Supervisory Oversight and Response System), RBSF (Risk-based Supervisory Framework), employed by respective banking regulators (e.g. the Financial Services Authority (UK), Australian Prudential Regulation Authority and Office of the Superintendent of Financial Institutions (Canada)). Despite various bank rating systems being introduced, the CAMEL rating system is still a useful indicator to understand the financial profiles of banks. Apart from banks' rating systems, different perspectives such as, ownership structure, bank specialisation and size are also employed to analyse the efficiency of Malaysian banks.¹⁸²

6.6.1 The CAMEL Rating system

This system is employed by many supervisory authorities around the world. CAMEL rating is based on five aspects that make up the acronym CAMEL. C stands for capital adequacy,

¹⁸² In this section, the cost efficiency and alternative profit efficiency were employed to analyse the differences in CAMEL ratings, ownership structure, bank specialisation and size. The use of alternative profit efficiency would be seen to best reflect the market structure of Malaysian banks and the effect of financial liberalisation over the period of 2000-2011.

A denotes asset quality, M for management, E is for earnings and L is liquidity. Banks are rated by supervisors using the composite CAMEL from 1 to 5 where 1 is satisfactory and 5 being unsatisfactory (DeYoung, 1998). CAMEL is employed as a decision-making tool by authorities to determine the level of supervision required on banks. BNM used to adopt CAMELS (where S represents sensitivity to interest rate), but changed to RBSF in 2006 (Bank Negara Malaysia, 2006b).

There are no standardised financial indicators employed for the specific aspects of acronyms (Gasbarro et al., 2002). Table 6.16 displays the common financial ratios that are utilised for CAMEL by supervisory authorities around the world.

Table 6.16 Descriptions of CAMEL Rating System

Factors	indicator	Financial ratios	Descriptions
Capital adequacy	Capital ratio	Capital / Loans	Capital adequacy shows how much capital is used to finance loans and advances according the Basel accord but without the risk weights being assigned to loans. There is still argument at what is the optimal capital required but higher capital ratio denotes lower risk (Saurina, 2004).
Asset Quality	Non-Performing Loans (NPL) ratio	Non-Performing Loans / Total loans	Asset quality indicates the level of non-performing loans over total loans (Ansari, 2006). A higher NPL ratio reflects higher risk.
Management	Cost-to-Income ratio (CIR) and TC/TR ratio	Cost / Income and Total Cost/ Total Revenue Ratio	This aspect of assessment usually involves qualitative examination relating to strategic and operational management of the banks. However, the nearest financial ratio being used to measure management effectiveness is the CIR and TC/TR ratio. A lower ratio denotes sound management of expenditure based on the income received.
Earnings	Return on Assets (ROA) and Return on Equity (ROE)	Profit / Total Assets and Profit / Capital	Widely used by various parties in assessing profitability of firms (Tian and Zeitun, 2007). A higher ratio indicates stronger earnings for banks and has lower risks.
Liquidity	Loans / Deposit (L/D) ratio and Deposit mix	Loans / Deposit And Customers' deposits / Total Deposits and inter-bank deposits	A highly important aspect of banks which was abundantly discussed during the global credit crisis and led to the introduction of Basel III. The L/D ratio indicates how much of deposits are used to finance loans. A lower L/D ratio implies higher liquidity. On the other hand, the deposit mix measures the level of dependence of banks on customer deposit rather than purchased funds. A higher deposit mix shows the ability of banks in relying on deposit rather than purchased funds and in return implies lower risk (Kashyap et al., 2002)

For the analysis of banks' characteristics using the CAMEL ratings system, Spong et al. (1995)'s study is referred to and followed. Banks' efficiency scores (both SFA cost- and profit-efficiency) are ranked and clustered into four quartiles. Similar to the Bauer et al. (1998) study, the first quartile represents the best performing banks, the fourth quartile the worst. Once these banks have been identified, they are then matched based on CAMEL's ratings.

Table 6.17 displays the result of the financial indicators profile of the most and least cost and profit efficient banks. First, the result on capital adequacy (C for capital adequacy) displays an average capital ratio of 10.4%. Both cost and profit efficiencies show a higher

capital ratio for worst quartile efficiency scores, registering 12.1% and 10.9% respectively. On the other hand, the capital ratio for best-practice banks' efficiency was low, at 9.4% and 8.9% respectively. It is normally anticipated that low risk banks will exhibit a greater capital ratio. However, in the case of Malaysia, these worst-performing banks hold more capital than the best-performing banks. The worst-practice banks are deemed riskier and therefore hold higher capital to maintain the confidence of shareholders, investors and regulators. Also, the worst-practice banks are generally subjected to closer monitoring by the banking supervisors from BNM; and due to their risky profile these banks are normally instructed by regulators to hold larger amounts of capital to absorb future potential losses (Awdeh et al., 2011).

Second, for asset quality (A for asset quality), mixed findings are observed for cost- and profit-efficiency. The worst-practice banks for cost efficiency exhibit a lower NPL ratio (6.4%) compared to the best (8.9%). The higher asset quality (i.e. lower NPL ratio) achieved by the former group is at the expense of higher operation costs, as exhibited by higher CIR (86.3%) and TC/TR (80.9%). On the other hand, the profit efficiency results are consistent with the asset quality; where the worst practice banks show higher NPL than the best practice banks.

Third, the management related financial ratios show consistent results regarding best and worst performance efficiency scores. The best efficiency score banks exhibit low CIR and TC/TR for both cost and profit efficiencies.

Fourth, based on the ROA and ROE (for E, earnings) of cost and profit efficiency scores, the worst-practice banks exhibited lower ROA (1.0%) and ROE (9.4%) than best-practice banks. This suggests that best-practice banks generate higher-than-average profits. Additionally, the earning assets for worst-practice banks for the profit efficiency model exhibit higher ratio than best-practice banks at 74.2% and 73.0% respectively.

Fifth, in terms of liquidity (L for liquidity), worst-practice banks are more liquid than the best-practice banks. In other words, the worst-practice banks exhibit lower L/D ratio (i.e. higher liquidity) than best-practice banks and are affected with higher expenditure to

manage the excess liquidity, as shown by CIR and TC/TR. At the same time; worst-practice banks rely fairly highly on interbank deposits (purchased deposits) which may be more expensive than retail and wholesale deposits and can affect their overall returns (e.g. ROA and ROE).

In general, from Table 6.17, the best-practice banks for cost- and profit-efficiency exhibit a lower level of capital ratio, marginally higher NPL (for cost efficiency only), lower operating costs, higher earnings and lower level of liquidity but with more deposits from customers. On the other hand, the worst-practice banks are characterised by a higher level of capital adequacy, lower NPL (i.e. higher level of asset quality but only for cost efficiency), higher operating costs, higher liquidity (i.e. lower L/D ratio), a lower level of earnings and more reliance on purchased funds from the interbank market. The result of this analysis may imply that BNM could adopt efficiency frontier analysis and compare best- and worst-practice banks as a tool that supports their decision for the bank rating system.¹⁸³

¹⁸³ A t-test is conducted to examine the differences between the best practice and worst practice banks based on CAMEL indicators. From the p-values, significant differences are found in capital ratio, CIR, TC/TR, ROA, ROE, L/D ratio and deposit mix. On the other hand there are no statistical differences found in NPL ratio and earnings assets.

Table 6.17 CAMEL indicators for Best and Worst Practice Banks

	Cost		Profit		Average
	Best-Practice Banks	Worst-Practice Banks	Best-Practice Banks	Worst-Practice Banks	
Capital Adequacy					
Capital Ratio ¹	9.41%	12.14%	8.87%	10.87%	10.42%
Asset Quality					
NPL Ratio ²	8.92%	6.40%	6.14%	6.94%	7.79%
Management					
CIR ³	78.99%	86.35%	79.98%	84.97%	80.92%
TC/TR ⁴	64.82%	80.88%	68.25%	75.69%	69.85%
Earnings					
ROA ⁵	1.29%	1.05%	1.39%	0.88%	1.22%
ROE ⁶	14.95%	9.44%	17.46%	8.57%	13.30%
Earnings Assets ⁷	73.05%	73.13%	73.02%	74.21%	73.50%
Liquidity					
L/D Ratio ⁸	68.75%	57.07%	68.30%	59.07%	64.73%
Deposit Mix ⁹	89.30%	84.44%	90.36%	86.99%	87.37%
Notes: (1) Equity to total assets (2) Non-performing loans to total loans (3) Non-interest expenses/ (non-interest income + net interest income) (4) Total cost over total revenue (5) Net profit to total assets (6) Net profit to total equity (7) Loans + investment + other earnings assets to total assets (8) Loans to deposits (9) Customers' deposits to total deposits.					

6.6.2 Ownership Structure

Following the analysis of the 'best' and 'worst' profiles of bank efficiency, this section provides the information regarding the differences between foreign and domestic bank structures that is useful to regulators and analysts (Spong et al., 1995). This analysis is often important in assessing the impact of foreign ownership on banking efficiency, particularly in developing countries that are opening up their banking industry to foreign penetration and competition (Kasman and Yildirim, 2006).

Claessens et al. (2001) and Kasman and Yildirim (2006) state that the involvement of foreign banks within a country can improve the overall operation of the financial services sector through increased competition. At the same time, efficiency can also be improved by

stimulating domestic banks to reduce cost, improve technology, increase efficiency and intensify diversity (Lensink and Hermes, 2004; Suffian and Habibullah, 2010). The participation of foreign banks could also pressure domestic banks to improve quality to retain their market share. The presence of foreign banks also provides spill over effects, such as the introduction of new financial services, improved bank regulation,¹⁸⁴ reduced government directed lending and an increased quality of human capital (Lansink and Hermes, 2004; Isik and Hassan, 2003).¹⁸⁵

From the earlier discussion (see Section 6.3) and hypothesis 2, foreign banks' impact is positive (1.09) and significant at 1% significance level, implying that foreign banks are less cost efficient when compared to domestic banks. Figure 6.6 and Table 6.18 exhibit the average cost efficiency scores for both domestic and foreign banks between 2000 and 2011. It was observed that the domestic banks are more cost efficient compared to the foreign banks, registering average cost efficiency scores of 0.83 and 0.69 respectively. This surprising result is different from other studies relating to ownership in developing/transition countries (see Batthacharya, Lovell and Sahay, 1997; Hassan and Marton, 2003; Isik and Hassan, 2002; Harvylchyk, 2006; Kasman and Yildirim, 2006; Matthews and Ismail, 2006). Nevertheless, Mohd Tahir, Abu Bakar and Haron (2008), Lensik et al. (2008) and Sufian and Habibullah (2010) also found that domestic banks are more cost efficient than foreign banks in developing countries.¹⁸⁶

¹⁸⁴ The involvement of foreign banks can also lead to better regulation due to the demand for an improved regulatory framework from both banks and regulator of home country.

¹⁸⁵ The foreign banks were always believed to bring in knowledge, technology and greater competition into the local banking market (Kasman and Yildirim, 2006). It is also found that the presence of foreign banks can result in cost cutting interest, improves efficiency and increase diversity of products and services (Sufian and Habibullah, 2010). They were deemed to have different operating strategies, organisation structures, regulatory requirements and support from home government (Isik and Hassan, 2003).

¹⁸⁶ The impact of foreign banks' entry into a country has been well documented in past literatures, but the empirical findings have been inconclusive (Sensarma, 2006; Sufian, 2010). In most developing or transition countries, it was noted that the foreign banks are more efficient than the foreign banks (Batthacharya, Lovell and Sahay, 1997; Hassan and Marton, 2003). For developing countries, it was observed that domestic banks were better than the foreign banks (Mohd Tahir, Abu Bakar and Haron, 2008). Similarly, for developed countries like the United States, most studies found that the domestic banks were more efficient than their foreign counterparts (De Young and Noelle, 1996; Hasan and Hunter, 1996). Nonetheless, there were also studies in developed countries other than the US that found that foreign banks are nearly as efficient as domestic banks (Vennet, 1996; Hasan and Lozano-Vivas, 1998).

These findings however are supported by Berger et al.'s (2000) hypothesis regarding the home field advantage. Foreign banks in this hypothesis are described as having higher costs and lower revenues while providing the same financial services. Foreign banks also faced the 'institutional distance' effect, where the distance between the home and host country can cause informational, agency and enforcement cost for foreign banks (Mian, 2006). From a different perspective, domestic banks have the advantages of operating in their own home, are supported by better information regarding the country's economy, language, laws, politics and different treatment from the regulators (Hymer, 1976). Lensink and Hermes (2004) state that the banking industry of developing countries normally has a lower level of economic development, which can result in higher operating costs for foreign banks. The competitive environment of developing nations is also normally low and dominant domestic banks may exercise market power to maximise profit (Lensink and Hermes, 2004; Sturm and Williams, 2004). Moreover, developing countries may also implement a defensive strategy by creating bigger domestic banks through mergers and acquisitions, creating an implicit barrier to entry for foreign banks (Sturm and Williams, 2004). The concentration of domestic banks that enjoy economies of scale can also increase the operational cost for foreign banks (Williams, 2003). In addition, regulators may also establish policies to inhibit the development of foreign banks, such as limiting the branches expansion and prohibiting in involving in certain financial services (Williams, 2003).

In the case of Malaysia, the foreign banks' cost efficiency was lower than the domestic banks due to the gradual liberalisation initiatives implemented via the FSMP. This was seen in the first phase of the FSMP, where the focus was on building the capacity of domestic institutions and strengthening the regulatory and supervisory framework (Bank Negara Malaysia, 2001b). During this first phase, Malaysian domestic banks underwent a consolidation exercise. A fragmented banking industry was turned into a more concentrated banking market and consequently, banks with higher market share enjoyed greater competitive advantages during the second- and third-phase of the FSMP. Hence, this resulted in higher (cost) disadvantages for foreign banks. This can also be evidenced from the discussion in Chapter 4 where BNM delayed the implementation of internet banking for foreign banks while allowing domestic banks to proceed. At the same time, the expansion of branches for foreign banks was also controlled, requiring approval from BNM. Another

point to note is that foreign banks in Malaysia originate from both developed and developing nations; therefore, the subset of foreign banks from less developed countries may influence the negative net impact on cost efficiency in foreign banks.

Additionally, for cost efficiency, foreign banks faced higher input prices compared to the domestic banks in producing outputs, which resulted in lower cost efficiency compared to domestic banks, particularly during the economic downturn in the third phase of the FSMP (between 2008 to 2011). This is due to foreign banks having to incur higher costs of funds from the interbank market, whereas cheaper consumer deposits were a better source for domestic banks. Furthermore, the de-novo foreign banks introduced during the second phase of the FSMP may have faced higher input prices from operating and establishment costs.

Figure 6.6 Cost Efficiency Scores of Domestic and Foreign Banks, 2000–2011

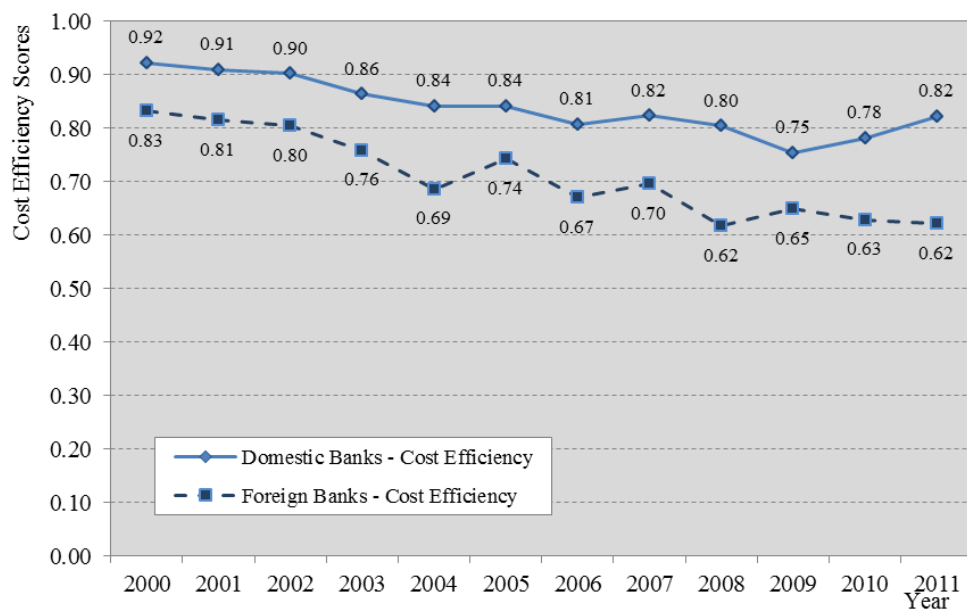


Table 6.18 Cost Efficiency Scores of Domestic and Foreign Banks, 2000–2011

Year	Cost Efficiency – Domestic Banks	Cost Efficiency – Foreign Banks
2000	0.9213	0.8307
2001	0.9088	0.8140
2002	0.9011	0.8035
2003	0.8625	0.7580
2004	0.8398	0.6855
2005	0.8404	0.7421
2006	0.8066	0.6692
2007	0.8222	0.6951
2008	0.8031	0.6165
2009	0.7539	0.6481
2010	0.7805	0.6282
2011	0.8208	0.6202
2000–2003	0.8986	0.8018
2004–2007	0.8256	0.6979
2008–2011	0.7885	0.6283
2000–2011	0.8305	0.6984

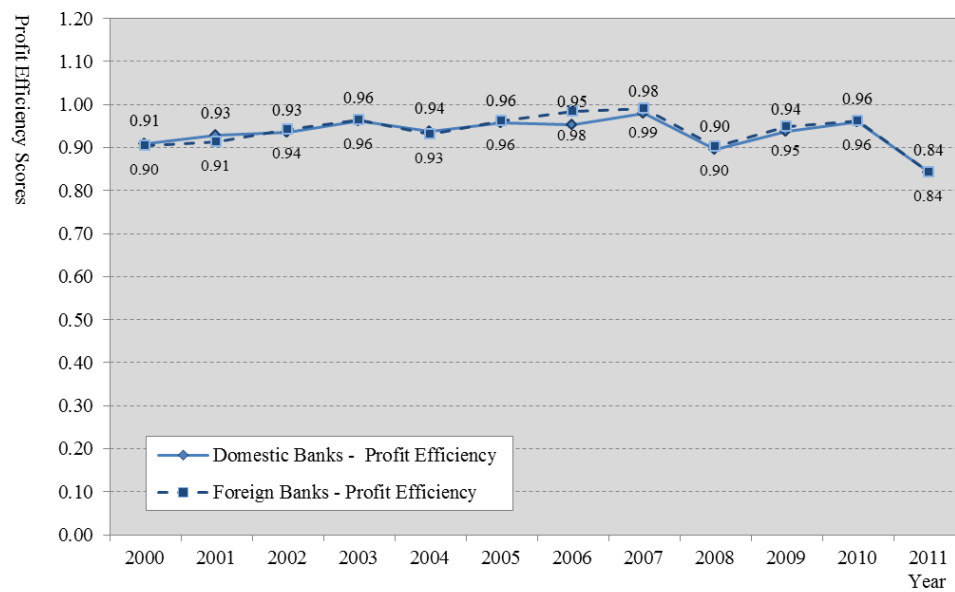
Different results however are observed for profit efficiency. Despite their lower cost efficiency, the profit efficiency of foreign banks are only very marginally higher than the domestic banks. Table 6.19 shows that the average profit efficiency for Malaysian banks between 2000 and 2011 was around 0.93. Whereas, the average profit efficiency for foreign banks was only slightly higher than domestic banks at 0.9345 and 0.9323 respectively. As shown earlier, based on hypothesis 2 and with reference to Table 6.3, there are no statistically significant differences between foreign and domestic banks (based on the t-value of the dummy variables assigned to foreign banks, which were negative and insignificant). This finding would questions the previous findings that the foreign banks normally have greater banking knowledge, better operating strategies, strong support from the home country's headquarters, technology and greater competition (Kasman and Yildirim, 2006; Isik and Hassan, 2003).

Foreign banks in Malaysia practiced 'cherry-picking' behaviour by choosing quality customers with niche banking activities that could give higher profits, rather than the 'mass'

banking activities of the domestic banks (Hassan, 2002). They also have not been strong in supporting the government's priority lending sectors – such as, housing loans for low cost housing – in which payment arrears has been high (Isik and Hassan, 2002). Additionally, whilst domestic banks were busy building their capacity during the first phase of the FSMP, the foreign banks had started to rely on the Central Credit Reference Information System (CCRIS) database and used their own credit scoring system in granting loans to customers. The use of credit scoring by the foreign banks generally improved their credit portfolio by reducing loan delinquency, and at the same time maximising retention of customers by cross-selling their products or granting higher credit limits (see Hamilton and Howcroft, 1995; Capon, 1982). Foreign banks also relied on this risk management tool as a basis for pricing their loans¹⁸⁷ according to borrowers' risks, when BNM first deregulated the interest rate in 2004 (Bank Negara Malaysia, 2004)).

Figure 6.7 displays the trend for the profit efficiency for domestic and foreign banks during the years 2000 to 2011. Based on the trends, the domestic and foreign banks demonstrated similar directions. There was a decline in profit efficiency for foreign banks in the third phase of the FSMP (2008–2011) , suggesting that the foreign banks were more cautious in lending when faced with higher costs from inter-bank lending during the global credit crisis (foreign banks have greater purchased funds compared to customers' deposits for their source of funds). The foreign banks in Malaysia were also susceptible to the international transmission channel, particularly when liquidity was tight in the home country (the effect of the global credit crisis) and their headquarters had to pull back liquidity from overseas branches, which created shocks in host countries (De Haas and Van Lelyveid, 2006; Kohn, 2008; Ceterolli and Goldberg, 2010;).

¹⁸⁷ Bochlinger and Leippold (2006) asserted that credit scoring can be used for loan pricing just as much on how it detect bad obligor during the application stage.

Figure 6.7 Profit Efficiency Scores of Domestic and Foreign Banks, 2000–2011**Table 6.19 Profit Efficiency Scores of Domestic and Foreign Banks, 2000–2011**

	Profit Efficiency – Domestic Banks	Profit Efficiency – Foreign Banks
2000	0.9094	0.9045
2001	0.9282	0.9131
2002	0.9347	0.9416
2003	0.9616	0.9635
2004	0.9379	0.9311
2005	0.9569	0.9612
2006	0.9535	0.9843
2007	0.9793	0.9900
2008	0.8965	0.9017
2009	0.9371	0.9478
2010	0.9597	0.9615
2011	0.8419	0.8423
2000–2003	0.9334	0.9303
2004–2007	0.9586	0.9671
2008–2011	0.9317	0.9133
2000–2011	0.9323	0.9345

6.6.3 Bank Specialisation

There are two types of banks analysed in this study: conventional and Islamic. As mentioned in Chapter 4, the investment banks were removed from this research due to significant differences in business operations and structures, and financial services offered to customers. Analysis by bank specialisation is necessary to provide policy feedbacks to the regulator. Theoretically, Islamic banking differs significantly to conventional banking.¹⁸⁸ However, in practice, the Islamic banks have been actively imitating their conventional counterparts while still following shariah principles. That is, products developed by Islamic banks resemble those in conventional banks (Beck, Kunt and Merrouche, 2012, Chong and Liu, 2009; Khan, 2010). Figure 6.8 and Table 6.20 report the cost efficiency scores for both conventional and Islamic banks for the years 2000 to 2011. The average cost efficiency scores for conventional banks is marginally higher than Islamic banks at 0.77 and 0.75, respectively. The trend was consistent with the bank type control variable which recorded a positive coefficient and significant at 10% level. This finding was also consistent with Saaid et al. (2003), Kamarudin, Safa and Mohamad (2008), Mokhtar, Abdullah and AlHabshi (2006), Batchelor and Wadud (2003), Srairi (2010) and Hassan (2005), who reported that Islamic banks had lower cost efficiency compared to conventional banks.

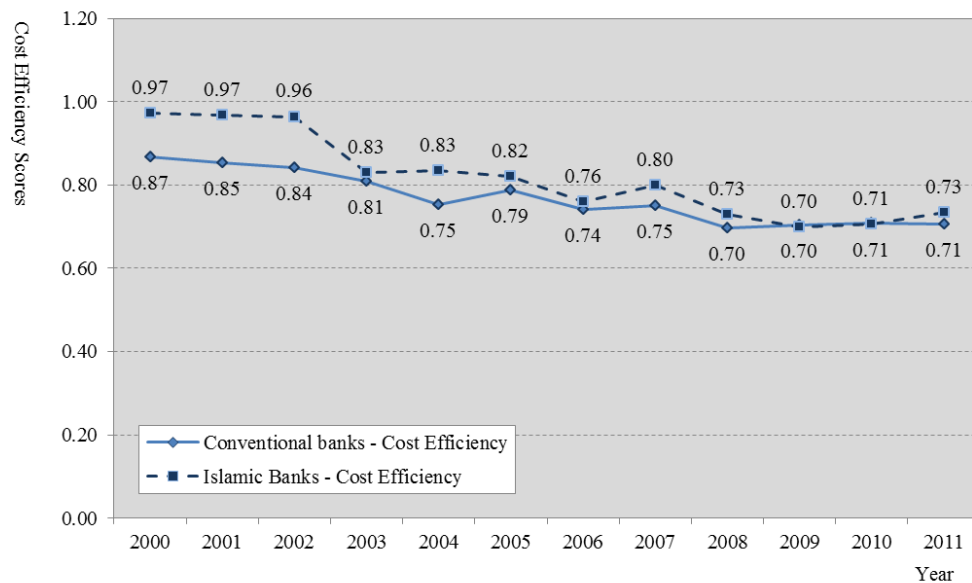
Based on the trends of cost efficiency of both types of bank, Islamic banks were more efficient in the first phase of the FSMP and declined as the timeline moved into the second and third phases of the FSMP. The greater efficiency during the first phase of the FSMP was due to higher market power in relation to Islamic financial products. During this period, there were only two Islamic full-fledged banks, which demonstrated their strength by exercising market power. They further demonstrated lower cost efficiency compared to conventional banks at the end of the second phase of the FSMP, with the introduction of new foreign full-fledged Islamic banks. The lower average cost efficiency scores for Islamic banks were influenced by higher costs incurred by these de-novo banks (new foreign

¹⁸⁸ In Islamic banking, the banks do not practice interest on financing (riba), require goods and services as underlying transactions, avoid speculation and prohibit financing illicit activities (Beck, Kunt and Merrouche, 2012).

Islamic banks) coupled with intense competition in the Malaysian banking scenario. In the case of Malaysia, Islamic banks face higher prices of funds and labour. These greater input prices for Islamic banks can be explained by the complex structure of Islamic products (following shariah principles) which can lead to a higher cost of funds. Islamic banks also require a more specialised workforce, which contributes to a greater cost of labour (Archer and Abdel Karim, 2002; Kamarudin, Safa and Mohamad, 2008). Additionally, with the nascent stage of implementation of Islamic banks compared to conventional banks, the regulations to support Islamic banks in Malaysia was also at its infancy stage, which could also result in a playing field that is not levelled between Islamic and conventional banks (Mokhtar, Abdullah and AlHabsi, 2006; Hassan, 2005).¹⁸⁹

Nevertheless, the difference in the average cost efficiency scores was small; at some point, Islamic banks achieved higher cost efficiency than conventional banks. For example, during the second phase of the FSMP (2004-2007), new foreign Islamic banks and Islamic bank subsidiaries were introduced in 2005–06. With this measure, Islamic banking assets and liabilities, via Islamic banking scheme (IBS) in the conventional banks, were separated. Despite being separated into different entities, these Islamic bank subsidiaries enjoyed savings from cost-sharing activities with their more established parent banks, and therefore experienced lower operational costs.

¹⁸⁹ This is consistent with Srairi (2010), who mentions that, due to their nascent age, Islamic banks did not have the economies of scale of the conventional banks, and were not ready to compete in the banking industry.

Figure 6.8 Cost Efficiency Scores for Conventional and Islamic Banks, 2000–2011**Table 6.20 Cost Efficiency Scores for Conventional and Islamic Banks, 2000–2011**

Year	Cost Efficiency – Conventional Banks	Cost Efficiency – Islamic Banks
2000	0.8672	0.9720
2001	0.8525	0.9674
2002	0.8421	0.9642
2003	0.8085	0.8296
2004	0.7531	0.8338
2005	0.7873	0.8215
2006	0.7404	0.7596
2007	0.7503	0.8002
2008	0.6974	0.7292
2009	0.7047	0.6992
2010	0.7073	0.7052
2011	0.7055	0.7339
2000–2003	0.8428	0.9333
2004–2007	0.7581	0.7929
2008–2011	0.7037	0.7166
2000–2011	0.7692	0.7545

With regards to profit efficiency, Figure 6.9 and Table 6.21 exhibit the profit efficiency for Islamic and conventional banks from 2000 to 2011. From Table 6.21, the average profit

efficiency of Islamic banks is lower than conventional banks at 0.83 and 0.96 respectively. The result is consistent with Hassan (2005) and Srairi (2010), who reported that Islamic banks are less profit efficient compared to conventional banks, due to size, under-developed Islamic banking regulation and overutilization of inputs in producing a given level of outputs (Srairi, 2010; Archer and Abdul Kareem, 2002; Mokhtar et al., 2006). For example, as discussed in Chapter 4, the Islamic banks were smaller in size and unable to compete with larger conventional banks. The Islamic banks also demonstrated higher input prices, with lower outputs and lower average ROA and ROE compared to conventional banks. This may suggest that Islamic banks have difficulty sourcing cheaper funds from customers' deposits and interbank funds, due to their adherence to strict shariah principles, as well as producing income from financing (loans) which could also be set off against various transaction costs, such as 'buying and selling' documentations while approving financing to customers (Mokhtar et al., 2006).

Figure 6.9 Profit Efficiency Scores for Conventional and Islamic Banks, 2000–2011

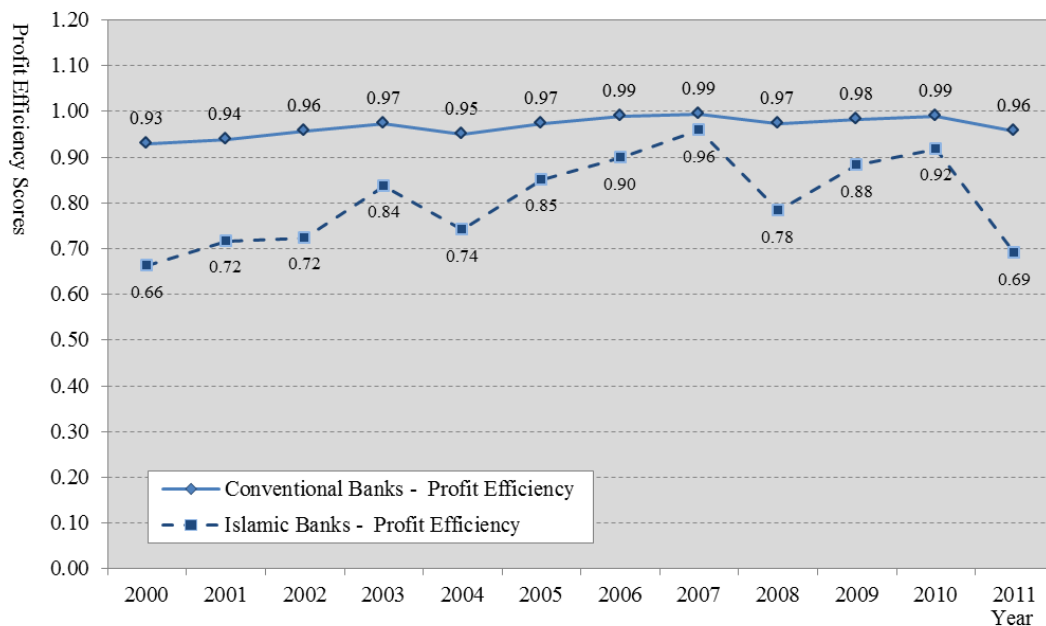


Table 6.21 Profit Efficiency Scores for Conventional and Islamic Banks, 2000–2011

Year	Profit Efficiency – Conventional Banks	Profit Efficiency – Islamic Banks
2000	0.9291	0.6625
2001	0.9376	0.7170
2002	0.9576	0.7240
2003	0.9740	0.8372
2004	0.9511	0.7418
2005	0.9730	0.8514
2006	0.9890	0.8991
2007	0.9946	0.9586
2008	0.9733	0.7825
2009	0.9835	0.8819
2010	0.9892	0.9185
2011	0.9563	0.6900
2000–2003	0.9493	0.7352
2004–2007	0.9766	0.9028
2008–2011	0.9760	0.8188
2000–2011	0.9672	0.8313

6.6.4 Bank Size

From previous literature regarding banking efficiency, larger banks generally have a number of advantages compared to smaller banks. The analysis on the relationship between size and bank efficiency provides imperative information for policymakers, as well as bank managers, to evaluate the optimal scale required for the operation of the banks. The larger banks normally have several advantages over smaller banks; such as, product diversity, larger outreach networks, greater technology, superior managerial performance, better cost control, higher profitability and greater economies of scale and scope (i.e. from both growth and joint production) (Evanoff and Israilevich, 1991; Casu and Girardone, 2002; Tsionas et al., 2003). Nevertheless, the larger banks are complex to manage and can result in lower efficiency (Delis and Papanikalou, 2009).

Table 6.22 reports the cost efficiency scores of Malaysian banks according to different bank sizes¹⁹⁰ and indicates that the largest bank group (more than MYR 160.0 billion asset size) has the highest cost efficiency compared to smaller banks (e.g. banks with less than MYR 40.0 billion of asset). However, smaller banks, with asset size in the region of MYR20.0 billion to MYR39.9 billion, show higher average cost efficiency than of those between MYR40.0 billion to MYR80.0 billion. In the Malaysian banking industry, this may indicate that smaller banks could also be more cost efficient than large banks. This can also be evidenced from the coefficient of bank size (ψ_6) in Table 6.1, which is positive at 2.45 and significant at 1% significance level. According to Akhigbe and McNulty (2003), in situations where the banking industry faces financial liberalisation, large banks were found to be less efficient than small banks. This could be attributed to large banks being required to invest in major restructuring and cost-cutting initiatives more than small banks. During the first phase of the FSMP (2000–2003), the domestic banks faced a consolidation exercise in which the new larger entities incurred greater costs compared to the smaller banks (see Chapter 4) (Suffian, 2004; Bonin et al., 2005). These new consolidated banks were publicly traded; and with diverse ownership come greater agency costs, which could lead to lower cost efficiency scores (Kwan, 2006).

Moreover, large banks normally have considerable market power, in which Berger and Hannan (1998) state that there are several reasons that can justify why larger banks demonstrate lower cost efficiency. First, banks could charge higher prices in excess of competitive level where managers do not have incentives to work as hard to keep costs under control and enjoy the ‘quiet life’. Second, with greater market power, managers are allowed to pursue objectives other than profit maximisation. Third, managers may devote resources to maintaining market power, which raises costs, which subsequently reduces cost efficiency (Berger and Hannan, 1998) (see also Chapter 2).

¹⁹⁰ In order to further investigate the relationship between cost- and profit-efficiency and bank size, banks are sub-divided into five classifications in terms of their total asset size.

Table 6.22 Cost Efficiency Scores of Malaysian Banks According to Asset Size, 2000-2011

Year	1.0-19.9 MYR Billions	20.0-39.9 MYR Billions	40.0-79.9 MYR Billions	80.0-159.9 MYR Billions	>160.0 MYR Billions	Average Cost Efficiency
2000	0.8518	0.9385	0.9434	0.9013		0.8760
2001	0.8297	0.8826	0.9611	0.9425		0.8614
2002	0.7836	0.9258	0.9315	0.8515		0.8523
2003	0.7284	0.8657	0.9186	0.8873		0.8103
2004	0.6414	0.8557	0.8942	0.9021		0.7596
2005	0.7165	0.8546	0.8643	0.8452		0.7913
2006	0.7057	0.7207	0.8117	0.8602	0.8692	0.7450
2007	0.7393	0.7963	0.7268	0.8360	0.9324	0.7648
2008	0.6867	0.7431	0.6771	0.8143	0.8950	0.7098
2009	0.6920	0.7168	0.6877	0.7742	0.7587	0.7025
2010	0.6678	0.7628	0.7112	0.8024	0.8112	0.7065
2011	0.6536	0.7398	0.7759	0.7200	0.9058	0.7177
Average	0.7228	0.8260	0.7811	0.8305	0.8651	0.7655

Table 6.23 displays the profit efficiency of Malaysian banks according to asset size. With regards to profit efficiency, large banks are found to be more profit efficient than small banks. This is in line with the negative coefficients (-9.9) for variable asset size (ψ_6) shown in Table 6.3. Larger banks generally demonstrate higher profit efficiency due to superior managerial performance, higher profitability, and greater market power from economies of scale and/or scope (Evanoff and Israilevich, 1991; Casu and Girardone, 2002; Tsionas et al., 2003).

Large banks generally possess sizeable market power, which allows them to enjoy greater profits, creating an incentive for them to behave prudently. As a result, this behaviour could lead to banks engaging in less risky activities with lower monitoring and operating costs (Evanoff and Israilevich, 1991). Furthermore, banks with market power face less pressure to increase the quality of banking services, which consequently reduces their operating costs with increasing profits (Berger, 2003). Large banks also use more efficient technology and

employ specialised staff for specialised functions to increase operational efficiency (Molyneux and Iqbal, 2005). Greater technology changes, such as advances in information-processing technology, could help improve the profit efficiency of large banks. With technology, larger banks could also enhance their monitoring of deposit and loan accounts, obtain ‘hard’ information about potential customers and improve the management of large branch networks that could increase their level of profit efficiency (Petersen and Rajan, 1995; Berger, 2003; Gilbert and Wheelock, 2013).

Table 6.23 Profit Efficiency Scores of Malaysian Banks According to Asset Size, 2000-2011

Year	1.0-19.9 MYR Billions	20.0-39.9 MYR Billions	40.0-79.9 MYR Billions	80.0-159.9 MYR Billions	>160.0 MYR Billions	Average Profit Efficiency
2000	0.8839	0.9599	0.9814	0.9882		0.9069
2001	0.8874	0.9691	0.9779	0.9894		0.9206
2002	0.8974	0.9753	0.9839	0.9923		0.9381
2003	0.9331	0.9853	0.9916	0.9956		0.9626
2004	0.8875	0.9738	0.9851	0.9875		0.9344
2005	0.9247	0.9867	0.9900	0.9939		0.9590
2006	0.9413	0.9941	0.9943	0.9974	0.9983	0.9673
2007	0.9722	0.9945	0.9975	0.9987	0.9988	0.9841
2008	0.8391	0.9676	0.9859	0.9932	0.9939	0.8991
2009	0.9091	0.9535	0.9915	0.9958	0.9964	0.9423
2010	0.9377	0.9721	0.9947	0.9972	0.9980	0.9605
2011	0.7489	0.8477	0.9412	0.9825	0.9897	0.8421
Average	0.8941	0.9686	0.9838	0.9926	0.9951	0.9334

6.7 Scale Economies and Technological Change

6.7.1 Scale Economies

Economies of scale are closely related to banks’ optimal behaviour to select the output levels corresponding to the minimum cost of a unit of output. The economies of scale of an industry may also have regulatory implications, particularly regarding banks’ mergers and

acquisitions, and antitrust enforcement. Chapter 5 discussed measurements of scale, stipulating that the measurement of economies of scale (SE) is computed by cost elasticity by differentiating the cost function with respect to outputs. The level of scale economies from the estimated stochastic frontier indicates whether a bank that had minimised cost of producing outputs could lower its costs by producing an alternative level of outputs (Mester, 1996). That is, the measurement of SE can be measured by the partial derivative of the estimated cost function, with respect to the outputs as follows:

$$SE = \frac{\partial \ln TC}{\partial \ln y_k} = \sum_{k=1}^3 \delta_{km} + \sum_{k=1}^3 \sum_{m=1}^3 \delta_{km} \ln y_m + \sum_{n=1}^2 \sum_{k=1}^3 \eta_{nk} \ln \left(\frac{W_n}{W_3} \right) + \sum_{k=1}^3 \kappa_{1k} T + \sum_{k=1}^3 \sum_{r=1}^2 \omega_{kr} \ln c_r \quad (5.15)$$

where SE is scale economies, $\partial \ln TC$ is the log of total cost and $\partial \ln y_k$ is the log of outputs. The measurement of SE refers to the value of point estimates of the scale elasticities, which are estimated using the average value of outputs, input prices and control variables such as capital adequacy, asset quality, liquidity and time, together with the coefficient of the estimated parameters from the preferred cost function (Mester, 1996; Altunbas et al., 2001). From this estimation, a scale efficient bank will generate constant returns to scale ($SE=1$). On the other hand, banks are considered scale inefficient if they are operating at increasing or decreasing returns to scale ($SE < 1$ or $SE > 1$).

Table 6.24 suggests that the average scale economies between 2000 and 2011 are about 10.3%. This is shown by the overall average scale economies estimates for Malaysian banks (2000–2011) that is less than 1 (i.e. 0.89), which suggests that economies of scale were present in the Malaysian banking industry. These economies of scale are more common for the banks with asset size of less than MYR40.0 billion, which ranges from 1.0% to 37.1%. This shows that an increase of 100% of outputs could result in total increase in cost between 62.9% and 99.0%, implying that small banks can potentially save operating costs by extending their production scale. That is, an equal proportionate rise in all outputs leads to a less than proportionate rise in total costs. For large category banks (banks with assets over MYR 80.0 billion), the values of scale economies are more than one. These large banks face

diseconomies of scale ranging from 1.0% to 30.6%. Hence, an increase of 100% in the level of outputs could increase the total costs by 101.0% to 130.6%, respectively. This result suggests that big banks experience diseconomies of scale and could reduce their average cost and gain efficiency by decreasing their scale of operations. The scale economies for banks with assets of MYR 20.0 billion and 80.0 billion are close to one, which implies that the optimal bank size is in the range of MYR 20.0 billion to MYR 80.0 billion as they exhibit constant return to scale. Generally, table 6.24 shows that, as the asset size increases the returns of scale demonstrate a pattern from increasing, to constant and to decreasing, in which the average cost curve displays a U-shape, with medium-sized banks being more scale efficient than very small or very large banks. This finding regarding Malaysian banks is consistent with previous studies performed by Altunbas et al. (2000) and Altunbas et al. (2001) on Japanese and European banking system.

Table 6.24 Scale Economies for Malaysian Banks, 2000–2011

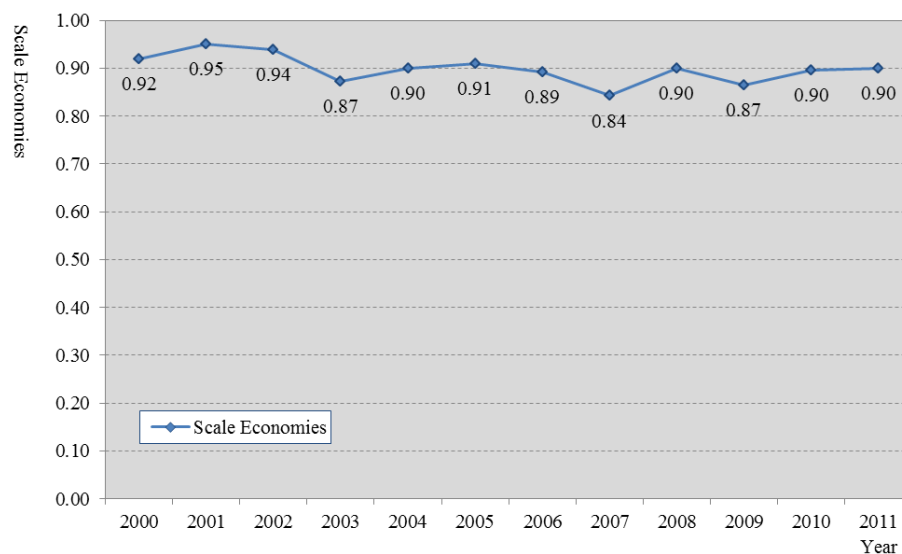
Year	1.0–19.9 MYR Billion	20.0–39.9 MYR Billion	40.0–79.9 MYR Billion	80.0–159.9 MYR Billion	>160.0 MYR Billion	Average
2000	0.8600*	0.9457*	1.1875	1.2854		0.9193
2001	0.8687*	0.9830*	1.1191*	1.5457		0.9500
2002	0.7875*	1.0542*	1.1420*	1.2394*		0.9395
2003	0.6864*	0.9897*	1.0880*	1.2096*		0.8722
2004	0.6774*	1.0718*	1.0561*	1.3074*		0.9001
2005	0.6296*	1.1145*	1.1316*	1.3260*		0.9090
2006	0.6386*	1.1327*	1.0705*	1.2697*	1.3854	0.8915
2007	0.6660*	1.0639*	0.9602*	1.0053*	1.1818*	0.8423
2008	0.8108*	0.9858*	0.9963*	1.1476*	1.0747*	0.8997
2009	0.7598*	0.9033*	0.9671*	1.1242*	1.1651*	0.8656
2010	0.7505*	0.9542*	1.0616*	1.2243*	1.2794*	0.8966
2011	0.7390*	0.9336*	1.0141*	1.0894*	1.3065*	0.9005
Average	0.7473*	1.0219*	1.0403*	1.2103*	1.2256*	0.8968

Notes: The scale economies measure is $\delta \ln TC / \delta \ln y_k$ and the estimates are evaluated at the mean of the data rather than mean estimate of scale economies calculated at each observation. * denotes the scale economies estimates are statistically different from 1 at 5% level for a 2-tailed test.

Figure 6.10 shows the trend of the scale economies of Malaysian banks from 2000 to 2011. It is observed that the Malaysian banks were experiencing economies of scale and the

average scale economies exhibit an upward trend (increasing economies of scale), particularly for the later years of the sample. This upward trend in later years may imply that Malaysian banks tend to achieve optimal operating efficiency (constant returns of scale) by gradually making changes in their scale of production. Looking at the trend, during the early years of the sample (i.e. 2000, 2001 and 2003) higher levels of scale efficiency is evident, reflecting the consolidation of domestic banks. This consolidation initiative has increased the size of individual domestic banks, which consequently expanded their scale of production. In this case, the Malaysian banks enjoyed the scale economies derived from the diversification of risk gained from a larger portfolio of loans and a larger base of deposits. Thus, this diversification allows new larger domestic banks to manage risks with relatively fewer resources. That is, a larger scale of operations could improve a bank's risk-return trade-off (Hughes and Mester, 2011).

Figure 6.10 Trends of Scale Economies for Malaysian Banks, 2000–2011



As mentioned in Evanoff and Israilevich (1995), scale efficiency should not be confused with scale elasticity. These are two different concepts because they measure different things. Scale elasticity can be measured as the proportionate change in cost associated with a proportionate change in outputs. On the other hand, scale efficiency refers to the measurement of the average production cost at the observed operation scale compared to what is obtainable at the optimal scale size. Therefore, one should not get confused with a

scale elastic measure near one as this does not necessarily indicate small-scale inefficiency, nor does a large difference imply substantial scale inefficiency (Evanoff and Israilevich, 1995). The scale inefficiency (SI) is measured as the proportion changes in the operating cost of the banks on constant returns to scale to the operating cost on economies of scale, which can be written as (Evanoff and Israilevich, 1995):

$$SI = e^{\left(\frac{0.5}{s}\right)(1-F_1)^2} - 1 \quad (6.1)$$

where SI is the scale inefficiency, s and F denotes the second derivative and first derivative of equation 5.15. Hence, the scale inefficiency is a function of first and second derivatives of the cost function.

Table 6.25 exhibit the scale efficiencies of SFA and DEA for Malaysian banks from 2000 to 2011. The scale efficiency using DEA can be computed from the cost efficiency scores of the constant returns to scale (CRS) to corresponding cost efficiency scores derived from cost efficiency of variable returns to scale (VRS) (SCALE= CRS/VRS, see Chapter 5 for details). The average scale efficiencies of SFA and DEA for Malaysian banks are 0.96 and 0.90, respectively. From the results, it is noted that the difference between scale inefficiency and scale economies is just a few percentage of each other, which is consistent with Berger et al. (1993, 1994) and Evanoff and Israilevich (1991). However, there are several issues relating to the measurement of scale efficiency (see Berger and Humphrey, 1994). For instance, the scale economies are normally measured using data of all the banks in the sample instead of using data on the most efficient banks on the production possibilities frontier, which could confound scale effects with differences in cost efficiency. Additionally, in comparison to scale economies (which is the ratio of marginal cost to average cost along a ray that holds outputs mix constant), scale efficiency accounts for the full difference in ray average cost between the point of evaluation and the scale efficient point (the bottom of U shaped average cost function), which might be far from the point of evaluation (Berger and Humphrey, 1994).

Table 6.25 Scale Efficiency for Malaysian Banks, 2000–2011

	SFA Scale Efficiency	DEA Scale Efficiency
2000	0.9762	0.9234
2001	0.9669	0.9285
2002	0.9680	0.9148
2003	0.9586	0.9076
2004	0.9588	0.8880
2005	0.9627	0.8970
2006	0.9387	0.8982
2007	0.9398	0.8787
2008	0.9284	0.9086
2009	0.9729	0.8878
2010	0.9632	0.9147
2011	0.9728	0.9101
average	0.9587	0.9043

6.7.2 Technological Change

A time trend is employed to control for changes in technology of Malaysian banks over the time period under study. The time variable is a ‘catch-all’ variable that is used to capture the technological effects, which is particularly important in the continuously changing business environment where bank technology may be different from time to time (Altunbas et al., 1999; Isik and Hassan, 2002). The time variable is also used to demonstrate that banks learn by performing organisational changes by efficiently utilising their inputs, outputs as well as other environmental factors (Altunbas et al., 1999). Therefore, technological change can be estimated by a ratio of changes in total cost to a change in technology, as shown from the measurement of partial derivative of the estimated cost function, with respect to the time trend (equation 5.29). This can be mathematically written as (Altunbas et al., 1999):

$$\frac{\partial \ln TC}{\partial \ln T} = \tau_1 + \tau_{11} \ln T + \sum_{k=1}^3 \delta_k \ln y_k + \sum_{n=1}^2 \eta_n \ln \left(\frac{w_n}{w_3} \right) \quad (6.2)$$

where $\partial \ln TC$ is the changes of log of total cost and $\partial \ln T$ denotes the changes in log technology. Table 6.26 reports the measurement of technological progress that exhibit the possible contribution of technical advances that could reduce the average costs for Malaysian banks between 2000 and 2011. For the case of Malaysian banks, Table 6.26 indicates that technological change could reduce annual production costs by approximately 11.3%. The technological change increased in the early years (2001 to 2003) during the introduction of internet banking to Malaysian domestic banks. Technological change also experienced a similar trend toward the end of the period of this study (2007–2011), mainly driven by heavy technology investment in Malaysian banks to improve their risk management practices and complying with the requirement of Basel II. For this, banks invested in various information technologies related tools, such as data warehouse and business intelligence solutions, to capture data, building new scoring models, enhance data quality, and improve new reporting structures.

Table 6.26 Technological Change of Malaysian Banks, 2000–2011

Year	Technological Change
2000	-0.0640
2001	-0.0602
2002	-0.0786
2003	-0.1035
2004	-0.1365
2005	-0.1175
2006	-0.0957
2007	-0.0738
2008	-0.1029
2009	-0.1569
2010	-0.1723
2011	-0.1411
Average	-0.1125

According to Altunbas et al. (1999), the technological change can be decomposed into pure technical change, scale augmenting technical change and non-neutral technical change. The pure technical change ($\tau_1 + \tau_{11} \ln T$) relates to time in equation 6.3. It accounts for the reduction in total costs over a period of time that may be influenced by technological changes. The scale augmenting technical change ($\sum_{k=1}^3 \delta_k \ln y_k$) that relates to outputs from equation 6.3, account for reduction in cost due to efficient scale of production. The non-neutral technical change ($\sum_{n=1}^2 \eta_n \ln(\frac{w_n}{w_3})$) is related to prices in equation 6.3, and considers the reduction in total costs to input prices (Baltagi and Griffin, 1998; Kasman and Kirbas, 2006).

From Table 6.27, the main source of technological change for Malaysian banks between 2000 and 2011 is pure technical change, at around 10.3%. This is particularly true as shown from the increasing trend of pure technological trend from 2000 to 2011, in which the improvement in technological change may have come about after the consolidation exercise of Malaysian domestic banks during the first phase of FSMP (2000-2003), where merged banks could take full advantage of new technologies by investing in cutting-edge technology and management systems (Khrishnasamy et al., 2003). The evidence could be seen from the enhanced network of delivery channels and branches throughout the country

after the consolidation exercise where these banks also offer a full range of banking services over the internet (Abdul Majid et al., 2010; Suffian, 2011). The higher pure technological change exhibited at the end of the study period may also imply the effects of new competition from foreign banks, and banks' compliance behaviour towards implementation of Basel II that require further enhancement in technological investments.

Table 6.27 Decomposition of Technical Changes

Year	Technological Change	Pure Technological change	Scale Augmentation Change	Non-Neutral Tech Change
2000	-0.0640	-0.0695	-0.0203	0.0258
2001	-0.0602	-0.0827	-0.0137	0.0362
2002	-0.0786	-0.0904	-0.0137	0.0255
2003	-0.1035	-0.0959	-0.0246	0.0171
2004	-0.1365	-0.1002	-0.0324	-0.0039
2005	-0.1175	-0.1036	-0.0385	0.0246
2006	-0.0957	-0.1066	-0.0474	0.0583
2007	-0.0738	-0.1091	-0.0514	0.0867
2008	-0.1029	-0.1114	-0.0413	0.0498
2009	-0.1569	-0.1134	-0.0472	0.0036
2010	-0.1723	-0.1152	-0.0389	-0.0182
2011	-0.1411	-0.1168	-0.0347	0.0104
Average	-0.1125	-0.1031	-0.0351	0.0258

6.8 Conclusions

This chapter has examined the cost- and profit-efficiency levels of Malaysian banks during the implementation of a ten-year financial liberalisation programme via its Financial Sector Master Plan (FSMP) from 2000 to 2011. The efficiency scores are generated from SFA, DEA and NDEA models. Following Resti (1997) and Bauer et al. (1998), SFA was used as the main methodology to measure the level of efficiency of Malaysian banks, and DEA and NDEA models were employed to test for consistency purposes. The results of the preferred SFA models show that the average cost efficiency is at 76.5% for the years 2000–2011. On the other hand, the standard profit- and alternative profit-efficiency exhibit average scores of

62.2% and 93.3%, respectively. Looking at the trend of the first phase of FSMP (2000–2003), the cost- and standard profit-efficiency were trending downward, mainly due to the impact of post-consolidation of domestic banks. During this period, consolidated banks were expected to incur higher costs, relating to merger and acquisition activities, such as rationalisation of branches, staff redundancy compensation and branding. Conversely, the alternative profit efficiency scores were on an increasing trend, suggesting that the effect of consolidation of domestic banks had resulted in increasing market power by these banks.

In the second phase of the FSMP (2004–2007), the cost- and profit-efficiency scores dropped marginally in 2004 due to the adjustments that the banks have to make upon the introduction of new interest rate framework (NIRF) by BNM, who liberalised the interest rate to the market players. Subsequently, the efficiency scores for all three economic efficiencies (i.e. cost-, standard profit- and alternative profit-efficiency) displayed marginally higher degree of efficiency scores in 2005, which suggests that deregulation of interest rates could result in higher efficiency. Nevertheless, as further liberalisation initiatives take place, which was the introduction of new foreign players, both cost- and standard profit-efficiency were trending downwards as these new banks' efficiency scores could affect the net impact of these efficiency scores. However, the alternative profit efficiency has increased steadily, indicating a stronger presence of imperfect competition (i.e. oligopoly market structure).

As the timeline moved to the third phase of FSMP (2008–2011), Malaysia was also affected by the impact of the global credit crisis. There were large influx of funds flowing into the country from Europe and US searching for higher yields in mid-2007. These large inflows affected the stability of liquidity, as seen in 2008, when banks in Europe and the US deleveraged, leading to high withdrawal of funds and volatility in the liquidity of Malaysian banks. As a result, average efficiency scores were trending downward. In general, the results of efficiency scores indicate alternative profit efficiency appears to be greater than cost-and standard profit-efficiency. This may imply that, in a more liberalised environment, banks have been more innovative and have offered a variety of products and services to customers, intending to maximise banks' revenue. In doing so, banks were affected by higher costs, but at the same time increased revenue in higher proportion (Berger and Mester, 2003).

Concurrently, banks may have incurred higher costs to capture more market share, which consequently resulted in banks facing more inherent risks (Vivas, 1997). Hence, banks with a greater market share have higher market power in pricing their products and services, as reflected by the higher scores of alternative profit efficiency.

In this chapter, several hypotheses developed in Chapters 4 and 5 have been tested. In terms of ownership, domestic banks were found to be more cost efficient than foreign banks, but were no better than foreign banks with regards to profit efficiency (hypothesis 2). This could be potentially attributed to various restrictions made by BNM on foreign banks to protect the domestic banks. However, the very marginally higher profit efficiencies may suggest that foreign banks practice ‘cherry-picking’, choosing quality customers with niche banking activities that could give higher profits, rather than the ‘mass’ banking activities of the domestic banks (Hassan, 2002).

In terms of specialisation, Islamic banks were more cost efficient than conventional banks, but were no better in terms of profit efficiency (hypothesis 3). This could be likely explained by BNM initiatives to increase the prominence of Islamic banks through the introduction of Islamic banks’ subsidiaries (which formerly operated as IBS, in a dual banking scheme) during the second phase of FSMP (2004–2007). Despite this initiative, these new Islamic banking subsidiaries still rely on their associated conventional banks within their respective groups. Hence, these new Islamic banks’ subsidiaries enjoy cost savings from the cost-sharing activities, which largely have to be borne by their parents or associated conventional banks. On the other hand, the lower profit efficiency of Islamic banks can be attributed to their shorter establishment age on the Malaysian banking scene compared to conventional banks, strict shariah rules in terms of what businesses can be conducted, limited product and services diversity, and unknown shariah ‘jargon’ that cannot be easily understood by customers (Kamaruddin et al, 2008; Abdul Majid et al., 2011).

With regards to size, large banks are less cost efficient but more profit efficient compared to small banks (hypothesis 4). This could be likely explained from the need of large banks to invest in major restructuring and cost-cutting activities more than smaller banks, particularly in the period where banks were facing financial liberalisation (Akhigbe and McNulty, 2003).

In addition, most of these large banks are publicly traded; and with diverse ownership comes greater agency costs, which could lead to lower cost efficiency (Kwan, 2006). At the same time, this result may also support the ‘quiet life’ hypothesis (Hicks, 1935) due to significant market power by larger banks, where managers may not have the incentives to work as hard to keep costs under control and enjoying ‘quiet life’ (Berger and Hannan, 1998). The large banks were more profit efficient than their smaller rivals because they enjoy sizeable market power, which could result in banks facing less pressure to increase the quality of banking services, subsequently reducing operating costs to maximise profits. Larger banks also benefitted from the greater technology changes (e.g. advances in information-processing technology), which could improve their profitability. They also enjoy extensive branch networks and have better opportunities to get new customers.

In terms of market structure under the traditional SCP hypothesis, it was observed that a higher concentration exhibits lower cost efficiency, but higher profit efficiency (hypothesis 6a). The effect of higher concentration was largely driven by the consolidation of domestic banks in the first phase of the FSMP, which consequently resulted in lower cost efficiency due to higher costs incurred during the post-consolidation period. On the other hand, banks with a larger market share were found to generate higher cost efficiency (hypothesis 6b), supporting the RMP hypothesis under which banks with a large market share, and well-differentiated products, exert more market power to achieve greater performance (Shepherd, 1982).

In term of banks’ inherent risks (hypotheses 7–9), for capital adequacy, there was no significant relationship between capital and cost efficiency but higher capital adequacy exhibits higher profits (hypothesis 7). This is particularly true for banks with a larger margin of safety and greater ability to absorb potential risks. Hence, a bank with a strong capital level normally enjoys reliable access to sufficient funds on favourable terms (Berger, 1995). For asset quality, higher asset quality (NPL ratio) exhibits lower costs and lower profits (hypothesis 8). The direction of the coefficient in the estimation of parameters supports hypothesis 8 because NPLs increase costs and reduce profit, but the t-ratios are insignificant. In terms of liquidity, higher liquidity exhibits lower costs, but higher profits

(hypothesis 9). This, therefore, indicates that banks that are more liquid require higher expenses and resources to manage excess liquidity and work harder to maximise profits.

This chapter has also reported and compared empirical results from different methodologies between parametric and nonparametric models. Following Bauer et al. (1998)'s five multilevel consistency tests, the results to some extent are consistent with past empirical literature, which in most cases indicate low and moderate compatibility across different methods (Bauer et al., 1998). The differences between efficiency scores obtained from different approaches are attributed mainly to the inherent advantages and disadvantages discussed earlier (Dong et al., 2014). From the five consistency conditions, low consistencies are shown in rank-order correlations of different methods and identification of best- and worst-practice banks (consistency condition test 2 and 3). On the other hand, there were moderate consistencies in the comparable means, stability over time and comparison against financial performance indicators (consistency condition test 1, 4 and 5). From these results, it is important for regulators and other interested parties to cautiously treat the results on Malaysian banks that seem to be sensitive to the methods selected for frontier estimation. Hence, the use of multiple frontier methods for consistency checking is recommended. In this chapter, the main characteristics of Malaysian banks, using attributes such as CAMEL, ownership structure, specialisation and size, were analysed and discussed. From the CAMEL rating system, based on the best-practice and worst-practice banks, the financial indicator used for CAMEL acronyms were consistent with the cost- and profit-efficiency scores.

Additionally, this chapter explores whether scale economies exist for the Malaysian banking industry, using both SFA and DEA models. In general, small banks face economies of scale, large banks display diseconomies of scale and medium-sized banks appear to be most scale efficient. Hence, the average cost curve display a U-shape, with medium-sized banks being more scale efficient than very small or very large banks. In term of scale efficiency, the result exhibits small difference of less than 5.0% between scale inefficiency and scale economies, which is consistent with Berger et al. (1993, 1994) and Evanoff and Israilevich (1991).

Finally, this chapter has examined the technological change of Malaysian banks. The contribution of technology towards Malaysian banks was positive and reduced real annual production costs by approximately 11.3%. In terms of composition, the main source of technological change for Malaysian banks during the period under study (2000–2011) was majorly influenced by pure technical change, at 10.3%, instead of scale augmentation with a change of 3.5%. This implies that Malaysian banks have taken full advantage of new technological investments, influenced by various regulatory initiatives implemented by BNM, such as implementation of Basel II, which requires further enhancement of technological investment in data management and the implementation of internet banking.

Chapter 7 Conclusions

7.1 Introduction and Summary of Findings

This chapter provides the conclusions and policy implications of this research, as well as limitations and suggestions for future research. The conclusions are reached after taking into account the results and discussion of the findings in previous chapters. This thesis examines the impact of gradual financial liberalisation via the FSMP (a ten-year liberalisation programme introduced in 2001) in the level of efficiency of Malaysian banks. Prior to the Asian financial crisis, the banking system in Malaysia was fragmented, with 33 domestic banking institutions. Banks at that time had a low capital level and were unable to face the macro pressures coming from economic vulnerabilities. In addition, the banks also over-relied on corporations for financing; less attention was given to SMEs and retail consumers, causing gaps in access to financing and increased concentration on banks' credit portfolios.

Before the Asian crisis, banks were found to have low risk management and corporate governance, resulting in a high level of fragility. Furthermore, the market conditions were very rigid, where prescriptive rules-based regulation and supervision was implemented in the financial sector. The pricing mechanism of banking products and services was also rigid, which did not encourage competition among financial players. These conditions made Malaysian banks more vulnerable to macroeconomic distress and the inability to withstand these pressures during the Asian financial crisis. A major response by the Malaysian government to these pressures has been a substantial consolidation measure, resulting in a reduction in the total number of banks to 10 domestic banking groups. At the same time, the FSMP was introduced to strengthen and further liberalise the Malaysian banking industry following the Asian financial crisis in 1997–98.

The implementation of the FSMP changed the financial landscape of the banking industry. Within the FSMP period, Malaysia witnessed a series of financial liberalisation measures, including; liberalisation of interest rate to market players, introduction of new foreign banks (both Islamic and conventional banks), branch liberalisation by allowing foreign banks to

increase their branches, de-pegging of Malaysian Ringgit to US Dollars, simplified product approval process, lifting of wage moratorium, and allowing of outsourcing of banks' non-core activities. With these initiatives, excessive government intervention in banks' operations was reduced.

Apart from investigating the impact of financial liberalisation on the Malaysian banking industry, this thesis seeks to test various contextual issues that influenced the level of efficiency of banks, such as the risk appetite of banks, ownership structure, specialisation, and market structure and control. To test these variables, this research employs parametric model's SFA and data of all Malaysian banks for the years 2000–2011. This research also employs three distinct economic efficiency concepts (cost-, standard profit- and alternative profit efficiency), using a number of different measurement methods, including translog functional forms and 1-stage analysis to a single data set. By performing 1-stage analysis, variables relating to internal bank-specific characteristics (e.g. credit risk, capital adequacy risk, liquidity risk, ownership, specialisation, bank size), macro-economic environment (e.g. liberalisation periods and the global credit crisis) and market structure (e.g. the market concentration and market share) were used to test the effect of heterogeneity. These control and environmental variables interact with input and output variables of cost-, standard profit- and alternative profit-efficiency functions, allowing efficiency scores to account for the differences between banks and explain what these differences mean to the costs and profits, and inefficiency (u_i) of Malaysian banks. Hence, various hypotheses were developed from the introduction of control and environmental variables. At the same time, a different set of combinations of control and environmental variables were tested in several stages to determine the best-fitting model (i.e. preferred models). Consequently, the preferred models from respective cost- and profit-efficiency models were employed to estimate the cost- and profit-efficiency scores.

Following Resti (1997) and Bauer et al. (1998), this research also estimates efficiency scores from nonparametric traditional DEA and new DEA (Tone, 2002) models to test the consistency between parametric and nonparametric approaches. Despite the differences between parametric and nonparametric approaches- attributed mainly to the inherent advantages and disadvantages (Dong et al., 2014)- five consistency condition tests were

performed. It was found that low consistencies were shown in rank–order correlations of different methods and identification of best- and worst-practice banks (consistency condition test 2 and 3) and moderate consistencies in the comparable means, stability over time and comparison against financial performance indicators (consistency condition test 1, 4 and 5).¹⁹¹ Therefore, it is important for regulators, and other interested parties, to treat the results on Malaysian banks cautiously; they seem to be sensitive to the methods selected for frontier estimation. Hence, the use of multiple frontier methods for consistency checking is recommended.

Also studied, was the impact of financial liberalisation on the cost- and profit-efficiency of the banks. The average cost, standard profit- and alternative profit-inefficiency, for the years 2000–2011, were 23.5%, 37.8% and 6.7% respectively. This research found that the cost- and profit-inefficiencies in Malaysian banks are substantial, where inefficient banks being approximately 20.0% to 40.0% less cost- and profit-efficient when compared to best-practice banks. Therefore, in order to produce the same level of outputs of the best-practice banks, these inefficient banks should improve their cost and profits by approximately 20.0% and 40.0%, respectively. The results stemming from cost efficiency scores appear to be greater than standard profit efficiency scores, but lower than alternative profit efficiency. The differences between cost- and profit-efficiency indicate that a cost efficient bank may not necessarily be profit efficient. Moreover, the higher alternative profit efficiency compared to standard profit efficiency implies imperfect competition. Despite various initiatives introduced to improve the degree of competition in the market (e.g. liberalising controlled interest rates regime, allowing foreign banks to increase branches) and reduce market concentration (e.g. introducing new foreign banks), these measures have yet to show any improvements due to their nascent or growing stages of implementation, particularly during the post-consolidation period of domestic banks.

In general, the results of efficiency scores indicate that alternative profit efficiency appears to be greater than cost efficiency. This implies that the Malaysian banks are more efficient

¹⁹¹ This result is similar to Bauer et al. (1998), who found mixed evidence and were unable to draw a policy conclusion based on the six consistency conditions proposed. For the first three conditions, they found that parametric and nonparametric approaches were not mutually consistent. However, a possible consistency is found in the final three conditions.

in generating profits than controlling costs. In a more liberalised environment, banks are equipped with control over their own interest rate pricing on loans and deposits, and become more innovative, offering a variety of products and services to customers to maximise profits. In doing so, banks were affected by higher costs, but at the same time increased higher proportion of revenue, (Berger and Mester, 2003). However, banks may have incurred higher costs to capture higher market share (Vivas, 1997). Hence, banks with greater market share would have higher market power in pricing their products and services, but faced less pressure to reduce costs (Srairi, 2010).

From the observation of trends in the first phase of the FSMP (2000–2003), this thesis found that cost- and standard profit-efficiency were trending downward, mainly due to the impact of post-consolidation of domestic banks. During this period, consolidated banks were expected to incur higher costs relating to merger and acquisition activities, such as rationalisation of branches, staff redundancy compensation and branding. Conversely, the alternative profit efficiency scores were on an increasing trend, suggesting that the effect of consolidation of domestic banks had resulted in greater market power by domestic banks.

In the second phase of the FSMP (2004–2007), the cost- and profit-efficiency scores declined marginally in 2004 due to adjustments that the banks had to make upon the introduction of the new interest rate framework (NIRF) by BNM, who liberalised the interest rate to the market players. However, cost and profit efficiencies were slightly higher in 2005, suggesting that deregulation of interest rates could result in higher efficiency. As further liberalisation initiatives took place in 2006, cost- and standard profit-efficiency declined with the introduction of Islamic foreign banks. These new banks had low efficiency scores, as they were ‘green’ in the market, which affected the net impact of the overall cost- and profit-efficiency scores. However, the alternative profit efficiency has increased steadily, indicating stronger presence of market power despite new competitive pressures.

In the third phase of the FSMP (2008–2011), Malaysia was also affected by the global credit crisis. In mid-2007, a large influx of funds flowed into the country from Europe and the US, searching for higher yields. These large inflows affected the stability in liquidity as reflected

in 2008 where banks in Europe and the US were deleveraging, leading to high withdrawals of funds and volatility in the liquidity of Malaysian banks. At the same time, BNM reduced the policy interest rates (i.e. the OPR) to stabilise the economy. As a result, the average efficiency scores were trending downward, mainly driven by banks allocating more resources to managing excess liquidity, as well as adjusting their input and output prices according to changes in policy interest rates and demand for credit.

Based on the additional research questions posed earlier in Chapter 1; coupled with hypotheses developed in Chapters 4 and 5, this research found mixed findings regarding the cost- and profit-efficiency of Malaysian banks. First, higher capital adequacy level has no impact on cost but exhibits higher profits. This is particularly true for banks with larger margin of safety and has greater ability to absorb potential risks. Hence, a bank with a strong capital level normally enjoys reliable access to sufficient sources of funds on favourable terms (Berger, 1995). In terms of asset quality, higher level of NPLs exhibits lower profits but has no impact to costs. Additionally, the higher the banks' liquidity was, the higher their profits, but with no impact on costs. As evidenced in 2007, the liquidity of Malaysian banks increased due to large inflows of foreign funds. Malaysian banks probably earned more income from conversion of excess funds into short term assets.

Second, domestic banks were found to be more cost efficient than foreign banks, but foreign banks were slightly better than domestic banks with regards to profit efficiency. This could be potentially attributed to various restrictions made by BNM on foreign banks to protect the domestic banks. However, the marginally higher profit efficiencies could be attributed to the foreign banks' practicing 'cherry-picking' – choosing quality customers with niche banking activities that could give higher profits – rather than following the 'mass' banking activities of the domestic banks (Hassan, 2002).

Third, Islamic banks were more cost efficient than conventional banks, but were no better in terms of profit efficiency. This can be likely explained from the effect of Islamic banking subsidiaries enjoying cost savings from cost sharing activities, which largely had been borne by their parents or associated conventional banks. On the other hand, the lower profit efficiency of Islamic banks was attributed to their shorter establishment age, strict shariah

rules on business activities, limited diversity in products, and obscure shariah ‘jargon’, which could not be easily understood by customers (Kamaruddin et al, 2008; Abdul Majid et al., 2011).

Fourth, large banks are less cost efficient, but more profit efficient, compared to small banks. During the period of liberalisation, the large banks have to invest in more major restructuring and cost-cutting activities than the smaller banks (Akhigbe and McNulty, 2003). In addition, these large banks are publicly traded; and with diverse ownership come with greater agency costs, which could lead to lower cost efficiency (Kwan, 2006). Also, due to significant market power by these larger banks, their managers might not have had the incentives to work at keeping costs under control and instead, enjoyed the ‘quiet life’ (Berger and Hannan, 1998). The large banks were more profit efficient than their smaller counterparts, because they enjoy sizeable market power, which could result in less pressure to increase the quality of banking services, reducing the overall operating costs to maximise profits. Larger banks in Malaysia also benefitted from the greater technology changes (e.g. advances in information-processing technology).

Fifth, a higher market concentration in the Malaysian banking industry exhibits in lower cost efficiency, but higher profit efficiency. The effects of higher concentration were largely driven by the consolidation of domestic banks in the first phase of the FSMP, which consequently resulted in lower cost efficiency, due to the higher costs of the post-consolidation period. On the other hand, banks with larger market share were found to generate higher profit efficiency, where banks with large market share and well-differentiated products exert more market power to maximise profits (Shepherd, 1982).

In terms of the characteristics of Malaysian banks, the CAMEL rating system showed that both best- and worst-practice banks were mostly consistent with the cost- and profit-efficiency scores. Apart from having lower capital, best-practice banks had slightly more NPLs and less liquidity, lower operating costs, higher earnings and less dependence on purchased funds. On the other hand, the worst-practice banks appear to have lower NPLs and higher capital (due to regulator’s interventions to maintain public confidence). Higher

operating costs were incurred by these banks, owing to greater resources in managing potential delinquent loans and excess liquidity.

This research also found that scale economies were prevalent across the Malaysian banking industry. Small banks face economies of scale, large banks display diseconomies of scale and medium-sized banks appear to be most scale efficient. Hence, the average cost curve displays a U-shape, with medium-sized banks being more scale efficient than very small or very large banks. In addition, from the result of technological change, the research found that the contribution of technology towards Malaysian banks was positive and could reduce the real annual production cost by approximately 11.3%. In terms of composition, the main source of technological change for Malaysian banks during the period under study (2000–2011) is majorly influenced by pure technical change, at 10.3%, instead of scale augmentation change, at 3.5%. This implies that the banks took full advantage of new technological investment, influenced by various regulatory initiatives implemented by BNM such as, Basel II in 2008, which requires further enhancement in technological investments for data management and implementation of internet banking in 2001.

7.2 Policy Implications

The objective of this research was to examine the impact of financial liberalisation on the efficiency of Malaysian banks. Understanding the impact of financial liberalisation on the Malaysian banking industry is of great interest to policymakers, bankers, analysts and economists. Such an understanding would provide some explanation of what has happened to these banks following the implementation of financial liberalisation via the FSMP in 2001. Generally, the implementation of financial liberalisation through various deregulation measures such as, deregulation of interest rates and removing entry barriers to foreign banks should increase the efficiency of banks. However, this has not been the case. The effect of banking consolidation during the first phase of the FSMP had a significant impact on market concentration in the Malaysian banking industry, and this appears to have led to imperfect competition. Although the market competition situation has improved, through the introduction of foreign Islamic banks and foreign conventional banks, the gradual liberalisations introduced in the second (2004-2007) and third phases (2008-2011) of the

FSMP have yet to realise their potential and are still at a nascent stage. Therefore, BNM ought to accelerate their liberalisation initiatives to ensure adequate competitive pressures on large domestic banks. Probably, greater participation of foreign ownership through equity can be exploited as a catalyst for more efficiency. Besides equity participation, foreign banks should be allowed to set up more branches (than what has been permitted by BNM) all over the country, especially in rural and suburban areas.

The market concentration has increased since the consolidation of domestic banks during the first phase of the FSMP (2000-2003). With more domestic banks merging in 2005 and 2010, these domestic banks have increased in size, changing the structure of the banking market. Thus, BNM should monitor the impact of these large banks in the market. A small number of large domestic banks could lead to collusive strategies, anti-competitive behaviour; and hence, can result in greater risks towards public welfare. Furthermore, market power may lead to lower efficiency in large banks, with managers enjoying the 'quiet life', and earning higher interest rates on loans and deposits. Despite the introduction of the Competition Act, 2012, BNM should play a more active role in monitoring the competitive conditions of the banking industry.

Additionally, this research has investigated the determinants (i.e. control and environmental variables) of banking inefficiency. Inefficient banks in Malaysia could be characterised as having high operating costs, low earnings and high dependence on purchased funds. Knowing this, the regulator can design appropriate policies with the aim of improving the operations and performance of inefficient banks. This research found consistent results against the widely-used CAMEL ratings and the efficiency scores derived from frontier measurements. Therefore, frontier measurement was found to be one of the most useful measurement tools which could be adopted by BNM to support policy decisions towards the Malaysian banking Industry.

With regard to technology, pure technological change was found to be the main contributor, apart from managerial improvement in efficiency. Technological improvements can shift the production function, assuming other factors remain the same. Therefore, the regulators might provide further incentives or relax some policies relating to technological innovation.

With this, banks are expected to be able to reduce their operating costs over a period of time and create greater access for customers (e.g. internet banking). Nonetheless, these initiatives should not compromise prudential regulation.

7.3 Limitations of Present Research and Future Research

In this research, the parametric SFA model was employed as the main approach to measure the efficiency of Malaysian banks; and nonparametric DEA models were used to test the consistency of efficiency scores derived from SFA. As discussed earlier in previous chapters, different methodologies of frontier measurement could result in different efficiency scores, due to their inherent advantages and disadvantages. This research only employs three types of efficiency measurement: SFA, DEA and NDEA. For benchmarking purposes, a combination of different methodologies (e.g. comparison of SFA with thick frontier analysis (TFA) and distribution free analysis (DFA); and between DEA and free disposal hull (FDH)) could give a better picture regarding inefficiency estimates of Malaysian banks. Nonetheless, regarding the inefficiency assumptions of SFA, this study employed a half-normal model, instead of gamma distribution, exponential and truncated-normal assumptions, which could result in different inefficiency estimates. A comparison of the results derived using these different inefficiency assumptions would provide stronger support to our findings. Therefore, in future, this research could be extended by employing the methods and techniques mentioned above.

In addition, different variables used as measurements of efficiency could derive different inefficiency estimates as well. For instance, off balance sheet items may also influence the operating costs and profitability of banks. However, due to data not available, this research does not include off balance sheet items. Similarly, data for the number of employees and the risk weighted capital ratio (RWCR) are also unavailable. Therefore, a complete study that included these data might result in different inefficiency estimates. Additionally, while the multiproduct nature of banks is widely recognised, it is not always possible to include all items or dimensions of a bank's output in model specifications. Thus, for this research, an intermediation approach was utilised, but a concurrent production approach could also be used for comparison purposes.

Prior to the introduction of Islamic banks' subsidiaries, conventional banks in Malaysia also offered Islamic banking products that were managed under a separate window via IBS. Since the data of Islamic windows and conventional banks are not separated, the efficiency of these banks was estimated as conventional banks until the disaggregation of Islamic windows into legally incorporated Islamic bank subsidiaries. This limitation warrants careful judgement regarding overall Malaysian Islamic banking's assets and liabilities, particularly for conventional banks with IBS during first (2000-2003) and second phase (2004-2007) of the FSMP.

Another shortcoming of the present research may relate to the sample size, which was only confined to Malaysian banks. Therefore, the research could be extended to a number of Southeast Asian countries (SEA; e.g. Indonesia, Singapore, Philippines, Vietnam, Myanmar, Brunei and Thailand) and banking sector efficiency across those countries. This would be an interesting analysis, especially as most SEA countries were affected by the Asian financial crisis of 1997–98 and different paths were taken by each country in terms of macro-economic policy frameworks, with some relying on the reform programmes of the International Monetary Fund (IMF) for economic recovery. However, the lack of public data from these countries makes it difficult to perform such research.

An additional limitation to this research may relate to the time considered for the empirical analysis. Although the whole period of the FSMP was covered -which involved a variety of transformations and liberalisation initiatives- these initiatives did not have an immediate effect on the Malaysian banking industry. Therefore, it would be interesting to carry out the same research beyond 2011, to particularly see the impact of competition by the newly introduced foreign banks in the late second and third phases of the FSMP.

Bibliography

- Abd-Kadir, H., Selamat, Z. & Idros, M. 2010, "Productivity of Malaysian Banks after Mergers and Acquisition", *European Journal of Economics, Finance and Administrative Sciences*, , no. 24.
- Abdul Majid, M., Saal, D.S. & Battisti, G. 2011, "The impact of Islamic banking on the cost efficiency and productivity change of Malaysian commercial banks", *Applied Economics*, vol. 43, no. 16, pp. 2033-2054.
- Abiad, A. & Mody, A. 2003, *Financial reform: What shakes it? What shapes it?* International Monetary Fund.
- Adams, R.M., Berger, A.N. & Sickles, R.C. 1999, "Semiparametric Approaches to Stochastic Panel Frontiers with Applications in the Banking Industry", *Journal of Business & Economic Statistics*, vol. 17, no. 3, pp. 349-358.
- Afriat, S.N. 1972, "Efficiency Estimation of Production Functions", *International Economic Review*, vol. 13, no. 3, pp. 568-598.
- Ahmad Mokhtar, H.S., AlHabshi, S.M. & Abdullah, N. 2006, "A Conceptual Framework for and Survey of Banking Efficiency Study", *UNITAR E-Journal*, vol. 2, no. 2, pp. 1-19.
- Aigner, D.J. & Chu, S.F. 1968, "On Estimating the Industry Production Function", *The American Economic Review*, vol. 58, no. 4, pp. 826-839.
- Aigner, D., Lovell, C.A.K. & Schmidt, P. 1977, "Formulation and estimation of stochastic frontier production function models", *Journal of Econometrics*, vol. 6, no. 1, pp. 21-37.
- Akhavein, J.D., Berger, A.N. & Humphrey, D.B. 1997, "The Effects of Megamergers on Efficiency and Prices: Evidence from a Bank Profit Function", *Review of Industrial Organisation*, vol. 12, no. 1, pp. 95-139.
- Akhigbe, A. & McNulty, J.E. 2003, "The profit efficiency of small US commercial banks", *Journal of Banking & Finance*, vol. 27, no. 2, pp. 307-325.
- Alam, I.M.S. 2001, "A Nonparametric Approach for Assessing Productivity Dynamics of Large U.S. Banks", *Journal of Money, Credit and Banking*, vol. 33, no. 1, pp. 121-139.
- Al-Jarrah, I. & Molyneux, P. 2007, "Efficiency in Arabian Banking", *Jordan Journal of Business Administration*, vol. 3, no. 3, pp. 373-390.

- Allen, L. & Rai, A. 1996, "Operational efficiency in banking: An international comparison", *Journal of Banking & Finance*, vol. 20, no. 4, pp. 655-672.
- Altunbas, Y. & Chakravarty, S.P. 2001, "Frontier cost functions and bank efficiency", *Economics Letters*, vol. 72, no. 2, pp. 233-240.
- Altunbas, Y., Evans, L. & Molyneux, P. 2001, "Bank Ownership and Efficiency", *Journal of Money, Credit and Banking*, vol. 33, no. 4, pp. 926-954.
- Altunbas, Y., Goddard, J. & Molyneux, P. 1999, "Technical change in banking", *Economics Letters*, vol. 64, no. 2, pp. 215-221.
- Altunbas, Y., Liu, M., Molyneux, P. & Seth, R. 2000, "Efficiency and risk in Japanese banking", *Journal of Banking & Finance*, vol. 24, no. 10, pp. 1605-1628.
- Aly, H.Y., Grabowski, R., Pasurka, C. & Rangan, N. 1990, "Technical, scale and allocative efficiencies in US banking: an empirical investigation", *Review of Economics and Statistics*, vol. 72, no. 2, pp. 211-218.
- Ang, J.B. 2009, "Financial development and the FDI-growth nexus: the Malaysian experience", *Applied Economics*, vol. 41, no. 13, pp. 1595-1601.
- Ang, J.B. 2007, "Are saving and investment cointegrated? The case of Malaysia (1965–2003)", *Applied Economics*, vol. 39, no. 17, pp. 2167-2174.
- Ang, J.B. 2008, "Determinants of foreign direct investment in Malaysia", *Journal of Policy Modeling*, vol. 30, no. 1, pp. 185-189.
- Ang, J.B. & McKibbin, W.J. 2007, "Financial liberalization, financial sector development and growth: Evidence from Malaysia", *Journal of Development Economics*, vol. 84, no. 1, pp. 215-233.
- Ansari, M.S. 2007, "An empirical investigation of cost efficiency in the banking sector of Pakistan", *SBP Research Bulletin*, vol. 3, no. 2.
- Arestis, P. 2005, "Financial Liberalisation and the relationship between Finance and Growth", *CEPP Working Paper 05/05, Centre for Economic and Public Policy, University of Cambridge*, vol. June 2005.
- Arestis, P. & Demetriades, P. 1997, "Financial Development and Economic Growth: Assessing the Evidence", *The Economic Journal*, vol. 107, no. 442, pp. 783-799.
- Arestis, P., Nissanke, M. & Stein, H. 2003, "Finance and Development: Institutional and Policy Alternatives to Financial Liberalization", *Economics Working Papers, The Levy Economics Institute of Bard College*, vol. WP 377.
- Ariff, M. & Can, L. 2008, "Cost and profit efficiency of Chinese banks: A non-parametric analysis", *China Economic Review*, vol. 19, no. 2, pp. 260-273.

- Assenmacher-Wesche, K. & Gerlach, S. 2010, "Monetary policy and financial imbalances: facts and fiction", *Economic Policy*, vol. 25, no. 63, pp. 437-482.
- 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*, vol. 98, no. 2, pp. 300-313.
- Athukorala, P. 2002, "Malaysian Trade Policy and the 2001 WTO Trade Policy Review", *World Economy*, vol. 25, no. 9, pp. 1297-1317.
- Athukorala, P. & Menon, J. 1995, "Developing with Foreign Investment: Malaysia", *Australian Economic Review*, vol. 28, no. 1, pp. 9-22.
- Attaulah, A., Cockerill, T. & Le, H. 2004, "Financial Liberalization and Bank Efficiency: A Comparative Analysis of India and Pakistan", *Applied Economics*, vol. 36, no. 17, pp. 1915-1924.
- Avkiran, N.K. 2000, "Rising productivity of Australian trading banks under deregulation 1986–1995", *Journal of Economics and Finance*, vol. 24, no. 2, pp. 122-140.
- Avkiran, N.K. 1999, "The evidence on efficiency gains: The role of mergers and the benefits to the public", *Journal of Banking & Finance*, vol. 23, no. 7, pp. 991-1013.
- Awdeh, A., El-Moussawi, C. & Machrouh, F. 2011, "The effect of capital requirements on banking risk", *International Research Journal of Finance and Economics*, vol. 66, pp. 133-146.
- Bain, J.S. 1951, "Relation of Profit Rate to Industry Concentration", *Quarterly Journal of Economics*, vol. 65, no. August, pp. 293-324.
- Baker, M. & Wurgler, J. 2013, "Do Strict Capital Requirements Raise the Cost of Capital? Banking Regulation and the Low Risk Anomaly", *NBER Working Paper No 19018*, , pp. 1-51.
- Ball, C.A. & Tschoegl, A.E. 1982, "The decision to establish a foreign bank branch or subsidiary: an application of binary classification procedures", *Journal of financial and Quantitative Analysis*, vol. 17, no. 3, pp. 411-424.
- Ball, C.A. & Tschoegl, A.E. 1982, "The Decision to Establish a Foreign Bank Branch or Subsidiary: An Application of Binary Classification Procedures", *The Journal of Financial and Quantitative Analysis*, vol. 17, no. 3, pp. 411-424.
- Baltagi, B.H. & Griffin, J.M. 1988, "A General Index of Technical Change", *Journal of Political Economy*, vol. 96, no. 1, pp. 20-41.
- Bank Negara Malaysia 2011c, *Financial Sector Blueprint 2011 - 2020*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2011b, *Financial Stability and Payment System Report 2011*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2011a, *Annual Report 2011*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2010b, *Financial Stability and Payment System Report 2010*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2010a, *Annual Report 2010*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2009b, *Financial Stability and Payment System Report 2009*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2009a, *Annual Report 2009*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2008b, *Financial Stability and Payment System Report 2008*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2008a, *Annual Report 2008*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2007b, *Financial Stability and Payment System Report 2007*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2007a, *Annual Report 2007*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2006b, *Financial Stability and Payment System Report 2006*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2006a, *Annual Report 2006*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2005, *Annual Report 2005*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2004, *Annual Report 2004*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2003, *Annual Report 2003*, Bank Negara Malaysia Publication, Kuala Lumpur.

Bank Negara Malaysia 2002, *Annual Report 2002*, Bank Negara Malaysia Publication, Kuala Lumpur.

- Bank Negara Malaysia 2001b, *Financial Sector Master Plan 2001-2010*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 2001a, *Annual Report 2001*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 2000, *Annual Report 2000*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 1999b, *The Central Bank and The Financial System in Malaysia - A Decade of Change 1989 - 1999*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 1999a, *Annual Report 1999*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 1998, *Annual Report 1998*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Bank Negara Malaysia 1997, *Annual Report 1997*, Bank Negara Malaysia Publication, Kuala Lumpur.
- Banker, R.D., Charnes, A. & Cooper, W.W. 1984, "Some Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis", *Management Science*, vol. 30, no. 9, pp. 1078-1092.
- Barr, R.S., Kilgo, K.A., Siems, T.F. & Zimmel, S. 2002, "Evaluating the Productive Efficiency and Performance of U.S. Commercial Banks", *Managerial Finance*, vol. 28, no. 8, pp. 3-25.
- Barth, J.R., Caprio, G. & Levine, R. 2004, "Bank regulation and supervision: what works best?", *Journal of Financial Intermediation*, vol. 13, no. 2, pp. 205-248.
- Basel Committee on Banking Supervision 2010, *Basel III: International Framework for Liquidity Risk Measurement, Standards and Monitoring*, Bank of International Settlements, Basel.
- Basel Committee on Banking Supervision 2005, *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework*, Bank of International Settlements, Basel.
- Basel Committee on Banking Supervision 1999, *A New Capital Adequacy Framework*, Bank of International Settlements, Basel.
- Batchelor, V.B., Kuppusamy, K. & Allen, D.E. 2005, "The Technological Progress of Malaysian Banks: An Empirical Investigation", *School of Accounting, Finance and Economics and FIMARC Working Paper Series, Edith Cowan University*, , no. Working Paper 0502, pp. 1-13.

- Batchelor, V.B. & Wadud, I.K.M.M. 2004, "Technical and Scale Efficiency of Islamic Banking Operations in Malaysia: An Empirical Investing With A Longitudinal Perspective", *Labuan Bulletin of International Business & Finance*, vol. 2, no. 1, pp. 52-69.
- Battese, G.E. & Coelli, T.J. 1995, "A model for technical inefficiency effects in a stochastic frontier production function for panel data", *Empirical Economics*, vol. 20, no. 2, pp. 325-332.
- Battese, G.E. & Coelli, T.J. 1992, "Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India", *Journal of Productivity Analysis*, vol. 3, no. 1, pp. 153-169.
- Battese, G.E. & Corra, G.S. 1977, "Estimation of a Production Frontier Model: with Application to the Pastoral Zone of Eastern Australia", *Australian Journal of Agricultural Economics*, vol. 21, no. 3, pp. 169-179.
- Bauer, P.W., Berger, A.N., Ferrier, G.D. & Humphrey, D.B. 1998, "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods", *Journal of economics and business*, vol. 50, no. 2, pp. 85-114.
- Baumol, W.J. 1958, "On the Theory of Oligopoly", *Economica*, vol. 25, no. 99, pp. 187-198.
- Baumol, W.J., Panzar, J.C. & Willig, R.D. 1988, *Contestable markets and the theory of industry structure*, Revised Edition edn, Harcourt Brace Jovanovich, San Diego.
- Bayraktar, N. & Wang, Y. 2004, "Foreign bank entry, performance of domestic banks, and sequence of financial liberalization", *World Bank Policy Research Working Paper*, , no. 3416.
- Beck, T., Demirgüç-Kunt, A. & Levine, R. 2006, "Bank concentration, competition, and crises: First results", *Journal of Banking & Finance*, vol. 30, no. 5, pp. 1581-1603.
- Beck, T., Demirgüç-Kunt, A. & Merrouche, O. 2013, "Islamic vs. conventional banking: Business model, efficiency and stability", *Journal of Banking & Finance*, vol. 37, no. 2, pp. 433-447.
- Beck, T. & Hesse, H. 2006, "Bank Efficiency, Ownership and Market Structure: Why are Interest Spreads so High in Uganda?", *Oxford University Department of Economics Discussion Paper Series*, vol. No 277, no. September 2006.
- Beck, T. & Levine, R. 2004, "Stock markets, banks, and growth: Panel evidence", *Journal of Banking & Finance*, vol. 28, no. 3, pp. 423-442.
- Belsley, D.A. 1984, "Collinearity and forecasting", *Journal of Forecasting*, vol. 3, no. 2, pp. 183-196.

- Bencivenga, V.R. & Smith, B.D. 1991, "Financial Intermediation and Endogenous Growth", *The Review of Economic Studies*, vol. 58, no. 2, pp. 195-209.
- Benston, G.J. & Kaufman, G.G. 1996, "The Appropriate Role of Bank Regulation", *The Economic Journal*, vol. 106, no. 436, pp. 688-697.
- Berg, S.A., Førsund, F.R., Hjalmarsson, L. & Suominen, M. 1993, "Banking efficiency in the Nordic countries", *Journal of Banking & Finance*, vol. 17, no. 2-3, pp. 371-388.
- Berg, S.A. & Kim, M. 1998, "Banks as Multioutput Oligopolies: An Empirical Evaluation of the Retail and Corporate Banking Markets", *Journal of Money, Credit and Banking*, vol. 30, no. 2, pp. 135-153.
- Bergendahl, G. 1998, "DEA and benchmarks—an application to Nordic banks", *Annals of Operations Research*, vol. 82, pp. 233-250.
- Berger, A.N. 1993, "“Distribution-free” estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions", *Journal of Productivity Analysis*, vol. 4, no. 3, pp. 261-292.
- Berger, A.N. & di Patti, E.B. 2006, "Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and An Application to the Banking Industry", *Journal of Banking & Finance*, vol. 30, no. 4, pp. 1065-1102.
- Berger, A.N. & di Patti, E.B. 2003, "Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and an Application to the Banking Industry", *FEDS Working Paper No. 2002-54*, .
- Berger, A.N. & Humphrey, D.B. 1992, "Measurement and Efficiency Issues in Commercial Banking" in *Output Measurement in the Service Sectors*, ed. Z. Griliches, National Bureau of Economic Research Vol. 56 edn, University of Chicago Press, Chicago, pp. 245-300.
- Berger, A.N., DeYoung, R., Genay, H. & Udell, G.F. 2000, "Globalization of financial institutions: Evidence from cross-border banking performance", *Brookings-Wharton papers on financial services*, vol. 2000, no. 1, pp. 23-120.
- Berger, A.N. & Hannan, T.H. 1998, "The efficiency cost of market power in the banking industry: A test of the “quiet life” and related hypotheses", *Review of Economics and Statistics*, vol. 80, no. 3, pp. 454-465.
- Berger, A.N. 1995, "The Profit-Structure Relationship in Banking--Tests of Market-Power and Efficient-Structure Hypotheses", *Journal of Money, Credit and Banking*, vol. 27, no. 2, pp. 404-431.
- Berger, A.N., Clarke, G.R.G., Cull, R., Klapper, L. & Udell, G.F. 2005, "Corporate governance and bank performance: A joint analysis of the static, selection, and dynamic effects of domestic, foreign, and state ownership", *Journal of Banking & Finance*, vol. 29, no. 8-9, pp. 2179-2221.

- Berger, A.N., De Young, R., Genay, H. & Udell, G.F. 2000, "Corporate Governance and Bank Performance: A Joint Analysis of the Static, Selection and Dynamic Effects of Domestic, Foreign and State Ownership", *Brookings Wharton Papers on Financial Services*, vol. 3, pp. 23-158.
- Berger, A.N. & De Young, R. 1997, "Problem loans and cost efficiency in commercial banks", *Journal of Banking & Finance*, vol. 21, no. 6, pp. 849-870.
- Berger, A.N. & DeYoung, R. 2001, "The Effects of Geographic Expansion on Bank Efficiency", *Journal of Financial Services Research*, vol. 19, no. 2, pp. 163.
- Berger, A.N., Hancock, D. & Humphrey, D.B. 1993, "Bank efficiency derived from the profit function", *Journal of Banking & Finance*, vol. 17, no. 2-3, pp. 317-347.
- Berger, A.N. & Hannan, T.H. 1997, "Using Efficiency Measures to Distinguish Among Alternative Explanations of the Structure-Performance Relationship in Banking", *Managerial Finance*, vol. 13, no. 1, pp. 6-31.
- Berger, A.N., Hanweck, G.A. & Humphrey, D.B. 1987, "Competitive viability in banking : Scale, scope, and product mix economies", *Journal of Monetary Economics*, vol. 20, no. 3, pp. 501-520.
- Berger, A.N., Hasan, I. & Zhou, M. 2009, "Bank ownership and efficiency in China: What will happen in the world's largest nation?", *Journal of Banking & Finance*, vol. 33, no. 1, pp. 113-130.
- Berger, A.N. & Humphrey, D.B. 1997, "Efficiency of financial institutions: International survey and directions for future research", *European Journal of Operational Research*, vol. 98, no. 2, pp. 175-212.
- Berger, A.N. & Humphrey, D.B. 1994, "Bank Scale Economies, Mergers, Concentration and Efficiency: The U.S. Experience", *The Wharton School, University of Pennsylvania*, .
- Berger, A.N. & Humphrey, D.B. 1991, "The dominance of inefficiencies over scale and product mix economies in banking", *Journal of Monetary Economics*, vol. 28, no. 1, pp. 117-148.
- Berger, A.N., Leusner, J.H. & Mingo, J.J. 1997, "The efficiency of bank branches", *Journal of Monetary Economics*, vol. 40, no. 1, pp. 141-162.
- Berger, A.N. & Mester, L.J. 2003, "Explaining the dramatic changes in performance of US banks: technological change, deregulation, and dynamic changes in competition", *Journal of Financial Intermediation*, vol. 12, no. 1, pp. 57-95.
- Berger, A.N. & Mester, L.J. 1997, "Inside the black box: What explains differences in the efficiencies of financial institutions?", *Journal of Banking & Finance*, vol. 21, no. 7, pp. 895-947.

- Bertus, M., Jahera, J.S. & Yost, K. 2007, "A Relation between Bank Regulation and Economic Performance: A Cross-Country Analysis", *Banks and Bank System*, vol. 2, no. 3, pp. 32-45.
- Bhattacharyya, A., Lovell, C.A.K. & Sahay, P. 1997, "The impact of liberalization on the productive efficiency of Indian commercial banks", *European Journal of Operational Research*, vol. 98, no. 2, pp. 332-345.
- Bhattacharyya, A., Lovell, C.A.K. & Sahay, P. 1997, "The impact of liberalization on the productive efficiency of Indian commercial banks", *European Journal of Operational Research*, vol. 98, no. 2, pp. 332-345.
- Blöchlinger, A. & Leippold, M. 2006, "Economic benefit of powerful credit scoring", *Journal of Banking & Finance*, vol. 30, no. 3, pp. 851-873.
- Bonin, J.P., Hasan, I. & Wachtel, P. 2005, "Bank performance, efficiency and ownership in transition countries", *Journal of Banking & Finance*, vol. 29, no. 1, pp. 31-53.
- Borger, B. D., Ferrier, G. D., & Kerstens, K. 1998, "The choice of a technical efficiency measure on the free disposal hull reference technology: A comparison using US banking data". *European Journal of Operational Research*, vol. 105(3), pp. 427-446.
- Bos, J.W.B., Koetter, M., Kolari, J.W. & Kool, C.J.M. 2009, "Effects of heterogeneity on bank efficiency scores", *European Journal of Operational Research*, vol. 195, no. 1, pp. 251-261.
- Bos, J.W.B. & Kool, C.J.M. 2006, "Bank efficiency: The role of bank strategy and local market conditions", *Journal of Banking & Finance*, vol. 30, no. 7, pp. 1953-1974.
- Bos, J.W.B. & Schmiedel, H. 2007, "Is there a single frontier in a single European banking market?", *Journal of Banking & Finance*, vol. 31, no. 7, pp. 2081-2102.
- Boyd, J.H., Chang, C. & Smith, B.D. 2002, "Deposit insurance: a reconsideration", *Journal of Monetary Economics*, vol. 49, no. 6, pp. 1235-1260.
- Brown, K. 2003, "Islamic Banking Comparative Analysis", *The Arab Bank Review*, vol. 5, no. 2, pp. 43-50.
- Buch, C.M. 2003, "Information or Regulation: What Drives the International Activities of Commercial Banks?", *Journal of Money, Credit and Banking*, vol. 35, no. 6, Part 1, pp. 851-869.
- Caminal, R. & Matutes, C. 2002, "Market power and banking failures", *International Journal of Industrial Organization*, vol. 20, no. 9, pp. 1341-1361.
- Campus, A., Eatwell, J., Milgate, M. & Newman, P. 1987, "Marginalist economics", *The New Palgrave: A Dictionary of Economics, II, London: Macmillan, .*

- Canbas, S., Cabuk, A. & Kilic, S.B. 2005, "Prediction of commercial bank failure via multivariate statistical analysis of financial structures: The Turkish case", *European Journal of Operational Research*, vol. 166, no. 2, pp. 528-546.
- Capon, N. 1982, "Credit Scoring Systems: A Critical Analysis", *Journal of Marketing*, vol. 46, no. 2, pp. 82-91.
- Caprio, G. 1997, *The role of long term finance: theory and evidence*, World Bank Publications.
- Carter, E.E. 1971, "The Behavioral Theory of the Firm and Top-Level Corporate Decisions", *Administrative Science Quarterly*, vol. 16, no. 4, pp. 413-429.
- Castelli, L., Pesenti, R. & Ukovich, W. 2010, "A Classification of DEA Models when the Internal Structure of the Decision Making Units is Considered", *Annals of Operations Research*, vol. 173, no. 1, pp. 207-235.
- Casu, B., Girardone, C. & Molyneux, P. 2006, *Introduction to banking*, Pearson Education.
- Casu, B. & Girardone, C. 2002, "A comparative study of the cost efficiency of Italian bank conglomerates", *Managerial Finance*, vol. 28, no. 9, pp. 3-23.
- Casu, B. & Girardone, C. 2006, "Bank Competition, Concentration and Efficiency in the Single European Market", *Manchester School (14636786)*, vol. 74, no. 4, pp. 441-468.
- Catena, M. 2000, "Efficient Structure Hypothesis: An application to the Argentine Banking Sector", *Banco Central de la Republica Argentina*, vol. Working Paper 12, pp. 1-12.
- Cebenoyan, A.S., Cooperman, E.S., Register, C.A. & Hudgins, S.C. 1993, "The relative efficiency of stock versus mutual S&Ls: A stochastic cost frontier approach", *Journal of Financial Services Research*, vol. 7, no. 2, pp. 151-170.
- Cerutti, E. "Dell'Ariccia, Giovanni. Peria, Maria Soledad Martinez.(2007). How banks go abroad: Branches or subsidiaries", *Journal of Banking & Finance*, vol. 31, pp. 1669-1692.
- Cerutti, E., Dell'Ariccia, G. & Martínez Pería, M.S. 2007, "How banks go abroad: Branches or subsidiaries?", *Journal of Banking & Finance*, vol. 31, no. 6, pp. 1669-1692.
- Cetorelli, N. & Goldberg, L.S. 2011, "Global banks and international shock transmission: Evidence from the crisis", *IMF Economic Review*, vol. 59, no. 1, pp. 41-76.
- Chamberlin, E. 1933, "The Theory of Monopolistic Competition", *Harvard University Press, Cambridge MA*, .
- Chambers, R.G. 1988, *Applied Production Analysis*, Cambridge University Press, Cambridge.

- Chang, E.C., Hasan, I. & Hunter, W.C. 1998, "Efficiency of multinational banks: an empirical investigation", *Applied Financial Economics*, vol. 8, no. 6, pp. 689-696.
- Charnes, A., Cooper, W.W. & Rhodes, E. 1978, "Measuring the efficiency of decision making units", *European Journal of Operational Research*, vol. 2, no. 6, pp. 429-444.
- Chong, B.S. & Liu, M. 2009, "Islamic banking: Interest-free or interest-based?", *Pacific-Basin Finance Journal*, vol. 17, no. 1, pp. 125-144.
- Claessens, S. & van Horen, N. 2011, "Foreign Banks: Trends, Impact and Financial Stability", *De Nederlandsche Bank Working Paper*, , no. No. 330.
- Claessens, S., Demirgüç-Kunt, A. & Huizinga, H. 2001, "How does foreign entry affect domestic banking markets?", *Journal of Banking & Finance*, vol. 25, no. 5, pp. 891-911.
- Clark, J.A. 1988, "Economies of Scale and Scope at Depository Financial Institutions: A Review of Literature", *Federal Reserve Bank of Kansas City Economic Review*, , pp. 16-33.
- Clark, J.A. 1996, "Economic Cost, Scale Efficiency, and Competitive Viability in Banking", *Journal of Money, Credit and Banking*, vol. 28, no. 3, Part 1, pp. 342-364.
- Coase, R.H. 1937, "The Nature of the Firm", *Economica*, vol. 4, no. 16, pp. 386-405.
- Cobb, C.W. & Douglas, P.H. 1928, "A Theory of Prouction", *The American Economic Review*, vol. 18, no. 1, pp. 139-165.
- Coelli, T., Perelman, S. & Romano, E. 1999, "Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines", *Journal of Productivity Analysis*, vol. 11, no. 3, pp. 251-273.
- Coelli, T.J. 1996, "A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation", *Centre for Efficiency and Productivity Analysis (CEPA) Working Papers*, , no. 7/96.
- Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. & Battese, G.E. 2005, "An Introduction to Efficiency and Productivity Analysis", *Springer, New York*, .
- Coelli, T.J. 1995, "Recent Developments In Frontier Modelling and Efficiency Measurement", *Australian Journal of Agricultural and Resource Economics*, vol. 39, no. 3, pp. 219-245.
- Cook, W.D., Hababou, M. & Roberts, G.S. 2001, "The effects of financial liberalization on the Tunisian Banking Industry: a non-parametric approach", .
- Cook, W.D. & Seiford, L.M. 2009, "Data envelopment analysis (DEA) – Thirty years on", *European Journal of Operational Research*, vol. 192, no. 1, pp. 1-17.

- Cooper, W.W., Ruiz, J.L. & Sirvent, I. 2007, "Choosing weights from alternative optimal solutions of dual multiplier models in DEA", *European Journal of Operational Research*, vol. 180, no. 1, pp. 443-458.
- Cuesta, R.A. & Orea, L. 2002, "Mergers and technical efficiency in Spanish savings banks: A stochastic distance function approach", *Journal of Banking & Finance*, vol. 26, no. 12, pp. 2231-2247.
- Cummins, J.D. & Weiss, M.A. 1998, "Analysing Firm Performance in the Insurance Industry Using Frontier Efficiency Methods", *The Wharton School, University of Pennsylvania*, .
- Cyert, R.M. & George, K.D. 1969, "Competition, Growth, and Efficiency", *The Economic Journal*, vol. 79, no. 313, pp. 23-41.
- Cyert, R.M. & Hedrick, C.L. 1972, "Theory of the Firm: Past, Present, and Future; An Interpretation", *Journal of Economic Literature*, vol. 10, no. 2, pp. pp. 398-412.
- Cyert, R.M. & March, J.G. 1963, *A Behavioural Theory of the Firm*, Prentice-Hall, New Jersey.
- Dacanay, S.J.O. 2010, "The evolution of cost and profit efficiency of Philippine commercial banks", *Philippine Review of Economics*, vol. 47, no. 1, pp. 109-146.
- de Haan, L. & van den End, J.W. 2013, "Banks' responses to funding liquidity shocks: Lending adjustment, liquidity hoarding and fire sales", *Journal of International Financial Markets, Institutions and Money*, vol. 26, no. 0, pp. 152-174.
- de Haas, R. & van Lelyveld, I. 2006, "Foreign banks and credit stability in Central and Eastern Europe. A panel data analysis", *Journal of Banking & Finance*, vol. 30, no. 7, pp. 1927-1952.
- De Young, R. 1997, "Measuring Cost Efficiency: Don't Count on Accounting Ratios", *Financial Practice & Education*, vol. 7, no. 1, pp. 20-31.
- Debreu, G. 1951, "The Coefficient of Resource Utilization", *Econometrica*, vol. 19, no. 3, pp. pp. 273-292.
- Delis, M.D., Koutsomanoli-Fillipaki, A., Staikouras, C.K. & Katerina, G. 2009, "Evaluating cost and profit efficiency: a comparison of parametric and nonparametric methodologies", *Applied Financial Economics*, vol. 19, no. 3, pp. 191-202.
- Delis, M.D. & Papanikolaou, N.I. 2009, "Determinants of bank efficiency: evidence from a semi-parametric methodology", *Managerial Finance*, vol. 35, no. 3, pp. 260-275.
- Demirguc-Kunt, A. & Detragiache, E. 1999, "Monitoring banking sector fragility: a multivariate logit approach with an application to the 1996-97 banking crises", *No. 2085, The World Bank*, .

- Demirgüç-Kunt, A. & Levine, R. 1999, "Bank-based and market-based financial systems: Cross-country comparisons", *World Bank Policy Working Paper*, , no. 2143.
- Demsetz, H. 1997, "The Firm in Economic Theory: A Quiet Revolution", *The American Economic Review*, vol. 87, no. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, pp. pp. 426-429.
- Demsetz, H. 1982, "Barriers to Entry", *The American Economic Review*, vol. 72, no. 1, pp. pp. 47-57.
- Demsetz, H. 1973, "Industry Structure, Market Rivalry, and Public Policy", *Journal of Law and Economics*, vol. 16, no. 1, pp. 1-9.
- Demsetz, H. 1959, "The Nature of Equilibrium in Monopolistic Competition", *Journal of Political Economy*, vol. 67, no. 1, pp. pp. 21-30.
- Denizer, C., Desai, R.M. & Gueorguiev, N. 2000, "Financial Liberalization and Financial Repression in Formerly Socialist Economies", *Czech Journal of Economics and Finance (Finance a uver)*, vol. 50, no. 1, pp. 17-39.
- Department of Statistics Malaysia 2010, *The 2010 Population and Housing Census of Malaysia*, Department of Statistics Malaysia, Kuala Lumpur.
- Deprins, D., Simar, L. & Tulkens, H. 1984, "Measuring labor-efficiency in post offices", *M. Marchand, P. Pestieau, H. Tulkens, Editors , The performance of public enterprises: Concepts and measurement, North-Holland, Amsterdam, .*
- Detragiache, E. & Gupta, P. 2006, "Foreign banks in emerging market crises: Evidence from Malaysia", *Journal of Financial Stability*, vol. 2, no. 3, pp. 217-242.
- DeYoung, R. & Nolle, D.E. 1996, "Foreign-owned banks in the United States: Earning market share or buying it?", *Journal of Money, Credit and Banking*, vol. 28, no. 4, pp. 622-636.
- DeYoung, R. 1998, "The Efficiency of Financial Institutions: How Does Regulation Matter?", *Journal of economics and business*, vol. 50, no. 2, pp. 79-83.
- DeYoung, R. 1997, "A diagnostic test for the distribution-free efficiency estimator: An example using U.S. commercial bank data", *European Journal of Operational Research*, vol. 98, no. 2, pp. 243-249.
- DeYoung, R. & Hasan, I. 1998, "The performance of de novo commercial banks: A profit efficiency approach", *Journal of Banking & Finance*, vol. 22, no. 5, pp. 565-587.
- Diamond, D.W. 1984, "Financial intermediation and delegated monitoring", *The Review of Economic Studies*, vol. 51, no. 3, pp. 393-414.
- Diamond, D.W. & Rajan, R.G. 2000, "A Theory of Bank Capital", *The Journal of Finance*, vol. 55, no. 6, pp. 2431-2465.

- Diewert, W.E. 1974, "Applications of Duality Theory" in *Frontiers of Quantitative Economics*, eds. M.D. Intriligator & D.A. Kendrick, Vol. II edn, North-Holland, Amsterdam, pp. 106-176.
- Dogan, E. & Fausten, D.K. 2002, "Productivity and Technical Change in Malaysian Banking ", *Department of Economics Discussion Papers, Monash University*, vol. 5, no. 2, pp. 1-57.
- Dong, Y., Hamilton, R. & Tippett, M. 2014, "Cost efficiency of the Chinese banking sector: A comparison of stochastic frontier analysis and data envelopment analysis", *Economic Modelling*, vol. 36, no. 0, pp. 298-308.
- Drake, L. & Hall, M.J.B. 2003, "Efficiency in Japanese banking: An empirical analysis", *Journal of Banking & Finance*, vol. 27, no. 5, pp. 891-917.
- Drake, L., Hall, M.J.B. & Simper, R. 2006, "The impact of macroeconomic and regulatory factors on bank efficiency: A non-parametric analysis of Hong Kong's banking system", *Journal of Banking & Finance*, vol. 30, no. 5, pp. 1443-1466.
- Dunham, D. & Kelegama, S. 1994, "Economic Liberalization and Structural Reform: The Experience of Sri Lanka, 1977-93", *Working Paper Series No 163, Institute of Social Studies Colombo*, .
- Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S. & Shale, E.A. 2001, "Pitfalls and protocols in DEA", *European Journal of Operational Research*, vol. 132, no. 2, pp. 245-259.
- Eisenbeis, R.A., Ferrier, G.D. & Kwan, S.H. 1999, "The Informativeness of Stochastic Frontier and Programming Frontier Efficiency Scores: Cost Efficiency and Other Measures of Bank Holding Company Performance", *Federal Reserve Bank of Atlanta, Atlanta*, .
- Eisenhardt, K.M. 1989, "Agency Theory: An Assessment and Review", *The Academy of Management Review*, vol. 14, no. 1, pp. 57-74.
- Elyasiani, E. & Mehdiian, S. 1995, "The comparative efficiency performance of small and large US commercial banks in the pre- and post-deregulation eras", *Applied Economics*, vol. 27, no. 11, pp. 1069-1079.
- Elyasiani, E. & Mehdiian, S. 1990, "Efficiency in the commercial banking industry, a production frontier approach", *Applied Economics*, vol. 22, no. 4, pp. 539.
- Elzahi Saaid, A.E., Rosly, S.A., Ibrahim, M.H. & Abdullah, N. 2003, "The X-Efficiency of the Sudanese Islamic Banks", *International Journal of Economics, Management and Accounting*, vol. 11, no. 2.
- Evanoff, D.D. & Israilevich, P.R. 1995, "Productive Efficiency in Banking", *Federal Reserve Bank of Chicago Economic Perspective*, .

- Evanoff, D.D. & Israilevich, P.R. 1991, "Productive Efficiency in Banking", *Federal Reserve Bank of Chicago Perspective*, , pp. 11-32.
- Fama, E.F. 1980, "Agency Problems and Theory of the Firm", *Journal of Political Economy*, vol. 88, no. 2, pp. 288-307.
- Fama, E.F. & Jensen, M.C. 1983, "Separation of Ownership and Control", *The Journal of Law and Economics*, vol. 26, no. 2, pp. 301-325.
- Farell, M.J. 1957, "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society. Series A (General)*, vol. 120, pp. 253-281.
- Farrell, M.J. 1957, "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society. Series A (General)*, vol. 120, no. 3, pp. 253-290.
- Favero, C.A. & Papi, L. 1995, "Technical efficiency and scale efficiency in the Italian banking sector: a non-parametric approach", *Applied Economics*, vol. 27, no. 4, pp. 385-395.
- Ferreira, C. 2012, "Bank Efficiency, Market Concentration and Economic Growth in European Union", *Working Papers, School of Economics and Management, Technical University of Lisbon*, .
- Ferrier, G.D. & Lovell, C.A.K. 1990, "Measuring cost efficiency in banking : Econometric and linear programming evidence", *Journal of Econometrics*, vol. 46, no. 1-2, pp. 229-245.
- Fethi, M.D. & Pasiouras, F. 2010, "Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey", *European Journal of Operational Research*, vol. 204, no. 2, pp. 189-198.
- Fiordelisi, F. & Mare, D.S. 2013, "Probability of default and efficiency in cooperative banking", *Journal of International Financial Markets, Institutions and Money*, vol. 26, no. 0, pp. 30-45.
- Fiordelisi, F., Marques-Ibanez, D. & Molyneux, P. 2011, "Efficiency and risk in European banking", *Journal of Banking & Finance*, vol. 35, no. 5, pp. 1315-1326.
- Fiorentino, E., Karmann, A. & Koetter, M. 2006, "The Cost Efficiency of German Banks: A comparison of SFA and DEA", *Deutsche Bundesbank, Frankfurt*, .
- Fombrun, C. & Shanley, M. 1990, "What's in a Name? Reputation Building and Corporate Strategy", *The Academy of Management Journal*, vol. 33, no. 2, pp. 233-258.
- Frantz, R. 1992, "X-Efficiency and Allocative Efficiency: What Have We Learned?", *The American Economic Review*, vol. 82, no. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, pp. 434-438.

- Freixas, X. & Rochet, J.C. 2008, *Microeconomics of Banking*, The MIT Press, Cambridge, Massachussets.
- Fries, S. & Taci, A. 2005, "Cost efficiency of banks in transition: Evidence from 289 banks in 15 post-communist countries", *Journal of Banking & Finance*, vol. 29, no. 1, pp. 55-81.
- Fry, M.J. 1995, *Money, Interest and Banking in Economic Development*, 2nd edn, John Hopkins University Press, London.
- Fu, X.(. & Heffernan, S. 2007, "Cost X-efficiency in China's banking sector", *China Economic Review*, vol. 18, no. 1, pp. 35-53.
- Furfine, C.H. 2000, "Interbank payments and the daily federal funds rate", *Journal of Monetary Economics*, vol. 46, no. 2, pp. 535-553.
- Gallant, A.R. 1981, "On the bias in flexible functional forms and an essentially unbiased form : The fourier flexible form", *Journal of Econometrics*, vol. 15, no. 2, pp. 211-245.
- Gasbarro, D., Sadguna, I.G.M. & Zumwalt, J.K. "The Changing Relationship Between CAMEL Ratings and Bank Soundness during the Indonesian Banking Crisis", *Review of Quantitative Finance and Accounting*, vol. 19, no. 3, pp. 247-260.
- Gilbert, R.A. & Wheelock, D.C. 2013, "Big banks in small places: are community banks being driven out of rural markets?", *Federal Reserve Bank of St.Louis Review*, , no. May, pp. 199-218.
- Gilbert, R.A. & Wilson, P.W. 1998, "Effects of Deregulation on the Productivity of Korean Banks", *Journal of economics and business*, vol. 50, no. 2, pp. 133-155.
- Girardone, C., Molyneux, P. & Gardener, E.P. 2004, "Analysing the determinants of bank efficiency: the case of Italian banks", *Applied Economics*, vol. 36, no. 3, pp. 215-227.
- Glaessner, T.C. & Oks, D. 1998, "NAFTA, Capital Mobility and Mexico's Financial System", *World Bank Policy Research Working Paper*, , no. September 1998.
- Goddard, J.A., Molyneux, P. & Wilson, J.O.S. 2001, *European Banking: Efficiency, Technology and Growth*, John Wiley and Sons Ltd, Chichester.
- Goh, S., Alias, M. & Olekalns, N. 2003, "New evidence on financial openness in Malaysia", *Journal of Asian Economics*, vol. 14, no. 2, pp. 311-325.
- Goh, S., Alias, M. & Olekalns, N. 2003, "New evidence on financial openness in Malaysia", *Journal of Asian Economics*, vol. 14, no. 2, pp. 311-325.
- Goldberg, L.S. 2002, "When is US bank lending to emerging markets volatile?" in *Preventing currency crises in emerging markets* University of Chicago Press, , pp. 171-196.

- Goodhart, C.A., Kashyap, A.K., Tsomocos, D.P. & Vardoulakis, A.P. 2013, "An integrated framework for analyzing multiple financial regulations", *International Journal of Central Banking*, vol. 9, no. 1, pp. 109-143.
- Grabowski, R., Rangan, N. & Rezvanian, R. 1993, "Organizational forms in banking: An empirical investigation of cost efficiency", *Journal of Banking & Finance*, vol. 17, no. 2-3, pp. 531-538.
- Greene, W.H. 1991, *The Econometric Approach to Efficiency Measurement*, Stern School of Business, New York University, Mimeo.
- Greene, W.H. 2010, *Econometric Analysis*, 7th edn, Prentice Hall, University of Michigan.
- Greene, W.H. 1990, "A Gamma-distributed stochastic frontier model", *Journal of Econometrics*, vol. 46, no. 1-2, pp. 141-163.
- Greene, W.H. 1990, "A Gamma-distributed stochastic frontier model", *Journal of Econometrics*, vol. 46, no. 1-2, pp. 141-163.
- Grifell-Tatje, E. & Lovell, C.K. 1996, "Deregulation and productivity decline: the case of Spanish savings banks", *European Economic Review*, vol. 40, no. 6, pp. 1281-1303.
- Griffiths, A. & Wall, S. 2004, *Applied Economics*, Tenth edn, Pearson Education, London.
- Grigorian, D.A. & Manole, V. 2002, "Determinants of Commercial Bank Performance in Transition - An Application of Data Envelopment Analysis", *Policy Research Working Paper Series, number 2850, World Bank*, .
- Gujarati, D. 2003, *Basic Econometrics*, McGraw-Hill, New York.
- Hamilton, R. & Howcroft, J.B. 1995, "A practical approach to maximizing customer retention in the credit card industry", *Journal of Marketing Management*, vol. 11, no. 1-3, pp. 151-163.
- Hamilton, R., Qasrawi, W. & Al-Jarrah, I. 2010, "Cost and Profit Efficiency in the Jordanian Banking Sector 1993-2006: A Parametric Approach.", *International Research Journal of Finance and Economics*, vol. November 2010, no. 56.
- Hamilton, J.D. 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, vol. 57, no. 2, pp. 357-384.
- Hao, J., Hunter, W.C. & Yang, W.K. 2001, "Deregulation and efficiency: the case of private Korean banks", *Journal of economics and business*, vol. 53, no. 2-3, pp. 237-254.
- Hardy, D.C. & di Patti, E.B. 2001, "Bank Reform and Bank Efficiency in Pakistan", *Working Paper No. WP/01/138, International Monetary Fund, Washington DC*, .
- Hasan, I. & Hunter, W.C. 1996, "Efficiency of Japanese multinational banks in the United States", *Research in finance*, vol. 14, pp. 157-174.

- Hasan, I. & Lozano-Vivas, A. 1998, "Foreign banks, production technology, and efficiency: Spanish experience", *Working Paper presented at the Georgia Productivity Workshop III, Athens, Georgia*.
- Hasan, I. & Marton, K. 2003, "Development and efficiency of the banking sector in a transitional economy: Hungarian experience", *Journal of Banking & Finance*, vol. 27, no. 12, pp. 2249-2271.
- Hasan, I. & Marton, K. 2003, "Development and efficiency of the banking sector in a transitional economy: Hungarian experience", *Journal of Banking & Finance*, vol. 27, no. 12, pp. 2249-2271.
- Hassan, I., Koetter, M. & Wedow, M. 2007, "The Quality of Banking and Regional Growth", *Discussion Paper Series 2: Banking and Financial Studies, Deutsche Bundesbank Research Centre*, vol. 10.
- Hassan, M.K. 2005, "The Cost, Profit and X-Efficiency of Islamic Banks", *Economic Research Forum, Cairo/Egypt*, .
- Havrylchyk, O. 2006, "Efficiency of the Polish banking industry: Foreign versus domestic banks", *Journal of Banking & Finance*, vol. 30, no. 7, pp. 1975-1996.
- Hawkins, C.J. 1970, "On the Sales Revenue Maximization Hypothesis", *The Journal of Industrial Economics*, vol. 18, no. 2, pp. pp. 129-140.
- Hellmann, T.F., Murdock, K.C. & Stiglitz, J.E. 2000, "Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?", *American economic review*, , pp. 147-165.
- Hellwig, M. 1991, "Banking, Financial Intermediation and Corporate Finance" in *European Financial Integration*, eds. A. Giovannini & C. Mayer, 1991st edn, Cambridge University Press, Cambridge, pp. 35-68.
- Herring, R. 2005, "Implementing Basel II: Is the Game Worth the Candle?", *Financial Markets, Institutions & Instruments*, vol. 14, no. 5, pp. 267-287.
- Hicks, J.R. 1935, "Annual Survey of Economic Theory: The Theory of Monopoly", *Econometrica*, vol. 3, no. 1, pp. pp. 1-20.
- Hon, L.Y., Tuck, C.E. & Yu, K.L. 2011, "Efficiency in the Malaysian Banking Industry", *ASEAN Economic Bulletin*, vol. 28, no. 1, pp. 16-44.
- Huang, T. & Wang, M. 2002, "Comparisons of Economic Inefficiency between Output and Input Measures of Technical Inefficiency Using the Fourier Flexible Cost Function.", *Journal of Productivity Analysis*, vol. 22, no. 1, pp. 123-142.
- Hughes, J.P. & Mester, L.J. 2011, "Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function", *Research Department, Federal Reserve Bank of Philadelphia*, vol. Working Paper No. 11-27, pp. 1-37.

- Hughes, J.P. & Mester, L.J. 2008, "Efficiency in Banking: Theory, Practice and Evidence", *Research Department, Federal Reserve Bank of Philadelphia*, .
- Hughes, J.P. & Mester, L.J. 1993, "A Quality and Risk-Adjusted Cost Function for Banks: Evidence on the Too-Big-To-Fail Doctrine", *Journal of Productivity Analysis*, vol. 4, no. 3, pp. 293-315.
- Humphrey, D.B. 1993, "Cost and Technical Change: Effects from Bank Deregulation", *Journal of Productivity Analysis*, vol. 4, no. 1/2, pp. 9-34.
- Humphrey, D.B. 1990, "Why Do Estimates of Bank Scale Economies Differ?", *Economic Review*, vol. 76, pp. 38-50.
- Humphrey, D.B. & Pulley, L.B. 1997, "Banks' Responses to Deregulation: Profits, Technology, and Efficiency", *Journal of Money, Credit and Banking*, vol. 29, no. 1, pp. pp. 73-93.
- Hymer, S. 1976, *The international operations of national firms: A study of direct foreign investment*, MIT press Cambridge, MA.
- Ibrahim, M. 2011, "Impact of the global crisis on Malaysia's financial system", *Bank of International Settlement Papers*, vol. 54, pp. 267-278.
- Igan, D. & Kang, H. 2011, "Do Loan-to-Value and Debt-to-Income Limits Work? Evidence from Korea", *IMF Working Papers WP/11/297*, , pp. 1-34.
- International Monetary Fund 2007, *Global Financial Stability Report - Market Developments and Issues*, International Monetary Fund, Washington DC.
- Iqbal, M. & Molyneux, P. 2005, *Thirty Years of Islamic Banking: History, Performance and Prospects*, Palgrave Macmillan, London.
- Isik, I. & Hassan, M.K. 2003, "Efficiency, Ownership and Market Structure, Corporate Control and Governance in the Turkish Banking Industry", *Journal of Business Finance & Accounting*, vol. 30, no. 9-10, pp. 1363-1421.
- Isik, I. & Hassan, M.K. 2002, "Cost and Profit Efficiency of the Turkish Banking Industry: An Empirical Investigation", *Financial Review*, vol. 37, no. 2, pp. 257-279.
- Ismail, M. & Rahim, H.A. 2009, "Impact of merger on efficiency and productivity in Malaysian commercial banks", *International Journal of Economics and Finance*, vol. 1, no. 2, pp. P225.
- Jensen, M.C. & Meckling, W.H. 1976, "Theory of the firm: Managerial behavior, agency costs and ownership structure", *Journal of Financial Economics*, vol. 3, no. 4, pp. 305-360.

- Jomo, K.S. & Chin, K.F. 2001, "Financial reform and crisis in Malaysia", *Financial Crisis and System Reform in Asia*, eds. M. Tsurumi, Hosei University Press, Tokyo, pp. 225-250.
- Jondrow, J., Knox Lovell, C.A., Materov, I.S. & Schmidt, P. 1982, "On the estimation of technical inefficiency in the stochastic frontier production function model", *Journal of Econometrics*, vol. 19, no. 2-3, pp. 233-238.
- K. S., J. & Hui, W.C. 2003, "The political economy of Malaysian federalism: economic development, public policy and conflict containment", *Journal of International Development*, vol. 15, no. 4, pp. 441-456.
- Kamaruddin, B.H., Safab, M.S. & Mohd, R. 2008, "Assessing Production Efficiency of Islamic Banks and Conventional Bank Islamic Windows in Malaysia", *International Journal of Business and Management Science*, vol. 1, no. 1, pp. 31.
- Kaparakis, E.I., Miller, S.M. & Noulas, A.G. 1994, "Short-Run Cost Inefficiency of Commercial Banks: A Flexible Stochastic Frontier Approach", *Journal of Money, Credit and Banking*, vol. 26, no. 4, pp. pp. 875-893.
- Kashyap, A.K., Stein, J.C. & Hanson, S. 2010, "An analysis of the impact of 'substantially heightened' capital requirements on large financial institutions", *Journal of Economic Perspectives*, .
- Kashyap, A.K., Rajan, R. & Stein, J.C. 2002, "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking", *The Journal of Finance*, vol. 57, no. 1, pp. 33-73.
- Kasman, A. & Kirbas-Kasman, S. 2006, "Technical change in banking: Evidence from transition countries", *International Journal of the Economics of Business*, vol. 13, no. 1, pp. 129-144.
- Kasman, A. & Yildirim, C. 2006, "Cost and Profit Efficiencies in Transition Banking: the Case on New EU Members", *Applied Economics*, vol. 38, no. 9, pp. 1079-1090.
- Khan, F. 2010, "How 'Islamic' is Islamic Banking?", *Journal of Economic Behavior & Organization*, vol. 76, no. 3, pp. 805-820.
- Khemani, R.S. 1990, "Merger Policy and Small Open Economies" in *Perspectives in Industrial Organization*, eds. B. Dankbaar, J. Groenewegen & H. Schenk, Kluwer Academic Publishers, Dordrecht, pp. 215-234.
- Klein, M.A. 1971, "A Theory of the Banking Firm", *Journal of Money, Credit and Banking*, vol. 3, no. 2, Part 1, pp. 205-218.
- Knox Lovell, C.A. & Pastor, J.T. 1997, "Target setting: An application to a bank branch network", *European Journal of Operational Research*, vol. 98, no. 2, pp. 290-299.

- Koetter, M. 2006, "Measurement Matters-Alternative Input Price Proxies for Bank Efficiency Analyses", *Journal of Financial Services Research*, vol. 30, no. 2, pp. 199-227.
- Kohn, D.L. 2008, "Global Economic Integration and Decoupling", *International Research Forum on Monetary Policy* Frankfurt, Germany, 26 June 2008.
- Kolari, J.W. & Zardkoohi, A. 1987, "Bank Costs, Structure, and Performance", *Lexington, Mass.*, .
- Koopmans, T.C. 1951, "Analysis of Production as an Efficient Combination of Activities", in *Koopmans, T.J., Activity Analysis of Production and Allocation*, Wiley and Son, New York, .
- Koutsoyiannis, A. 1979, *Modern Microeconomics*, 2nd Edition edn, Macmillan, London.
- Kraft, E., Hofler, R. & Payne, J. 2006, "Privatization, foreign bank entry and bank efficiency in Croatia: a Fourier-flexible function stochastic cost frontier analysis", *Applied Economics*, vol. 38, no. 17, pp. 2075-2088.
- Kraft, E. & Tirturoglu, D. 1998, "Bank Efficiency in Croatia: A Stochastic-Frontier Analysis", *Journal of Comparative Economics*, vol. 26, no. 2, pp. 282-300.
- Krishnasamy, G., Ridzwa, A.H. & Perumal, V. 2004, "Malaysian post merger banks' productivity: Application of Malmquist productivity index", *Managerial Finance*, vol. 30, no. 4, pp. 63-74.
- Kroszner, R.S. & Strahan, P.E. 1999, "What Drives Deregulation? Economics and Politics of the Relaxation of Bank Branching Restrictions", *The Quarterly Journal of Economics*, vol. November 1999, pp. 1437-1467.
- Kumbhakar, S.C. & Lovell, K.C.A. 2000, *Stochastic Frontier Efficiency*, Cambridge University Press, Cambridge.
- Kumbhakar, S.C. & Sarkar, S. 2003, "Deregulation, Ownership, and Productivity Growth in the Banking Industry: Evidence from India", *Journal of Money, Credit and Banking*, vol. 35, no. 3, pp. pp. 403-424.
- Kupiec, P.H. 2007, "Financial stability and Basel II", *Annals of Finance*, vol. 3, no. 1, pp. 107-130.
- Kwan, S.H. & Eisenbeis, R.A. 1996, "An Analysis of Inefficiencies in Banking: A Stochastic Cost Frontier", *FRBSF Economic Review*, vol. 2, pp. 16-26.
- Kwan, S.H. 2006, "The X-efficiency of commercial banks in Hong Kong", *Journal of Banking & Finance*, vol. 30, no. 4, pp. 1127-1147.
- Kwon, W.J. 2007, "Islamic Principle and Takaful Insurance: Re-evaluation", *Journal of Insurance Regulation*, vol. 26, no. 1, pp. 53-81.

- Lang, G. & Welzel, P. 1996, "Efficiency and technical progress in banking Empirical results for a panel of German cooperative banks", *Journal of Banking & Finance*, vol. 20, no. 6, pp. 1003-1023.
- Lastra, R.M. 2004, "Risk-based capital requirements and their impact upon the banking industry: Basel II and CAD III", *Journal of Financial Regulation and Compliance*, vol. 12, no. 3, pp. 225-239.
- Lee, J. 2011, "Promoting Financial Stability in Malaysia" in *Addressing Risks in Promoting Financial Stability*, ed. I. Gunadi, The South East Asian Central Banks Research and Training Centre, Kuala Lumpur, pp. 89-108.
- Leibenstein, H. 1979, "A Branch of Economics is Missing: Micro-Micro Theory", *Journal of Economic Literature*, vol. 17, no. 2, pp. 477-502.
- Leibenstein, H. 1975, "Aspects of the X-Efficiency Theory of the Firm", *The Bell Journal of Economics*, vol. 6, no. 2, pp. 580-606.
- Leibenstein, H. 1966, "Allocative Efficiency vs. "X-Efficiency"", *The American Economic Review*, vol. 56, no. 3, pp. 392-415.
- Leightner, J.E. & Lovell, C.A.K. 1998, "The Impact of Financial Liberalization on the Performance of Thai Banks", *Journal of economics and business*, vol. 50, no. 2, pp. 115-131.
- Lensink, R. & Hermes, N. 2004, "The short-term effects of foreign bank entry on domestic bank behaviour: Does economic development matter?", *Journal of Banking & Finance*, vol. 28, no. 3, pp. 553-568.
- Lensink, R., Meesters, A. & Naaborg, I. 2008, "Bank efficiency and foreign ownership: Do good institutions matter?", *Journal of Banking & Finance*, vol. 32, no. 5, pp. 834-844.
- Levine, R. 1996, "Foreign Banks, Financial Development, and Economic Growth" in *International Financial Markets*, ed. E.B. Claude, AES Press, Washington.
- Levine, R. 2002, "Bank-Based or Market-Based Financial Systems: Which Is Better?", *Journal of Financial Intermediation*, vol. 11, no. 4, pp. 398-428.
- Levine, R. 1997, "Financial Development and Economic Growth: Views and Agenda", *Journal of Economic Literature*, vol. 35, no. 2, pp. 688-726.
- Levine, R. 1991, "Stock Markets, Growth, and Tax Policy", *The Journal of Finance*, vol. 46, no. 4, pp. 1445-1465.
- Loke, A.Y.S. 2011, "Foreign Banks, Cross-Border Lending and Implications of Malaysia: Balance Sheet Perspectives" in *International and Cross-Border Bank Lending and Implications on SEACEN Economies: Balance Sheet Perspectives*, eds. R.Y. Siregar & V. Pontines, The South East Asian Central Banks Research and Training Centre, Kuala Lumpur, pp. 111-134.

- Lopez, R.E. 1982, "Applications of Duality Theory in Agriculture", *Western Journal of Agricultural Economics*, vol. 7, no. 2, pp. 353-366.
- Lovell, K.C.A. 1993, "Production Frontiers and Productive Efficiency", *Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (eds.), The measurement of Productive Efficiency: Techniques and Application*, Oxford University Press, New York, , pp. 3-67.
- Lusk, J.L., Featherstone, A.M., Marsh, T.L. & Abdulkadri, A. 2002, "Empirical Properties of Duality Theory", *The Australian Journal of Agricultural and Resource Economics*, vol. 46, no. 1, pp. 45-68.
- Machlup, F. 1967, "Theories of the Firm: Marginalist, Behavioral, Managerial", *The American Economic Review*, vol. 57, no. 1, pp. pp. 1-33.
- Mahajan, A., Rangan, N. & Zardkoohi, A. 1996, "Cost structures in multinational and domestic banking", *Journal of Banking & Finance*, vol. 20, no. 2, pp. 283-306.
- Malmquist, S. 1953, "Index Numbers and Indifference Surfaces", *Trabajos de Estadística*, vol. 4, pp. 209-242.
- Marimuthu, M. & Arokiasamy, L. 2011, "Conventional Banks and Productivity Level: The Malaysian Perspective", *African Journal of Business Management*, vol. 5, no. 6, pp. 2493-2500.
- Mason, E.S. 1939, "Price and Production Policies of Large-Scale Enterprise", *American Economic Review*, vol. 29, pp. 61-74.
- Matthews, K. & Ismail, M. 2006, "Efficiency and Productivity Growth of Domestic and Foreign Commercial Banks in Malaysia", *Cardiff Business School Working Paper Series*, Cardiff University, vol. 2006, no. 2, pp. 1-23.
- Matthews, K. & Thompson, J. 2005, *The economics of banking*, John Wiley & Sons Limited, West Sussex.
- Maudos, J. & Pastor, J.M. 2003, "Cost and profit efficiency in the Spanish banking sector (1985-1996): a non-parametric approach", *Applied Financial Economics*, vol. 13, no. 1, pp. 1-12.
- Maudos, J., Pastor, J.M., Pérez, F. & Quesada, J. 2002, "Cost and profit efficiency in European banks", *Journal of International Financial Markets, Institutions and Money*, vol. 12, no. 1, pp. 33-58.
- Maysami, R.C. & Kwon, W.J. 1999, "An Analysis of Islamic Takaful Insurance", *Journal of Insurance Regulation*, vol. 18, no. 1, pp. 109.
- Mazzeo, M.J. 2002, "Product Choice and Oligopoly Market Structure", *The Rand journal of economics*, vol. 33, no. 2, pp. pp. 221-242.

- McAllister, P.H. & McManus, D. 1993, "Resolving the scale efficiency puzzle in banking", *Journal of Banking & Finance*, vol. 17, no. 2-3, pp. 389-405.
- McCauley, R.N. & Seth, R. 1992, "Foreign bank credit to US corporations: the implications of offshore loans", *Federal Reserve Bank of New York Quarterly Review*, vol. 17, no. 1, pp. 52-65.
- McGuigan, J.R. & Moyer, R.C. 1989, "Managerial Economics", *West Publishing Company, St. Paul, .*
- McKinnon, R.I. 1973, *Money and Capital in Economic Development*, Brookings Institution, Washington DC.
- Meeusen, W. & Broeck, J.v.D. 1977, "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error", *International Economic Review*, vol. 18, no. 2, pp. pp. 435-444.
- Merton, R.C. & Bodie, Z. 2004, "The Design of Financial Systems: Towards a Synthesis of Function and Structure", *National Bureau of Economic Research, NBER Working Paper 10620*, , pp. 1-44.
- Mester, L.J. 1997, "Measuring efficiency at U.S. banks: Accounting for heterogeneity is important", *European Journal of Operational Research*, vol. 98, no. 2, pp. 230-242.
- Mester, L.J. 1996, "A study of bank efficiency taking into account risk-preferences", *Journal of Banking & Finance*, vol. 20, no. 6, pp. 1025-1045.
- Mester, L.J. 1993, "Efficiency in the savings and loan industry", *Journal of Banking & Finance*, vol. 17, no. 2-3, pp. 267-286.
- Mian, A. 2006, "Distance Constraints: The Limits of Foreign Lending in Poor Economies", *The Journal of Finance*, vol. 61, no. 3, pp. 1465-1505.
- Millon, M.H. & Thakor, A.V. 1985, "Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies", *The Journal of Finance*, vol. 40, no. 5, pp. 1403-1422.
- Mique, J. & Belanger, G. 1974, "Toward a General Theory of Managerial Discretion", *Public Choice*, vol. 17, pp. 27-43.
- Mitchell, K. & Onvural, N.M. 1996, "Economies of Scale and Scope at Large Commercial Banks: Evidence from the Fourier Flexible Functional Form", *Journal of Money, Credit and Banking*, vol. 28, no. 2, pp. pp. 178-199.
- Mobarek, A. & Kalonov, A. 2014, "Comparative performance analysis between conventional and Islamic banks: empirical evidence from OIC countries", *Applied Economics*, vol. 46, no. 3, pp. 253-270.

- Mohd Tahir, I., Abu Bakar, N.M. & Haron, S. 2010, "Cost and Profit Efficiency of the Malaysian Commercial Banks: A Comparison between Domestic and Foreign Banks ", *International Journal of Economics and Finance*, vol. 2, no. 1, pp. 186-197.
- Mohd Tahir, I., Abu Bakar, N.M. & Haron, S. 2009, "Evaluating Efficiency of Malaysian Banks Using Data Envelopment Analysis", *International Journal of Business and Management*, vol. 4, no. 8, pp. 96-106.
- Mohd Tahir, I., Abu Bakar, N.M. & Haron, S. 2008, "Technical Efficiency of Malaysian Commercial Banks: A Stochastic Frontier Approach", *Banks and Bank System*, vol. 3, no. 4, pp. 65-72.
- Mohd Zaini, A.K., Chan, S. & Hassan, S. 2010, "Bank Efficiency And Non-Performing Loans: Evidence From Malaysia And Singapore", *Prague Economic Papers*, vol. 2010, no. 2, pp. 118-132.
- Molyneux, P., Altunbas, Y. & Gardener, E.P.M. 1996, *Efficiency in European Banking*, John Wiley, Chichester.
- Mueller, D.C. 1983, "The Determinants of Persistent Profits", *Washington: Bureau of Economics, U.S. Federal Trade Commission*, vol. June.
- Nadiri, M.I. 1970, "Some approaches to the theory and measurement of total factor productivity: a survey", *Journal of Economic Literature*, vol. 8, no. 4, pp. 1137-1177.
- Napier, N.K. 1989, "Mergers and Acquisitions, Human Resource Issues and Outcomes: A Review and Suggested Typology*", *Journal of Management Studies*, vol. 26, no. 3, pp. 271-290.
- Njie, M. 2007, "Impact of financial liberalization on the productive efficiency of the banking industry in Malaysia", *Banker's Journal Malaysia*, vol. 131, pp. 49-67.
- Nohria, N. & Ranjay, G. 1996, "Is Slack Good or Bad for Innovation", *Academy of Management Journal*, vol. 39, no. 5, pp. 1245-1264.
- Noulas, A.G. 1997, "Productivity growth in the Hellenic banking industry: state versus private banks", *Applied Financial Economics*, vol. 7, no. 3, pp. 223-228.
- Omar, M.A., Abdul Rahman, A.R., Mohd Yusof, R., Abd. Majid, M.S. & Mohd Rasid, Mohd Eskandar Shah 2006, "Efficiency of Commercial Banks in Malaysia", *Asian Academy of Management Journal of Accounting and Finance*, vol. 2, no. 2, pp. 19-42.
- Panzar, J.C. & Rosse, J.N. 1987, "Testing for" monopoly" equilibrium", *The Journal of Industrial Economics*, , pp. 443-456.
- Panzar, J.C. & Willig, R.D. 1981, "Economies of Scope", *The American Economic Review*, vol. 71, no. 2, Papers and Proceedings of the Ninety-Third Annual Meeting of the American Economic Association, pp. pp. 268-272.

- Pastor, J.M. 2010, "Credit risk and efficiency in the European banking system: A three-stage analysis", *Applied Financial Economics*, vol. 12, no. 12, pp. 895-911.
- Petersen, M.A. & Rajan, R.G. 1995, "The effect of credit market competition on lending relationships", *The Quarterly Journal of Economics*, vol. 110, no. 2, pp. 407-443.
- Pettway, R.H. 1980, "Potential Insolvency, Market Efficiency, and Bank Regulation of Large Commercial Banks", *The Journal of Financial and Quantitative Analysis*, vol. 15, no. 1, pp. 219-236.
- Piloff, S.J. & Santomero, A.M. 1996, "The Value Effect of Mergers and Acquisitions" in *Bank Mergers and Acquisitions*, eds. Y. Amihud & G. Miller, 1st edn, Kluwer Academic Publisher, Dordrecht, Netherlands, pp. 59.
- Pitt, M.M. & Lee, L. 1981, "The measurement and sources of technical inefficiency in the Indonesian weaving industry", *Journal of Development Economics*, vol. 9, no. 1, pp. 43-64.
- Puig-Junoy, J. & Ortun, V. 2002, "Cost Efficiency in Primary Care Contracting", *A Stochastic Frontier Cost Function Approach. Barcelona (España): UPF Working Paper*, .
- Qasrawi, W. 2009, *The Impacts of Deregulation on the Jordanian Banking Sector 1993-2006: An Empirical Analysis Using Frontier Approaches*, Loughborough University.
- Qi, M. & Yang, X. 2009, "Loss given default of high loan-to-value residential mortgages", *Journal of Banking & Finance*, vol. 33, no. 5, pp. 788-799.
- Radam, A., Baharom, A.H., Dayang-Afizzah, A.M. & Ismail, F. 2009, "Effect of Mergers on Efficiency and Productivity: Some Evidence for banks in Malaysia", *MPRA Paper 12726, University Library of Munich, Germany*, , pp. 1-19.
- Rajan, R.G. & Zingales, L. 2001, "The Influence of the Financial Revolution on the Nature of Firms", *American Economic Review*, vol. 91, no. 2, pp. 206-211.
- Rajoo, U. 2008, "Problem Bank Identification, Intervention and Resolution in Malaysia" in *Comparison of Problem Bank Identification, Intervention and Resolution in the SEACEN Countries*, ed. S. Batunanggar, The South East Asian Central Banks Research and Training Centre, Kuala Lumpur, pp. 109-125.
- Ramanathan, R. 2003, *An Introduction to Data Envelopment Analysis: A Tool for Performance Management*, Sage Publication, Thousand Oaks, California.
- Rating Agency Malaysia 2010, *Credit Rating Rationale: Financial Institutional Ratings - Bank Islam Malaysia Berhad 2010*, Rating Agency Services Berhad, Kuala Lumpur.
- Ravenscraft, D. 1980, *Price-Raising and Cost Reducing Effects in Profit-Concentration Studies: A Monte Carlo Simulation Analysis*, Northwestern University.

- Ray, S.C. & Das, A. 2010, "Distribution of cost and profit efficiency: Evidence from Indian banking", *European Journal of Operational Research*, vol. 201, no. 1, pp. 297-307.
- Resende, M. 2006, "Efficiency Measurement and Regulation in U.S. Telecommunications: A Robustness Analysis", *EUI Working Papers*, vol. Eco No. 2006/15, pp. 1-40.
- Resti, A. 1997, "Evaluating the cost-efficiency of the Italian banking system: What can be learned from the joint application of parametric and non-parametric techniques", *Journal of Banking & Finance*, vol. 21, no. 2, pp. 221-250.
- Rezvanian, R. & Mehdian, S. 2002, "An examination of cost structure and production performance of commercial banks in Singapore", *Journal of Banking & Finance*, vol. 26, no. 1, pp. 79-98.
- Richmond, J. 1974, "Estimating the Efficiency of Production", *International Economic Review*, vol. 15, no. 2, pp. 515-521.
- Robins, J.A. 1987, "Organizational Economics: Notes on the Use of Transaction-Cost Theory in the Study of Organizations", *Administrative Science Quarterly*, vol. 32, no. 1, pp. 68-86.
- Rogers, K. 1998, "Product Mix, Bank Powers, and Complementarities at U. S. Commercial Banks", *Journal of economics and business*, vol. 50, no. 2, pp. 205-218.
- Roubini, N. & Sala-i-Martin, X. 1992, "Financial repression and economic growth", *Journal of Development Economics*, vol. 39, no. 1, pp. 5-30.
- Said, R.M., Nor, F.M., Low, S. & Rahman, A.A. 2008, "The Efficiency Effects of Mergers and Acquisitions in Malaysian Banking Institutions", *Asian Journal of Business and Accounting*, vol. 1, no. 1, pp. 47-66.
- Samad, A. & Hassan, M.K. 2000, "The performance of Malaysian Islamic bank during 1984-1997: an exploratory study", *Thoughts on Economics*, vol. 10, no. 1-2, pp. 7-26.
- Samuelson, P.A. 1966, "The Fundamental Singularity Theorem for Non-Joint Production", *International Economic Review*, vol. 7, no. 1, pp. 34-41.
- Santos, J.A.C. 2001, "Bank Capital Regulation in Contemporary Banking Theory: A Review of the Literature", *Financial Markets, Institutions & Instruments*, vol. 10, no. 2, pp. 41-84.
- Santos, J.A.C. 1999, "Bank capital and equity investment regulations", *Journal of Banking & Finance*, vol. 23, no. 7, pp. 1095-1120.
- Sathye, M. 2001, "X-efficiency in Australian banking: An empirical investigation", *Journal of Banking & Finance*, vol. 25, no. 3, pp. 613-630.

- Saurina, J. & Trucharte, C. 2004, "The impact of Basel II on lending to small-and medium-sized firms: A regulatory policy assessment based on Spanish credit register data", *Journal of Financial Services Research*, vol. 26, no. 2, pp. 121-144.
- Schmidt, P. & Sickles, R.C. 1984, "Production Frontier and Panel Data", *Journal of Business & Economic Statistics*, vol. 2, pp. 367-374.
- Schooner, H.M. & Taylor, M.W. 2010, *Global Bank Regulation: Principles and Policies*, Academic Press, Burlington.
- Schumpeter, J.A. 1912 [1934], *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*, Harvard University Press, Cambridge.
- Sealey, C.W., Jr. & Lindley, J.T. 1977, "Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions", *The Journal of Finance*, vol. 32, no. 4, pp. 1251-1266.
- Sensarma, R. 2006, "Are foreign banks always the best? Comparison of state-owned, private and foreign banks in India", *Economic Modelling*, vol. 23, no. 4, pp. 717-735.
- Shaw, E. 1973, *Financial Deepening in Economic Development*, Oxford University Press, New York.
- Shephard, R.W. 1953, *Cost and Production Functions*, Princeton University Press, New Jersey.
- Shepherd, W.G. 1985, "The Economics of Industrial Organization", *Englewood Cliffs, N.J.: Prentice-Hall*, .
- Shepherd, W.G. 1982, "Economies of scale and monopoly profits" in *Industrial Organisation, Antitrust, and Public Policy*, eds. J.V. Craven & K. Nijhoff, Springer, Boston, pp. 41-68.
- Sherlund, S.M., Barrett, C.B. & Adesina, A.A. 2002, "Smallholder technical efficiency controlling for environmental production conditions", *Journal of Development Economics*, vol. 69, no. 1, pp. 85-101.
- Sherman, H.D. & Zhu, J. 2006, "Benchmarking with quality-adjusted DEA (Q-DEA) to seek lower-cost high-quality service: Evidence from a US bank application", *Annals of Operations Research*, vol. 145, no. 1, pp. 301-319.
- Siems, T.F. & Barr, R.S. 1998, "Benchmarking the productive efficiency of US banks", *Financial Industry Studies Federal Reserve Bank of Dallas*, , pp. 11-24.
- Simon, H.A. 1955, "A Behavioral Model of Rational Choice", *The Quarterly Journal of Economics*, vol. 69, no. 1, pp. 99-118.

- Simon, H.A. 1959, "Theories of Decision-Making in Economics and Behavioral Science", *The American Economic Review*, vol. 49, no. 3, pp. 253-283.
- Solé, J. 2007, "Introducing Islamic Banks into Conventional Banking Systems", *IMF Working Papers*, , pp. 1-26.
- Spong, K., Sullivan, R.J. & DeYoung, R. 1995, "What makes a Bank Efficient? – A Look at Financial Characteristics and Bank Management and Ownership Structure", *Federal Reserve Bank of Kansas City Economic Review*, , pp. 1-9.
- Sraffa, P. 1926, "The Laws of Returns under Competitive Conditions", *The Economic Journal*, vol. 36, no. 144, pp. 535-550.
- Srairi, S.A. 2010, "Cost and profit efficiency of conventional and Islamic banks in GCC countries", *Journal of Productivity Analysis*, vol. 34, no. 1, pp. 45-62.
- Stevenson, R. 1980, "Likelihood Functions for Generalized Stochastic Frontier Estimation", *Journal of Econometrics*, vol. 13, pp. 58-66.
- Stigler, G.J. 1971, "The Theory of Economic Regulation", *The Bell Journal of Economics and Management Science*, vol. 2, no. 1, pp. 3-21.
- Stoughton, N.M. 1993, "Moral Hazard and the Portfolio Management Problem", *The Journal of Finance*, vol. 48, no. 5, pp. 2009-2028.
- Sturm, J. & Williams, B. 2004, "Foreign bank entry, deregulation and bank efficiency: Lessons from the Australian experience", *Journal of Banking & Finance*, vol. 28, no. 7, pp. 1775-1799.
- Sufian, F. 2011, "Banks Total Productivity Change in a Developing Economy: Does Ownership and Origins Matter?", *Journal of Asian Economics*, vol. 22, pp. 84-98.
- Sufian, F. 2010, "Productivity, technology and efficiency of De Novo Islamic banks: Empirical evidence from Malaysia", *Journal Financial Services Marketing*, , no. 3, pp. 258.
- Sufian, F. 2009, "Determinants of Bank Efficiency during Unstable Macroeconomic Environment : Empirical Evidence from Malaysia", *Research in International Business and Finance*, vol. 23, pp. 54-77.
- Sufian, F. 2006, "Size and returns to scale of the Islamic banking industry in Malaysia: foreign versus domestic banks", *International Journal of Economics, Management and Accounting*, vol. 14, no. 2.
- Sufian, F. 2004, "The Efficiency Effect of Bank Mergers and Acquisition in Developing Economy: Evidence from Malaysia", *International Journal of Applied Econometrics and Quantitative Studies*, vol. 1, no. 4, pp. 53-74.

- Sufian, F. & Habibullah, M.S. 2010, "Does Foreign Banks Entry Fosters Bank Efficiency? Empirical Evidence from Malaysia", *Engineering Economics*, vol. 21, no. 5, pp. 464-474.
- Sufian, F. & Habibullah, M.S. 2009, "Do mergers and acquisitions leads to a higher technical and scale efficiency? A counter evidence from Malaysia", *African Journal of Business Management*, vol. 3, no. 8, pp. 340-349.
- Sufian, F. & Majid, M.A. 2007, "Banks' efficiency and stock prices in emerging markets: evidence from Malaysia", *Journal of Asia-Pacific Business*, vol. 7, no. 4, pp. 35-53.
- Sufian, F. 2010, "The impact of the Asian financial crisis on bank efficiency: The 1997 experience of Malaysia and Thailand", *Journal of International Development*, vol. 22, no. 7, pp. 866-889.
- Suhaimi, R., Abdullah, F. & Saban, G. 2010, "Factors affecting profit efficiency of commercial banks in Malaysia", *Science and Social Research (CSSR), 2010 International Conference onIEEE*, , pp. 1059.
- Sulaiman, M. & Govindan, K. 2000, "Macroeconomic fundamentals of the Malaysian economy: before and after the crisis", *National Seminar on Strengthening the Macroeconomic Fundamentals of the Malaysian Economy, organized by UUM and UKM, June*, pp. 5.
- Tan, A.C.K. 2011, "Propagation of Real Shocks in the SEACEN Economies: Malaysia" in *Real Sector Propagation of Recent Global Financial Crisis*, eds. R.Y. Siregar & V.C.S. Lim, The South East Asian Central Banks Research and Training Centre, Kuala Lumpur, pp. 109-148.
- Terrell, H.S. 1993, "US branches and agencies of foreign banks: a new look", *Fed.Res.Bull.*, vol. 79, pp. 913.
- Thomas, L.C. 2000, "A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers", *International Journal of Forecasting*, vol. 16, no. 2, pp. 149-172.
- Tone, K. 2002, "A Strange Case of the Cost and Allocative Efficiencies in DEA", *The Journal of the Operational Research Society*, vol. 53, no. 11, pp. pp. 1225-1231.
- Tortosa-Ausina, E. 2002, "Bank Cost Efficiency and Output Specification", *Journal of Productivity Analysis*, vol. 18, no. 3, pp. 199-222.
- Tortosa-Ausina, E. 2003, "Nontraditional activities and bank efficiency revisited: a distributional analysis for Spanish financial institutions", *Journal of economics and business*, vol. 55, no. 4, pp. 371-395.
- Tsionas, E.G., Lolos, S.E.G. & Christopoulos, D.K. 2003, "The performance of the Greek banking system in view of the EMU: results from a non-parametric approach", *Economic Modelling*, vol. 20, no. 3, pp. 571-592.

- Tulkens, H. 1993, "On FDH Analysis: some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit", *Journal of Productivity Analysis*, vol. 4, no. 1, pp. 183-210.
- Unite, A.A. & Sullivan, M.J. 2003, "The effect of foreign entry and ownership structure on the Philippine domestic banking market", *Journal of Banking & Finance*, vol. 27, no. 12, pp. 2323-2345.
- Vennet, R.V. 2002, "Cost and Profit Efficiency of Financial Conglomerates and Universal Banks in Europe", *Journal of Money, Credit and Banking*, vol. 34, no. 1, pp. pp. 254-282.
- Vennet, R.V. 1998, "Cost and Profit Dynamics in Financial Conglomerates and Universal Banks in Europe", *SSRN eLibrary*, [Online], Available from: <http://ssrn.com/abstract=160493>.
- Vivas, A.L. 1997, "Profit efficiency for Spanish savings banks", *European Journal of Operational Research*, vol. 98, no. 2, pp. 381-394.
- Waldman, D.E. & Jensen, E.J. 2001, *Industrial Organization: Theory and Practice*, Second Edition edn, Addison-Wesley Longman, Boston.
- Weill, L. 2004, "Measuring Cost Efficiency in European Banking: A Comparison of Frontier Techniques", *Journal of Productivity Analysis*, vol. 21, no. 2, pp. 133-152.
- Weiner, S.E. 1992, "The changing role of reserve requirements in monetary policy", *Economic Review-Federal Reserve Bank of Kansas City*, vol. 77, pp. 45-45.
- Weitzman, M.L. 1973, "Duality Theory for Infinite Horizon Convex Models", *Management Science*, vol. 19, no. 7, Theory Series, pp. pp. 783-789.
- Weller, C.E. 2001, "Financial crises after financial liberalisation: Exceptional circumstances or structural weakness?", *Journal of Development Studies*, vol. 38, no. 1, pp. 98-127.
- Wheelock, D.C. & Wilson, P.W. 2001, "New Evidence to Returns to Scale and Product Mix among U.S. Commercial Banks", *Journal of Monetary Economics*, vol. 47, pp. 653-674.
- Wheelock, D.C. & Wilson, P.W. 1995, "Explaining Bank Failures: Deposit Insurance, Regulation, and Efficiency", *The review of economics and statistics*, vol. 77, no. 4, pp. pp. 689-700.
- Williams, J. & Nguyen, N. 2005, "Financial liberalisation, crisis, and restructuring: A comparative study of bank performance and bank governance in South East Asia", *Journal of Banking & Finance*, vol. 29, no. 8, pp. 2119-2154.
- Williams, B. 2003, "Domestic and international determinants of bank profits: Foreign banks in Australia", *Journal of Banking & Finance*, vol. 27, no. 6, pp. 1185-1210.

- Williamson, O.E. 1981, "The Economics of Organization: The Transaction Cost Approach", *American Journal of Sociology*, vol. 87, no. 3, pp. 548-577.
- Williamson, O.E. 1963, "Managerial Discretion and Business Behavior", *The American Economic Review*, vol. 53, no. 5, pp. pp. 1032-1057.
- Winsten, C.B. 1957, "Discussion on Mr. Farrell's Paper", *Journal of Royal Statistical Society, Series A (General)*, vol. 120, no. 3, pp. 282-284.
- Wong, E., Fong, T., Li, K. & Choi, H. 2011, "Loan-to-value Ratio as a Macro-prudential Tool: Hong Kong's Experience and Cross-country Evidence", *Hong Kong Monetary Authority Working Paper 01/2011*, .
- World Bank 2011, *World Development Indicators 2011*, World Bank, Washington DC.
- Yarrow, G.K. 1976, "On the Predictions of Managerial Theories of the Firm", *The Journal of Industrial Economics*, vol. 24, no. 4, pp. pp. 267-279.
- Yudistira, D. 2004, "Efficiency in Islamic Banking: An Empirical Analysis of Eighteen Banks", *Islamic Economics Studies*, vol. 12, no. 1, pp. 1-18.
- Zaim, O. 1995, "The effect of financial liberalization on the efficiency of Turkish commercial banks", *Applied Financial Economics*, vol. 5, no. 4, pp. 257-264.
- Zamani, A.G. 2006, "Re-engineering the Malaysian financial system to promote sustainable growth", *BIS Papers*, vol. 28, pp. 269-275.
- Zhu, J. 2002, "Imprecise data envelopment analysis IDEA: A review and improvement with an application", *European Journal of Operational Research*, vol. 144, no. 3, pp. 513-529.