Performance and Risk in the Chinese Regional Banking Sector: A Meta-Frontier Approach

YUZHU LI

Thesis submitted to the University of Nottingham

for the degree of Doctor of philosophy

Jan 2021

Abstract

This thesis examines the performance and risk of Chinese commercial banks, with a focus on the Chinese regional commercial banking sector from 2005 to 2015. To that end, cost efficiency, scale economies and loan loss provisioning practices are investigated. We begin our analysis by classifying Chinese provinces into three regions using a novel hybrid hierarchical and k-means clustering method based on provincial macroeconomic conditions. We then estimate pooled and individual cost frontiers, and we test whether Chinese regional commercial banks operate under different technologies. Our likelihood test results reject the common technology assumption in favour of technology heterogeneity. Building on the latter, we adopt a stochastic meta-frontier method to estimate a country meta-frontier that envelops regional frontiers. We find that cost efficiency for regional banks mainly stems from technological gaps between regional frontiers and the country meta-frontier. These gaps can be narrowed through increased foreign bank presence, increased banking sector competition and acceleration in regional urbanisation. Banks in the North-eastern and Coastal regions of China are found to be the most cost-efficient, while banks in the Inland region are found to be the least cost-efficient. In addition, we expand our analysis, and we obtain economies of scale from both the individual frontiers and the meta-frontier. Our comparison suggests that banks in the Inland area would operate at the optimal size if they adopt the meta-frontier technology, while regional banks in the core financial centres would start to enjoy economies of scale when they adopt the meta-frontier technology. We also find that investment banking activities, higher liquidity ratio, and lower credit risk would contribute to economies of scale, while banks' profitability from traditional lending activities does not seem to impact scale economies.

After examining the performance of Chinese regional commercial banks, we focus on loan loss provision practices of the Chinese banking system using a

series of system Generalized Methods of Moments dynamic panel data regressions developed by Arellano and Bover (1995) and Blundell and Bond (1998). Overall, we find that Chinese banks are generally backwards-looking in terms of provisioning, and there is strong evidence of income-smoothing. Furthermore, we find that Chinese banks engage more intensively in income-smoothing during the 2007-2008 global financial crisis, in line with the arguments of Curcio et al. (2016) and El Wood (2012) that stress that adverse financial conditions might incentivise bank managers to lower perceived risk of stakeholders and signal bank resilience by reducing earnings volatility. We also find that loan loss provisions mainly signal an expense rather than financial strength to investors. The higher level of loan loss provisions is also linked to lower cost efficiencies derived from the meta-frontier, supporting Berger and DeYoung's (1997) "bad management" hypothesis. Finally, following Beatty and Liao (2014), we first estimate the discretional component of loan loss provisions by obtaining residuals from regressions of loan loss provisions on their non-discretionary elements and following Jin et al. (2018), we estimate an equation of determination of discretionary loan loss provisions. We find consistent results that reliance on retail (core) deposits has a positive impact on the quality of reported earnings as banks with more reliance on core deposits for funding tend to engage less in discretional use of loan loss provisions.

This thesis contributes to the literature methodologically and empirically and highlights several important policy implications. Firstly, banks in different regions in China are found to operate under different technologies, and higher cost efficiencies can be achieved by narrowing the technological gaps between the regional frontiers and the country's meta-frontier. This thesis suggests that this can be achieved by encouraging more competition, more foreign bank presence in the banking sector and advancement of urbanisation to narrow the technological gaps while encouraging better bank management via better capital adequacy and lower levels of non-performing loans. With respect to size effects, our findings stress the importance of optimal technology adoption for banks operating in economic centres. In terms of bank loan loss provisioning in China, a dynamic, forward-looking counter-cyclical loan loss provisioning method may

be beneficial to avoid its cyclical nature that amplifies credit shortages during economic crises. Furthermore, our findings that banks engage more in income-smoothing during a crisis period should encourage regulators to monitor more closely bank accounting information during times of economic and financial stress. As banks' earnings quality is positively driven by retail deposits, regulators should also pay more attention to regulating banks that rely heavily on wholesale funding in order to enhance market discipline.

Acknowledgements

I would like to express my sincere gratitude towards my supervisors Professor Richard Simper and Dr Aristeidis Dadoukis, for their kindness, patience, generous support, and expert guidance. I would also like to express my gratitude to my prior supervisor Professor Robert Webb for his kind advice and insight. My appreciation also goes to several other faculty members and staff in the Nottingham University Business School for their support academically and otherwise.

Contents

Table of Contents

Chapter 1. Introduction *13*

1.1 Research questions 14

1.2 Research motivation and contribution 14

1.3 Overview of Chapters 16

Chapter 2. A Review of the Chinese Banking System *17*

2.1 Introduction 18

2.2 Chinese Banking Reform 18

2.3 Structure of the Chinese Banking System 22

2.4 An overview of the Chinese Banking Sector from 2005 to 2015 26

2.5 Conclusion 35

Chapter 3. A Literature Review on Bank Performance Measurements - A Justification for Using Cost Efficiency *37*

3.1 Chapter Summary 38

3.2 A Literature Review on the Theories and Examples of Financial Ratio Analysis 38

3.3 Statistical Problems of Financial Ratios 43

3.4 Modelling Bank Efficiency 49

3.4.1 Different Approaches to Efficiency Modelling 50

3.4.1.a Technical Efficiency 51

3.4.1.b Cost Efficiency 53

3.5 Modelling Efficiency 55

3.5.1 Deterministic and Distribution Free Models 55

3.5.1.a Corrected Ordinary Least Squares (COLS) Regression Estimation 55

3.5.1.b TFA (Thick Frontier Approach) 58

3.5.1.c Schmidt and Sickles (1984) Distribution Free Approach 58

3.6 SFA (Stochastic Frontier Approach) 60

3.6.1 Battese and Coelli (1995) Model 61

3.6.2 Greene (2005)'s True Fixed Effects Model (TFE) 64

3.6.3 A summary of efficiency models 66

3.7 Literature Review on the Theories of the Banking Firm 66

3.7.1 Production Approach (Banks as Producers of Loans) 67

3.7.2 Intermediation Approach (Banks as Intermediaries) 68

3.7.3 Value-added Approach (Banks as Value-Adders) 70

3.7.4 Profit Approach (Banks as Profit Maximisers) 71

3.5 Conclusion 74

Chapter 4. Environmental Impact on Cost Efficiency of Chinese Regional Commercial Banks *75*

4.1 Chapter Summary 76

4.2 Introduction 76

4.3 Motivation 78

4.4 A Review on the Regional Studies of Commercial Banking Systems 79

4.5 Methodology 83

4.5.1 Cluster Analysis - Hybrid Hierarchical K-means Clustering 83

4.5.2 Stochastic Meta-frontier (SMF) Approach 95

4.6 The implementation of Methodology 103

4.6.1 The Implementation of Cluster Analysis 103

4.6.2 The Implementation of Stochastic Meta-frontier (SMF) 104

4.7 Data Description 115

4.8 Empirical Results - Cluster Analysis 117

4.8.1 Descriptive Statistics 117

4.8.2 Results of Cluster Analysis 119

4.9 Empirical Results – Stochastic Meta-frontier Analysis 131

4.9.1 Descriptive Statistics 131

4.9.2 Pooled results 143

4.9.3 Cluster-Specific Frontier Results 148

4.9.4 Meta-frontier Results 150

4.9.5 Cost Efficiencies and Cost Gap Ratios 153

4.9.6 Trends of CE_pooled, CE_meta, CGR and CE_I 160

4.9.7 Rankings 162

4.10 Determinants of Economies of Scale 164

4.10.1 Literature review and hypotheses development 164

4.10.2 Methodology 166

4.10.3 Result Analysis 169

4.11 Conclusion and Policy Implications 175

Chapter 5. Loan Loss Provisioning Behaviour of Chinese Regional Banks *180*

5.1 Introduction 181

5.2 Literature Review and Hypotheses Development 182

5.2.1 Procyclicality/ Countercyclicality 183

5.2.2 Income Smoothing 185

5.2.3 Capital Management 188

5.2.4 Signalling 189
5.2.5 The impact of funding structure on earnings quality 191
5.2.6 The Relationship between Loan Quality and Bank Efficiency 193

5.3 Methodology 194
5.3.1 Baseline Model specification: 194
5.3.2 Testing of Signalling 199
5.3.3 Estimating discretionary LLPs 200
5.3.4 Alternative test of Signalling 202
5.3.5 Testing the Impact of Retail Deposit on Earnings Quality 203
5.3.6 Variable descriptions 204

5.4 Empirical Results 207
5.4.1 Data Description 207
5.4.2 Descriptive Statistics 208
5.4.3 A test of the Loan loss provisioning hypotheses 215
5.4.4 Test on Signalling theory 223
5.4.5 Estimation of Discretionary LLPs 225
5.4.6 Alternative tests on signalling theory 227
5.4.7 Test of Impact of Retail Deposits on Earnings Quality 229

5.5 Robustness Check 233

5.6 Conclusion and Policy Implications 236

Chapter 6. Conclusion ***237***

6.1 Thesis Summary 238

6.2 Answers to the research questions 239

6.3 Policy and Practical Implications 241

References ***246***

Appendix 1 ..... ***259***

Appendix 2 ..... ***263***

List of Abbreviations

ABC	Agricultural Bank of China
AMC	Asset Management Company
BOC	Bank of China
BC	Battesse and Coelli
CAR	Capital Adequacy Ratio
CBRC	China Banking Regulatory Comission
CNCB	China Citic Bank
CCB	China Construction Bank
CEB	China Everbright Bank
CMB	China Merchant Bank
CRS	Constant Returns to Scale
CPI	Consumer Price Index
COLS	Corrected Ordinary Least Squares
CE	Cost Efficiency
CGR	Cost Gap Ratio
DRS	Decreasing Returns to Scale
EBTP	Earnings Before Tax and Provisions
ES	Economies of Scale
ETA	Equity to Total Asset Ratio
EU	European Union
FEs	Fixed Effects
FBS	Foreign Bank Market Share
FX	Foreign Exchange
GMM	Generalized Method of Moments
CFC	Global Financial Crisis
GDP	Gross Domestic Product
GRP	Gross Regional Product
GDB	Guangdong Development Bank
HHI	Herfindahl-Hirschman Index
HAC	Hierarchical Agglomerative Clustering
HDC	Hierarchical Divisive Clustering
HXB	Hua Xia Bank
IRS	Increasing Returns to Scale
ICBC	Industrial and Commercial Bank of China
IPO	Initial Public Offering
IFRS	International Financial Reporting Standards
JSCB	Joint-Stock Commercial Bank
LSDV	Least Square Dummy Variable Method
LLP	Loan Loss Provision

LBRA	Logarithm of Branch Numbers
LTA	Logarithm of Total Assets
LTC	Logarithm of Total Cost
ML	Maximum Likelihood
MCE	Meta-Frontier Cost Efficiency
NBSC	National Bureau of Statistics of China
NIM	Net Interest Margin
NPL	Non-Performing Loan
OLS	Ordinary Least Squares
OECD	Organisation for Economic Co-operation and Development
PBC	People's Bank of China
PAB	Pin An Bank
PML	Pseudo-Maximum Likelihood
ROAA	Return of Average Assets
ROA	Return On Asset
ROE	Return On Equity
RTS	Return to Scale
SPDB	Shanghai Pudong Development Bank
SOCB	State-Owned Commercial Bank
SOE	State-Owned Enterprise
SFA	Stochastic Frontier Analysis
SMF	Stochastic Meta-Frontier
TGR	Technology Gap Ratio
TFA	Thick Frontier Approach
TSS	Total Sum of Squares

List of Figures

Figure 2.1. Structure of the Chinese Banking System 23
Figure 2.2. Total Assets and Liabilities of Banking Institutions (2005-2015) 27
Figure 2.3. Loan and Deposit share of State-owned Commercial Banks (2005-2015) 28
Figure 2.4. Asset Share breakdown of China's Banking sector 28
Figure 2.5. Asset Share breakdown of China's Commercial banking sector .. 29
Figure 2.6. Return on Average Assets of Chinese Banks 31
Figure 2.7. Return on Average Assets by Commercial Bank Types 31
Figure 2.8. Return on Average Equity by Commercial Bank Types 32
Figure 2.9. NPLs of Chinese Domestic Commercial Banks 33
Figure 2.10. Capital Adequacy of Chinese Commercial Banks 35
Figure 3.1. A disproportional relationship 45
Figure 3.2. A nonlinear relationship 45
Figure 3.3. Output oriented efficiency 51
Figure 3.4. Input oriented efficiency 52
Figure 3.5. Production Approach 68
Figure 3.6. The Intermediation Approach 69
Figure 3.7. Value-Added Approach source: Humphrey (1992) 71
Figure 4.1. Regional divisions of Mainland China by NBSC (excluding Hong Kong, Macao and Taiwan) 88
Figure 4.2. Meta-frontier Cost model 101
Figure 4.3. Dendrogram for HAC 120
Figure 4.4. Regional Division in Mainland China (excluding Hong Kong, Macao and Taiwan) – Results from Cluster Analysis 128
Figure 4.5. Determination of Optimal No. of Clusters-Elbow Method 129
Figure 4.6. Cross-validation of Optimal No. of Clusters- Silhouette Widths 130
Figure 4.7. Microeconomic and Macroeconomic Environmental Variables. 136
Figure 4.8. Histogram of Cost Efficiencies 153
Figure 4.9. Bar Chart of Meta-frontier Efficiency Scores, Cost Gap Ratio and Individual Frontier Efficiency Scores 155
Figure 4.10. Trends of Cost efficiency from Pooled models, Individual Frontiers and Meta-frontier with Cost Gap Ratio 161
Figure 4.11. The trend of Economies of scale 172

List of Tables

Table 4.1. Clustering Macroeconomic Variables 103
Table 4.2. List of Inputs and Outputs 105
Table 4.3. Microeconomic Variables 107
Table 4.4. Summary stats of Macroeconomic Variables 118
Table 4.5. Summary stats of Standardized Macroeconomic Variables 119
Table 4.6. Calinski/Harabasz pseudo-F Index 122
Table 4.7. Duda/Hart Index 123
Table 4.8. Cluster Analysis Results 125
Table 4.9. Economic Regions by National Bureau of Statistics China Database (NBSC) 126
Table 4.10. Observations, Provinces and Clusters 131
Table 4.11. Observation Counts 132
Table 4.12. Summary Statistics of Frontier Variables 134
Table 4.13. Summary Statistics of Environmental Variables 135
Table 4.14. Correlation Table of Environmental Variables 143
Table 4.15. Pooled Frontier Estimates 146
Table 4.16. Cluster Specific Frontier Estimates 148
Table 4.17. Meta-frontier Estimates 151
Table 4.18. Cost Efficiencies and Cost Gap Ratios 154
Table 4.19. Rankings Similarity 163
Table 4.20. Estimates of Scale Economies 170
Table 4.21. Bank-Specific Economies of Scale for each Province and each Cluster 171
Table 4.22. Summary Statistics for Determinants of Economies of Scale 173
Table 4.23. Determinants of Scale Economies 175
Table 5.1. List of variables and expected sign 205
Table 5.2. Descriptive Statistics – Listed vs non-listed and National vs Regional Banks 210
Table 5.2. Descriptive Statistics Listed vs Unlisted -continued 211
Table 5.3. Descriptive Statistics- Crisis vs Non-crisis period 212
Table 5.4. Correlation Coefficient 213
Table 5.4. Correlation Coefficient - Continued 214
Table 5.5. Determinants of Loan Loss Provisions-All Banks 218
Table 5.6. Determinants of Loan Loss Provisions- Regional Banks Panel A 220
Table 5.6. Determinants of Loan Loss Provisions- Regional Banks Panel B 222
Table 5.7. Test on signalling 224
Table 5.8. Estimation of discretionary Loan loss provisions (DLLP) 226
Table 5.9. Alternative test on Signalling 228
Table 5.10. Descriptive Statistics for Test of Impact of Retail Deposit on Earnings Quality 230
Table 5.11. Test of Impact of Retail Deposit on Earnings Quality 231
Table 5.12. Robustness check- all banks 234
Table 5.13. Robustness check- National vs Regional Banks 235
Table A.1. Bank Performance Studies using Profitability Ratios 259
Table A.2. Summary of literature on SFA studies in Chinese Banking Sector 263

Chapter 1. Introduction

1.1 Research questions

This thesis attempts to answer the following main research questions:

(1) To examine whether regional commercial banks in China with a large disparity in regional economic environments operate under the same technology.

(2) Does cost inefficiency for Chinese commercial banks mainly come from managerial inefficiency or their technological gaps?

(3) What affects the deviation between each bank and its regional frontier? What are the determinants of the technological gaps?

(4) To investigate regional differences in terms of cost efficiency and the main drivers for these differences.

(5) Do regional banks in China enjoy economies of scale under the current technology? Does adopting a better technology allow them to enjoy greater economies of scale?

(6) What are the determinants of economies of scale for Chinese regional banks?

(7) What are the current practices of loan loss provisioning in China? Are loan loss provisions in China procyclical or countercyclical? Do Chinese banks use loan loss provisions (LLPs) as a mechanism for income-smoothing and capital management?

(8) Do LLPs in China signal good news or bad news to the investors?

(9) Does retail deposit have an impact on the discretionary use of LLPs?

1.2 Research motivation and contribution

To the best of our knowledge, regional disparities in the performance of Chinese banks have drawn little attention thus far (Sun et al., 2013). So far, the primary focus has been on large national-branching state-owned and joint-stock banks (Yang et al., 2015). Within the literature, bank efficiencies in China have been estimated with one frontier under the assumption that they operate under the

same technology[1] (Fu and Heffernan, 2007, Kumbhakar and Wang, 2007, Berger et al., 2009, Jiang et al., 2009, Du and Girma, 2011). However, regional differentiation, especially the coast-inland gap, is considered a fundamental characteristic of the Chinese economy. (Naughton, 2006). Thus, it is of great interest to capture the regional economic differences when modelling Chinese bank performance. Furthermore, regional banks offer a more homogenous sample, including mainly city and rural commercial banks that operate primarily within their own provinces or regions. It would also be interesting to know whether adopting a better technology would allow banks to enjoy greater economies of scale. With their rapid growth in recent years, this study would provide insight for regulators to find the best strategies to maximise both cost efficiency and economies of scale for these regional banks, also to create competition for the largest banks. In terms of risks in the Chinese regional banking sector, there has been a lack of studies on loan loss provisioning practice in China, considering it is a relatively large bank accrual. Moreover, we wanted to know if Chinese banks manipulate LLPs to practice earnings management and capital management, even with increasing market discipline and increasing audit quality over the past decade. It would also be of interest for investors to know whether the information contained in LLPs signals financial strength as good news or an expense as bad news.

As far as I am aware, this study is the first one in the literature that applies the latest stochastic meta-frontier (SMF) method to the Chinese regional banking sector. This study offers insight into the macroeconomic factors that contribute to technological gaps. Moreover, this study also links cost efficiency with scale economies and quantifies the effect of better technology on scale economies and offers bank-specific factors that impact economies of scale. This thesis also offers strong evidence of income-smoothing for Chinese commercial banks, especially during a financial crisis period, and shows evidence that loan loss

[1] Huang and Fu (2013) recognise the differences in financial market structure, institutional environment and political development and estimate a meta-frontier for mainland China and Taiwan under the assumptions that they operate under different technologies. However, they do not study provincial/regional differences within mainland China.

provisions do not signal good news to investors. Lastly, we offer insight for policy makers that the banks relying on wholesale funding may need closer monitoring in terms of accounting information in order to enhance market discipline.

1.3 Overview of Chapters

Chapter 2 gives an overview of the Chinese banking system. Chapter 3 provides a literature review on the main theories on banking firms and provides the foundation for the methodology on bank efficiency and stochastic frontier analysis. Chapter 4 firstly uses a novel hybrid clustering method to group Chinese provinces into three regions based on regional macroeconomic conditions, then adopts SMF methods to investigate the contributors to both technological gap and managerial inefficiency. It also provides estimates on economies of scale and determinants of economies of scale using a series of regressions. Chapter 5 studies the behaviour and determinants of LLPs and tests a list of hypotheses. The cost efficiencies obtained from Chapter 4 are also included as an explanatory variable in Chapter 5, offering the link between cost efficiency and LLPs. Finally, Chapter 6 concludes the thesis.

Chapter 2. A Review of the Chinese Banking System

2.1 Introduction

As the world's second-largest economy, China also has the third-largest territory in terms of landmass, with considerable regional disparities in terms of economic development (Naughton, 2006). The Chinese banking sector plays an important role in the Chinese economy. As of the end of 2016, the total assets of the Chinese banking sector totalled RMB 232 trillion, which was approximately three times as big as China's GDP.

This chapter provides an overview of the Chinese banking sector. In section 2.2, we review the four stages of the Chinese banking reform from 1978 until the present. In section 2.3, the structure of the Chinese banking system after banking reforms is examined. An overview of the Chinese banking sector during our study period from 2005 to 2015 is given in section 2.4. Section 2.5 provides the conclusion of the chapter.

2.2 Chinese Banking Reform

The Chinese banking system started as a monobank system in the planned economy in the 1940s. That is, a government bank was in control of all the financial businesses in China. The central bank at that time, the People's Bank of China (PBC), was established in 1948 as a result of consolidating Huabei Bank, Beihai Bank and Xibei Peasant Bank. Other banks are either restructured into or administrated by PBC. Therefore, they served as part of the monobank system to facilitate the nation's production plans without any competition among each other.

The first phase of reform started in 1978 to increase competition and improve the performance of the Chinese banking system. During the 1980s, four commercial banks with designated functions known as the 'big four' are carved out of the monobank taking charge of different business areas. As a result, a two-tier banking system has been created. The first Tier is the PBC, with the function of maintaining price stability, supervising financial institutions,

conducting clearance and issuing banknotes. The second Tier consists of the newly established specialised state-owned commercial banks, operating as commercial banks. Most of the Deposit and Lending business was taken charge of by the Industrial and Commercial Bank of China (ICBC), the bank with the largest assets, in the cities and by Agricultural Bank of China (ABC), the bank with the largest personnel number and more than 44000 branches, in the rural areas. The China Construction Bank (CCB) was responsible for financing infrastructure projects, and the Bank of China (BOC) was responsible for foreign trade and transactions. The 'big four' banks were all accountable for lending to state-owned enterprises and set up local branches and offices to help local governments fulfil production plans without considering either the default risk of these loans or their profitability. Due to state-owned enterprises being highly inefficient and making losses, the loan default rates were very high, leading to a high volume of non-performing loans. All of the 'big four' banks are troubled by unrecoverable loans due to lending on demand to local and central governments accompanied by the problem of skill-lacking staff and a lack of profit-oriented objectives (Naughton, 2006).

The 'big four' were created to serve their designated areas of the economy, and local branches and offices are under the administrative control of their respective local government. Therefore, bank branches have little competition across regions or sectors. The second phase of the reform to further deepen the institutional restructuring started in 1984, with the objectives of increasing competition among state banks and improving the efficiency of the banking system. Since the early 1980s, foreign banks and domestic joint-stock commercial banks (JSCBs) were allowed to enter the market and nationwide and regional JSCBs were launched with shareholdings ownership structures. In order to change from policy-oriented towards profit-oriented, the restrictions of having to serve a designated sector for state-owned banks were removed during the first step of the reform in 1985. State-owned banks were encouraged to enlarge their business scopes and compete with one another as well as JSCBs. However, the competition was still rather limited due to regular government interventions and control of local branches by local governments. As of 1994, 'big four' state-owned banks still played a dominant role in maintaining

economic growth and stability (Yao et al., 2007).

The third phase of the reform aimed to commercialise the banking sector. Three policy banks, namely, China Development Bank, Agriculture Development Bank and Import and Export Bank of China, were established to take over the policy lending activities from 'Big Four' in 1994. In 1995, the law of the People's Republic of China on Commercial Banks was enacted, symbolising the beginning of bank commercialisation. The law defines the 'Big Four' as state-owned commercial banks, which are profit-driven and operationally independent. Nonetheless, due to a lack of branch networks and capital, the newly established policy banks did not have the capacity to fulfil the demand of policy lending. Therefore, state-owned banks still participated in policy lending and were under pressure by central and local authorities to lend to their favoured organisations or sectors. As a result, large volumes of non-performing loans (NPLs) have been accumulated in the 1990s. By the end of 1998, major state-owned banks were deemed insolvent, and the banking system and the total NPLs amounted to RMB 3.3 trillion under the four-category loan classification system, accounting for 41 percent of GDP for the year.

To strengthen the banking system and prevent further accumulation of NPLs, PBOC decided to prevent local government interference by merging local branches into nine large regional branches. Local governments can no longer appoint senior officials in their respective branches of PBOC.

The first round of bailouts of state-owned commercial banks (SOCBs) was during 1998-1999. To resolve the bad debt problem, RMB 270 billion was injected into SOCBs in 1998 via bond issuing and four state-owned asset management companies (ACMs) were established to purchase NPLs. These are Cinda Asset Management Company for CCB, China Great Wall Asset Management Company for ABC, Oriental Asset Management Company for BOC, and China Huarong Asset Management Corporation for ICBC.

Apart from JSCBs, 90 profit-oriented city commercial banks with shareholding ownership structures were also set up to replace the former urban cooperatives. They were only able to operate within their respective regions and were expected to compete with local branches of state-owned banks.

The last phase of the Chinese banking reform from 2003 onwards initiated by the central government mainly comprised of radical capital injections into SOCBs, stripping off non-performing loans, as well as further privatisation by inviting foreign strategic investment into domestic banks and encouraging Initial Public Offerings (IPOs) in order for them to be listed on the stock market (Hung, 2009). Since entry to WTO in 2001, to comply with WTO rules of no restrictions on foreign currency businesses, China made the plan to fully open up its banking sector by 2006. Therefore, the pressure to improve the performance and efficiency of domestic banks, especially state-owned banks, became a pressing issue. To further strengthen the banking system, the government further injected US$22.5 billion each into BOC and CCBC in 2003 and US$15 billion into ICBC in 2005 and further dumping of NPLs of a total of RMB 475 billion in 2004 and RMB 705 billion in 2005 executed by the AMCs. At the end of 2008, the average NPLs have been reduced to 2.4% of total loans (USCC, 2013).

A number of strategic investors have obtained shares in SOCBs as well as a large number of JSCBs. For example, Bank of America acquired a 9% stake in CCBC, and Goldman Sachs, Allianz and American Express acquired a 10 % stake in ICBC; Around 10% of the shares in BOC were bought by Royal Bank of Scotland together with Merrill Lynch and Hong Kong tycoon Li Ka-Shing (Kwan, 2009).

All domestic banks were encouraged to list their shares on the stock market to privatise further and boost their corporate governance. In 1991, Shenzhen development bank became the first bank in China that went public and was listed on the Shenzhen Stock Exchange. Bank of Communications became the first domestic bank in china to be listed overseas, making their IPO on the Hong Kong Stock Exchange in 2005. Subsequently, three of the 'big four' SOCBs, namely CCB, BOC and ICBC, were listed on the Hong Kong Stock Exchange from 2005 to 2006, as the government believes that stricter disclosure regulations of offshore markets could help further accelerate the structural reform of the domestic banking system and improve its efficiency. These three SOCBs were further listed on the Shanghai Stock Exchange as well from 2006 to 2007. By the end of 2017, five state-owned banks, one Postal Bank, four joint-stock banks (JSCBs), and more than ten city/rural commercial banks were listed on the H-

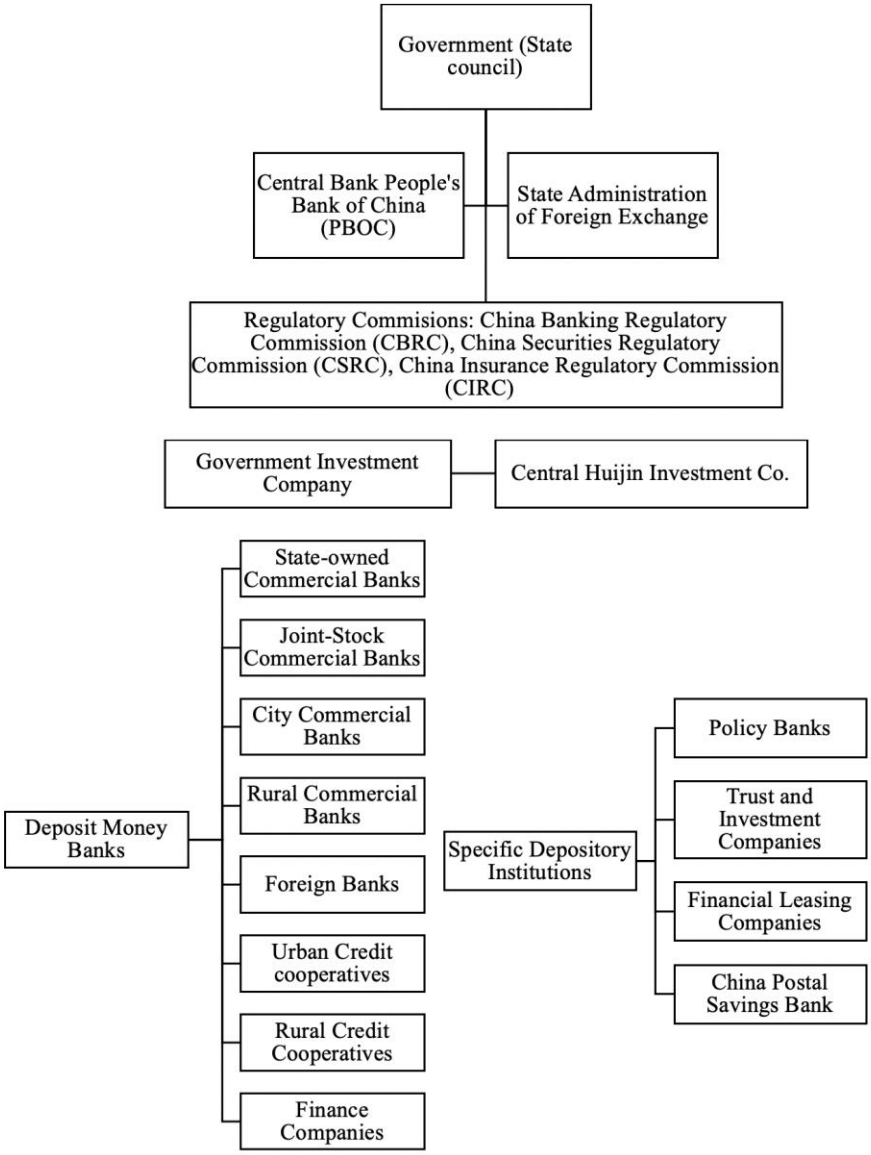
share market. Another four JSCBs and more than ten city/rural commercial banks were listed on the A-share market(Wu, 2018).

To increase banking supervision and regulation, the China Banking Regulatory Commission (CBRC) was established, allowing PBOC to focus on macroeconomic stability and growth. Two important laws on the banking sector, The Law of the People's Republic of China on The People's Bank of China and The Law of the People's Republic China on Commercial banks, came into effect in 2005, which set up the legal foundations for the stability of the Chinese banking system. In 2004, Regulation Governing Capital Adequacy of Commercial Banks, the regulation which set up the new capital adequacy requirements for Chinese commercial banks, is in compliance with both 1998 Basel I and 2002 Basel II, the international regulatory accords. The regulation set the calculation methods for capital adequacy, which accompany the law that requires all commercial banks to hold a minimum of 8 per cent capital adequacy rating. The regulation also set up new rules for internal rating and credit risk management.

2.3 Structure of the Chinese Banking System

According to Wu (2018), PBOC officially became the central bank of China in 1984 with the original four state-owned banks, namely Industrial and Commercial Bank of China (ICBC), China Construction Bank(CCB)，Agricultural banks of China (ABC) and Bank of China (BOC) established/converted as specialized banks during 1978-1984. They were fully converted to full-commercial banks when the three policy banks, China Development Bank, Agriculture Development Bank and Import and Export Bank of China, were established to take charge of policy lending activities in 1994. Bank of Communications, which was a joint-stock bank, was re-classified as a state-owned bank in 2015. Both policy banks and the big commercial banks are regulated by the China Banking Regulatory Commission (USCC, 2013) (see figure 2.1).

Figure 2.1. Structure of the Chinese Banking System



As of 2020, China's banking sector can be mainly categorized into three tiers. The first two tiers are five state-owned commercial banks and 12 joint-stock commercial banks, both having a national presence. The state-owned banks and joint-stock banks account for 56% of the asset market share in the Chinese banking sector in 2018. The third tier consists of city and rural commercial banks.

They are established to serve the local economy, and the vast majority of them operate regionally (Sun et al., 2013, Zhu and Chen, 2016).

The state-owned Chinese banks had a severe bad debt problem in the 1990s that resulted from lending to poorly run projects by SOEs. The NPLs had surged to RMB 2.5 trillion, amounting to 39% of total loans in Chinese banks and 31% of China's GDP (USCC, 2013). To resolve the bad debt problem, the Chinese government has taken four major steps to move bad loans out of the banking system. Firstly, four state-owned asset management companies were established in 1999 to purchase NPLs, and a series of NPL-disposal happened from 1999 to 2008. From 2003 to 2008, four state banks each received a capital injection from China's FX reserves. At the end of 2008, the average NPLs have been reduced to 2.4% of total loans (USCC, 2013). State-owned banks also received a significant amount of investment from foreign banks from 2004 to 2006, resulting in state banks each having a few foreign banks on the list of their top 5 shareholders. During 2005-2010, the state banks had also been listed on the stock market to further boost their capital base (Kwan, 2009).

Joint-stock commercial banks (JSBs) emerged in 1987, and their market share in the Chinese bank sector had risen from below 12% in 2006 to 19% in 2016. Because of less reliance on government support and excessive branch network and because they are mostly owned by commercial entities such as SOEs, they operate in a more competitive and market-oriented manner. JSBs can be classified into categories according to their background. For example, Shanghai Pudong Development Bank (SPDB) is based in Shanghai City and Guangdong Development Bank (GDB) is based in Guangdong Province, and they are both supported by the local government and enjoying the low cost of funding from local governments' deposits. Some banks are industry-based, such as China Merchant Bank (CMB) and Hua Xia Bank (HXB), which mainly serve corporations in the transportation sector and infrastructure sector, respectively. Some JSBs belong to their financial holding companies, such as China CITIC Bank (CNCB) from China CITIC Group and China Everbright Bank (CEB) from China Everbright Group. These JSBs enjoy the advantage of a readily available platform for diversifying their services. The last category contains banks

supported by private corporations and investors. For instance, Min Sheng Bank (MSB) has well-diversified ownership from different private sectors. Pin An Bank (PAB) is owned by its parent company Pin An Group, which is primarily an insurance company.

Because state-owned banks dominate the Chinese banking sector and the credit is mainly supplied to state-owned enterprises and state-led projects, there is limited credit available from the banking sector to SMEs and entrepreneurs. This credit gap and underdevelopment of the capital market in China lead to a significant presence of shadow banking activities (USCC, 2013).

The third tier consists of city and rural commercial banks and rural cooperative banks, which operate primarily regionally. These banks are usually only licensed for operation in one particular city/area and owned mainly by the local government, enterprises or employees. Therefore, although there have been some recent expansions into nearby cities or areas among these banks, the vast majority of their operations are still concentrated on the regional level (Wu, 2018).

City commercial banks are originated from urban credit cooperatives. In 1995 due to poor management and the heavy burden of NPLs, People's Bank of China arranged to merge all urban credit cooperatives and some rural credit cooperatives and local financial institutions into joint-stock companies and city commercial banks, with new shareholders ranging from local government to local collective and private firms. City commercial banks were formed based on one-city-one-commercial bank rule and were only allowed to set up branches outside their own administrative regions after 2006. Since 2007, a few city commercial banks have expanded their operation outside their province rapidly, although, to this date, the majority of them still operate regionally (Sun et al., 2013). Similar to city commercial banks, rural commercial banks are formed as joint-stock companies from mergers of rural credit cooperatives during the second and third phases of the Chinese banking reform. Both the size and number of rural commercial banks have developed rapidly (Zhu and Chen, 2016). As can be seen in table 2.5 in section 2.4 in this chapter, the total assets for both

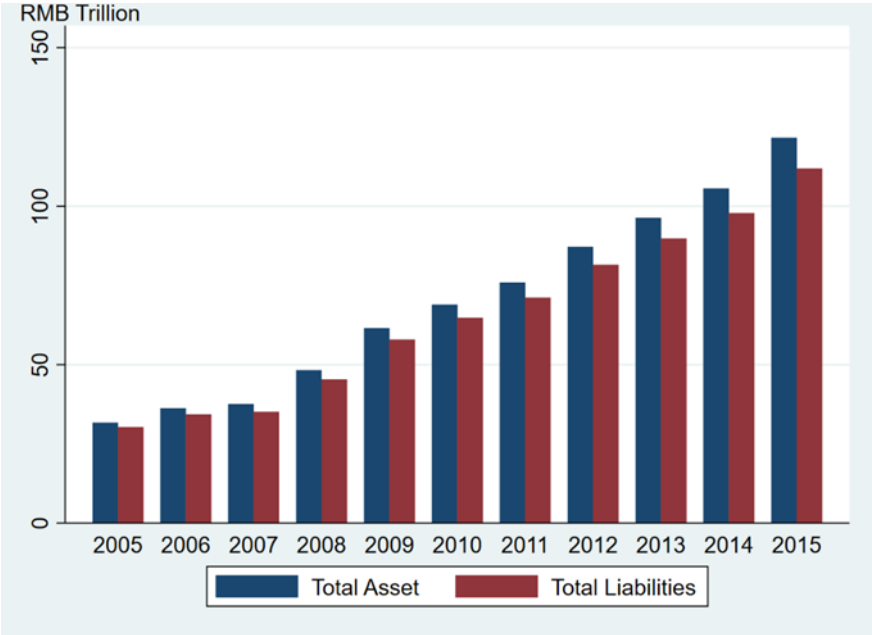
city and rural commercial banks have reached 20% of total commercial banking assets.

City and rural commercial banks are used as our main sample in this thesis to investigate the impact of regional disparities on bank performance. This is because for each bank, only data on the aggregate level are available to us, and we do not have access to branch level data. Even if branch level data are available, branches are usually allowed to transfer deposits and loans to branches in other provinces/regions, thus not reflecting each branch's actual performance. The rationale for this thesis to investigate the link between regional disparity and bank performance is presented in Chapter 4, section 4.3 Motivation.

2.4 An overview of the Chinese Banking Sector from 2005 to 2015

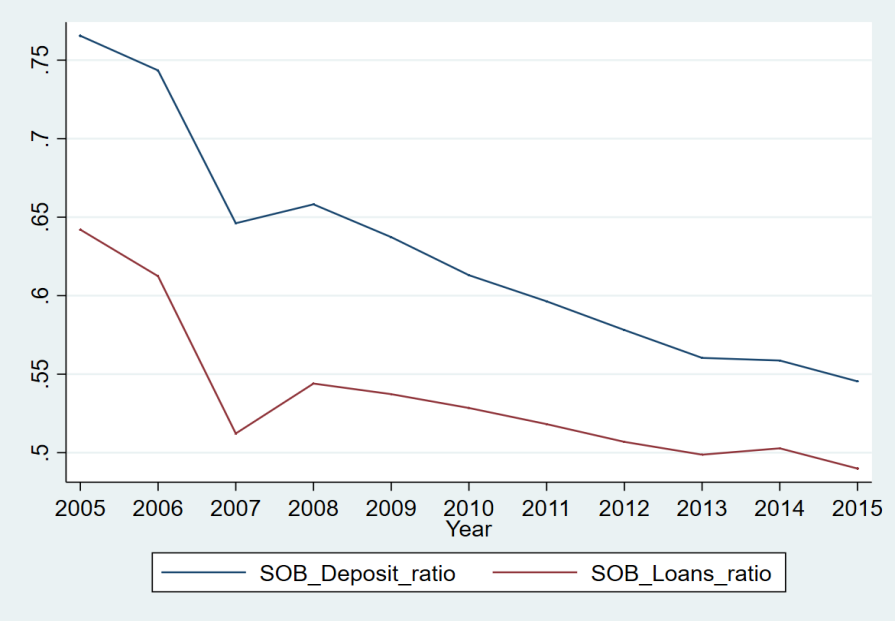
As of 2015, as in Figure 2.2, the total assets of the Chinese banking sector reached RMB 121.58 trillion in 2015 and a total of RMB 111.88 Trillion in terms of total liabilities (BankScope). The total assets have increased approximately 284 per cent from the beginning of the period, and the total liabilities have increased by around 270 per cent. Although market share is still dominated by the biggest five state-owned banks, there has been a steady trend of decentralization of deposit and loan market share in the last decade (Figure 2.2).

Figure 2.2. Total Assets and Liabilities of Banking Institutions (2005-2015)



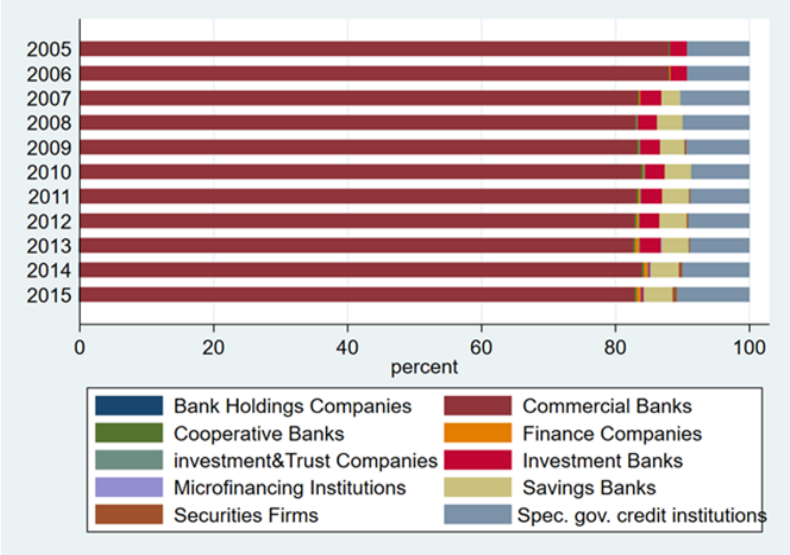
Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.3. Loan and Deposit share of State-owned Commercial Banks (2005-2015)



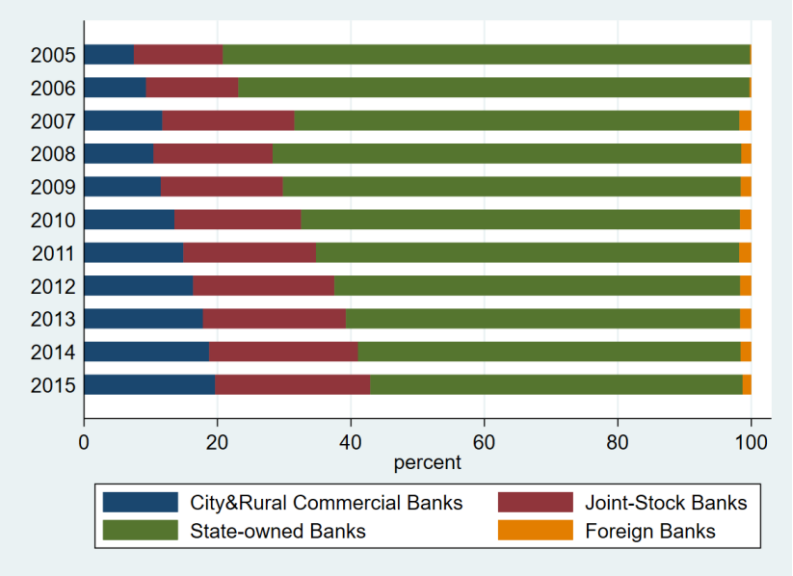
Data Source: BankScope by Bureau Van Dij, and author's own calculation

Figure 2.4. Asset Share breakdown of China's Banking sector



Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.5. Asset Share breakdown of China's Commercial banking sector



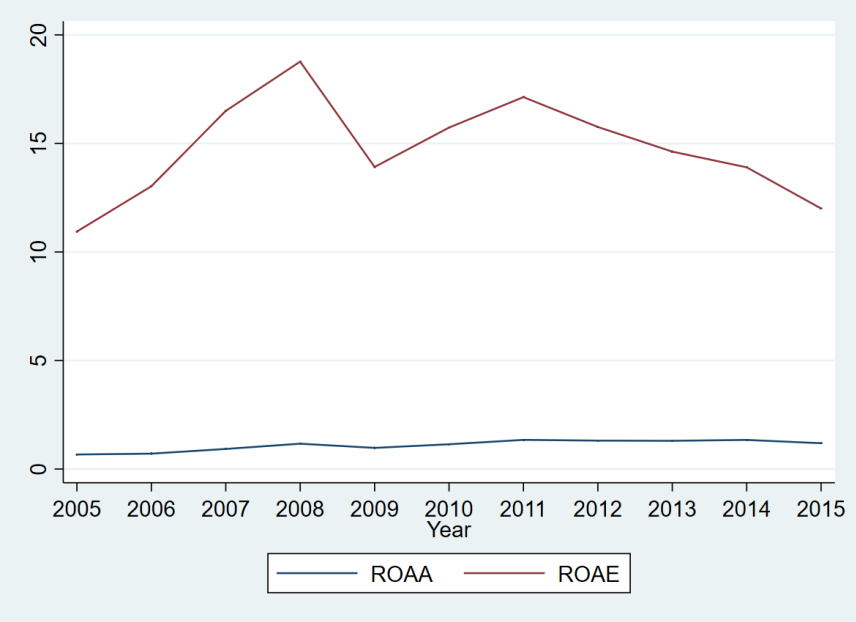
Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.4 displays the market shares for each specialisation in the Chinese banking sector, according to BankScope's classifications. Commercial banks have dominated the banking sector throughout the study period, while we can also observe an increasing number of other financial institutions emerging during the period. Figure 2.5 indicates that, within the commercial banking sector, although the big five state-owned commercial banks dominate the commercial banking sector, we can observe a steady and evident trend of decrease of their market shares over the 11 years. Both joint-stock commercial banks and city and rural commercial banks show a clear trend of expanding their asset shares in the market. Especially for city and rural commercial banks, the market share went from less than 10 per cent to nearly 20 per cent during the period. Foreign bank asset share increased significantly during the beginning of the period from 2005 to 2007 and remained stable during the majority of the period with a slight decrease in the more recent years.

The profitability of the Chinese Banking sector can be measured by return on asset (ROA) and Return on equity (ROE). In terms of ROA, figure 2.7 indicates that foreign banks appear to be the most efficient before the third quarter of 2006,

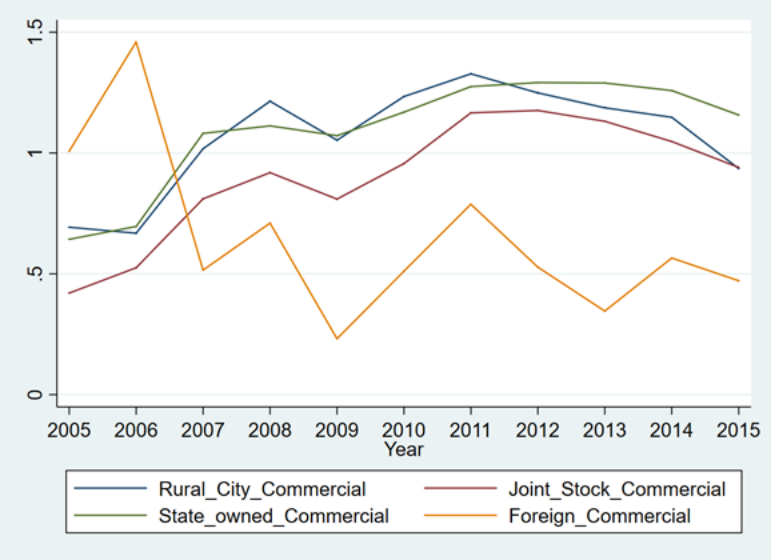
and their profitability drops dramatically from 2006 to 2007, then fluctuates over the rest of the period and remains below the other bank types. On the contrary, Chinese domestic banks perform well, represented by a steadily increasing ROA over the 11-year period. Especially, we can observe that Chinese banks continue to thrive during and post 2007-2008 global financial crisis, whilst foreign banks perform significantly worse during and after the crisis period. One reason is that Chinese domestic banks were mainly under the administration of the government and were much less affected than foreign banks during the crisis. The capital injections and stripping of NPLs during the period also contributed to the increasing healthiness of the domestic banks. Among domestic banks, rural and city commercial banks generally performed better than joint-stock banks, although the difference was not large (less than 0.5 per cent).

Figure 2.6. Return on Average Assets of Chinese Banks



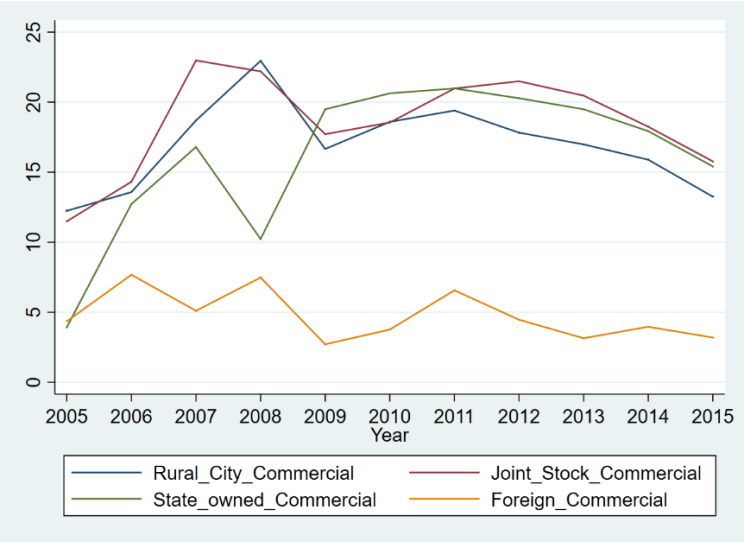
Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.7. Return on Average Assets by Commercial Bank Types



Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.8. Return on Average Equity by Commercial Bank Types



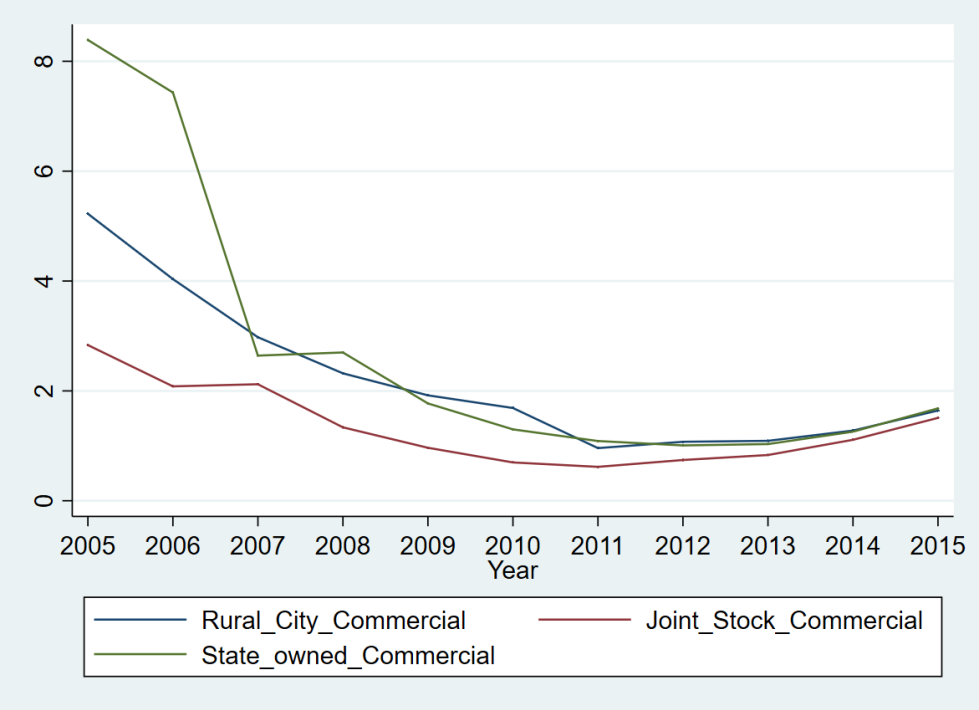
Source: BankScope by Bureau Van Dijk, and author's own calculation

In terms of profitability measured by return on equity (ROE), all banks witnessed a sharp drop in profitability during or after the 2008-2009 global financial crisis. Figure 2.8 indicates that joint-stock and rural and city commercial banks performed better than state-owned banks during the pre-crisis and crisis periods. However, state-owned commercial banks were able to bounce back and reach a similar level of ROE as other types of commercial banks within a few years after the crisis. Joint-stock banks appear to be more profitable than all other types of commercial banks from 2011 onwards. This may be partially due to the more extensive branch network and thus bigger deposit base and easier access to cheap funding.

Chinese banks have a long history of the problem of non-performing loans, and the problem is still yet to be solved. According to USCC (2013), The NPL problem is not only especially prevalent in state-owned commercial banks but is also putting strains on the rural cooperative system, including rural cooperative banks and rural commercial banks, as they are requested to support projects and enterprises owned by the local government. The loan allocations of rural commercial banks and rural cooperative banks are also under the influence of the local government. The rural cooperative system has been managed by the provincial credit union and reports to the provincial government since 2003 and

they contributed significantly to the 2008-2009 stimulus program through excessive lending.

Figure 2.9. NPLs of Chinese Domestic Commercial Banks



Data Source: BankScope by Bureau Van Dijk, and author's own calculation

Figure 2.9 indicates the status of NPLs of Chinese domestic commercial banks during 2005-2015. The state-owned commercial banks seem to have the highest level of NPLs amongst other types of domestic banks.

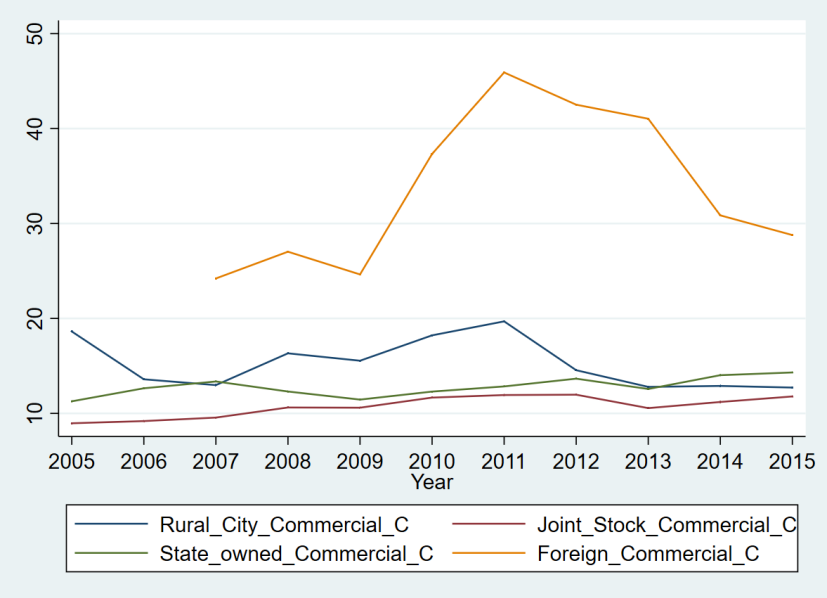
We can also see a steady drop in NPL from the beginning of the period until 2011, reaching the lowest point of less than 1% on average around 2012-2013, despite the estimated total loan issuance of RMB 9.6 trillion in 2009, which was more than double the amount in 2008 and around RMB 8 trillion in 2010(KPMG, 2013). USCC (2013) suspect that the government-owned asset management companies may have played an important role in the disposal of bad loans that resulted from this large amount of lending in the stimulus program. As mentioned before, rural commercial banks are serving local government, and their loan allocation is largely influenced by the local government rather than

being commercially decided. They also invested heavily in the stimulus program during the crisis period. It is not surprising that rural commercial banks seem to have more bad loans than the rest of the categories.

It is clear that NPL has been putting strains on the performance of Chinese commercial banks, especially the bank performance measured by x-efficiency, where banks are considered as intermediators to use deposits and capital to produce loans and other earning assets. This gives more evidential support that the non-performing loans problem is region-specific and Chinese banks, especially considering the burden of lending to the local government-led projects, should be examined regionally in order to obtain the most accurate picture of their performance (USCC, 2013).

Figure 2.10 presents the total regulatory capital adequacy ratios (CARs) of Chinese commercial banks. Foreign banks hold the highest level of total regulatory capital throughout the period. Since 2009, CARs of Chinese banks have been above the level of 10%, in compliance with Basel Accords, in spite of a disparity among different banks. For example, JSCBs have been maintaining the lowest level of CARs above Basel requirement, compared to the rest of the banks in China. With the fact that JSCBs also appear to be the most profitable during the period, we can infer that JSCBs are the most commercial-oriented banks that hold the least amount of eligible capital to allow for more investments.

Figure 2.10. Capital Adequacy of Chinese Commercial Banks



Data Source: BankScope by Bureau Van Dijk, and author's own calculation

2.5 Conclusion

This chapter gives an overview of the four phases of Chinese banking reform since 1979 and the subsequent banking structure following the reforms, followed by an analysis of the current Chinese banking sector. Chinese banking reform succeeded in increasing competition, reducing NPLs and increasing efficiency, corporate governance as well as profitability, especially for SOCBs. During our study period, the size of the banking sector increases steadily in terms of total assets and total liabilities while the loan to deposit ratio decreases. Although state-owned commercial banks still dominate the Chinese banking sector, we can observe a steady increase in the market share of joint-stock commercial banks and city and rural commercial banks (regional banks). The profitability of the banking sector has an increasing trend at the beginning of the period and a decreasing trend towards the end of the period, with a sharp drop during or

shortly after the crisis period. Overall, foreign banks appear to be less profitable than domestic banks during 2005-2015. Chinese banks have also been increasing their regulatory capital levels and remain above the minimum capital adequacy requirement set by Basel Accords.

In the following chapters, we will first present a justification for the performance measure used in the thesis, i.e., cost-efficiency. Then, we will provide empirical evidence on how regional regulatory, economic, and industry factors affect Chinese bank performance and finally test a series of hypotheses that may explain loan loss provisioning behaviours presented in the Chinese banking sector.

Chapter 3. A Literature Review on Bank Performance Measurements - A Justification for Using Cost Efficiency

3.1 Chapter Summary

This chapter provides a justification for using cost efficiency as a preferred measurement for bank performance. We first present studies using financial ratios to measure bank performance and the inherent mathematical and statistical issues involved in the dynamics between the numerator and denominator of financial ratios. Then we introduce the concept of efficiency and common models used in the literature. We only present common parametric approaches in the literature as this thesis uses parametric approaches. Finally, we present choices of inputs and outputs for efficiency modelling under different bank theories.

3.2 A Literature Review on the Theories and Examples of Financial Ratio Analysis

Return on Asset (ROA) and Return on Equity (ROE) are the most commonly used financial ratios to measure the performance of firms in both the financial and non-financial sectors. The banking industry is highly regulated. Basel regulations and the effects of bank failures on the economy make the most important financial ratios for banks likely differ from nonfinancial firms. For example, banks are more concerned with ROE rather than ROA. This is because loans contribute to the majority of assets for banking firms. Big banks tend to practice Originate-To-Distribute when dealing with assets, i.e. selling off assets on their balance sheet, while small banks still tend to hold the loans (Originate-To-Hold). Therefore, big banks might appear to have a higher return on assets by simply selling off loans and obtaining fees and income from securitisation. This is unlikely to happen in non-financial firms where they originate to distribute core investment assets. With the improvement of technology and digitalisation, the fixed assets of financial firms are moving from branches to the internet. Therefore, it casts doubt on the validity of using ROA as a measurement for the profitability of financial firms. In terms of ROE, there has

been evidence in the press releases that both financial and non-financial firms can temporarily boost ROE by buying back shares [2].

Rose and Hudgins (2012:172) define ROE as "a measure of the rate of return flowing to shareholders. It approximates the net benefits that the stockholders have received from investing their capital in the financial firm (i.e. placing their fund at risk in the hope of earning a suitable profit)". They interpret ROA as "an indicator of managerial efficiency; It indicates how capable management has been in converting assets into net earnings"(Rose and Hudgins, 2012). Kosmidou (2008) uses ROA as an indicator for a study on the performance of the Greek banking industry. They define ROA as "net profit after tax divided by total assets and indicates the returns generated from the assets financed by the bank". The study on the profitability of the Turkish banking industry by Sufian and Habibullah (2009) uses both ROA and ROE as measurements of profitability. They regard ROA as "a general measure for bank profitability reflecting bank's ability to achieve a return on its sources of fund to generate profits" and ROE as "net profit divided by shareholders 'equity'". The table in Appendix A summarizes the recent studies which use financial ratios as indicators for bank performance.

2 The chief executive of Credit Suisse declared in 2015 that they will no longer use ROE as targets and will consider growth targets instead. This is mainly because equity is largely dependent on regulations and the fact that banks shrinking equity to boost ROE contributes to the severity of financial crisis (Jenkins, 2015). Thus, it brings ROA as potentially the preferred ratio.

The analysis of ratios can offer insight into the health of financial institutions as a ‘snap-shot’. Moreover, these analysis ratios can also have additional statistical benefits in analysing banks. That is, Wiseman (2009) summarizes these as follows:

(1) Ratios can be used to deflate or scale for size effects. This is because many variables taken from financial statements, such as net income and debt, are correlated with size. Therefore, these variables are usually divided by a measure of firm size. Deflated variables can be classified into two kinds: proportions and rates. Proportions require that the ratio's numerator must have a linear relationship with the denominator with zero intercept. That is, when the denominator is zero, the numerator must be zero. Studies on administrative intensity have shown that failure to meet proportionality conditions can lead to biased results (Feinberg and Trotta, 1984, Freeman and Kronenfeld, 1973). On the other hand, rates measure the frequency of occurrence. An example of a rate is employee turnover. Wiseman (2009) argues that proportions are less prone to biased results than rates. Deflating ratios by size variables has been commonly practised in the banking literature because the size of banks varies considerably within almost all banking industries. The most interested variables, such as net income, net interest income, and non-interest income, are usually divided by total operating revenue or total assets to make it possible to compare profitability between big and small banks (Rose and Hudgins, 2012).

(2) Ratios can be used to correct for heteroscedasticity. Heteroscedasticity happens when the variance of the error term is not constant across the sample, causing inefficiency in parameter estimates. This is because firm size usually varies considerably in the sample studied. Studies which fail to take size into consideration are likely to have problems of heteroscedasticity. This is again true in the banking industry. For example, McAllister and McManus (1993) point out that specification problems often occur when we do not take size into account when we

model large and small banks together in one model. Therefore, caution should always be exercised when we have both small banks and large banks in the sample as it is very likely to cause statistical problems, including heteroscedasticity.

(3) The third purpose for using ratio is to capture a multi-dimensional theoretical construct (Bollen and Ward, 1979). In other words, ratios can help to use one single measure to capture a complex idea with multiple dimensions. An example is Tobin's Q, which measures a firm's competitive advantage calculated as the firm's market value divided by its total asset value. These ratios can also be called conceptual ratios, where one variable is only considered theoretically meaningful with reference to another variable. For example, ROA can be both regarded as a deflated ratio and also a theoretical construct as one can see ROA as a proxy of asset productivity. Another example is the Herfindahl index, which is commonly used in banking studies to capture market concentration and to measure the degree of competition in a particular industry. In this case, this ratio is constructed based on theory to use a single measurement to capture a complex status.

There are also disadvantages of using ratios, which has been mentioned in the literature:

(1) Financial ratios can only reflect the profitability at the current period, without any power to indicate how the current investment or management decision can affect the future net income (Sherman and Gold, 1985). This characteristic of short-termism can be misleading when looking at the performance of banks in the long run. They gave an example that a cut down of expense on marketing and new product development might inflate the current accounting profitability ratio, but this action might hinder the capability of the bank to generate profit in the future (ibid.).

(2) Sherman and Gold (1985) also point out that financial ratios might not accurately reflect the state of the business. Especially, we cannot tell which aspect of the business is performing well or poorly by purely looking at the ratio figures. He mentions that accounting ratios are aggregate figures of all dimensions of a business, and 'a bank may appear to be performing well even if it is poorly managed on certain of these dimensions, as long as it compensates by performing particularly well on other dimensions' (ibid:298).

(3) Mukherjee et al. (2002) also point out that financial ratios cannot reflect the interaction between various inputs and outputs. It is stated that 'financial ratios… ignores the multidimensional aspects of bank performance. They give only a restricted, incomplete picture of the process and fail to account for the interaction between the different factors, leading to contradictory results'.

(4) They are not able to reveal how efficiently resources were being used. The study on bank branches by Sherman and Gold (1985) suggests that a bank branch being more profitable than others does not necessarily mean that the branch is utilizing resources more efficiently. For example, 'the branch with more complex and more resource consuming transactions may appear less efficient than one with less complex transactions'(Sherman and Gold, 1985: 298). They propose that Data Envelopment Analysis (DEA) can help conquer this problem. The results of their study suggest that 'DEA can provide insights into branch operating efficiency beyond that available from accounting ratios' (ibid.: 298). In other words, results from DEA can show which inputs and outputs to increase or decrease relative to competitors to reach the frontier.

(5) Financial ratios fail to consider the multiple inputs and multiple outputs of a business (ibid.). This is particularly common in the banking sector, where banks use multiple resources such as labour and capital to produce

a variety of outputs such as loans and securities. Therefore, financial ratios are not accurate measurements of bank operational efficiency. By just looking at the financial ratio figures, one cannot tell how efficiently banks utilize and allocate resources. Mukherjee et al. (2002) argue that Stochastic Frontier Analysis and Data Envelopment Analysis (DEA) can help resolve the problems of financial ratios mentioned above.

3.3 Statistical Problems of Financial Ratios

Even though we have seen a growth in the use of econometric specification to analyse the profitability of financial firms using ratios, there are still potentially hidden problems of which little is discussed in the literature. These problems of non-proportionality and inter-temporal instability are usually considered independently (see the seminal study by Sudarsanam and Taffler (1995)). Interestingly, these problems have had little to no exposure in the banking literature, even though ratios such as capital requirements under Basel III are prominent in central bank regulations.

Hence my aim is to re-analyse the statistical nature of bank financial ratios, following Sudarsanam and Taffler (1995) by bringing together the theories and discussions concerning these issues.

There are mainly two concerns in the literature. The first one is whether the functional form of the numerator and denominator allows a comparison of ratios to be meaningful. The second aspect is that since the primary purpose for using ratios instead of their denominators is to control for size effects, if ratios fail to control for size, in other words, if ratios turn out to be correlated with size, then this purpose cannot be served.

Firstly, we will look at the functional form of financial ratios. Several authors have pointed out that the relationship between numerator and denominator should be linear and proportional. The majority of the literature, such as Lev and

Sunder (1979) and Whittington (1980), support this argument stating that the relationship between the numerator and denominator should satisfy a strict proportional relationship. If we define the numerator as y and the denominator as x, then the relationship between y and x can be expressed in the equation below:

$$y = \alpha + \beta x + \mu \tag{3.1}$$

The strict proportionality requires that the power of x equals to 1 and the intercept term α equals to zero. Namely, linearity and zero intercept, which can also be simply expressed as

$$Y = \beta x \tag{3.2}$$

A violation of these properties makes comparisons of ratios invalid. Figure 3.1 indicates a disproportional relationship between return and equity. The line r, which represents the industry norm, shows a disproportional relationship between return and equity. It is clear from the graph that although A shares the same ROE and C and B share the same ROE as D, firms A and B are able to obtain a higher amount of return than average firms with the same level of equity. Similarly, firms C and D can only have returns that are lower than average firms in the industry with similar equity levels. Therefore, we can say that firms A and B outperform the industry norm, and C and D underperform the industry norm. Therefore, when ratios are not proportional, it does not make sense to use ratios as an indicator for firm performance. The reason for a disproportional relationship can be due to the fact that some part of the return of the company might depend on the size of the company, represented by F. These returns may come from grants, rents and some extraordinary items (De Andrés et al. (2009))

Figure 3.1. A disproportional relationship

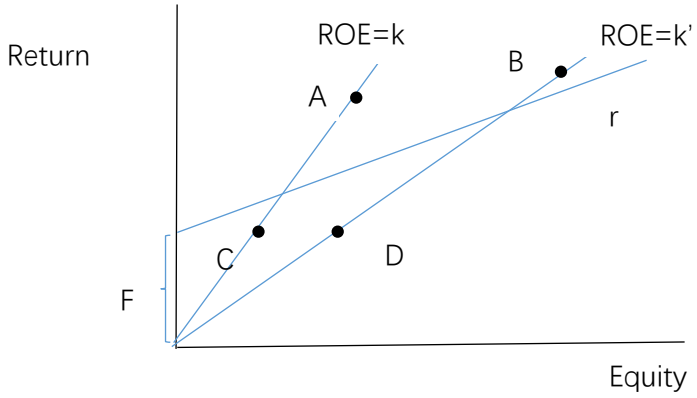
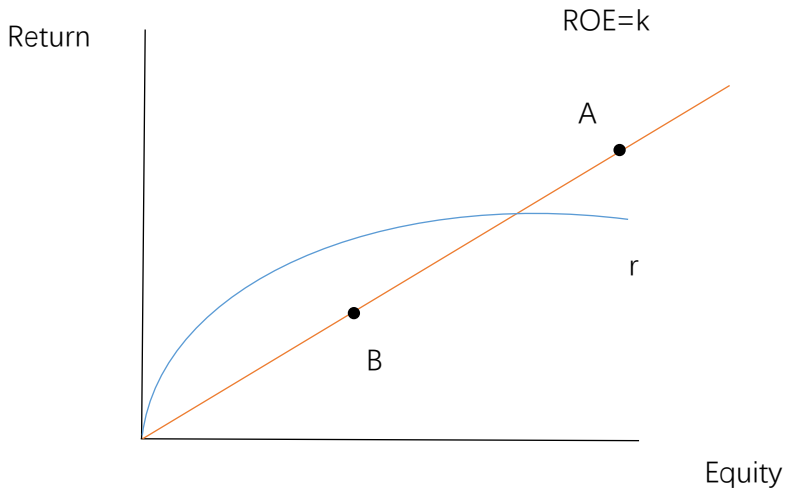


Figure 3.2 below shows a nonlinear relationship between the numerator and denominator of the ROE ratio. When the industry norm curve r is nonlinear, even though firms A and B have the same ROE, firm A outperforms the industry norm, and firm B underperforms the industry norm.

Figure 3.2. A nonlinear relationship



Accounting figures are rarely clean and subject to manipulations. De Andrés et al. (2009) provide three possible explanations for the nonlinearity of the relationship between ratio components, namely, economies of scale, small firm effect and earnings management practice. Firstly, there is likely to be a reduction in the average production costs as firm size increases, which will cause a relationship between return and assets or equity to be nonlinear; Secondly, small

firms have more growth opportunities, thus higher risk-adjusted returns than large firms. This has also been observed by Hopwood and McKeown (2003) in terms of the extremeness of returns in small firms. When firms grow larger, this extremeness seems to decline. De Andrés et al. (2009) presents evidence for both positive extremeness and negative extremeness and find that the small low-technology firms seem to be more innovative than big ones. However, in some cases, small firms face financial constraints and lack professional managers, which puts them in a disadvantaged position. Therefore, when small firms face limitations and there exists economies of scale, the relationship follows a convex curve; when small firms have more advantages than big firms in an industry and/ or diseconomies of scale exist, the ratio components relationship follows a concave curve. Thirdly, earnings management caused by the manipulation of accounting figures on financial statements of firms will cause nonlinearity in the return to asset relationship. Large firms tend to practice more earnings management than small firms, therefore, pushing the curve nonlinear and concave. In conclusion, extraordinary items, economies of scale, small firm effects and earnings management can all contribute to the distortion and nonlinearity of financial ratios.

Moreover, McLeay and Trigueiros (2002) argue that the above proportionality condition is insufficient to make ratio comparison valid. In addition to the requirement that y/x is a constant, they believe that "when Y is proportional to X, the rate of change of Y with respect to X remains constant and similar to ratio itself", which can be shown as

$$\frac{Y}{X} = \frac{dY}{dX}, \tag{3.3}$$

This can be rearranged as

$$\frac{dY}{Y} = \frac{dX}{X}. \tag{3.4}$$

This indicates that the relative change in Y equals the relative change in X.

This condition is also called the scale invariance condition. The intuitive explanation is that the relative changes in the numerator and denominator should be similar. For example, in ROA ratio, when the size of the company has grown three times, we expect the return to grow approximately three times as well. McLeay and Trigueiros (2002) call it a multiplicative process, meaning that the numerator and denominator grow at a similar speed. We also need to make sure that the relative change in numerator and denominator are similar across different firm sizes. If we cannot guarantee that the relative change of the numerator is similar to the relative change of the denominator across different sizes of firms, we cannot know whether a ratio above the norm is due to, for instance, better profitability or liquidity, or just merely a characteristic of larger firms.

There has also been empirical evidence indicating that the proportionality and linearity relationship rarely exist in the real world. Sudarsanam and Taffler (1995) use Box-Cox transformation to test the violation of proportionality condition on 24 commonly used ratios using a large sample of 500 companies from six industries. The results indicate that the proportionality condition is not satisfied for any of the 24 ratios and log linearity is a better description of the ratio component relationship. The variation of the ratio components relationship across industries and over time indicates that great caution needs to be exercised when comparing ratios in different industries and different time periods.

In terms of the relationship between ratio and size, both De Andrés et al. (2009) and Hopwood and McKeown (2003) find that the ratio is indeed affected by firm size. In other words, they find that the variability of the relationship between ratio components is less severe in big firms than in small firms. As they both focus on the profitability ratios, it essentially means that the variability of earnings/net income of big firms is less vulnerable to external shocks or non-financial influences than smaller firms. If the level and the variability of ratios are affected by size, it will cause statistical spuriousness when using ratios in statistical analysis. If the results from ratio analysis are spurious, the

conclusions made previously about the determinants of performance using ratios as performance indicators should be re-examined (Wiseman, 2009).

In banking research or financial research in general, we often compare ratios across firms of different sizes or use regression analysis to determine how the independent variables on the right-hand side of the equation affect the dependent variable on the left-hand side. A question as simple as 'whether non-interest income will determine the returns of a particular banking sector' will incur the consideration of bank size, as it is likely that the larger the bank is, the higher the return and the non-interest income will be. Therefore, we usually divide both variables by size with the hope to eliminate the effects of size. This leads to one of the fundamental reasons for using financial ratios in market research: that is, to isolate the effect of size.

However, if a ratio fails to control for size, in other words, if this ratio is not independent of size, and if we do not include this ratio on the right-hand side of the equation, we will face a standard omitted variables problem and the coefficients we obtained for the right-hand side variables will be biased. On the other hand, including size in the equation would also be problematic if size is correlated with the variance of the independent variables.

By conducting a spearman correlation analysis on the relationship between size and seven commonly used financial ratios as well as the size and standard deviations of the ratios using a sample of all firms on Compustat from 1975 to 1994, Hopwood and McKeown (2003) find that size is highly correlated with decile mean and decile standard deviations of the ratios. The ratios tested include net income/Total assets, current asset/sales, long-term debt/Total assets etc. More precisely, size is strongly negatively correlated with the decile standard deviations of ratios, indicating that the larger the firms, the less cross-sectional financial variability between firms. The explanation by Hopwood and McKeown (2003) is that, for earning-type ratios, the extremeness in earnings will decline as firms grow larger because small firms generally have higher bankruptcy rates and likelihood of extreme growth. During the process of small firms growing

larger, unprofitable firms will either drop out or be acquired, and the rest will tend to diversify, which makes the earnings structure of these firms become more similar.

Sherman and Gold (1985) emphasise the importance of examining the operating efficiency of banks instead of merely focusing on financial ratios. The operating efficiency of banks is different from profitability. Evidence shows that an increase in operating efficiency is likely to increase profitability to a larger extent. Therefore, it is of greater importance and relevance to study how efficiently banks utilise resources to produce services and obtain financial outcomes.

To sum up, the abovementioned problems of non-proportionality and inter-temporal instability are prevalent in banking data, prompting me to seek more suitable measurements such as cost efficiency to analyse the performance of banking firms to enable inter-bank and inter-temporal comparisons.

3.4 Modelling Bank Efficiency

There are two primary methodologies for efficiency estimation in the efficiency literature: parametric approach and non-parametric approach. The main non-parametric approaches are Data Envelopment Approach (DEA) and Free Disposable Hull (FDH). As mentioned before, this thesis uses a parametric approach -the Stochastic Frontier Approach (SFA) approach- to estimate cost efficiency and economies of scale. Therefore, we will not discuss DEA in this methodological review. The rationale for using the parametric approach is to investigate the impact of regional economic factors on inefficiency in a meta-frontier model. Using the DEA model would require significant algorithmic model development to achieve this, which is beyond the scope of this study as this thesis mainly focuses on empirical implications.

Within the parametric realm, current approaches can be generally classified into

two categories based on the distribution of the error component: Approaches that do not impose any assumptions on the error component and approaches that make specific assumptions and use maximum likelihood estimation on the error component.

In this section, we discuss the main parametric approaches. We give an emphasis to the stochastic frontier approach, especially the Battese and Coelli (1995) model where our stochastic meta-frontier (SMF) builds upon. In section 3.4.1, we present the two concepts of efficiency: technical efficiency and economic efficiency (cost efficiency). Section 5 presents the two categories of modelling techniques: deterministic and distribution approaches and stochastic frontier approaches. Within deterministic and distribution approaches in section 3.5.1, we present corrected ordinary least square approach, thick frontier approach and Schmidt and Sickles 1984 model; Within stochastic frontier approach in section 3.6, we review the two most widely used approaches on panel data studies: Battese and Coelli (1005) Model and Greene (2005)'s true fixed effects model. Finally, in section 3.6, we discuss input and output choices for building efficiency models based on four different bank theories: production approach, intermediation approach, value-added approach, and profit approach.

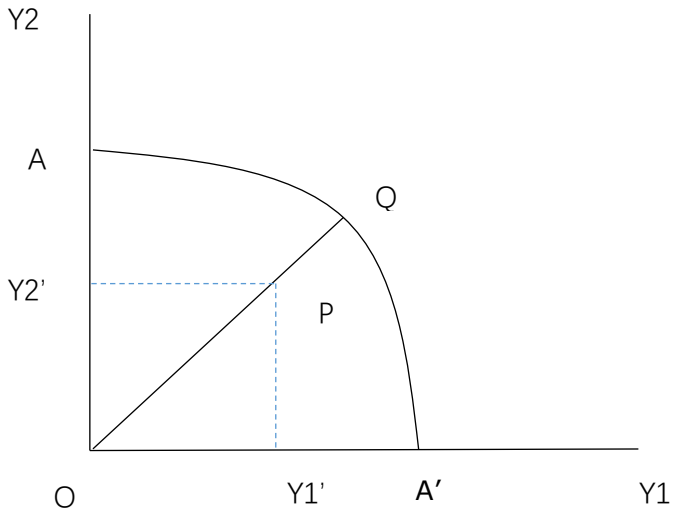
3.4.1 Different Approaches to Efficiency Modelling

Efficiency measures a producer's deviation from the optimal utilisation of resources. In other words, if a producer maximizes the outputs given a certain level of inputs, we could say that the producer is efficient. Efficiency is commonly measured as either technical efficiency or economic efficiency. Technical efficiency focuses on the production technology, while economic efficiency considers both technical efficiency and allocative efficiency.

3.4.1.a Technical Efficiency

The producer has to have a certain type of technology to transform or redistribute in order to convert the inputs into outputs. A production unit is technically efficient if the reduction in one input requires the increase in another input or reduction in at least one output, or an increase in output requires at least a decrease in one other output or an increase in one input (Koopmans, 1951). Technical efficiency can be measured from either output or input-oriented perspectives. The output-oriented technical efficiency originally comes from the production possibility frontier (See figure 3.3). The curve AA' represents the possible outputs when resources are fully utilized under the current technology, assuming there are only two outputs Y1 and Y2, in this case. Any producer producing under the curve is utilizing the resources in a sub-optimal way, therefore indicating technical inefficiency. The ratio of OP to OQ can then be used to measure technical efficiency. When P moves from O towards Q, the producer becomes more efficient, and the ratio becomes larger.

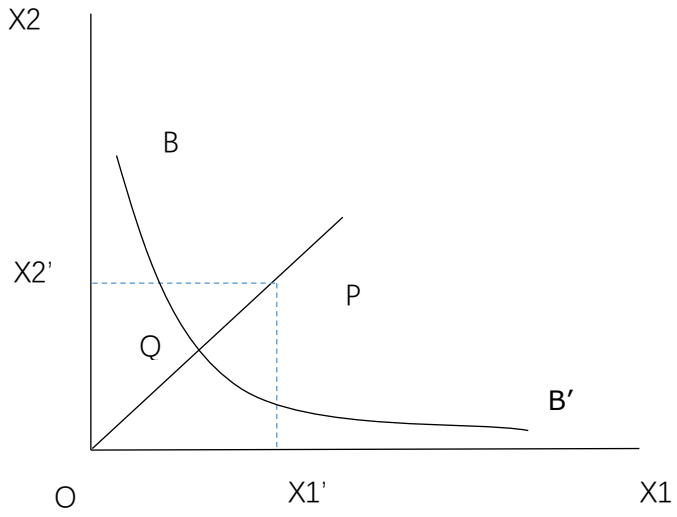
Figure 3.3. Output oriented efficiency



On the other hand, input-oriented efficiency can be demonstrated by the isoquant in figure 3.4. The curve BB' represents the combination of input X1 and X2 that

can produce the same amount of output Y. The input mix will lie on the curve BB' if the producer is fully efficient, such as point Q. The ratio of OQ to OP can be used to measure input-oriented technical efficiency. When a producer is producing at point P, more resources of X1 and X2 are used than the most efficient producer at point Q. Therefore, the less efficient the producer is, the lower the ratio of OQ to OP. Therefore, the inefficient producer at point Q can move towards the frontier by adopting a better technology such as improving production facilities so that at least one of the inputs (X1 or X2) can be reduced while maintaining the same level of output.

Figure 3.4. Input oriented efficiency



The technical efficiency can be estimated using a translog functional form. Below is an example of modelling an output-oriented production function. Let f(x) be the function of the production frontier. Then the output-oriented technical efficiency can be measured as:

$$TE(\mathrm{y}, \mathbf{x}) = \frac{y}{f(x)} \qquad (3.5)$$

Therefore,

$$y_i = f(\text{x_i}, \boldsymbol{\beta})TE_i \quad (3.6)$$

where $0 < TE(y_i, \boldsymbol{x_i}) \leq 1$ and $\boldsymbol{\beta}$ is the vector of the parameters of the production function, and *i* is the *ith* firm in the sample. The production function is usually in the log-linear form, therefore,

$$lny_i = lnf(X_i, \beta) + lnTE = Lnf(X_i, \beta) - u_i \quad (3.7)$$

Where $u_i \geq 0$ measures technical inefficiency, and $u_i = -lnTE_i$.

Therefore,

$$TE_i = \exp(-u_i) \quad (3.8)$$

3.4.1.b Cost Efficiency

Cost, standard profit and alternative profit efficiency are the fundamental concepts of economic efficiency. Cost efficiency is defined by Berger and Mester (1997) as "how close a bank's cost is to what a best practice bank's cost would be for producing the same output bundle under the same conditions. It is derived from a cost function in which variable costs depend on the prices of variable inputs, the quantities of variable outputs and any fixed inputs or outputs, environmental factors, random error, and efficiency" (p898).

Bank cost efficiency can be measured by how close the bank's cost is to the industry's best practice, which is the ratio of the minimum cost to observed cost (Kumbhakar and Lovell, 2003). The potential to lower the cost while still producing the same output is referred to as cost inefficiency or x-inefficiency.

Cost inefficiency can be seen as a product of technical inefficiency and allocative inefficiency. According to Greene (2008), allocative inefficiency comes from a suboptimal choice of input mix given prices and outputs.

In terms of advantages of cost efficiency, Berger and Mester (2003) believe that economic efficiency is a superior efficiency measure for financial institutions compared to technical efficiency and provides a better economic foundation because it is "based on economic optimization in reaction to market prices and competition, rather than being based solely on the use of technology.

Figure 3.5. Cost efficiency for two input and two output Case

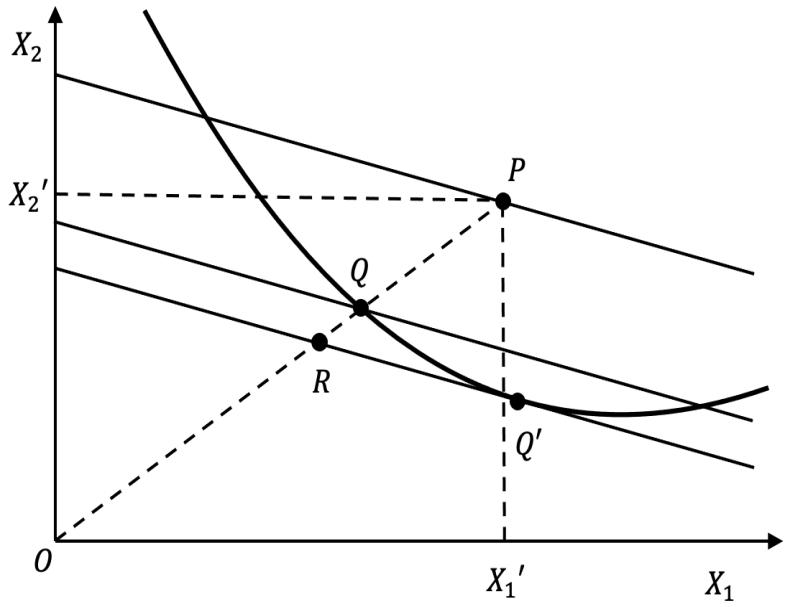


Figure 3.5 presents the two-inputs and one-output scenario for cost efficiency. A firm uses two inputs X_1 and X_2 to produce output Y. The isoquant QQ' represent the best practice in the industry. A firm at point P can minimise its cost by shifting the isocost towards the frontier until it is tangent to QQ' (see isocost RQ'). In this figure, OQ/OP is technical efficiency, and OR/OQ is allocative efficiency. The economic efficiency (cost efficiency) is $CE = OR/OP$, which is the product of technical efficiency and allocative efficiency $(OQ/OP) \times (OR/$

OQ). A fully efficient firm will have a CE value of 1, and the inefficiency represents the maximum cost a bank can reduce in order to produce the same quantity of the output mix (DeYoung, 1997, Kumbhakar and Lovell, 2003).

This thesis adopts a cost function in the estimation for efficiency in Chapter 4. Firstly, as previously mentioned, due to issues relating to accounting-based profit and expense ratios, cost frontier analysis is deemed a more appropriate performance measure to allow cross-sectional and inter-temporal comparisons. That is, for cost frontier analysis, each individual bank is compared to a hypothetical best practice bank that is simulated to be similar to be the bank it is compared to with the same output combination and under the same market conditions. This solves the inherent issues relating to accounting-based ratios where different sizes, product mix, and market conditions can directly affect banks' revenue and expenses (DeYoung, 1997). Secondly, the cost frontier paints a more realistic picture of the Chinese banking sector, where Chinese banks choose the input mix that minimises the cost of producing output. The majority of the efficiency studies on the Chinese Banking sector has adopted the cost function for this reason (Fu and Heffernan, 2007, Berger et al., 2009, Du and Girma, 2011).

3.5 Modelling Efficiency

3.5.1 Deterministic and Distribution Free Models

3.5.1.a Corrected Ordinary Least Squares (COLS) Regression Estimation

In deterministic frontier models, the deviation from the theoretical maximum is attributed solely to the inefficiency without considering any external luck or unexpected shocks.

Corrected OLS was first proposed by Winsten (1957). The deterministic frontier

production function can be formulated as

$$lny_i = lny_i^* - u_i, \qquad u_i \geq 0,$$
$$lny_i^* = f(x_i; \beta). \tag{3.9}$$

The model does not have any random error. We separate the intercept from the rest of the model, and it can be reformulated as

$$lny_i = \alpha + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} - u_i \tag{3.10}$$

Where $\boldsymbol{x_i}$ is a vector of inputs with an undefined functional form. In other words, the variables may be in log forms for a log-linear function or may include cross products for a translog function. We assume the distribution of $-u_i$ is independent of all variables in the model thus $E(-u_i|\boldsymbol{x_i}|) \leq 0$. In this setting, the constant is the only parameter of our interest.

The model can be then written as

$$lny_i = (\alpha + E[-u_i]) + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} + (-\boldsymbol{u_i} - E[-u_i]) = \alpha^* + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} + \boldsymbol{e_i} \tag{3.11}$$

Where e_i are the residuals we will estimate. However, because $E(u_i) \neq 0$, the $\hat{\alpha}$ estimated for α is biased because OLS regression will produce consistent slope coefficients $\boldsymbol{\beta}^{\mathbf{T}}$ and a biased intercept. Therefore, the only problem with the OLS estimates is the biased intercept. We can simply shift the estimated production function upward until all residuals except the one on the frontier are negative. All the OLS residuals after shifting are

$$e_{i,COLS} = e_i - maxe_i \tag{3.12}$$

The Cobb-Douglas production function can be written as

$$Y = bX_{1i}^{\beta 1} X_{2i}^{\beta 2} U_i \quad (3.13)$$

Where

$$0 < U_i \leq 1.$$

Taking logs, we can obtain

$$lnY_i = lnb + \beta_1 lnX_{1i} + \beta_2 lnX_{2i} + lnU_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} - u_i \quad (3.14)$$

Where $\alpha = lnb, x_{1i} = lnX_1, x_{2i} = lnX_2$ and $-u_i = LnU_i$. It is deterministic since the only stochastic element $-u_i$ is wholly attributed to efficiency. Therefore, to compute the parameters, we should constrain u_i to be non-negative. That is

$$lny_i - \alpha - \boldsymbol{\beta}^{\mathbf{T}} \boldsymbol{x_i} \leq \mathbf{0} \quad (3.15)$$

In terms of empirical applications, to the best of our knowledge, the majority of them are outside of banking studies. For example, Cornwell et al. (1990) implement a production function without imposing strong assumptions on the distribution using panel data of US airlines between 1970 and 1981. They use the generalised least squares (GLS) estimator, the within estimator and the efficient instrumental variables estimator and find that the first estimator generates different results than the third one, although they all indicate a significant change in efficiency over time. Tandon et al. (2000) utilise the fixed effects panel data model to estimate the frontier and overall efficiency of the health system for 191 countries. They define a set of health system goals such as responsiveness, fairness in financing and health inequality, in addition to traditional health goals and find that health expenditure is not the only way to achieve health goals and countries at all levels of health expenditures exhibit inefficiency.

The main advantage of COLS is that it does not impose any specific assumption on the distribution of inefficiency, and it is easy to compute. The major disadvantage is that the frontier function is assumed to be deterministic.

Therefore, the deviation from the frontier is assigned to inefficiency without considering any random shocks. Moreover, The COLS model is also very sensitive to outliers (Kumbhakar et al., 2015). As a result, the inefficiency estimated using COLS may have an upward bias, meaning that the inefficiency can be overestimated.

3.5.1.b TFA (Thick Frontier Approach)

TFA (Thick Frontier Approach) was proposed by Berger and Humphrey (1991). It is also one of the distribution-free approaches since it does not impose any assumption regarding the distribution of the error term. The sample is divided into N quantiles according to a predetermined efficiency indicator such as average output. Therefore, the firms with the largest average output are assumed to be the most efficient. Then the efficiency frontier will be estimated using the sample from the most and least efficient groups, respectively. The difference evaluated at the mean values between the groups is considered only a result of market factors and inefficiency. The comparison of the mean values is considered to eliminate the influence of noise in the data. The advantage of the TFA approach is that it allows for an error term within each quantile and the between-quantile difference is solely due to market factor and inefficiency. The drawback of this approach is that the criteria for stratifying the sample are not unique, especially when there are multiple inputs and outputs. Furthermore, if the sample is too small, there would be a shortage of observations in each quantile.

3.5.1.c Schmidt and Sickles (1984) Distribution Free Approach

To capture bank-specific heterogeneity, fixed effects and random-effects models can be estimated via:

$$y_{it} = f(\boldsymbol{x_{it}};\ \boldsymbol{\beta}) + \varepsilon_{it} = \beta_0 + \boldsymbol{x_{it}}\,\boldsymbol{\beta} + \varepsilon_{it} \qquad (3.16)$$

$$\varepsilon_{it} = v_{it} - u_i \ u_i \geq 0, i = 1, \ldots, N; t = 1, \ldots, T$$

Because u_i does not change over time, it can capture this feature of the panel dataset. As there is no assumption in terms of the distribution of u_i, these models are considered as distribution-free approaches (Schmidt and Sickles, 1984).

The fixed-effect model can be written as

$$y_{it} = \alpha_i + \boldsymbol{x_{it}}\,\boldsymbol{\beta} + v_{it} \tag{3.17}$$

Where

$$\alpha_i \equiv \beta_0 - u_i \tag{3.18}$$

As for the fixed-effect model, the advantage is that it allows u_i to be freely correlated with $\boldsymbol{x_{it}}$. This feature is especially useful when we assume inefficiency is correlated with the input choices. There are some obvious disadvantages as well. One is that these models do not let the inefficiency and individual heterogeneity be separated. The other problem is that it is impossible to include other time-invariant variables in the model since these variables and the α_i together will cause perfect multicollinearity. This model is usually estimated by the fixed effect model using within estimators (Kumbhakar et al., 2015). Because the number of dummies is equal to the number of cross-sections, if the number of cross-sections is large, it will cause difficulties in the calculation. This difficulty can be overcome by a first-difference or a within transformation. Then the estimated $\widehat{\alpha_i}$ will be adjusted to obtain the estimated $\widehat{u_i}$.

The problem with the time-invariant models is that they assume that the management of firms does not learn over time, and thus the efficiency of a firm does not change over time, which can be unrealistic, especially within a highly competitive market such as the banking industry[3].

[3] There are time-varying models in the literature that accommodate the change in productivity and efficiency. Similarly to time-invariant models, time-varying models can be either fixed-effects or random-effects models. In the fixed-effects model, the time-varying efficiency term is non-stochastic, and it is just a function of time. In the time-varying random

The main advantages of distribution-free approaches are that the estimation does not depend on the distributional assumptions of the inefficiency term u_i and they are easy to estimate. On the other hand, the disadvantage is that there is a lack of information in terms of the statistical properties of the estimators of u_i. To distinguish inefficiency effects from random errors, we need to add a random error term and impose a distributional assumption for both inefficiency and statistical errors, the parameters of which can be estimated using maximum likelihood estimators. Section 3.6 will discuss this method in detail.

3.6 SFA (Stochastic Frontier Approach)

The Stochastic Frontier Approach is proposed by Aigner et al. (1977) and Meeusen and van Den Broeck (1977). Unlike distribution-free approaches, SFA imposes specific assumptions in terms of the distribution of error terms. This approach is also called the parametric approach. The idea is that the deviation from the frontier is not solely due to the inefficiency of firms. There are other factors such as luck and unexpected events that can affect this deviation. Moreover, the measurement errors as well as the model misspecification can also otherwise lead to change in efficiency levels. It assumes that each firm has its own frontier and each frontier is placed under the influence of all the stochastic disturbances. The stochastic frontier can be formulated as

$$y_i = f(\boldsymbol{x_i})TE_i e^{v_i} \tag{3.19}$$

Taking logs for both sides and reformulating the above equation, we can obtain

$$lny_i = \alpha + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} + v_i - u_i = \alpha + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} + \boldsymbol{\varepsilon_i} \tag{3.20}$$

Where u_i is the positive error and v_i is the unrestricted error. $\boldsymbol{\beta}^{\mathbf{T}}$ are the slope

effect models, the inefficiency is either a random term or a composition of a time-invariant stochastic term and a time-varying deterministic term (Kumbhakar et al., 2015). There are a few models that suit this purpose such as Cornwell Schmidt and Sickles model (Lee and Schmidt, 1993, Cornwell et al., 1990).

coefficients. The statistical representation is that

$$u_i \sim i.i.d.N^{+}(0, \sigma_u^2)$$
$$v_i \sim i.i.d.N(0, \sigma_v^2) \quad (3.21)$$

In this example, the distribution of inefficiency term u_i is non-negative, following a half-normal distribution. Therefore the $\alpha + \boldsymbol{\beta}^{\mathbf{T}}\boldsymbol{x_i} + v_i$ represents the stochastic frontier and $-u_i$ is the inefficiency term.

There also exist other distributions, such as truncated normal distribution and exponential distribution, based on assumptions of the inefficiency term. Stevenson (1980) proposes a truncated normal distribution that allows the distribution of the inefficiency to have a non-zero mode, under the assumption that there may be some degree of inherent inefficiency within the industry. As the truncated model has one more parameter to estimate, it adds significant difficulty in the maximum likelihood estimation. Meeusen and van Den Broeck (1977) propose an exponential distribution for inefficiency u_i , the density function of which can be expressed as $f(u_i) = \left(\frac{1}{\lambda}\right) * \exp\left(-\frac{u_i}{\lambda}\right)$, where $u_i \geq 0$ and $\lambda \geq 0$. Similar to half-normal distribution, it has mode of 0, mean equal to $\lambda \geq 0$ and standard deviation of λ^2.

3.6.1 Battese and Coelli (1995) Model

Before going into the details of the Battese and Coelli (1995) model, we firstly discuss existing techniques to measure the impact of exogenous variables on efficiency. There are mainly three ways to examine exogenous factors that drive efficiency. The first approach is the two-stage approach, and the second and third approaches are both one-stage approaches, with different assumptions on how exogenous variables interact with the models.

The two-stage approach is fairly straightforward. The first stage is to estimate

efficiency using one of the models mentioned above. The second stage is simply regressing the factor which could potentially influence efficiency on the efficiency term obtained from the first stage. The second stage regression can be shown as below:

$$EFF_i = \sigma_o + \sum_n \sigma_n z_n + \epsilon_i \quad (3.22)$$

Where z_n are the environmental variables that, in theory, can influence efficiency. We can use an OLS or a Tobit model to conduct the above regression. The results from the two-stage model are considered to be biased because if the efficiency term is assumed to be identically and independently distributed in the SFA model, which is the case in most of the SFA models, then if the second stage regression proves that the efficiency is dependent, the assumptions of the first stage estimation are violated (Wang and Schmidt, 2002).

Within the literature where the one-step approach is used, there are mainly two ways to include the exogenous effects. The first group of studies include exogenous variables as regressors in the main equation for frontier estimation. These studies assume that exogenous variables affect the shape of the frontier. Such studies include Good et al. (2013) on European and U.S. air carriers, Abdul-Majid et al. (2011) on Malaysian commercial banking industry, Berger and Mester (2001) on US banks and Goddard et al. (2014) on banks in Latin America.

The second group of studies assume that, instead of affecting the shape of the frontier, exogenous variables directly affect the magnitude of the efficiency effects only. This is first modelled by Battese and Coelli (1995) and tested on the data from Indian paddy farmers for a ten year period from 1975-76 to 1984-85. The parameters of the stochastic frontier and the model for the technical inefficiency are estimated simultaneously. The model assumes that exogenous variables have a linear relationship with the mean of the distribution of inefficiency effects u_{it}. The mean of the distribution can be expressed as m_{it} of

a function of $\boldsymbol{z_{it}}$

$$m_{it} = \boldsymbol{z_{it}}\boldsymbol{\theta} + W_{it}. \tag{3.23}$$

We can see from the subscripts that the mean of the distribution is observation specific. That is, for each observation, it has its own distribution, thus its own expected value. Where $\boldsymbol{z_{it}}$ are the set of exogenous variables that may affect the inefficiency term u_{it}. W_{it} is a random variable that has the distribution of $N^{-z_{it}\theta}(0, \sigma^2)$. The truncation point is $-z_{it}\theta$, so that u_{it} has the distribution of $N^{+}(z_{it}\theta, \sigma^2)$. That is, a positive truncation at 0 with a normal distribution with mean $z_{it}\theta$ and variance σ^2. This model is commonly called BC(95) model. $\boldsymbol{\theta}$ is a vector of parameters to be estimated. One can look at $\gamma = \frac{\sigma_u^2}{\sigma_u^2+\sigma_u^2}$ to see whether there is a sign of model misspecification, although this is not conclusive. Nevertheless, a value of γ close to 0 indicates that the model is close to purely stochastic, and a value of 1 indicates a strictly deterministic model.

This model has been widely used in the efficiency literature, which incorporates exogenous variables, such as Lozano-Vivas et al. (2002) on banks in 84 countries, Kasman and Yildirim (2006) on new EU countries during the transition and Xiang et al. (2013) on a comparison studies on Australian, Canadian and UK banks.

As discussed in chapter 2, as regional disparity being our main research interest in terms of the efficiency estimation process, we adopt a SMF approach to estimate efficiency scores in chapter 4. The Battese and Coelli (1995) model is the foundation that our SMF builds upon[4].

[4] We do not use Greene (2005)'s true fixed effects model as this model has significantly more parameters to be estimated, and our data being unbalanced, making convergence in the maximum likelihood iteration process a major difficulty.

3.6.2 Greene (2005)'s True Fixed Effects Model (TFE)

Although the Battese and Coelli (1995) model has been used widely in empirical studies based on panel datasets, it does not separate individual effects from inefficiency. This problem is solved by Greene (2005) with a model formulated as

$$y_{it} = \alpha_i + \boldsymbol{x_{it}}\,\boldsymbol{\beta} + v_{it} \mp u_{it} \qquad (3.24)$$

where $\boldsymbol{x_{it}}$ is a vector of inputs and other explanatory variables, and $\boldsymbol{\beta}$ is a vector of unknown parameters to be estimated. α_i are the fixed parameters that represent individual fixed effects that are not part of inefficiency. This model is also called the "true fixed-effects" stochastic frontier model. The only difference between the "true fixed-effects" and the standard fixed effects model is that Greene (2005)'s model contains one additional time-varying efficiency term $\mp u_{it}$, where the sign of u_{it} depends on if it is a cost (+) or a production/profit (-) frontier. α_i are fixed parameters representing unobserved individual effects that do not contain inefficiency. When α_i are treated as random variables, the above model becomes the "true random-effects" (TRE) stochastic frontier model. The sign of u_{it} depends on whether we are estimating a production/profit function (-) or cost function(+). Different distributional assumptions can be imposed on the inefficiency term u_{it} for both TFE and TRE, including half-normal, truncated normal and exponential distributions. Similar to Battese and Coelli (1995) Model, the inefficiency term can be expressed as a function of a list of exogenous variables.

There are several limitations of Greene (2005) 's TFE model. The first problem is the vast number of parameters to be estimated, especially when the number of firms in the sample is large. The second issue is the incidental parameters problem resulting from fixed effects α_i in the frontier model. The incidental

parameters problem arises when the number of parameters increases with the number of cross-sections, rendering the estimates of vector β to be biased and inconsistent. Greene (2005) find that in a sample with small and relatively large group sizes, there indeed exist evident biases both in estimates of coefficients and inefficiency. However, when T is large, there seems to be less bias in the parameter estimates. He also finds that the random-effects model produces more distorted results, and it is less consistent, especially with the assumption of efficiency being independent of explanatory variables. To resolve the incidental parameters problem, Wang and Ho (2010) updates Greene(2005)'s models through within and first-difference transformation of the model while the closed-form likelihood function is maintained. As α_i s are dropped from the model, consistent estimates can be obtained, meaning that the incidental parameter problem would be eliminated. As the number of parameters needing estimation is reduced, the difficulty of estimating too many parameters is also reduced.

Another limitation regarding Greene(2005)'s fixed effects models relates to their inability to separate persistent inefficiency from unobserved heterogeneity. Although Greene (2005) 's TFE and TRE models can separate individual heterogeneity from time-varying inefficiency, they do not consider persistent technical inefficiency, which can represent effects of inputs such as management which does not vary over time, unless there is a change in policy or ownership (Kumbhakar et al., 2015). This problem is solved by the more recent development of models that overcome the limitation of earlier models by separating firm effects, persistent inefficiency and time-varying inefficiency. For example, Kumbhakar et al. (2014) split error term into four components: firm-specific heterogeneity, short-run time-varying efficiency, persistent time-invariant inefficiency and random shocks.[5]

[5] We do not go into details of the more recent models that separate persistent inefficiency from time-varying inefficiency, such as Kumbhakar et al. (2014), as these models are outside the scope of this Thesis.

3.6.3 A summary of efficiency models

Empirical studies using models such as the Battese and Coelli (1995) model often simply pool the panel data and treat them as cross-sectional. That is, the same units observed at different points in time are treated as separate units. However, these models do not take into account the time-invariant and individual-specific characteristics contained in panel data. Among the abovementioned models, including the distribution-free models and stochastic models, only Schmidt and Sickles (1984) and Greene (2005) 's True Fixed Effects Model (TFE) account for the time-invariant individual heterogeneity within the panel data. Schmidt and Sickles (1984) distribution-free model assumes inefficiency to be individual-specific and time-invariant. However, the unobserved heterogeneity is contained within the inefficiency term, meaning that the model cannot distinguish inefficiency effects from unobserved time-invariant heterogeneity. Therefore, the inefficiency estimated using the time-invariant inefficiency models such as Schmidt and Sickles (1984) might partially and entirely be attributable to individual heterogeneity (Greene, 2005). Another drawback of time-invariant inefficiency models is that it might not be reasonable to assume that the inefficiency of a firm does not over time, especially when the frontier is estimated over an extended period. In contrast, Greene (2005) 's TFE model allows for the separation of individual effects from inefficiency. However, it suffers from the drawbacks of estimation difficulty and incidental parameter problems. Furthermore, it is unable to separate persistent inefficiency from individual heterogeneity. More recent models, such as Kumbhakar et al. (2014)'s four-error-component model, are aimed to tackle these issues.

3.7 Literature Review on the Theories of the Banking Firm

There are different inputs and outputs choices to model bank operations. This is mainly due to the different objectives in different banking industries and the evolution of the role of banking firms over time.

There are four main approaches to model banking firms in the literature: the production approach, the intermediation approach, the value-added approach, and the profit approach. Here we will review these models of the banking firm in chronological order to illustrate how changes of bank models correspond to and adapt to changes in bank functions.

3.7.1 Production Approach (Banks as Producers of Loans)

The production approach treats banks as firms that use capital and labour to produce different kinds of deposits and loan accounts (figure 3.5). In this approach, a '***flow'*** measurement – the number of deposits and loan transactions per unit of time, rather than the financial values, i.e. a '***stock***' value of deposits and loans – is used to measure outputs.

According to Benston (1972), measuring inputs and outputs of banking firms is less problematic than other types of firms. This is mainly due to the homogeneity of the inputs and outputs of financial institutions. When measuring costs, the problems of externalities (social costs) and a poor reflection of economic cost (or opportunity cost) on accounting figures are less severe in financial institutions compared to non-financial firms.

In terms of what to include and what to exclude in the cost of production, Benston (1972) argues that the majority of costs of financial institutions come from salaries and supplies; In addition, the opportunity cost of different reserve requirements and varying taxation on income should also be excluded because these are not from operations. He also emphasises that annual data or changes in data should be used to avoid the problem of mismatch between expenses and outputs.

In terms of the output of financial institutions, Benston (1972) argues that the

number of deposit and loan accounts made is the best measurement for bank outputs. He explains that this is because, from a bank production point of view, it is most closely associated with bank operations. By using this measurement, bank outputs can be directly related to operating costs. Thus the marginal cost for a particular type of output can be calculated, which is beneficial for bank managers in decision making.

Benston and Smith (1976) provide a transaction cost approach to understand the production of financial intermediation. They view financial institutions as firms that use inputs such as capital and labour to produce specialised financial commodities, and they use the revenues obtained from selling these commodities to cover all the costs. The existence of financial intermediaries is to reduce transaction costs. They emphasize that the main reasons for capital and labour costs in financial institutions are documentation, information and monitoring.

Figure 3.5. Production Approach

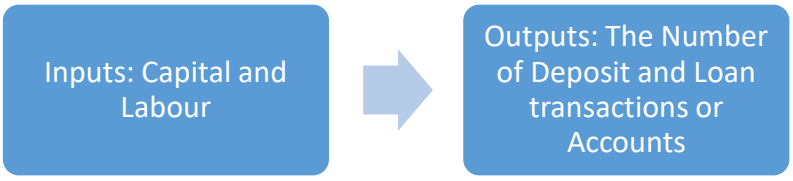


Figure 3.5 presents the inputs and outputs choices for a production approach, where banks are firms that use capital and labour to produce different quantities of deposits and loans.

3.7.2 Intermediation Approach (Banks as Intermediaries)

As opposed to the production approach in which outputs are treated as a '*flow*', the intermediation approach regards outputs as a '***stock***', which is measured by the value of loans and investment (Heffernan, 2005). Costs include both operating costs and interest costs. The intermediation approach sees bank production as a multistage process, where loanable funds act as an intermediate product that serves as an output from labour and capital and an input for

producing earning assets.

Figure 3.6. The Intermediation Approach

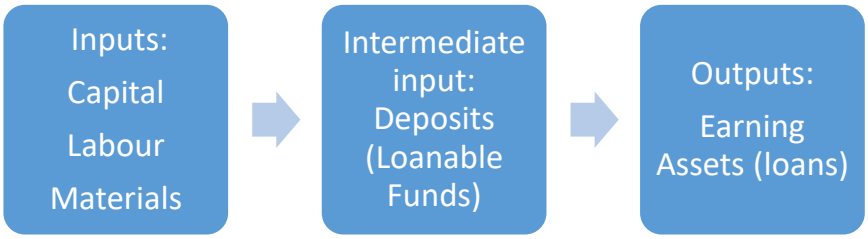


Figure 3.6 presents an intermediation approach for banking where deposits are seen as an output from labour and capital and an input for other earnings assets such as loans.

Sealey and Lindley (1977) apply the theory of firms to financial institutions and propose that financial firms act as agents for funds transformation, from surplus spending units to deficit spending units, in order to achieve financial intermediation. They argue that economic outputs rather than technical outputs should be used because economic outputs are more closely related to financial firms' decision making. That is, only the services provided by financial firms that are valued higher than the original inputs can satisfy the requirement for outputs. Thus, earning assets should be treated as outputs, and deposits, which incur positive costs without yielding any direct revenue, should be treated as inputs. The labour, capital and material costs together with interests paid by financial firms to provide services for depositors are a way to pay for the use of funds. Thus, these services act as payment for the inputs. They also argue that if our purpose is to apply the general theory of firms to financial firms, banks should be treated as profit maximisers rather than rational investors. They argue that although demand deposits services provide benefits to society, from a profit-maximising point of view, these social benefits are not banks' concern. Only the market valuation of output should be considered as output, which should be measured by earning assets.

3.7.3 Value-added Approach (Banks as Value-Adders)

Under the value-added approach, deposits are treated as a vital part of bank outputs. Unlike the intermediation approach, for which the deposits are seen as an intermediate stage product used to produce final output (loans), the value-added approach incorporates both input and output functions of deposits (Humphrey, 1992). That is, the cost of obtaining deposits is considered as an input, while the safe-keeping service that the bank provides is considered an output.

More specifically, in terms of inputs, Humphrey (1992) argues that if only labour and capital are included as costs, it would ignore banks' costs in raising funds to support their operations. This is because large banks often use purchased funds to complement self-raised deposits. Hence, if only labour and capital costs are included, and funding costs are ignored, the model would omit an important cost required to provide standard banking services. That is, if inputs are measured as a proportion of labour and capital cost as part of the total cost, then the dynamics between numerator (labour/capital cost) and denominator (total cost) will be less reliable and less stable, as this input measure can be significantly impacted by the cost of funding (Sudarsanam and Taffler, 1995, Humphrey, 1990). Thus, total costs, including both operating expenses and interest expenses, are a more comprehensive representation of inputs.

In terms of outputs, because the production of deposit services accounts for nearly half of the capital and labour inputs, only including loans and other earning assets as outputs seems to be incomplete. Humphrey (1992) developed a total factor model which includes both loans and deposits in the production function of the bank. He uses both stock and flow measurements for deposits and finds that the results differ between these two measurements. The model can be summarized in figure 3.7 below.

Figure 3.7. Value-Added Approach source: Humphrey (1992)

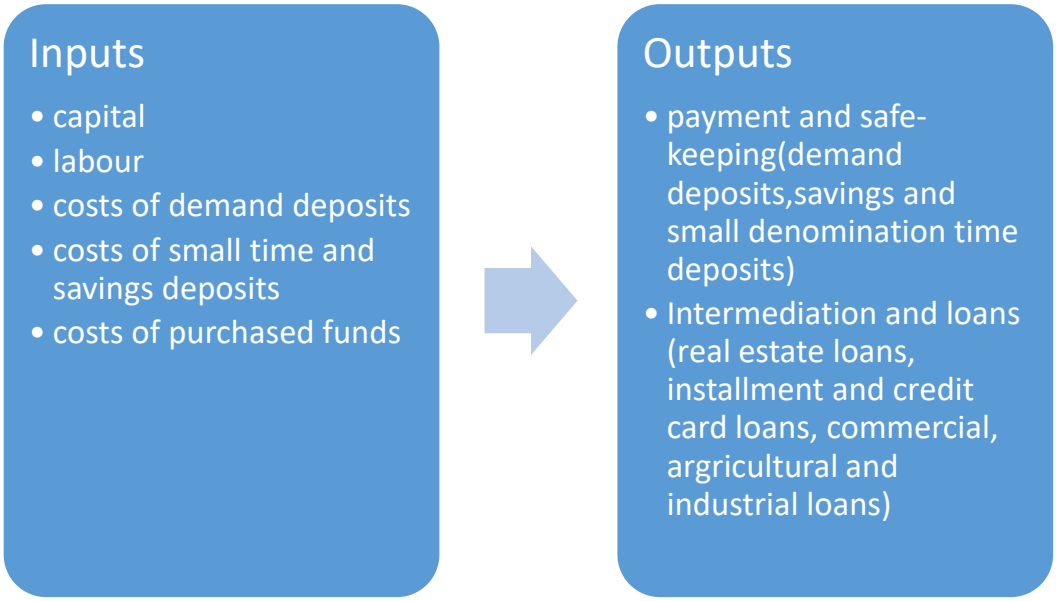


Figure 3.7 presents a value-added approach where banks incur costs such as capital and labour costs as well as the cost incurred in obtaining funding to provide services to society, such as payments and safekeeping and intermediation and lending. The costs of obtaining deposits and other funding are regarded as inputs, while the amount of the book-keeping services relating to deposits are outputs.

3.7.4 Profit Approach (Banks as Profit Maximisers)

The profit approach views banks as profit maximisers rather than merely cost-minimizers. Mullineaux (1978) argues that regulators should not only be concerned with technical efficiency but also allocative efficiency. Because if a bank is profit-inefficient, it means that the bank failed to utilise the resources to their highest value. The theory of profit function was originally developed by MacFadden (1966). It assumes banks are price takers and allows consideration for economic efficiency, which can be decomposed into technical and allocative efficiency. Mullineaux (1978) applies this profit function to the banking industry. He defines profit function as the expression of '*maximized* profit as a function of *prices* of output, variable factor inputs and *quantities* of fixed factors in a competitive market' (Mullineaux, 1978: 260). The profit efficiency increases when output prices increase, input prices decrease, and the quantities of fixed factors increase. A bank is considered more technically efficient if it constantly produces more output than other banks given the same amount of

input. Likewise, a bank is considered more profit-efficient when it maximizes output prices and minimizes input prices and quantities. The main benefit of using the profit approach as opposed to the production approach is that the level of output is not included in the profit function. Since the production and cost function estimates are sensitive to the definition of output, the profit function can avoid this problem. The profit approach can be summarized as:

$$\pi = R'Y - P'X \tag{3.26}$$

Subject to a technology constraint

$$t(Y, X, F) = 0, \tag{3.27}$$

where Y is a vector of bank output including a variety of loans and services, and X is a vector of variable inputs such as labour, materials and deposits. R' is a vector of exogenous prices for bank loans and services, and P' is a vector of given input prices and interest rates. F is the level of fixed input that remains the same during the production period.

Berger and Mester (2003) also point out the main advantages of the profit approach. They argue that the standard cost function can be misleading when examining the performance of banks. In their study of US banks during the period of 1991-1997, they find that their cost efficiency decreased while profit efficiency increased. This result proves the hypothesis that during the era of technological change and an increasing price for raising funds and deposits, the costs of inputs are likely to increase significantly. However, banks use these additional costs to improve the quality and increase the variety of services, which in turn makes banks generate more revenue. Thus, profit efficiency increases despite the decrease in cost efficiency.

As mentioned before, unlike the cost minimization approach, which focuses on minimising variable costs subject to given prices of inputs, quantities of output and environmental factors, the profit approach focuses on profit maximization. It is regarded as superior to the cost minimization approach because it suits the managerial goals, which take both revenue and costs into account. Thus, if a managerial strategy will incur additional costs but will increase revenue to a larger extent, this strategy should be considered appropriate since it will improve overall profit (Berger and Mester, 2003). The standard profit maximization model can be specified as:

$$\ln(\pi + \theta) = f_{\pi}(X_{\pi}) + lnu_{\pi} + ln\varepsilon_{\pi} \qquad (3.28)$$

In this model, output prices will replace output quantities in the cost function, with all other business conditions remaining the same. Therefore, output prices are considered a major factor that influences the choice of outputs when maximizing profits.

The alternative profit maximization model specifies logged output quantities in the X vector, rather than in logged output prices *lnp*:

$$\ln(\pi + \theta) = f_{a\pi}(X_{C}) + lnu_{a\pi} + ln\varepsilon_{a\pi} \qquad (3.29)$$

The profit approach implies that when we measure the productivity and efficiency of a banking system, we should always choose the most appropriate model. For example, mutual or thrift banks might focus more on cost-minimizing rather than profit-maximizing because their primary function is to provide services rather than to maximize profits. State-owned banks may also have different business models than private banks.

3.5 Conclusion

In this chapter, we have explored issues of non-proportionality and inter-temporal instability associated with accounting-based financial ratios and provided a justification for using cost efficiencies obtained from a cost frontier function as a preferred measure for bank performance. We then reviewed the concepts of technical efficiency and cost efficiency and reviewed both deterministic and distribution-free approaches and stochastic frontier approaches. We examined two commonly used stochastic frontier models: The Battese and Coelli (1995) model and Greene (2005)'s true fixed effects model. We justified our choice of Battese and Coelli (1995) to be more suitable for our efficiency estimation due to its ability to allow us to specify the functional form of the inefficiency term as well as higher data requirements needed for Greene (2005)'s model in terms of convergence. Finally, we reviewed the four approaches to the choices of inputs and outputs. We choose the intermediation approach as it is most suitable in terms of our data availability, and previous efficiency studies on Chinese banks have chosen mainly the intermediation approach (Berger et al., 2009, Du and Girma, 2011).

In the next chapter, we will adopt a stochastic meta-frontier (SMF) approach, which is built upon the Battese and Coelli (1995) model, and an intermediation approach to inputs and outputs choices is implemented.

Chapter 4. Environmental Impact on Cost Efficiency of Chinese Regional Commercial Banks

4.1 Chapter Summary

This chapter is arranged as follows: Section 4.2 is the introduction. In section 4.3, we present the motivation of the research and section 4.4 gives a brief review of studies on the regional banking sector, especially in China. Section 4.5 introduces the main methodologies for the entire chapter, including the methodology for cluster analysis and Stochastic Meta-frontier (SMF) approach. Section 4.6 demonstrates the implementation of these methodologies in this study, including model specifications. Section 4.7 offers the data description. In section 4.8, we present the result analysis from cluster analysis, and section 4.9 presents the results from estimations using a pooled frontier, cluster-individual frontiers and the meta-frontier. Cost efficiencies and cost gap ratios are also analysed. Section 4.10 offers results from economies of scale estimation from both individual frontiers and the meta-frontier and regression results on the determinants of economies of scale. Finally, 4.11 gives the conclusion of the chapter and policy implications.

4.2 Introduction

This chapter investigates regional differences in cost efficiencies in the Chinese commercial banking system. Previous studies which have examined the efficiency of Chinese banks have predominantly viewed and studied the Chinese commercial banking sector as a whole (Fu and Heffernan, 2007, Kumbhakar and Wang, 2007, Berger et al., 2009, Jiang et al., 2009, Du and Girma, 2011). That is, the commercial banks in China have been modelled using one single frontier, under the assumption that they are operating under a homogeneous technology. This study shows strong evidence which is against this assumption. In other words, we find that due to differences in geographical, demographical, economic, and financial development, regional commercial banks in China operate under different technologies and estimating efficiencies of them using one single frontier may be inaccurate. To the best of our knowledge, this study is the first one in the literature which employs the stochastic meta-frontier (SMF)

technique in modelling regional disparities in the commercial banking industry in mainland China.

Firstly, we adopt a new algorithm - a two-stage hybrid hierarchical k-means clustering method. This method is considered more accurate than the traditional hierarchical or k-means clustering as our method maximises the advantage of both by forcing starting centres of second-stage k-means clustering to be determined by the results obtained from first-stage hierarchical clustering (Chen et al., 2005). The second superiority of this algorithm is that we determine the optimal number of clusters in an ad hoc manner by using the sum of squares of final cluster results instead of deciding on the optimal clusters based on partial results from first-stage hierarchical clustering.

We then use the results from cluster analysis to estimate a stochastic meta-frontier (SMF) approach developed by Huang et al. (2014). Moreover, this study also investigates a set of microeconomic factors and macroeconomic factors which potentially influence individual cost efficiencies as well as technology gap ratios (TGRs) or cost gap ratios (CGRs). We find that, for all three clusters, the primary source of cost inefficiency comes from a substantial technological gap between the cluster-specific frontier and country meta-frontier. We find that banks in different regions defined by the cluster analysis are affected by different microeconomic variables. Cost efficiency in cluster 1 (the Inland Region) is found to be positively affected by higher capital requirements and lower levels of non-performing loans. In terms of macroeconomic conditions, we find that the deviation from the best available technology is mainly affected by the inferior economic and industrial conditions, including foreign bank presence, banking industry concentration and competition, as well as urbanization. More specifically, banks can narrow down technology gaps through increased foreign bank presence, increased banking competition and reduced concentration, as well as acceleration in regional urbanization. Banks in cluster 3 - the Northeastern and Coastal Region - appear to be the most cost-efficient, and banks in cluster 1 - the Inland Region - are found to be the least cost-efficient

compared to meta-frontier. In terms of the Inland Region, it is found that the deviation from meta-frontier mainly comes from a low urbanization rate, while the technology gap between cluster 2 - the Eastern Municipalities - and meta-frontier mainly comes from an overly concentrated regional banking sector. Although there exhibits managerial inefficiency in the banks in all regions, the technology gaps provoked by the regional disparity in economic conditions seem to be an overriding factor that diverges them from the best practice. Thus, this study gives an excellent insight into how banks in different regions perform and gives vital information for government and regulatory bodies as well as banks themselves to make region-tailored strategies to improve banking efficiency.

Lastly, we estimate economies of scale (ES) from both individual frontiers and the meta-frontier and examine how the adoption of meta-frontier technology would affect scale economies in each region. We then further investigate the determinants of economies of scale, including a list of bank-specific characteristics such as investment banking activities, liquidity, credit risk, capital, and insolvency risk.

4.3 Motivation

Existing literature primarily estimates Chinese commercial banks under one frontier. A table of literature review on a list of selected efficiency studies of Chinese commercial banks is given in appendix 1.

The geography of banking systems has been an interesting yet under-researched area in the literature. The studies which examine the impact of regional economic conditions on bank performance are focused on the US banks (Meyer and Yeager, 2001, Daly et al., 2003, Emmons et al., 2004, Furlong and Krainer, 2007), and they find mixed results in terms of whether regional risk is an important factor for bank risk. The focus on the US banking sector may be due to the public availability of branch-level data via central bank websites.

Regional disparity in economic development within China has been confirmed in a wide range of literature(Litwack and Qian, 1998, Sun and Yamori, 2009, Sun et al., 2013). However, in the literature where the performance of Chinese banks is examined, banks are always treated as if they operate under the same economic environment (Berger et al., 2009, Du and Girma, 2011, Fu and Heffernan, 2007, Jiang et al., 2009), even though theory and evidence suggest that financial development across regions, especially for a country with a vast area and economic disparity, are never 'space-neutral' (Crocco et al., 2010, Klagge and Martin, 2005). That is, financial development across different regions heavily relies on geographical, economic, social and political factors.

Therefore, it is of very much interest as well as necessity to re-examine the performance of Chinese commercial banks from a regional perspective. More importantly, we aim to use a more appropriate model, i.e. meta-frontier model, to avoid false assumptions of a homogenous environment for all banks. Finally, it would be of interest to find out how Chinese banks utilise regional economic development and how they react to regional economic risks, compared to the results from comparable studies in the US and Brazil, especially when Chinese banks' lending behaviour is influenced by government decisions.

4.4 A Review on the Regional Studies of Commercial Banking Systems

There have been established studies that discuss the geographic disparity of financial development in a country. Klagge and Martin (2005) provide theoretical and empirical evidence that financial development does not operate in a 'space-neutral' way in that financial markets in a country are usually not fully integrated and the demand and supply of funding locally and their proximity to financial centres highly restrict the health and growth of regional financial systems. Alessandrini et al. (2009) also recognise the existence of

regional economic disparities and a potential development gap between global banking markets and a local banking environment. Thus, the efficiency of standardised financial products may be constrained by their local economic situations. They propose that banking products and services need to be strategized locally in order to meet the need for their specific requirements for local development. Dow and Montagnoli (2007) find that the effects of monetary policy differ across different regions in the UK, and its efficacy lessens in the peripheral regions due to them lagging behind in the economic cycle. Ashton and Gregoriou (2014) further confirm that regional banks in the UK react to monetary policy differently. They find that regionally and nationally branching banks in the UK realize deposit interest rates differently. Therefore, they suggest that policymakers should take regional development requirements into account when setting monetary policy.

Crocco et al. (2010) examine the regional structure of Brazil's financial system. They argue that regional development affects financial development and bank operations in many ways. Especially, regional economic conditions and financial development have a reinforcing effect on each other. Since peripheral regions present higher requirements for liquidity in both banks and consumers, due to limited investment opportunities and higher credit risk due to disadvantaged economy, they tend to charge higher interest rates and curb the maturity as well as the volume of loans issued, affecting banks' loan and deposit generating ability directly. From their argument, we should expect a disadvantage of efficiency in terms of generating loans and deposits for peripheral banks, holding other factors equal. Therefore, peripheral regions become less and less attractive for banks to establish branches or headquarters, which means that the financial system of a country is usually highly concentrated in the centre areas, leading to what Crocco et al. (2010) call a 'core-periphery' structure. In the case of China, Beijing as a political Centre and Shanghai as a financial centre can be regarded as the central part of the 'core-periphery' structure compared to the rest of the regions. Indeed, this 'core-periphery' structure can be hierarchical, which means that regions where the centres are

situated, such as eastern areas where Beijing and Shanghai are located, can be regarded as the 'core' areas compared to the western part of China. Crocco et al. (2010) define a region as central by population density, level of wages and diversity of the secondary and tertiary sector. They divide Brazil into four regions (northeast, southeast, south, centre-west) and by using a spatial panel model, they find that municipalities with higher level of financial development, measured by tax paid by bank branches and credit supply, tend to be surrounded by municipalities where banks have a low level of financial activity. They also find the number of employees, average wage and number of bank branches to be good predictors of financial development.

There are also several studies that use US data to examine whether regional economic shocks have an impact on regional bank performance. The results are mixed, however. Studies such as Daly et al. (2003) and Furlong and Krainer (2007) find evidence of systematic effects of economic conditions on bank performance, while Meyer and Yeager (2001) find that regional economic conditions are not important risk factors for bank performance. Emmons et al. (2004) also find that regional risks for small community banks are not so important as these banks can always diversify their risks away through portfolio management.

Although there are studies that noticed the disparity of regional development of the Chinese economy and financial development (Litwack and Qian, 1998, Sun and Yamori, 2009), there have been very few studies on the regional economic condition with banking sector specifically. Within the limited literature, we only review the most relevant ones as below. Yang et al. (2015) demonstrate that socio-economic and demographic features greatly impact the performance of bank branches. They use a spatial econometric model to analyse and compare local branch performances of the world's biggest bank – Industrial and Commercial Bank of China (ICBC) from 2000-2010 across six different regions. The authors follow the value-added approach and proxy bank performance using the deposit quantity, and they find that ICBC relied on the Yangtze River Delta

and Bohai Sea region to generate the majority of the output while central regions in China were under-performed. They also find that the population of workers in the region is negatively related to the deposit performance of bank branches. They interpret these results as an indication that bank branches do not rely on the local population but depend on transferring from worker-populated regions and other sources such as government or corporations to obtain deposits. One of the biggest problems with this study is that by using deposits as the sole indicator of bank performance, it ignores another major output - banks-loans. This may hinder the accuracy of the results, therefore, making their conclusions questionable. Furthermore, without considering the dynamic relationship between inputs and outputs, the deposits alone may not be able to reflect the managerial side of the bank branch performance at all.

To the best of our knowledge, the only study that examines the impact of regional development on bank efficiency of Chinese regional commercial banks is (Sun et al., 2013). Due to a lack of branch-level data, they select city commercial banks as a sample since city commercial banks are believed to operate majorly within one province. They use both translog and fourier flexible form of a profit function to estimate profit efficiencies of city commercial banks during 2002-2009. In the second stage, they use ordinary least squares (OLS) regression to examine the effect of regional development measured by provincial GDP per capita on profit efficiency. They find that regional development has a positive impact on the profit efficiency of city commercial banks. However, it hinders the positive effect of strategic investors, which can be an indication of the ineffectiveness of governmental control.

Since this is the main study we build upon, we offer a critique by stressing the limitations of the study as follows. Firstly, although the study uses the provincial level data of regional economic conditions, it assumes banks in different regions operate under the same technology by using one frontier for all regions. This gives us one of the motivations to change this assumption and re-examine this hypothesis by estimating banks using a meta-frontier approach. Secondly, the study uses a two-step approach by first estimating profit efficiency and then

regressing environmental variables on profit efficiency scores in the second step. It is well-established in the literature that the two-step approach is considered inferior to one-step in many ways statistically (Wang and Schmidt, 2002). Last but not least, this study only uses one GDP per capita as the sole indicator of economic development, which may not be sufficiently comprehensive as other economic and monetary condition indicators may also play an important role influencing bank performance, such as competition and concentration in the regional banking sector, inflation rates, urbanization etc. (Bos and Schmiedel, 2007, Goddard et al., 2014).

In conclusion, different regional requirements in terms of monetary and banking policies are well-established in the literature, and there has been some evidence of regional impacts on bank performance in China. This prompts us to apply a new methodology that addresses potential biases in previous Chinese studies and assess whether regional impacts persist in a more recent time, especially after the financial crisis. As far as I am aware, this is the first study in the literature that clusters Chinese provinces into regions and models them separately using different frontiers enveloped by country meta-frontier.

4.5 Methodology

The primary methodologies used in this study, i.e., cluster analysis and the stochastic meta-frontier (SMF) approach, will be discussed in this section.

4.5.1 Cluster Analysis - Hybrid Hierarchical K-means Clustering

To conduct our meta-frontier analysis, we first need to define the different technology groups. Recall Chapter 3, we define technology as a decision unit's ability to transform inputs into outputs. The combination of inputs and outputs a firm chooses is constrained by the opportunities and resources available to them,

depending on in which industry, region or country the production occurs. The differences can range from the availability of capital, labour, quality of resources to the social and economic environment (O'Donnell et al., 2008). The most common methods adopted to define the technology groups in the meta-frontier banking literature are either by ownership types (Kontolaimou and Tsekouras, 2010, Goyal et al., 2019, Abdul-Majid et al., 2017) or geographical locations(Bos and Schmiedel, 2007, Huang and Fu, 2013, Lee and Huang, 2017). Among the limited number of Chinese banking meta-frontier studies, to our knowledge, they all focus on the heterogeneity of ownership types rather than geographical or economic environments(Lee and Huang, 2019, Zhu et al., 2020, Yin et al., 2018). In particular, they assume state-owned banks, joint-stock banks, city commercial, and foreign banks operate in different market segments due to different economic, financial, political and geographical restrictions. This is indeed a plausible method to classify technology groups since these studies include all banks with different ownership types in the sample. As our study includes city and rural commercial banks only, there is minimal heterogeneity in terms of ownership types in our sample. Instead, the primary heterogeneity in our sample comes from geographical and economic environments.

Traditionally, the boundaries of banking markets are mainly based on geographical locations as the boundaries of banking markets are believed to be defined by the demand side. Consumers of financial services, including households and businesses, define the markets, meaning that banking markets are initially local and small. Kwast et al. (1997) find that financial services markets are predominantly geographically local for both households and businesses in the US banking system. This is especially true for depository institutions where most customers including both households and businesses are local. They also find that both asset and credit services cater to local customers within each state in the US. Radecki (1998), however, argues that due to the relaxation of branching restrictions and reorganisation of bank holding companies, banking markets are larger and more competitive in the US. It is also driven by the fact that products provided by financial institutions are becoming more homogenous. Jackson and Eisenbeis (1997) employ a multivariate

cointegration analysis using a sample of consumer deposit rates from 29 different local banking markets to test if consumer deposits rates in distinct geographical locations are genuinely independent. They find strong evidence that deposits rates are well-integrated in a national market, proving that interstate restrictions do not restrict the market locally. On the other hand, regarding the UK banking market, Ashton (2009) find that there still exist regional deposit markets. Utilising an OLS regression model to test the impact of branch overlap on dissimilarity of interest rates between bank pairs, substantial differences were found in interest rates between banks with regionally focused branch networks. In addition, branch overlap does not have a significant impact on interest rate setting.

We believe that regional markets do exist in the Chinese banking system, especially for city and rural commercial banks, which are primarily built upon the one-city-one-bank rule. Liu and Wu (2008) define city and rural commercial banks as endogenous institutions. This is because city and rural commercial banks are developed locally and serve the local market. However, the literature is inconsistent in terms of how to group provinces into larger regions. According to the National Bureau of Statistics in China (NBSC), the most common divisions are the Western, the Intermediate and the Eastern zone, based on geographical differences and ease of policy implementation (NBSC, 2021). Yang et al. (2015) adopts an Exploratory Spatial Data Analysis (ESDA) approach and divide Chinese regions into Yangtze River, Delta Bohai Sea region, Pearl River Delta, Northeast region. Central region and Western region, based on economic development and reform, urbanisation and workers population. They find that the performance of Industrial and Commercial Bank of China is significantly better in developed Yangtze River Delta and Bohai Sea regions and underperform in the other four areas. Liu and Wu (2008) divide the mainland Chinese banking industry into three types of local banking systems: the developed systems, the developing systems, and the underdeveloped systems, based on bank assets per capita, bank branches per million people, and bank ownership diversification in 2006. However, to the best of our knowledge, it is based on arbitrary thresholds rather than statistical or econometric methods.

Furthermore, it does not consider any exogenous environmental factors such as macroeconomic conditions that may impact bank performance.

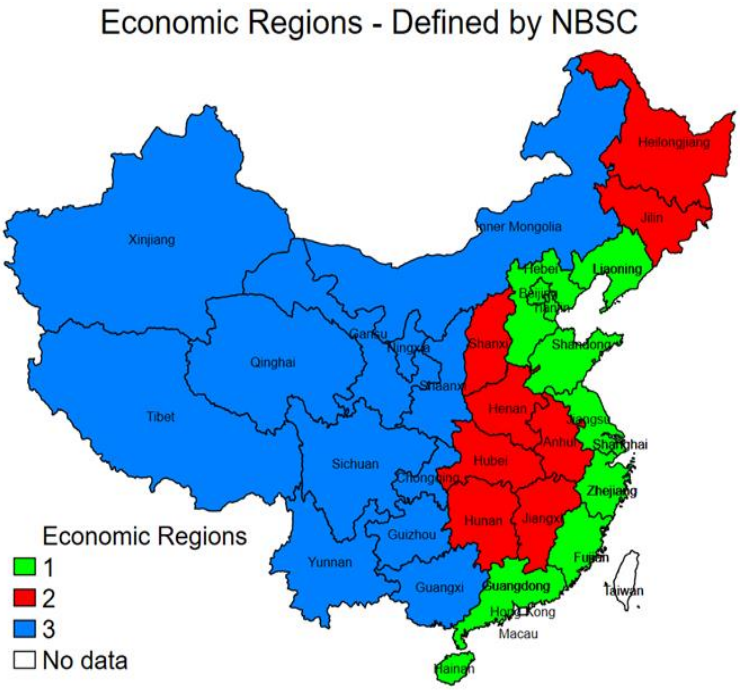
We focus on differences in technology and cost structures and define markets as where banks have a similar level of resources such as capital and labour and a similar level of economic and physical infrastructure development. It means that due to these conditions, banks in each cluster have access to the same level of optimal combination of inputs and outputs, and the deviation from regional technology frontier solely comes from managerial inefficiency. In other words, banks choose different combinations of inputs and outputs in order to compete in their respective market, and the optimal combination is largely restricted by their local economic environment.

Provincial data from NBSC indicates that during our study period from 2005-2015, there is a significant regional disparity in terms of macroeconomic indicators. For example, as of 2015, gross regional product (GRP) per capita is much higher in the Eastern Coastal region compared to other parts of the country. Similarly, unemployment tends be to evidently higher, reaching an average of 4% in the Central and Northeast region compared to around 2.5% in the South-Eastern region. The inflation rate is higher in the Western and Central regions in China compared to the Eastern region. The data not only confirms our understanding of regional disparity in terms of the economic environment but also prompts us to adopts more advanced statistical methods to group provinces into larger regions as provincial macroeconomic indicators follow vastly different patterns.

Cluster Analysis is used to determine the classifications of regions for the operations of Chinese commercial banks in this chapter. According to the National Bureau of Statistics of China, the provinces are divided into three economic regions: Eastern Zone, The Intermediate Zone and Western Region Zone (NBSC, 2020). This classification mainly focuses on geographical locations, historical development and jurisdictional conveniences. The regions are indicated by different colours on the map of China below. Although this

categorising method has many advantages, such as straightforwardness and convenience, it might not be strictly statistically driven or up-to-date as its relevance to the Chinese economy in the most recent decade (i.e. our study period) is unclear.

Figure 4.1. Regional divisions of Mainland China by NBSC (excluding Hong Kong, Macao and Taiwan)



Data Source: National Bureau of Statistics of China. Map source: Merryman (2008)

In this study, we propose using cluster analysis to regroup Chinese provinces according to the most recent statistics in macroeconomic indicators.

Cluster Analysis is a statistical method to find similarities and dissimilarities among individuals in order to find natural grouping solutions. There are many clustering methods, but they can mainly be summarised as one of the two types: partition or hierarchical. Partition clustering splits observations into different groups with distinguished characters, while hierarchical methods create clusters that are hierarchically related.

As for the partition method, algorithms such as k-means or k-medians are commonly used as iteration techniques to partition the data into k groups. The partitioning process starts with k groups, and observations closest to each group centre are then assigned to each group. Then the mean or median of the observations for each group is recalculated. This process repeats until all

observations are assigned to one of the groups. One distinct feature of the partition cluster analysis is that it allows the group centre to be recomputed. Therefore, the observations are allowed to move among groups during the clustering process.

As for the hierarchical clustering method, it can either be agglomerative or divisive. At the beginning of hierarchical clustering, each observation is treated as an individual group. Observations are then merged into fewer groups according to similarities among those observations. According to Kaufman and Rousseeuw (1990), the agglomerative methods go through $N(N-1)/2$ possible combinations of observations to find the closest two observations. The method to compare groups that contain more than two observations is called the linkage method. They are used to determine the distance between two clusters. There are many linkage methods that can be used in cluster analysis. We will choose the average linkage method in this study because of its robustness in most cases, according to Kaufman and Rousseeuw (1990). The average-linkage clustering uses the average similarity or dissimilarity of observations between the groups as the measure to distinguish between two groups. Weighted-average linkage is one variation of the average linkage methods. Unlike the average linkage approach, which takes group size into account when combining two groups, resulting in each observation taking the same weight, the weighted average linkage approach gives groups of unequal size the same weighting. Therefore, observations coming from groups of different sizes will carry different weights. As opposed to partition clustering, during the hierarchical clustering process, once an observation is assigned to a group, it does not allow the observation to move to a different group.

According to Bouvatier et al. (2014) and Husson et al. (2017), there are generally three ways to obtain the clusters. The first way is simply to keep the clusters obtained from the hierarchical cluster analysis. The second way is to use HAC (hierarchical agglomerative clustering) or HDC (hierarchical divisive clustering) to only determine the optimal number of clusters and then use the partition

method to obtain the clusters. The third way is to not only use the first-step hierarchical clustering to determine the optimal number of clusters but also to use the cluster centres obtained from the hierarchical cluster analysis as the starting point for the partition clustering in the second step. That is, the mean or median of the k groups obtained from HAC or HDC will be used as the group centres at the start of the partitioning process. This method is also called hybrid hierarchical k-means clustering, a type of unsupervised machine learning.

In this study, we will choose hybrid hierarchical k-means clustering following Bouvatier et al. (2014). That is, we will use hierarchical agglomerative clustering (HAC) with the Ward linkage method in the first step to determine the optimal number of clusters. In the second step, k-means partition clustering is used to consolidate and obtain the final clusters, where the k starting centres are obtained from the first step HAC. The main reason for using a hybrid is two-fold. Firstly, the weakest link of hierarchical clustering is at the end since observations cannot be reassigned during the clustering process (Bouvatier et al., 2014). Usually, partition clustering seems to result in more distinct clusters. On the other hand, the disadvantage of partition clustering is at the beginning, mostly due to the oversensitivity of clustering outcomes resulting from the randomly assigned initial starting centres. Therefore, it is natural to combine these two methods so that the weakness of one can be compensated by the other. More precisely, using HAC first will give more suitable starting centres for the k-means, and k-means will consolidate the final clusters to make the results more robust.

There are several ways to determine the optimal number of clusters depending on which clustering method we choose. In the following section, we will discuss the method for hierarchical clustering, k-means clustering as well as the hybrid method.

For hierarchical cluster analysis, the most commonly used ones are various

stopping rule indexes. The most frequently used methods are Calinski–Harabasz pseudo-F stopping-rule index and Duda–Hart Je(2)/Je(1) stopping-rule index (Caliński and Harabasz, 1974, Duda et al., 2012). The Calinski–Harabasz pseudo-F stopping-rule index for g groups and N observations is

$$\frac{\frac{trace(B)}{g-1}}{\frac{trace(W)}{N-g}} \tag{4.1}$$

where B is the between-cluster sum of squares and cross-products matrix, and W is the within-cluster sum of squares and cross-products matrix. The larger the values of the Calinski–Harabasz pseudo-F stopping-rule index, the more dissimilar the clusters are.

Duda–Hart Je(2)/Je(1) stopping-rule index is calculated as Je(2) divided by Je(1). Je(1) is the sum of squared errors within the group that is to be split, and Je(2) is the sum of squared errors in the two subgroups after the division. Unlike Calinski–Harabasz pseudo- F stopping-rule index, which can be used in both partition and hierarchical clustering, Duda–Hart Je(2)/Je(1) index requires the clustering methods to be hierarchical only. The pseudo-T-squared value for Duda and Hart Je(2)/Je(1) index is usually calculated to help with finding the optimal number of clusters. The pseudo-T-squared value can be calculated using the following equation.

$$\frac{1}{\frac{Je(2)}{Je(1)}} = 1 + \frac{T^2}{N_1 + N2 - 2} \tag{4.2}$$

Where N_1 and N_2 denote the numbers of observations in the resulted two subgroups.

According to Caliński and Harabasz (1974), larger values of the Calinski–Harabasz pseudo-F index indicate more distinctive clusters. With respect to Duda–Hart Je(2)/Je(1) stopping rule, as a rule of thumb, distinct clusters are usually indicated by the point where one of the largest Duda–Hart Je(2)/Je(1) values correspond to a small Duda–Hart pseudo-T-squared value with much larger pseudo-T-squared values.

For k-means cluster analysis, the method to choose the optimal number of clusters is relatively more subjective. The elbow method is the most widely used approach in the literature. The elbow method finds the optimal number of clusters by plotting the within sum of squares (WSS) or Log(WSS) of all cluster solutions and detecting a kink (elbow) of the curves. A kink indicates a most efficient point for WSS reduction, and any solutions beyond the elbow point have a negligible reduction in WSS compared to the previous solutions (Makles, 2012). Other ways to find the optimal point for WSS reduction are to use proportional reduction in WSS with respect to the total sum of squares (TSS) represented by η_k^2 or proportional reduction in WSS (k clusters) with respect to the WSS from the previous solution (k-1 clusters) represented by PRE_k.

$$\eta_k^2 = 1 - \frac{WSS(k)}{TSS}, \qquad k \in K \tag{4.3}$$

$$PRE_k = \frac{WSS(k-1) - WSS(k)}{WSS(k-1)}, \qquad k > 1 \tag{4.4}$$

Another method that has gained popularity in the literature since its introduction

is the Silhouette method. This is a very useful tool, especially for helping you choose between the cluster solutions obtained from different clustering algorithms. It also helps to detect any outliners in any of the clusters and, in return, offers a solution to optimise results by moving individual items between the clusters. According to Rousseeuw (1987), the most useful feature of the silhouette method is to cross-validate the existing solution and to aid the interpretation of the results.

The silhouette width is defined by Rousseeuw (1987) as:

$$\mathrm{s(i)} = \frac{\big(b(i) - a(i)\big)}{\{Max\ a(i), b(i)\}} \tag{4.5}$$

where $a(i)$ is the average Euclidian distance of object i from all other objects within cluster A and $b(i)$ includes the neighbours of object i, which is defined by the minimum distance between i and objects in its neighbour cluster B, where cluster B constitutes the objects which represent the smallest average Euclidian distances between objects i and any other cluster C which is different from cluster A. The denominator will take the value of $a(i)$ if $a(i)$is bigger than $b(i)$, and will take the value of $b(i)$ otherwise. So, we can easily deduct from the specification that $-1 \leq s(i) \leq 1$. When $s(i)$ is close to 1, it means that the smallest between dissimilarity is much bigger than average within dissimilarity. This means that object i is well-clustered that even i's neighbour cluster B, which is considered the best choice for i apart from A, is nowhere near as good fit as cluster A. On the contrary, if $s(i)$ is very close to -1, it indicates that the Euclidean distance between i and cluster A is much bigger than the minimum distance between i and cluster A, indicating that cluster B would be a much better destination for object i rather than cluster A. Therefore, object i should be regrouped into cluster B.

For the hybrid hierarchical and k-means cluster analysis, the most commonly used method to decide on the optimal number of clusters is to use the stopping rule indexes for the first-step hierarchical clustering. Then for the second step, we will use the pre-defined number of groups k obtained from the first step to further partition and consolidate our results using the k-means clustering method. However, after second-step partition clustering, there might be an even better solution which we could not know during the first step if we do not utilise the results from our second-step k-means clustering. Therefore, we propose a new method that takes into account the results from both steps when deciding the optimal number of clusters. The algorithm of this method is summarized as below:

1) Set k=1
2) Use hierarchical clustering method to produce k number of clusters, where k=1…20
3) Use the k number of clusters resulted from cutting the hierarchical cluster tree at k as the starting centres for the second-step k-means clustering, where the number of clusters for the second-step is pre-set as k
4) Use k-means clustering to produce clustering results for k number of clusters
5) Set k=k+1 and repeat the previous three steps until k reaches a pre-defined number, such as 20 in our case, then exit the loop
6) Store the values of WSS, log(WSS), η_k^2 and PRE_k from results for each k number of clusters (k=1…20). Draw graphs for these 20 solutions. Finally, use the elbow method to determine the optimal number of clusters.

In this study, we will use hybrid hierarchical k-means clustering to determine the classification of regions using macroeconomic conditions as our clustering criteria. Then we will utilise all the abovementioned three methods to decide on the optimal number of clusters. That is, we firstly use the hierarchical clustering stopping rules to determine the optimal number of clusters. Then, we use our

proposed algorithm with the elbow method to position the optimal number of clusters using the final cluster solutions from both steps of the hybrid method. Then we compare the decisions on the number of clusters from the above two different approaches and check if they are consistent. If they are, we have a conclusive result in the optimal number of clusters. If not, we need to further examine both results.

To further validate the decision on the number of clusters, we can use *silhouette* as a cross-checking tool. We will examine both the average silhouette width and if there are any outliners in each cluster. If the previous two approaches give inconclusive results, i.e., a different optimal number of clusters, then we will use the result which gives higher silhouette width with fewer or no mis-assigned objects. If the two approaches agree on the optimal number of clusters, we can solely use silhouette as a cross-validation method to check whether the solution has a positive average silhouette width with a minimal number of mis-grouped objects.

4.5.2 Stochastic Meta-frontier (SMF) Approach

Often, we want to compare the efficiencies of banks operating under different technologies. For example, we want to examine if banks in a particular area are more or less efficient than banks operating in another area. If the banks in different areas have different technologies, pooling all data and estimating the efficiency scores of all banks using a single frontier is not appropriate. Furthermore, if these banks are modelled with different frontiers, the efficiency scores of these areas are not comparable (Battese et al., 2004). To solve this problem, the meta-frontier approach is developed to enable the possibility to compare the efficiency performance of banks operating under different technologies.

Meta-frontier was first introduced by Hayami (1969) and Hayami and Ruttan

(1970, 1971) in studies concerning agricultural productivities in different countries. These studies point out that producers in different regions might not be able to use the best technology available to the whole industry due to the specific regulations and environmental restrictions each region faces. Therefore, each region might choose a sub-optimal technology, creating a gap between the industry-wide technology (meta-frontier) and group-specific technology (individual frontiers).

The major difficulty for estimating the technology gap between the meta-frontier and individual frontiers is the requirement for the meta-frontier to envelop the individual frontiers. A single frontier using pooled data cannot guarantee the envelopment. To solve this difficulty, Battese et al.(2004) and O'Donnell et al. (2008) propose a mixed-method two-step approach. In the first step, individual frontiers are estimated using a standard stochastic frontier method, and in the second step, a mathematical programming method is used to estimate the meta-frontier. Particularly, in the second step, a linear or quadratic programming problem is solved with a constraint. Huang et al. (2014) pointed out the limitations of the mathematical programming method, including the inability to take random shock into consideration and the lack of statistical properties in the second stage. They propose a stochastic meta-frontier (SMF) method where the stochastic frontier method is used in both the first and second stages to account for the random error during the estimation of meta-frontier.

We will, therefore, briefly compare these two major meta-frontier methods and demonstrate the differences between these two methods. Although meta-frontier has been applied in many different industries, we will use the banking industry as an example to illustrate.

In the first step, the stochastic cost function can be expressed as:

$$C_{jit} = f_t^j(X_{jit})e^{V_{jit}-U_{jit}} \quad (4.6)$$

where C_{jit} denotes the observed total cost and $f_t^j(.)$ is the group-specific frontier for group j. X_{ijt} is a vector of bank-specific input prices and outputs and V_{jit} is the random noise with a two-sided error $v_{jit} \sim N(0, \sigma_v^2)$ with a mean of 0 and variance of σ_v^2, and U_{jit} is the inefficiency term with truncated normal distribution $U_{jit} \sim N^+(\mu_{jit}(Z_{jit}), \sigma_u^2)$, where the mean of the term is expressed as a function of a list of environmental variables Z_{jit}s (Battese and Coelli, 1995). That is,

$$\mu_{ijt} = Z_{ijt}\theta + W_{it} \quad (4.7)$$

We can see from the subscripts that the mean of the distribution is observation specific. That is, for each observation, it has its own distribution, thus its own expected value. z_{ijt} is a set of exogenous variables that may affect the inefficiency term u_{ijt}. W_{it} is a random variable that has the distribution of $N^{-z_{it}\theta}(0, \sigma^2)$. The truncation point is $-z_{it}\theta$, so that u_{it} has the distribution of $N^+(z_{it}\theta, \sigma^2)$. That is, a positive truncation at 0 with a normal distribution with mean $z_{it}\theta$ and variance σ^2.

A bank's cost efficiency can then be derived as

$$CE_{it}^j = \frac{C_{jit}}{f_t^j(X_{jit})e^{V_{jit}}} = e^{-U_{jit}} \quad (4.8)$$

The mixed approach proposed by Battese et al. (2004) and O'Donnell et al. (2008) uses a standard maximum likelihood (ML) estimation technique for the first stage estimation of group-specific frontiers. The cost efficiency scores are estimated using the conditional mean $\widehat{CE} = \hat{E}(e^{-U_{jit}} \big| \widehat{\varepsilon_{jit}})$ proposed by Battese

and Coelli (1988). In the second stage, the mixed method estimates meta-frontier by solving the linear programming problem using the estimates obtained from the individual frontiers in the first stage. Minimizing the deviations from meta-frontier to group-specific frontiers can either be approached using the minimum sum of absolute deviations,

$$\min \mathrm{L} \equiv \sum_{i=1}^{N}\sum_{t=1}^{T} |lnf_t^M(X_{jit}) - ln\hat{f}_t^j(X_{jit})| \qquad (4.9)$$

$$\text{s.t. } f_t^M(X_{jit}) \geq ln\hat{f}_t^j(X_{jit}) ;$$

or minimum sum of squares of deviations.

$$\min \mathrm{L} \equiv \sum_{i=1}^{N}\sum_{t=1}^{T} \left(lnf_t^M(X_{jit}) - ln\hat{f}_t^j(X_{jit})\right)^2 \qquad (4.10)$$

$$\text{s.t. } f_t^M(X_{jit}) \geq ln\hat{f}_t^j(X_{jit}) ;$$

There are several drawbacks of the above two-step approach. Firstly, there is no statistical property in second step estimation, as it is calculated using linear or quadratic programming algebra, which relies on simulations or bootstrapping. Secondly, it is unable to incorporate exogenous environmental variables that impact inefficiency into the estimation. Thirdly, it does not take into consideration of any idiosyncratic shocks.

Because of these inherent problems of the mixed two-step approach, we will use a stochastic meta-frontier (SMF) method in this study proposed by Huang et al. (2014), which considers random shock and provides statistical properties in the

second stage. In the first stage estimation, the SMF approach is no different from the mixed approach. That is, group-specific frontiers are obtained using the standard stochastic frontier approach using maximum likelihood estimation methods. In the second stage of SMF, we define estimation errors from the first stage SFA as:

$$ln\hat{f}_t^j(X_{jit}) - lnf_t^j(X_{jit}) = V_{jit}^M = \varepsilon_{jit} - \hat{\varepsilon}_{jit} \quad (4.11)$$

From the structure of the above equation, we can clearly see that the stochastic frontier approach using the maximum likelihood method can be applied. The term $ln\hat{f}_t^j(X_{jit})$ represents the predicted values of the logarithm of total cost from group-specific frontiers and $lnf_t^M(X_{jit})$ represents meta-frontier. U_{jit}^M represents the cost gap term which is a function of a list of environmental variables with the distribution $U_{jit}^M \sim N^+(\mu_{jit}(Z_{jit}), \sigma_u^2)$. It can be estimated using the Battese and Coelli (1995) model. Although V_{jit}^M may follow a normal distribution with zero mean $V_{jit}^M \sim N(0, \sigma_v^2)$ and may be considered following a i.i.d. distribution, because it contains the error from the individual frontiers, the variance of the V_{jit}^M might not be constant, which results in heteroscedasticity. In this case, the maximum likelihood may not produce a consistent estimation. Therefore, we will use robust standard error following White (1982) in the estimation procedure of the above equation by computing the sandwich form for the covariance matrix of the estimates. This modification is also called the pseudo-maximum likelihood (PML) method.

The meta-frontier, which envelops all the group-specific individual frontiers, can then be expressed as

$$f_t^M(X_{jit}) = f_t^j(X_{jit})e^{-U_{jit}^M} \quad (4.12)$$

Where $U_{jit} \geq 0$. This equation guarantees that meta-frontier will lie below the individual frontiers, i.e., $f_t^M(X_{jit}) < f_t^j(X_{jit})$.

Substituting equation (4.11) into equation (4.12), we can get

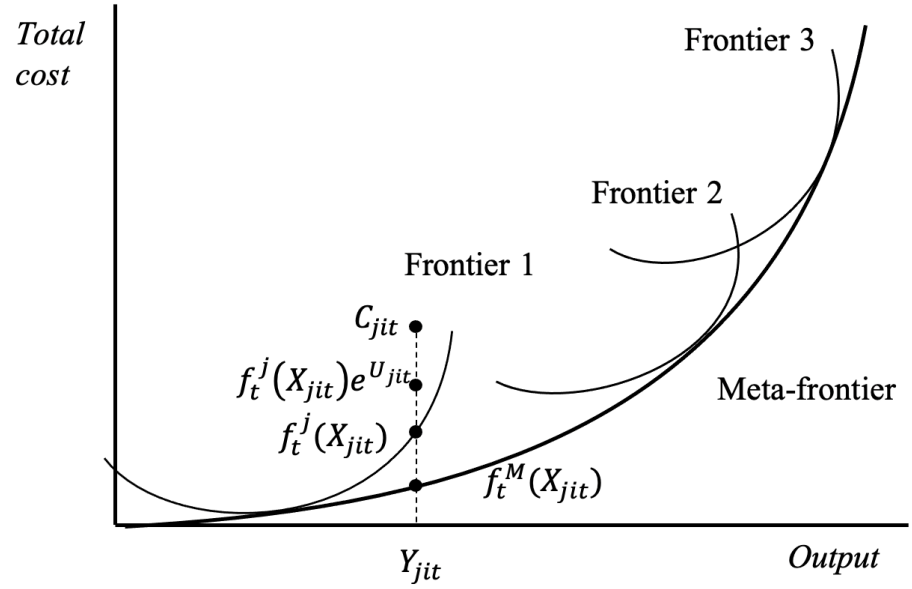
$$ln\hat{f}_t^j(X_{jit}) = lnf_t^M(X_{jit}) + U_{jit}^M + V_{jit}^M \quad (4.13)$$

The term U_{jit}^M , in this case, indicates the technology gap between group-specific frontier and meta-frontier. The technology gap term can also be expressed as a ratio of meta-frontier to the group-specific frontier. Since this is a cost function, we will call it the cost gap ratio (CGR). It measures how far the group individual frontiers deviate from the meta-frontier. The larger the value of CGR_{it}^j, the closer the distance between the individual frontier and meta-frontier, the more advanced technology the group adopts.

$$CGR_{it}^j = \frac{f_t^M(X_{jit})}{f_t^j(X_{jit})} = e^{-U_{jit}^M} \leq 1 \quad (4.14)$$

A graphical illustration of the relationship between the meta-frontier cost function and individual frontiers can be seen as below, where Frontier 1 and Frontier 2 are individual group frontiers that are enveloped in meta-frontier.

Figure 4.2. Meta-frontier Cost model



The x-axis represents the amount of output and the y-axis represents the total cost. C_{jit} is the observed total cost and and Y_{jit} is the observed total output, $f_t^j(X_{jit})e^{U_{jit}}$ represents the splitting point separating the distance between observed cost and meta-frontier cost, into inefficiency and random shock component; $f_t^j(X_{jit})$ is the group specific frontier, and $f_t^M(X_{jit})$ represents the meta-frontier, thus, the technological gap between them is measured as $\frac{f_t^M(X_{jit})}{f_t^j(X_{jit})}$.

As illustrated in the graph, at a given level of Y_{jit}, the ratio of meta-frontier to observed cost consists of three parts: cost gap ratio $CGR_{it}^j = \frac{f_t^M(X_{jit})}{f_t^j(X_{jit})}$, cost efficiency $CE_{it}^j = \frac{f_t^j(X_{jit})}{f_t^j(X_{jit})e^{U_{jit}}} = e^{-U_{jit}}$ and a random error term $\frac{f_t^j(X_{jit})e^{U_{jit}}}{C_{jit}} = e^{-V_{jit}}$. Putting them together, the equation can be written as:

$$\frac{f_t^M(X_{jit})}{C_{jit}} = \frac{f_t^M(X_{jit})}{f_t^j(X_{jit})} \times \frac{f_t^j(X_{jit})}{f_t^j(X_{jit})e^{U_{jit}}} \times \frac{f_t^j(X_{jit})e^{U_{jit}}}{C_{jit}} = CGR_{it}^j \times CE_{it}^j \times e^{-V_{jit}} \quad (4.15)$$

It can also be expressed as the formulation of meta-frontier efficiency

MCE_{jit} with random error taken into account. MCE_{jit} represents individual bank's cost efficiency with respect to meta-frontier.

$$MCE_{jit} = \frac{f_t^M(X_{jit})e^{V_{jit}}}{C_{jit}} = TGR_{it}^j \times CE_{it}^j \qquad (4.16)$$

Where $0 \leq TGR_{it}^j \leq 1$ and $0 \leq CE_{it}^j \leq 1$.

To sum up, this study will adopt the stochastic meta-frontier (SMF) approach proposed by Huang et al. (2014) as the mathematical programming in the second step of the mixed methods proposed by Battese et al. (2004) and O'Donnell et al. (2008) suffers from the problems of lack of statistical properties and inability to take random shocks into consideration. In the SMF approach, we will first use the stochastic frontier method to estimate individual frontiers and obtain the linear predictions of the dependent variable log total cost. In the second stage, we will use SFA pseudo-maximum likelihood (PML) to estimate meta-frontier by using the linear predictions of log total cost from first stage group-specific frontiers instead of the observed values of log total cost as the dependent variables, where the technology gap term U_{jit}^M is a function of environmental variables following Battese and Coelli (1995)[6].

[6] The inefficiency model we adopt here, i.e., the Battese and Coelli (1995) model, has the limitation that it is not a 'true' panel data model, meaning that it does not consider the unobserved individual heterogeneity in the estimation process. We adopt Battese and Coelli (1995) model here, instead of Greene (2005) 's true fixed effect model, is mainly due to the significant increase in the number of parameters in Greene (2005)''s TFE model, making convergence during the maximum likelihood estimation process a technical difficulty. Moreover, adopting Battese and Coelli (1995) also avoid the potential incidental parameters problems which are likely to be present if Greene (2005)'s TFE model is estimated, rendering coefficients biased.

4.6 The implementation of Methodology

4.6.1 The Implementation of Cluster Analysis

As mentioned previously, the cluster analysis method applied in this study is a hybrid two-step approach that combines HAC (hierarchical agglomerative clustering) and k-means partition clustering. In the first step, we obtain the optimal k number of clustering using Calinski–Harabasz pseudo- F stopping-rule and Duda–Hart Je(2)/Je(1) stopping-rule. In the second stage, we build on the first stage by using the cluster centres obtained from the first stage as the starting point of k-means clustering in the partition clustering process. As mentioned in the methodology section, this combined method is chosen to utilise the best features of both hierarchical clustering and partition clustering. According to Husson et al. (2017), since the k-means clustering is built on the first stage results, the second step can be seen as a process of consolidation to improve and optimize the clusters obtained from the first step.

In our study, we use the same set of environmental variables as the criteria for clustering in both first stage HAC and second stage partition clustering. Please note that these variables are on the provincial level, meaning that the value of these variables varies across provinces and years. These same variables will be also be used as the macroeconomic variables which affect the cost gap ratio (CGR) when estimating meta-frontier to achieve consistency. The variables and are listed below.

Table 4.1. Clustering Macroeconomic Variables

Variable name	Description
$GRPGR_{pt}$	Gross regional product (GRP) growth rate in each province, $GRPGR_{pt}$= [(GRP current year-GRP preceding year) / GRP preceding year] *100

$GRPPC_{pt}$	Real gross regional product per capita in each province (nominal gross regional product adjusted by CPI with 2005 as the base year) $GRPPC_{pt}$=current year Real GDP /current year population,
$Urban_{pt}$	Urbanisation Rate
INF_{pt}	Inflation rate in each province, Consumer Price Index (preceding year=100) as a proxy

There are a total of 34 provincial-level administrative divisions in China. However, the provincial-level data on macroeconomics provided by the national bureau of statistics is only available for 31 mainland provinces. Therefore, only 31 mainland provinces will be included in the cluster analysis.

4.6.2 The Implementation of Stochastic Meta-frontier (SMF)

In this study, we will use the translog cost function to implement the stochastic meta-frontier (SMF) approach.

The first stage group-specific cost function can be expressed as equation (4.17) below:

$$\ln(TC)_{it}{}^{j} = \alpha_0 + \sum_{m=1}^{3} \beta_m lny_{mit} + \sum_{n=1}^{2} \rho_n lnp_{nit} + \frac{1}{2}\left[\sum_{m=1}^{3}\sum_{n=1}^{3} \tau_{mn} lny_{mit} lny_{nit} + \sum_{m=1}^{2}\sum_{n=1}^{2} \varphi_{mn} lnp_{mit} lnp_{nit}\right] + \sum_{m=1}^{3}\sum_{n=1}^{2} \omega_{mn} lny_{mit} lnp_{nit} + \epsilon_1 t + \frac{1}{2}\epsilon_2 t^2 + \epsilon_3 lneq_{it} + \frac{1}{2}\epsilon_4 eq_{it}^2 + lnU_{it} + lnV_{it} \quad (4.17)$$

where j represents the clusters obtained from cluster analysis. Therefore, the group-specific translog function will be estimated for each cluster using the

equation above. Maximum likelihood (ML) is used as the estimation technique.

TC is the total cost; y_{mit} and p_{nit} represent a list of outputs and input prices, respectively. Foreign is the control variable for foreign banks where it takes the value of 1 if the bank is a foreign bank and 0 otherwise.

Table 4.2. List of Inputs and Outputs

Inputs		Input prices		Outputs	
X1	Total Non-interest expenses	P1	Total Non-interest expenses/ Total Asset	Y1	Net Loans
				Y2	Other Earnings Assets
X2	Total Interest Expenses	P2	Total Interest Expense/ Customer Deposits		
				Y3	Total Non-Interest Operating Income

From the inputs and outputs, we can see that we use the intermediation approach where banks are considered as intermediators using labour, capital and deposits to produce loans and other earnings assets. Unfortunately, due to the significant proportion of missing data in personnel expenses, the reason to use net loans as outputs instead of gross loans is that non-performing loans have been a rather significant problem in China. The time trend variable t is the year and has values from 1 to 11 since we have 11 years in our study period. The coefficient of t indicates the rate of neutral technological change. Equity is also included as a quasi-input to act as a risk control measurement in this model. To model the non-linearity of the effect of time and equity, we also include the squared terms of them in this study.

The importance of the inclusion of risk control variables has been discussed in the literature. Hughes and Mester (1993) claim that the capital requirement set by regulators is very likely limiting the banks' ability to choose the optimal capital level that minimises costs. Park and Weber (2006) further argue that, as the level of equity affects the risk-taking behaviour, bank managers can choose to manipulate inputs and outputs in an attempt to preserve equity level in compliance with capital regulation. When bank capital is low, bank managers can spend more on labour to monitor risky loans, whereas when bank capital level is more than sufficient, bank managers can choose to lower the labour expenses on monitoring loans to increase cost-efficiency. Alternatively, moral hazard behaviour may occur in banks with a low level of equity as managers in these banks may take high-risk, high-return activities, which may boost x-efficiency in the short run while damaging the bank's efficiency and solvency in the long run, suggested by Lozano-Vivas et al. (2002). In either case, it is evident that the equity level of banks very likely alters bank manager's behaviour. Bryce et al. (2015) further tested the essentiality of the inclusion of risk control variables in modelling x-efficiency by testing the relationship between loan loss provisions, x-efficiency and capital ratio. They find that only by including risk control variables when estimating x-efficiency can the capital management behaviour be found in the second stage GMM regressions. They argue that this is due to the fact that risk preferences of bank managers have significant influence on bank operations and only by taking the risk appetite of bank managers into account when estimating x-efficiency can the true efficiency of banks be revealed.

In the group-specific frontiers, the mean of the inefficiency term U_{it} is a function of environmental variables specified below. We follow Huang et al. (2014) and Lee and Huang (2017) by only including microeconomic variables in the estimation of inefficiency in the first stage. The translog function and the equation below are estimated simultaneously using the maximum likelihood method.

$$\mu_{it}^{j} = \theta_0 + \theta_1 ETA_{pt} + \theta_2 ROAA_{pt} + \theta_3 NPLLOAN_{pt} + W_{it}, \qquad (4.18)$$

Where p denotes province and t denotes the year. We can see that the microeconomic variables here are province and year specific.

Table 4.3. Microeconomic Variables

Variable Name	Description	Expected sign
ETA_{pt}	Equity to total assets ratio, ETA_{pt}= mean (Total Equity/Total Assets). The mean value of ETA_{pt} for each province each year is used to measure the capital level per province per year.	+/ _
$ROAA_{pt}$	Return of average assets, $ROAA_{pt}$ = mean (Net Income /Average Assets). The mean value of $ROAA_{pt}$ for each province each year is used to measure the profitability of banks per province per year.	_
$NPLLOAN_{pt}$	Non-performing loan to gross loan ratio, $NPLLOAN_{pt}$= mean (Non-performing loan/gross loan) The mean value of $NPLLOAN_{pt}$ for each province each year is used to measure the level of Non-performing loan per province per year.	+

ETA, a capital adequacy ratio, is widely used in literature to capture the regulatory condition in a specific area. There are generally two opposing views on the relationship between efficiency and capital ratio in the literature. On the one hand, Kasman and Yildirim (2006) and Berger and DeYoung (1997) argue that banks which are better equipped with capital are less likely to incur loan losses, therefore, less likely to spend extra expenditures on recovering loan losses. Conversely, a lower level of equity indicates higher risks within the bank, which makes borrowed funds more expensive. Therefore, the banks with a higher level of capital tend to be more efficient. On the other hand, Lozano-Vivas et al. (2002) argue that low levels of equity might be a sign of moral hazard

behaviour, where banks take high-risk, high-return investments, which make them appear to be more efficient in the short term. However, in the long term, these banks may suffer from greater losses.

H1: Banks with a higher level of capital may have either a higher or lower level of cost efficiency

ROAA is a measure of bank profitability, and it is commonly believed that there is a positive relationship between bank efficiency and profitability. The average profitability on the provincial level measures the competitive environment in that province. The more competitive in terms of profitability the area is, the more efficient the banks are. Lozano-Vivas et al. (2002) and Andries (2011) find that banks with higher profitability ratios correspond to a higher level of efficiency.

H2: More profitable banks are more cost-efficient.

NPLLOAN is essentially a measurement of the riskiness of loans. A higher level of risky loans may incur additional costs in labour or other resources on monitoring risky loans and recovering loan losses. For instance, Berger and DeYoung (1997) find that non-performing loans lead to a reduction in cost efficiency. However, they also point out that if the skimping hypothesis holds, banks may intentionally withdraw investment in labour and other resources on underwriting and monitoring loans in order to boost efficiency in the short term. In this case, the NPL level may not change much in the short term, but efficiency increases temporarily. Since Chinese banks had completed the dumping of non-performing loans to asset management companies at the beginning of the study period, we can assume that NPLs can largely reflect the true riskiness in loans in these banks (Elliott and Yan, 2013).

H3: Chinese banks with a higher level of non-performing loans are less cost-efficient.

The second stage meta-frontier is specified as below:

$$\ln(\widehat{TC})^J{}_{it} = \alpha_0 + \sum_{m=1}^{3} \beta_m lny_{mit} + \sum_{n=1}^{2} \rho_n lnp_{nit} + \frac{1}{2}[\sum_{m=1}^{3}\sum_{n=1}^{3} \tau_{mn} lny_{mit} lny_{nit} + \sum_{m=1}^{2}\sum_{n=1}^{2} \varphi_{mn} lnp_{mit} lnp_{nit}] + \sum_{m=1}^{3}\sum_{n=1}^{2} \omega_{mn} lny_{mit} lnp_{nit} + \epsilon_1 t + \frac{1}{2}\epsilon_2 t^2 + \epsilon_3 lneq_{it} + \frac{1}{2}\epsilon_4 eq_{it}^2 + lnU_{it}^M + lnV_{it}^M \quad (4.19)$$

where $\ln(\widehat{TC})^J{}_{it}$ are predicted values of lnTC from group-specific frontiers. The term U_{it}^M obtained from this equation measures the technological gap between meta-frontier and group-specific frontiers. Meta-frontier parameters are estimated using pseudo-maximum likelihood (PML).

Apart from the difference in the values for the dependent variable, another difference between the individual frontier and meta-frontier is the function of the term U_{it}^M. Please note that U_{it}^M here represents the technology cost gap ratio CGR, which can be expressed as a list of environmental variables. This is another feature of SMF that make it superior to the mathematical programming technique. Again, we follow Huang et al. (2014) and Lee and Huang (2017) in the second stage by including macroeconomic variables only as it is believed that the technological gap between individual frontiers and meta-frontiers are mainly caused by economic and regulatory conditions the groups are facing which are outside of the control of the management.

The function of the mean of U_{it}^M can be expressed by the equation below:

$$\mu_{it}^{M} = \lambda_0 + \lambda_1 GRPGR_{pt} + \lambda_2 GRPPC_{pt} + \lambda_3 Urban_{pt} + \lambda_4 Inf_{pt} + \lambda_5 HHI_{ct} + \lambda_6 FAS_t + W_{it} \quad (4.20)$$

where p denotes provinces, c denotes clusters, and t indicates the year. Therefore, all macroeconomic environmental variables in this equation are either on province or cluster level, which makes this study different from most studies on environmental influences where only macroeconomic variables on the national level are used.

We can notice that apart from HHI_{ct} and FAS_t, the rest of the macroeconomic variables are exactly the same as the variables used for cluster analysis. Please refer to section 4.1 for details of the description and calculation of these macroeconomic variables.

Macroeconomic conditions are the most commonly used environmental variables in the BC(95) model. $GRPGR_{pt}$ and $GRPPC_{pt}$ are both proxies for macroeconomic conditions. Maudos et al. (2002) find that banks operating in expanding economies, proxied by GDP growth rate, tend to appear more efficient. Lozano-Vivas et al. (2002) point out that regions with higher GDPPC tend to have ample savings that stimulate both demand for loans as well as supply of funding to meet the demands, boosting bank outputs and profitability. Higher GDPPC is also an indicator that banks may operate in a more mature banking system with more advanced technology. Here, both the gross regional product growth rate and gross regional product per capita are used to measure the expansion and general health of the economy in each province. We expect that their behaviour will follow the same pattern as the national level macroeconomic variables.

H4: Banks in regions with higher GDP levels and growth rates operate closer to meta-frontier

Kasman and Yildirim (2006) and Lozano-Vivas and Pasiouras (2014) argue that inflation may alter banks' behaviour. They claim that higher inflation may induce banks to use excessive branch networks to compete in the market. Demirguc-Kunt et al. (2003) find a positive relationship between inflation and

bank overhead cost. Therefore, we expect the sign of Inf to be positive.

H5: A higher level of inflation rate widens the technological gap between regional frontier and the national meta-frontier

Herfindahl–Hirschman Index (HHI), calculated as the sum of squares of market shares of all firms in a market, is one of the most widely used measurements for concentration. The value of HHI can range from 0 to 10,000, indicating from a market with a large number of firms to monopoly. It is determined by both the number of banks in the market and the inequality of the distribution of asset shares. A low HHI Index usually indicates low concentration, and a market with an HHI higher than 1800 points is generally considered to be highly concentrated. The relationship between HHI and the technology gap can be in either direction. If Hicks (1935)'s quiet life hypothesis holds, then the increase in concentration is caused by market power, which decreases the performance of banks as monopoly power reduce bank managers' incentive to improve bank performance. On the other hand, a higher concentration can be caused by better management, which in term may cause better performance in banks(Demsetz, 1973). Berger (1995) further argues that because of the structure-conduct-performance model, a concentrated market structure may encourage banks to collude, which will lead to higher productivity[7].

H6: Market concentration may have either positive or negative impacts on technological gaps.

FAS_t is calculated as the total assets of all foreign banks divided by the total assets of all banks in China for each year. Since this variable only has subscript

[7] We do not include HHI in the clustering process because the HHIs on the provincial level are significantly more inaccurate than the HHIs on the cluster level. This is mostly due to the data limitation. On the provincial level, although it is most likely that more than one bank is operating in each province during our sample period, due to data restriction, certain provinces in specific years only contain one observation, which does not reflect the true competitive nature of such province at a given time. However, on the cluster level, HHIs tend to be much less biased due to the fact that missing data is only a relatively small fraction of the total available data. The HHIs are calculated as the sum of squared asset shares in each cluster each year. The subscription of HHI_{ct} indicates that it is cluster and time specific, whereas the rest of the macroeconomic variables are province and time specific.

t, it is an indication that it is on the national level, not cluster or province-specific.

This chapter also aims to shed some light on the impact of foreign exposure, especially foreign bank presence, on the Chinese banking system. The foreign exposure of Chinese banks can be seen in the form of the rapid growth of cross-border loans and an increase of cross-border capital inflow in the form of foreign investment, foreign ownership and the establishment of foreign banks. As of mid-2018, Chinese banks' cross-border lending represents around 7% of all cross-border lending globally and 24% of all cross-border lending to merging market and developing economics (EMDE) borrowers. Chinese banks lend to 135 out of 143 EMDEs and 30 out of 31 advanced economies (AEs) (Cerutti et al., 2020). Moreover, China has also witnessed an acceleration of cross-border capital inflow in the form of direct and indirect investments in recent years. Zhao and Xu (2021) suggest that direct and indirect investment can lead to bank credit growth. They also suggest that since foreign banks have the advantage of more advacned financial innovation and a more mature business model, they may create competition and attract high-quality customers away from domestic banks. This will force local banks to accept customers with lower credit standards, thus leading to a higher level of non-performing loans.

Levine (1997) believes that foreign bank entry can boost financial development directly via providing high-quality banking services and indirectly through three channels. Firstly, foreign banks increase pressure for domestic banks to cut costs and increase efficiency in order to stay competitive in the banking market. Secondly, foreign bank presence promotes more rigorous accounting, auditing, and ratings by relevant institutions. Thirdly, they create pressure on the local governments to enhance local financial infrastructures such as supervisions, legislation and regulations in the financial market, accelerating domestic financial development. This is especially prevalent in developing economies where governments are under pressure to match international supervisory and regulatory standards. In addition, Levine (1997) also adds that lowering the barrier of foreign bank entry and alleviating the restrictions may allow more

comprehensive access to foreign capital markets and improve the quality and availability of financial services of domestic banks via intensifying competition and facilitating the implementation of new management, skills, and technologies.

There are mainly two ways foreign banks enter a domestic market. One way is directly entering the market by opening branches. Another way of foreign bank entry is via shareholding as strategic investors. In terms of the empirical literature, the results are mixed regarding the impact of bank entry on domestic bank performance. On the one hand, empirical studies show that foreign bank entry generally has a positive impact on the efficiency and profitability of domestic banks. On the other hand, it is found that foreign bank entry may destabilise the domestic banking market via excessive competition. Sturm and Williams (2004) find that allowing foreign banks access to the Australian banking market through deregulation prompts improvement in local bank efficiency. However, Barajas et al. (2000) find that profitability and credit quality may decrease due to higher quality customers switching from domestic to foreign banks. Claessens et al. (2001) also find evidence for a reduction in profitability for domestic banks as a result of foreign bank entrance. Unite and Sullivan (2003) find that foreign bank entry reduces profits for domestic banks but improves their operating efficiency. Lensink and Hermes (2004) find that for a less developed banking market, foreign bank entry raises costs for domestic banks in the short term and make them less efficient.

H7: Foreign Bank presence promotes narrowing of technological gaps.

An important parameter that is worth noticing is the variance parameter γ. It indicates the variance of the inefficiency term as a proportion of the total variance of the composed error. The γ parameter can be expressed as

$$\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \tag{4.21}$$

The value of γ is between zero and one by definition. However, for stochastic frontiers, γ should not equal zero or one. If γ is one, there is no existence of random error term v and the model collapses to a deterministic model. A γ value of zero indicates that there is a nonexistence of the one-sided error term and the composed error term is purely random error. Therefore, no inefficiency is explained by the model.

Log likelihood ratio test will act as the post-estimation test for the appropriateness of the models used. The log-likelihood ratio follows a mixed chi-squared distribution, and its test statistics can be calculated as

$$\chi\sim -2*(LL_R - LL_U), \qquad (4.22)$$

To test the existence of the one-sided error term u_{it} (whether γ equals to zero) and the joint significance of exogenous variables in the inefficiency model, the null hypotheses are

$$H_{01}: \gamma = \theta_1 = \theta_2 = \theta_3 = 0 \qquad (4.23)$$

and

$$H_{02}: \gamma = \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 0 \qquad (4.24)$$

where the restricted model is the OLS and the unrestricted model is the stochastic frontier function with an inefficiency model. The rejection of the null hypotheses indicates the rejection of restricted models and preference of the unrestricted models.

It is important to examine whether the data can support the regularity conditions. We impose a set of constraints to ensure properties of a cost function are satisfied. Symmetry conditions can be expressed as

$$\tau_{ij} = \tau_{ji} \qquad and \qquad \varphi_{ij} = \varphi_{ji} \qquad (4.25)$$

and restrictions to ensure linear homogeneity in input prices are as below:

$$\sum_i \rho_i = 1 \sum_i \varphi_{ij} = 0 \sum_j \omega_{ij} = 0 \quad (4.26)$$

4.7 Data Description

The dataset contains a unique panel dataset of 136 regional Chinese commercial banks during the 11-year period from 2005 to 2015 with a total of 1050 observations as well as a set of province-level macroeconomic variables for all the provinces these banks operate in. The unbalance in the dataset is resulted from the deletion of any observation which contains any missing value for any of the variables to be used in the frontier estimation[8]. The bank-level data are obtained from the balance sheet and income statement information from the *Fitch BankScope* database *(*now *Bureau van Dijk Orbis Bank Focus)*. All accounts comply with IFRS standards, and all bank data are deflated by GDP deflator using 2005 as the base year to take away the price effects of inflation.

The province-level macroeconomic variables come from the official website of the *National Bureau of Statistics of China* and are similarly deflated with rebased GDP deflator to real values before further analysis. We only include regional

[8] Due to frontier estimation requirement, if one variable for a bank is missing for a particular year, all data during that year are deemed missing. Our panel data set is unbalanced due to the unavailability of data and entry of new banks. To be best of our knowledge, there is minimum (if any) exit of a bank in our sample caused by merger or bankruptcy. Peng and Han (2021) report that only two Chinese banks have gone bankrupt over the past few decades. As there is a minimum occurrence of mergers, acquisitions and bankruptcies, it is likely that the incompleteness of our panel data is primarily due to random missing data and new bank entry. Baltagi and Chang (1994) find that making incomplete panel data balanced by dropping observations may worsen the performance of estimators. Dionne et al. (1998) also confirm that entry/exit of incomplete panel data may not cause biases to regression estimates when the reasons for the missing data are not endogenous to the frontier model. That is, the causes of data incompleteness are not determinants of the cost structure, such as an exit that is caused by lower productivity and efficiency of the bank. Thus bank exit seems to be more of a problem than bank entry. As there is minimum occurrence of bank mergers and bankruptcies in our sample period, it is feasible to not consider entry/exit biases in our estimation process.

commercial banks to ensure the homogeneity of the objectives of the banks in the dataset. More precisely, these regional commercial banks only include city and rural commercial banks. The rationale is to ensure that majority of the banks in our sample only operate largely and locally within one province or region. According to Wu (2018), historically, city and rural commercial banks were only licenced to operate within one region. Until 2005, the Bank of Shanghai became the first bank allowed to set up a branch in Beijing. City and rural commercial banks have tried to expand their operations outside their regions since 2005. However, due to a lack of business contacts and governmental support in the new provinces or regions they expand to, the majority of the operations of these banks still depends on their home regions. Sun et al. (2013). (2013) investigate the effects of strategic investors and regional economic development on the efficiency of Chinese commercial banks. They also argue that city commercial banks were initially created with the "one-city-one-bank" rule. They use city commercial banks only in their sample in order to investigate the regional impact on bank efficiency due to a lack of branch-level data on national banks. Although with China Banking Regulatory Commission's (CBRC) permission in 2005, several banks expanded rapidly outside their home-city since 2007, this number is still very limited, with an overwhelming majority still operating within their own province. As our study also focuses on the effects of regional economic development on commercial banks, we utilise a similar rationale to exclude large banks which operate extensively outside their headquarter regions. In addition to city commercial banks, we also include rural commercial banks in our sample. We find it feasible to include rural commercial banks in our sample because a nationwide consolidation of rural credit unions and rural cooperatives into commercial banks has happened since 2004, and the operations of rural banks have become more similar to that of city commercial banks. One reason Sun et al. (2013) did not include rural commercial banks could be due to the fact that their sample period is 2002-2009, a period which includes a rapid transitioning of rural banks from credit cooperatives to commercial banks. Therefore, the inclusion of these banks before and after transition could introduce a greater degree of heterogeneity to the sample, and the effects of transitioning could bring more noise to the estimation. As our study period spans

from 2005 to 2015, we only include fully transitioned rural commercial banks in our sample. Another practical reason is that as we need to estimate an individual frontier for each cluster in order to estimate the meta-frontier, the sample size for each individual frontier needs to be appropriate for a maximum likelihood estimation. By including rural commercial banks, we can increase our total sample size and in turn maximise the observations available for each individual frontier. Furthermore, we exclude foreign banks from our sample for a similar reason. Foreign banks usually are headquartered in one of the several large cities and have operations across regions in China. Their annual data is also reported on the aggregated level. Therefore, it would also be difficult to examine their efficiency on the regional level. In conclusion, we only include city and rural commercial banks in our sample for regional investigation.

4.8 Empirical Results - Cluster Analysis

4.8.1 Descriptive Statistics

The table below gives the summary statistics of the province-level environmental variables used in cluster analysis. We can see that these variables are not on the same scale. As mentioned before, GDP and GDP per capita are rebased using GDP deflator rebased to 2005. We choose Consumer Price Index (CPI) here as GDP deflator, and similar to GDPGR, GDPPC, URBAN and INF, CPI also varies across province and year, which makes our deflating a province and year specific process as well.

Table 4.4. Summary stats of Macroeconomic Variables

Variable	$GDPGR_{pt}$	$GDPPC_{pt}$	$URBAN_{pt}$	INF_{pt}
Obs.	341	341	341	341
Mean	11.6258	34052.4000	50.8714	2.9035
Std. Dev.	2.7169	21295.3800	14.7926	2.0113
Min	3.0000	5052.0000	20.7143	-2.3000
Max	23.8000	107747.7000	89.6083	10.1000

GDPGR is the Gross regional product (GRP) growth rate in each province, GDPPC is the real gross regional product per capita in each province, URBAN is the urbanisation rate in each province, and INF is the inflation rate in each province

Therefore, we use the z-score method to standardize the values of the below variables before proceeding with cluster analysis. The standardization procedure of variable x can be expressed as below:

$$x^* = \frac{x - m}{sd} \tag{4.27}$$

Where m is the mean of x, and sd is the standard deviation of x. This procedure will transform the means to be zeros and standard deviations to be ones. After standardization, the summary statistics of macroeconomic variables are as below.

Table 4.5. Summary stats of Standardized Macroeconomic Variables

	Obs	Mean	Std.Dev.	Min	Max
$sGDPGR_{pt}$	341	0	1	-3.17487	4.480906
$sGDPPC_{pt}$	341	0	1	-1.36182	3.460624
$sURBAN_{pt}$	341	0	1	-2.03867	2.618671
$sINF_{pt}$	341	0	1	-2.58718	3.578071

sGDPGR is the standardized gross regional product (GRP) growth rate in each province, sGDPPC is the standardized real gross regional product per capita in each province, sURBAN is the standardized urbanisation rate in each province, and sINF is the standardized inflation rate in each province

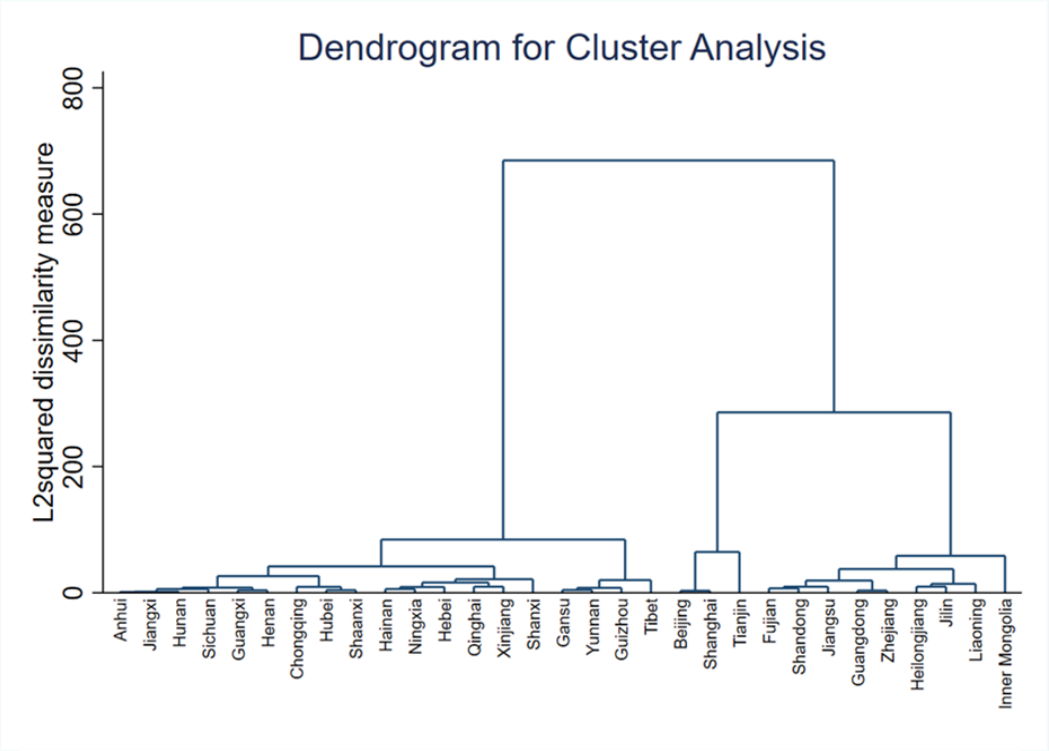
4.8.2 Results of Cluster Analysis

Since our macroeconomic data is a panel dataset that varies across provinces and years, we wish to incorporate the difference across provinces as we take changes over a year into account when clustering these variables. This is usually a difficult thing to do since cluster analysis cannot identify a panel dataset and is not able to differentiate between changes across year and changes across provinces. If we take the average of each variable for each province across all years, it will eliminate the effects of changes over years, thus, neglect the impact of time effects. Therefore, the best way to conquer this problem is to reshape the data wide so that each variable in each year becomes a new variable. As a result, cluster analysis will group provinces following similar time-trends of regional economic development into the same cluster[9].

[9] The cluster analysis results reflect an overall picture of macroeconomic conditions in 31 provinces over 2005-2015. Provinces with similar macroeconomic traits and following similar trend of economic change over the period are grouped into the same cluster. We do not allow clusters to move across years, in order to have fixed technology groups to construct a meta-frontier for the sample later on. Moreover, to our best

The first stage HAC cluster analysis results can be visually presented in a dendrogram. Ward linkage is used as the linkage method for hierarchical clustering.

Figure 4.3. Dendrogram for HAC



The x-axis on the dendrogram indicates the provinces and y-axis measures the dissimilarities among clusters

knowledge, it is unlikely for any bank in our sample to move the location of their headquarter and operation over the period, given that we only include city and rural commercial banks which are established to serve their local market. To the best of knowledge, our sample does not include any bank that has merged into another bank during the period, as we only select banks that were still active as of 2015. Given that we only include city and rural commercial banks that primarily operate within their own local market, any potential mergers and restructuring would be likely on the regional level, especially when central and provincial governments are largely involved in such decisions. For example, in 2021, Liaoning province decides to merger 12 local banks into one provincial level bank (Reuters, 2021). Such merger would not impact the cluster thus would still be included in the same technology group in the construction of meta-frontier, since the merged bank still operate within the province.

The x-axis on the dendrogram indicates the provinces, and the y-axis measures the dissimilarities among clusters. Each observation starts as its own cluster, and then they combine with each other until all groups are combined. Longer vertical lines indicate a more distinct separation between the groups. The dendrogram hints that clusters are clearly distinct when cutting the cluster tree at the point which gives us 4 clusters. This is confirmed by the statistics of two stopping rule indexes calculated below. As mentioned in the methodology section, in order to obtain distinct clusters, we would like to find one of the highest values of Calinski/Harabasz pseudo-F or one of the highest Duda/Hart Je(2)/Je(1) index values with a low corresponding pseudo-T-squared value which have much larger pseudo-T-squared values next to it. From both tables below, we can see that both indexes consistently point to 3 clusters as the most distinct solution (the optimal number of clusters and their corresponding indexes are underlined in the tables below).

Table 4.6. Calinski/Harabasz pseudo-F Index

No. of Clusters	Calinski/ Harabasz pseudo-F
2	24.92
3	26.59
4	22.26
5	20.11
6	19.42
7	18.67
8	18.47
9	17.9
10	17.34
11	17.05
12	17.07
13	17.05
14	17.04
15	16.52

Table 4.7. Duda/Hart Index

No. of Clusters	Duda/Hart	
	Je(2)/Je(1)	pseudo T-squared
1	0.5378	24.92
2	0.4438	12.53
3	0.7018	7.22
4	0.0478	19.91
5	0.6341	4.04
6	0.748	4.38
7	0.6288	3.54
8	0.5683	5.32
9	0.6562	2.1
10	0.3814	3.24

Therefore, we will choose 3 clusters as the optimal number of clusters based on the first-stage HAC results. Then the centres of these 3 clusters are used as the starting points for k-means clustering, where k is pre-defined to be 3. After k-means consolidation, we find that the provinces that belong to each cluster resulted from the first step do not change after second-step clustering. In other words, provinces have not moved between clusters after consolidation from the second-step k-means clustering process. Since both steps give exactly the same cluster solution, we will present the final results in table 4.8[10]. We can see that

[10] The cluster analysis results reflect an overall picture of macroeconomic conditions in 31 provinces over 2005-2015. Provinces with similar macroeconomic traits and following similar trends of economic change over the period are grouped into the same cluster. We do not allow clusters to move across years, in order to maintain fixed technology groups for the construction of a meta-frontier later on. Moreover, to our best knowledge, it is unlikely for any bank in our sample to move the location of their headquarter and operation over the period, given that we only include city and rural commercial banks which are established to serve their local market. Given that we only include city and rural commercial banks that primarily operate within their own local market, any potential mergers and restructuring would be likely on the regional level, especially when central and provincial governments are largely involved in such decisions. For

provinces are not equally split into clusters. Cluster two has the least number of provinces, cluster one has the greatest number of provinces, and cluster three is in-between. We attach the table of regional economic divisions by NBSC as well for comparison. We notice that only 31 mainland provincial-level divisions are included in both tables, reflecting the availability of mainland data from NBSC.

example, in 2021, Liaoning province decides to merge 12 local banks into one provincial-level bank (Reuters, 2021). Such a merger would not impact the cluster as they would still be included in the same technology group to construct the meta-frontier.

Table 4.8. Cluster Analysis Results

Cluster Analysis Results			
	Cluster 1	**Cluster 2**	**Cluster 3**
	Anhui	Beijing	Fujian
	Chongqing	Shanghai	Guangdong
	Gansu	Tianjin	Heilongjiang
	Guangxi		Inner Mongolia
	Guizhou		Jiangsu
	Hainan		Jilin
	Hebei		Liaoning
	Henan		Shandong
	Hubei		Zhejiang
	Hunan		
	Jiangxi		
	Ningxia		
	Qinghai		
	Shaanxi		
	Shanxi		
	Sichuan		
	Tibet		
	Xinjiang		
	Yunnan		
No. of Provinces	19	3	9
Total No. of Provinces	31		

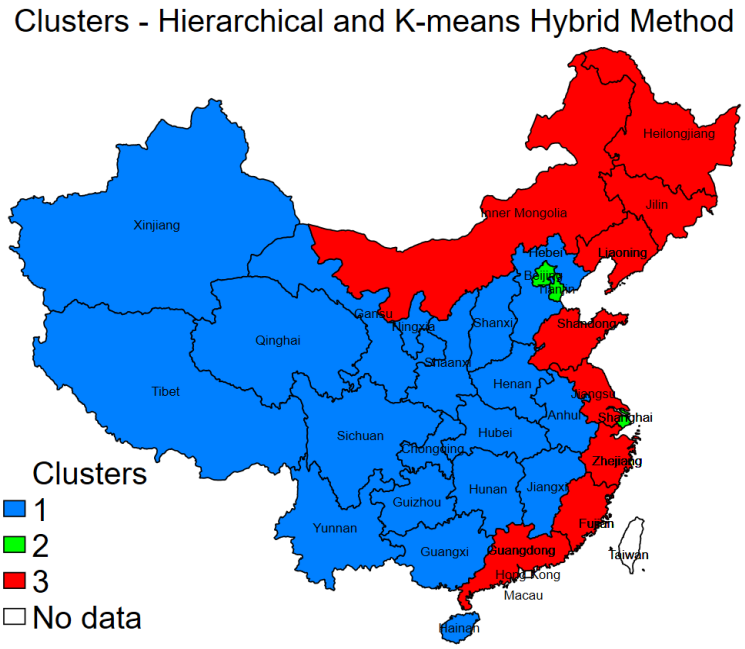
Table 4.9. Economic Regions by National Bureau of Statistics China Database (NBSC)

Economic Regions	Eastern Zone	Intermediate Zone	Western Zone
	Beijing	Anhui	Chongqing
	Fujian	Heilongjiang	Gansu
	Guangdong	Henan	Guangxi
	Hainan	Hubei	Guizhou
	Hebei	Hunan	Inner Mongolia
	Jiangsu	Jiangxi	Ningxia
	Liaoning	Jilin	Qinghai
	Shandong	Shanxi	Shaanxi
	Shanghai		Sichuan
	Tianjin		Tibet
	Zhejiang		Xinjiang
			Yunnan
No. of Provinces	11	8	12
Total No. of Provinces		31	

To visualise the geographical relationship of our clusters, we indicate them with different colours on the map below, where blue, green and red indicate clusters 1, 2 and 3, respectively and areas without data are not filled with colour. Please note that this map only contains provincial-level administrative divisions, and divisions of other levels are not included in the map. We can notice that, compared to the official division of economic regions by NBSC, cluster analysis has picked Beijing, Shanghai and Tianjin out of Eastern Zone and grouped them into a separate cluster (cluster 2). This result agrees with the fact that these three municipalities are economic and financial centres in China. They are of unique

economic status, so it is indeed expected that they are grouped together as a standalone cluster. The rest of the provinces which belong to the Eastern Zone still largely remain in the same cluster after conducting cluster analysis, with a few provinces such as Heilongjiang and Jilin entering the group while provinces such as Hainan and Hebei exiting the group. We can also observe a general trend of intermediate zone and western zone merging together into one big cluster. This might be due to the results of a series of economic reforms and investment in infrastructure, which narrowed the gap between the intermediate region and western regions over time. We have also attempted cluster analysis using different combinations of variables representing regional economic development. The resulting clusters solutions appear to be of very similar structure. However, some combinations give non-distinct clusters while others do not yield enough observations for each cluster to allow frontier analysis later on. Therefore, after exhausting all possible tuples of regional economic variables, we find one combination which gives acceptable numbers of observations to allow maximum likelihood estimation of cost frontiers later on.

Figure 4.4. Regional Division in Mainland China (excluding Hong Kong, Macao and Taiwan) – Results from Cluster Analysis

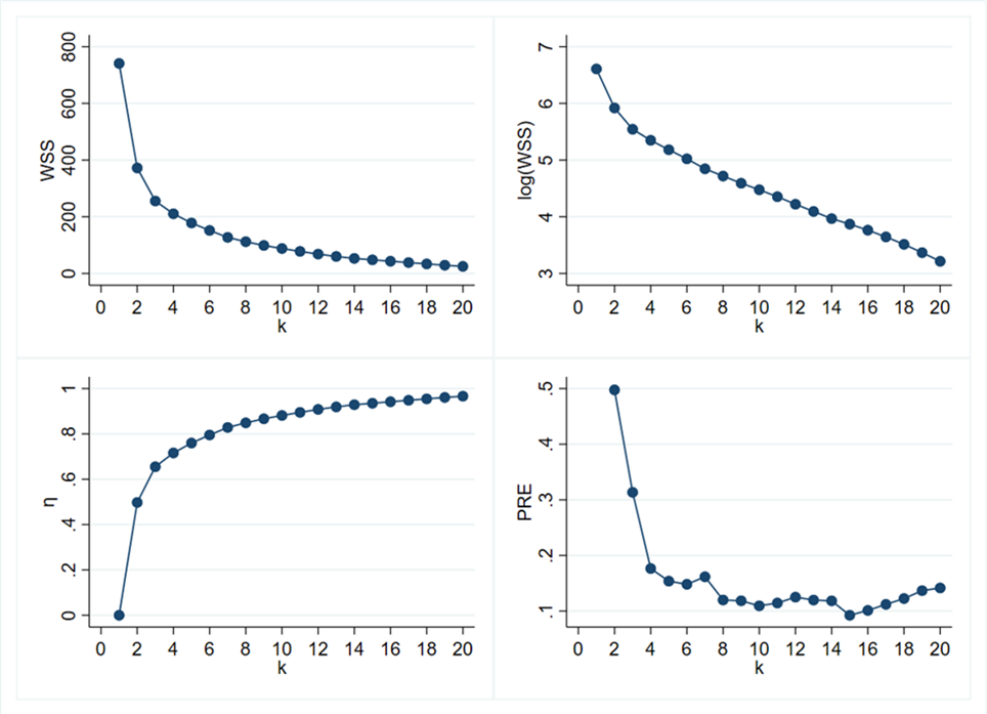


Blue, green and red indicate clusters 1, 2 and 3, respectively and areas without data are not filled with colour. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region. Data source: National Bureau of Statistics of China. Map source: Merryman (2008)

We can see that the way we determined the optimal number of clusters from the above method was by examining the dendrogram and two indexes of hierarchical clustering. Three was found to be the optimal number of clusters. Now we will use the second method - our proposed algorithm to determine the number of clusters. By running the first- and second-step clustering in one setting for 1 to 20 clusters, we plot the within sum of squares (WSS) from each final cluster solution for each chosen number of clusters, which is presented in the graph below. As we are looking for a kink in the curve for WSS, log(WSS) and η_k^2, we find that number three seems to be the optimal number of clusters. In other words, when the number of clusters is over 3, the marginal reduction in WSS is negligible compared to previous solutions. For the proportion of reduction in WSS with respect to the previous solution, indicated by PRE, the result is less clear. The proportional reduction of WSS from 2 to 3 clusters seems to be less than the proportional reduction of WSS from 1 to 2 clusters. Nevertheless, from

the first three graphs, we can still conclude that majority of the evidence from the reduction in WSS points clearly to 3 clusters as the optimal solution.

Figure 4.5. Determination of Optimal No. of Clusters-Elbow Method

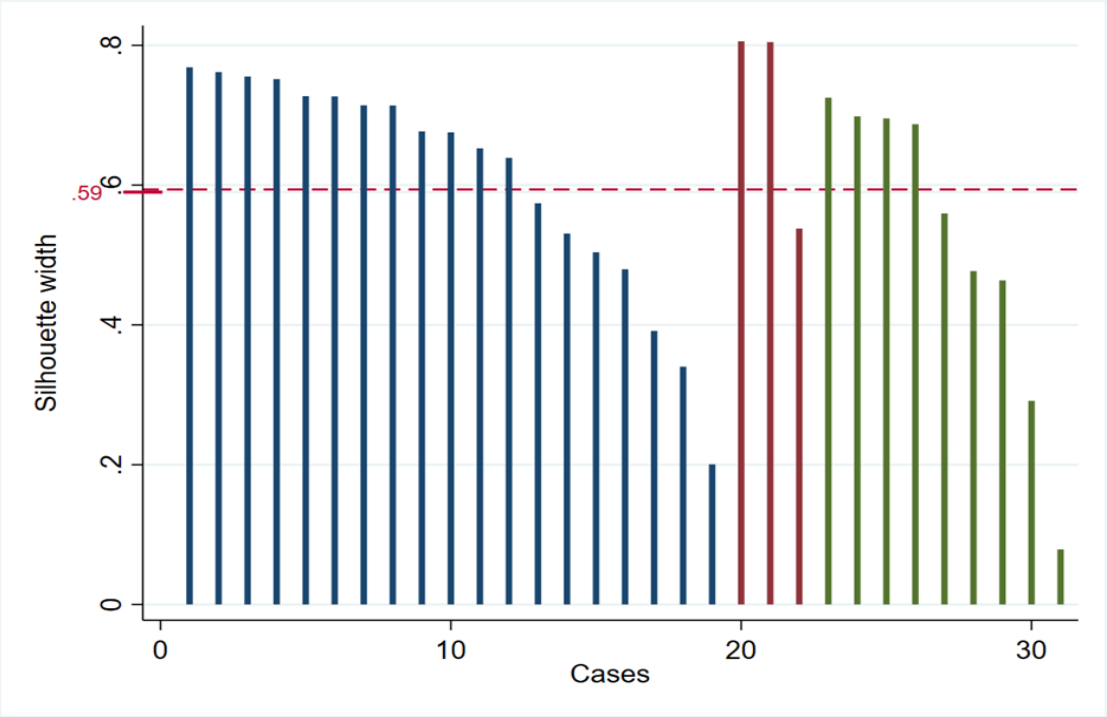


The x-axis indicates the number of clusters and the y-axis indicates the values of the within sum of squares WSS, the logarithm of the within sum of squares log(WSS), η, and the proportional reduction PRE

To further cross-validate this choice, we draw the silhouette graph. As mentioned in the methodology section, silhouette widths measure the ratio of dissimilarity within an object's own cluster over the dissimilarity between an object and its neighbour cluster. Therefore, it measures if one object is correctly assigned to its own cluster. A positive silhouette width indicates that the object is correctly assigned, while a negative silhouette width indicates that the object is closer to the neighbour cluster rather than the assigned cluster. Therefore, a

reassignment of the object to its neighbour cluster should be attempted. In our case, we can see that all provinces are correctly assigned to one of the three clusters as there is no negative silhouette width present in the graph, and the average silhouette width is well above zero at 0.59. Therefore, we can conclude that the cluster solution we have obtained is valid, and there is no outliner detected.

Figure 4.6. Cross-validation of Optimal No. of Clusters- Silhouette Widths



The x-axis indicates the province, the colour indicates the cluster, the y-axis indicates the silhouette width and the dashed line indicates the average silhouette width.

4.9 Empirical Results – Stochastic Meta-frontier Analysis

4.9.1 Descriptive Statistics

Table 4.10. Observations, Provinces and Clusters

Province	Clusters			Total
	1	2	3	
Sichuan	47	0	0	47
Hebei	38	0	0	38
Henan	37	0	0	37
Hubei	33	0	0	33
Guangxi	31	0	0	31
Chongqing	23	0	0	23
Shanxi	19	0	0	19
Shaanxi	14	0	0	14
Hunan	12	0	0	12
Anhui	11	0	0	11
Yunnan	11	0	0	11
Ningxia	10	0	0	10
Guizhou	7	0	0	7
Jiangxi	7	0	0	7
Xinjiang	7	0	0	7
Gansu	4	0	0	4
Shanghai	0	37	0	37
Tianjin	0	25	0	25
Beijing	0	22	0	22
Zhejiang	0	0	161	161
Guangdong	0	0	139	139
Shandong	0	0	123	123
Jiangsu	0	0	98	98
Liaoning	0	0	57	57
Fujian	0	0	36	36
Inner Mongolia	0	0	17	17
Heilongjiang	0	0	12	12
Jilin	0	0	12	12
Tibet	0	0	0	0
Qinghai	0	0	0	0
Hainan	0	0	0	0
Total	311	84	655	1,050

This table presents the number of observations in each province in our commercial banking data and the total number of observations in each cluster. Cluster 1- the Inland Region Cluster 2- the Eastern municipalities Cluster 3- the Northeastern and Coastal Region

Table 4.10 brings in our commercial banking data in each province. It presents the number of observations in each province and the total number of observations in each cluster. We can see that cluster 3 has the most observations and cluster 2 has the least. We can observe that Zhejiang, Guangdong and Shandong possess the highest volume of reported data in our sample. This could be either due to the higher number of regional commercial banks or the higher reporting rate of these banks.

Table 4.11 counts the observations in each cluster and its percentage of total observations each year. We can observe a general increase in the number of observations from 2005 to 2010, and the number becomes relatively stable from 2010 onwards. The reason for the increase can either be due to the period of transitioning of rural credit unions to rural commercial banks beginning from 2005, or it could be simply due to an increased rate of reporting or a decrease in missing values over time.

Table 4.11. Observation Counts

Year/Clusters	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
Cluster 1	13	26	25	19	26	32	35	35	35	35	30	311
Cluster 2	7	7	7	7	7	8	9	8	8	8	8	84
Cluster 3	31	49	52	53	53	64	69	71	71	75	67	655
Cluster 1	25%	32%	30%	24%	30%	31%	31%	31%	31%	30%	29%	30%
Cluster 2	14%	9%	8%	9%	8%	8%	8%	7%	7%	7%	8%	8%
Cluster 3	61%	60%	62%	67%	62%	62%	61%	62%	62%	64%	64%	62%
Total	51	82	84	79	86	104	113	114	114	118	105	1050

This table presents the number of observations in each cluster and its percentage of total observations each year. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

Table 4.12 on the following page offers summary statistics of inputs, input prices and outputs, as well as the risk control variable equity in each cluster.

Table 4.12. Summary Statistics of Frontier Variables

Input 1(X1)	Input 2 (X2)	Price1(P1)	Price2(P2)	Price3(P3)	Output 1(Y1)	Output 2(Y2)	Equity
				Cluster 1			
658.8217	1424.3400	0.0105	0.0264	27981.8700	30771.6300	243.5396	4571.3310
697.9180	1812.8790	0.0041	0.0135	28390.6000	38599.0600	321.8345	5129.0830
5179.3780	12649.0000	0.0400	0.0809	175691.4000	277207.8000	2044.9650	32941.3700
4.6196	2.7164	0.0023	0.0040	70.4224	11.4251	0.0679	10.4292
				Cluster 2			
2005.0310	5738.6280	0.0093	0.0296	113825.3000	116952.3000	837.1570	15214.2700
1605.7870	6552.0420	0.0037	0.0154	105796.8000	122493.6000	1018.2520	15841.4400
7380.5950	29199.0200	0.0247	0.0942	493963.7000	628169.9000	5307.4250	79422.4100
1.5230	0.0307	0.0047	0.0032	16.8741	30.3653	0.0460	61.7040
				Cluster 3			
791.1904	1776.8940	0.0111	0.0281	35964.2300	31698.6600	380.8415	5186.2110
1479.0290	3755.2720	0.0036	0.0141	63567.2800	58137.3900	1108.1030	8425.4760
20542.0500	44926.3500	0.0306	0.1189	809670.2000	618508.1000	20477.9300	110173.4000
3.3686	0.2150	0.0022	0.0016	23.2912	147.3511	0.6000	88.2244
				Total			
849.0913	1989.4090	0.0108	0.0277	39828.8200	38244.3700	376.6792	5806.3340
1353.3690	3795.8310	0.0038	0.0141	64284.1200	65376.7500	948.9063	8925.7000
20542.0500	44926.3500	0.0400	0.1189	809670.2000	628169.9000	20477.9300	110173.4000
1.5230	0.0307	0.0022	0.0016	16.8741	11.4251	0.0460	10.4292

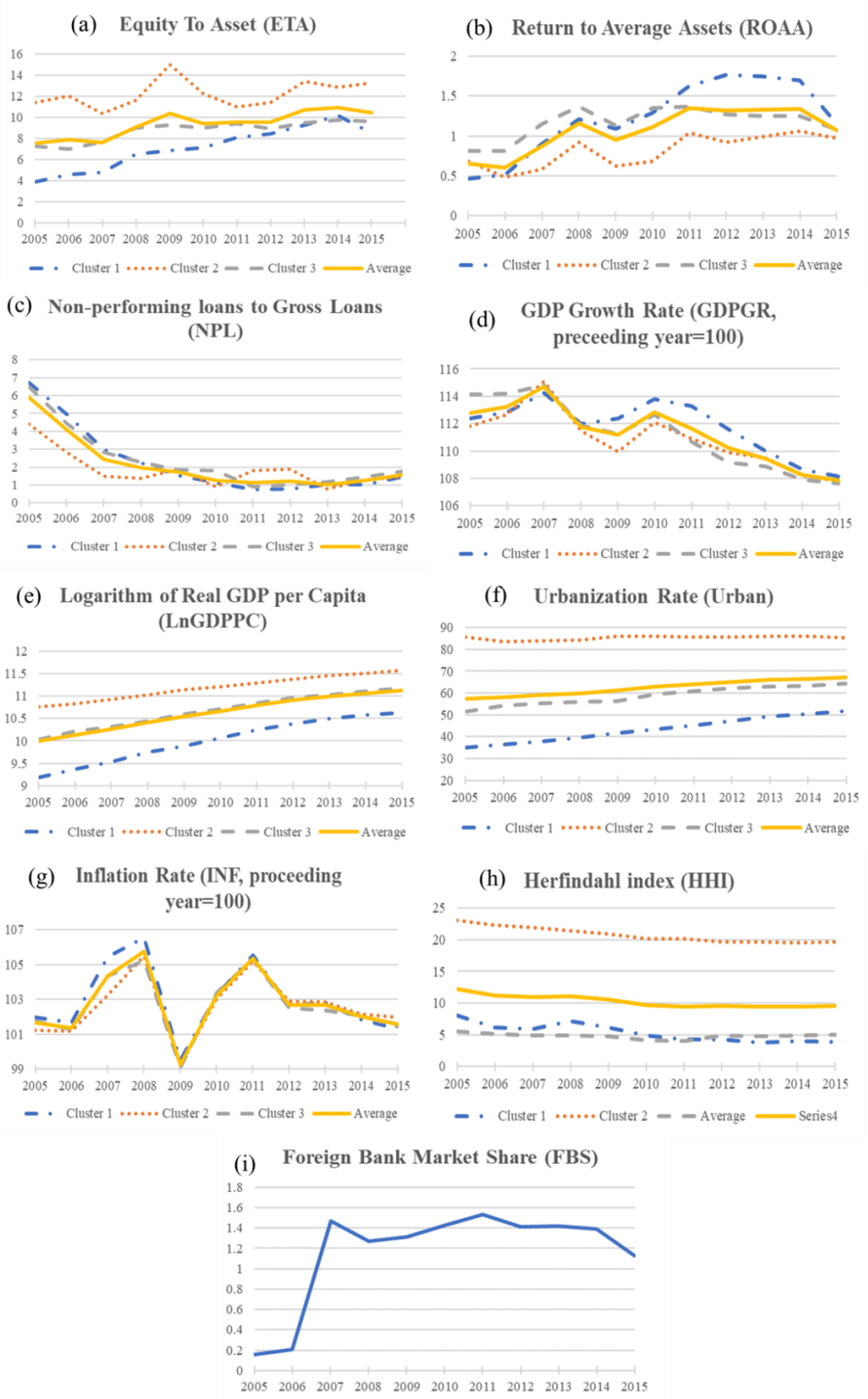
The first to fourth figures represent mean, standard deviation, maximum and minimum for each cluster and for the whole sample. Notes: X1= Personnel Expenses; x2= Total Interest Expense; x3= Other Operating Expenses; p1= Personnel Expenses/ Total; p2= Assets Total Interest Expense/Customer Deposits; p3=Other Operating Expenses/ Total Assets; y1=Net Loans; y2= Other Earnings Assets; y3=Total Non-Interest Operating Income. . Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

Table 4.13. Summary Statistics of Environmental Variables

Stats	ETA_{pt}	$ROAA_{pt}$	NPL_{pt}	$GDPGR_{pt}$	$LnGDPPC_{pt}$	$Urban_{pt}$	INF_{pt}	HHI_{ct}	FAS_t
					Cluster 1				
Obs	311	311	311	311	311	311	311	311	311
Mean	7.5315	1.3120	1.8582	111.6026	10.1088	44.5245	102.9148	4.9983	1.2330
Std. Dev.	5.2754	1.1350	1.9153	2.4690	0.4781	7.1073	1.9654	1.2074	0.4097
Min	43.9825	8.8020	14.5800	117.1000	10.8602	60.9416	108.5000	8.0684	1.5335
Max	1.7530	-0.0020	0.0300	103.1000	8.6500	27.4526	97.9000	3.8060	0.1546
					Cluster 2				
Obs	84	84	84	84	84	84	84	84	84
Mean	12.2568	0.8283	1.7976	110.7762	11.2054	85.3423	102.6452	20.6887	1.1778
Std. Dev.	3.4241	0.3025	1.4015	3.1573	0.2731	4.1354	1.7342	1.1708	0.4611
Min	18.8325	1.4447	6.9300	117.4000	11.5876	89.6083	105.8000	23.1036	1.5335
Max	4.1310	-0.3900	0.5517	106.9000	10.4852	75.1440	98.5000	19.5405	0.1546
					Cluster 3				
Obs	655	655	655	655	655	655	655	655	655
Mean	8.9310	1.1950	2.0502	110.7802	10.7478	59.6273	102.6689	4.7573	1.2313
Std. Dev.	2.4961	0.2767	2.5842	2.7292	0.3722	6.4268	1.7708	0.4081	0.4046
Min	21.2736	1.6938	29.2333	123.8000	11.3890	68.7068	105.8000	5.5445	1.5335
Max	2.8867	-0.1660	0.4300	103.0000	9.7008	45.0043	97.7000	3.9731	0.1546
					Total				
Obs	1050	1050	1050	1050	1050	1050	1050	1050	1050
Mean	8.7825	1.2003	1.9731	111.0234	10.5951	57.2112	102.7398	6.1032	1.2275
Std. Dev.	3.8054	0.6711	2.3266	2.7158	0.5238	12.5205	1.8296	4.3785	0.4108
Min	43.9825	8.8020	29.2333	123.8000	11.5876	89.6083	108.5000	23.1036	1.5335
Max	1.7530	-0.3900	0.0300	103.0000	8.6500	27.4526	97.7000	3.8060	0.1546

The table shows summary statistics of environmental variables in each cluster. ETA is the equity to total assets ratio, ROAA is the return of average assets, NPL is the non-performing loan to gross loan ratio, LnGDPPC is the logarithm of the real GDP per capita, Urban is the urbanisation rate, INF is the inflation rate, HHI is the Herfindahl index, and FAS is the foreign asset share. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

Figure 4.7. Microeconomic and Macroeconomic Environmental Variables



To analyse chosen industry-specific and macroeconomic variables, plots of mean values over the years give us a better picture of economic and industry-wide trends than just looking at summary statistics. For the majority of the environmental variables above, we can observe identifiable gaps among clusters. This further validates our cluster analysis results in the sense that these clusters do have distinct characteristics in terms of the environment. In terms of the banking sector-specific environment, they are obtained by calculating the averages of each variable for all banks within each cluster. In terms of the capital adequacy level, we can observe a general trend of increase in capital levels in the past decade. Cluster 1 has the lowest capital adequacy, cluster 2 has the highest capital adequacy, and cluster 3 lies in the middle. According to Kosmidou (2008), capital acts as a buffer for unexpected shocks. A high level of capital adequacy would signal investors a lower risk of bankruptcy, leading to a potentially lower cost of funding and an increase in cost efficiency. We can first observe that the overall capital level peaked slightly in 2009 in the midst of the global financial crisis and then flattened out afterwards, whilst still maintaining steady growth. Cluster 2 and cluster 1 have the highest level and lowest level of capital adequacy, respectively, while cluster 3 lies in the middle. We know that cluster 2 represents the most developed financial centres in China, and cluster 1 contains the less developed provinces. We find that banks in more developed regions tend to hold higher levels of capital, which may be due to more stringent enforcement of capital regulations in these areas.

In terms of the profitability level measured by Return on Average Assets (ROAA), we find that cluster 2 has a lower level of capital than clusters 1 and 3. According to Saunders and Cornett (2018), banks tend to reduce the level of risky assets in their asset profile in order to achieve a higher level of capital, thus potentially affecting bank profitability.

The levels of non-performing loans (NPLs) in the three clusters are of a similar scale and have a decreasing trend in the past 11 years. The highest level of NPLs is at the beginning of the period of around 6%, and it gradually decreases to around 1% in 2010 and remains at 1% level till 2013 with a slight rise from 2013 to 2015. The NPL level in cluster 2 has a slightly lower level than average prior

to 2009-2010 but slightly higher than average in the few years after, which could be linked to an increased loan default rate during the global financial crisis, which feeds into the non-performing loan level in the following few years. The effect of the global financial crisis, however, does not affect cluster 1 and cluster 3 as much as it does on cluster 2.

In terms of macroeconomic conditions, the trends of six indicators are presented in the diagrams above. GDP Growth rates (GDPGR) fluctuate over the study period with some degree of dispersion across clusters. GDPGRs increase from 2005 to 2007 then drop to the lowest level in 2009. There is a slight increase briefly from 2009 to 2010, but GDPGRs declined continuously from 2010 onwards. The cluster with the highest level of GDP growth rate is cluster 3 before 2007 but becomes cluster 1 after 2007.

Regarding the logarithm of real GDP per capita (LnGDPPC), there is a steady and continuous trend of increase. Cluster 2 remains the area with the highest LnGDPPC, and cluster 1 remains the area with the lowest LnGDPPC over the whole period, where cluster 3 lies in-between at the average level.

Urbanisation rates exhibit large dispersions among three clusters. Cluster 2 has an urbanisation rate of more than 80 percent, which is much higher than clusters 2 and 3. Cluster 1 has the lowest level of urbanisation rate, and cluster 3 lies slightly below the national average. The urbanisation rate remains stable for cluster 2 over the 11 years but increases steadily for clusters 1 and 3.

Inflation rates, measured by yearly changes in Consumer Price Index, do not deviate from each other much in terms of yearly means for each cluster. The inflation rates go up from 2006 to 2008 then drop sharply to the lowest point during the study period in 2009, where deflation was found. Inflation rates go up continuously from 2009 to 2011 to around 5%, then decline to around 3% and continue a gradual decline reaching around 1%-2% in 2015. We can observe a small level of difference among these 3 clusters from 2005 to 2008, where cluster 1 has the highest and cluster 2 has the lowest inflation rate. There is ample literature on the impact of inflation on bank profitability (Revell, 1979, Athanasoglou et al., 2008, Flamini et al., 2009, Kosmidou, 2008, Adelopo et al., 2018) and limited theory and literature on how inflation affects bank efficiency.

It is generally believed that high inflation induces banks to compete through excessive branch networks, thus reducing efficiency. Berger (1995) supports that excessive expansion and opening of branches in areas with insufficient demand and few customers is a source of inefficiency. Demirguc-Kunt et al. (2003) find evidence that high inflation increases bank costs. According to Pasiouras et al. (2009) and Kasman and Yildirim (2006), a higher inflation rate causes higher costs and lower profits. In the literature which included inflation as a determinant for bank profitability, the theory suggests that whether inflation facilitates or hinder profitability depends on whether bank wages and cost increases faster than inflation. Staikouras and Wood (2004) and Kosmidou (2008) argue that inflation can affect cost through price of labour demand for financial services from banks. For example, unanticipated inflation may cause difficulties in loan repayment for borrowers which may lead to precipitate ending of loan contracts. High and uncertain inflation may also cause difficulties in loan assessment. Adelopo et al. (2018), on the other hand, argue that higher inflation might encourage entrepreneurial behaviour, thus may encourage lending, thus positively affect banks' loan business. However, the authors did not mention the quality of these potential entrepreneurial loans. Staikouras and Wood (2004) believe that inflation also affects bank's investments in property markets in which the profit and loss are contingent on monetary policy. Athanasoglou et al. (2008), Flamini et al. (2009) and Adelopo et al. (2018) believe that the impact of inflation on bank profitability depends on if inflation can be correctly predicted and passed on to customers. Uncertainty will negatively affect banks' ability to predict inflation precisely thus increasing the chance that banks have to bear higher operating cost. Since the performance measure in this study is cost efficiency with intermediation model regarding banks as intermediaries using capital, labour and deposits to produce loans and other earnings assets, where loans are measured using net loans as gross loans subtracted by bad loans, it is expected that inflation will more likely have a significant and positive impact on cost inefficiency.

Herfindahl index (HHI_{ct}) is calculated as the sum of squared assets shares. The subscript of HHI_{ct} indicates that HHI is cluster and time specific. The reason that this variable is not province-specific is due to the fact that HHI would be

much less accurate on the provincial level since branch-level data is not available; therefore, the data on total assets is only available on the aggregate level for each bank. Kasman and Yildirim (2006) use HHI as a proxy for concentration and believe that high concentration can be linked to either higher or lower cost depending on the cause. In other words, the relationship between cost and market concentration can go in either the same or opposite direction, according to literature. If a high concentration results from an exercise of market power, the banks are likely to operate under a non-optimal condition, and higher concentration is associated with higher cost (Leibenstein, 1966). When higher concentration is associated with better performance and efficiency, higher concentration could be linked with lower cost (Demsetz, 1973). HHI has an overall decreasing trend, which indicates increasing competition in the Chinese regional commercial banking sector. Cluster 2 has the highest level of competition, with HHI starting from 23% and gradually declining to around 20%. HHI in cluster 1 drops from 8% to around 4%, while in cluster 3, it stays around 5% throughout the period. The overall decline of HHI may be due to entry of new banks into their respective local market during the period.

Foreign bank market share (FBS) witnessed a sharp increase at the beginning of the period then remains relatively stable at a later stage of the period. It is defined as the percentage asset share of foreign banks with respect to the total assets in the banking sector for each year. Foreign banks here are defined as banks with foreign investment of more than 25%. This measure proxies the level of competition coming from a foreign bank in the market. The sharpest increase can be seen from 2006-2007 from around 0.2% to 1.4% followed by a slight decline from 2007-2008. After that, we can observe a gradual but steady increase from 2008, reaching the highest foreign bank share at 1.5% in 2011. From 2011 to 2015, foreign bank market share slowly declines to reach 1.1% at the end of the period. Since China opened its banking market to foreign banks in 1979, the restrictions on foreign banks are being relaxed over time. Most of these restrictions imposed on foreign banks are lifted within five years' China's admittance to WTO (Kwan, 2009). According to Shen et al. (2009), there are

four main ways foreign banks choose to enter the Chinese banking market. The first way is to establish foreign banks branches and subsidies. The second way is to establish a new bank fully funded by foreign banks. The third way is for foreign banks to launch joint-ventures with Chinese banks where they hold more than 25% of total shares. The last method is to through foreign strategic investors (FSI), a policy established in 2001. The policy originally restricted a foreign bank's investment to be no more than 15% in a single bank, and then this restriction was relaxed in 2003 to 20% by CBRC in the updated guidance, with aggregate total foreign investment limited to less than 25%. The 25% limitation only applies to unlisted banks, while listed banks with more than 25% foreign shareholdings are still regarded as Chinese banks (Shen et al., 2009). Further reforms on relaxation of the regulation of Administration of Foreign-funded banks was conducted in 2015. According to Luo et al. (2017), the presence of foreign banks in China is highly unevenly distributed, as a disproportionally high number of foreign banks chose to establish branches in the eastern costal region, especially in China's financial centre, Shanghai[11].

It is generally believed that foreign banks will bring in healthy competition into the Chinese banking industry. However, some believe that foreign banks can also bring in a degree of threat to the stability of the Chinese financial market. Some are concerned that with more diverse products, wider international network and state-of-the-art risk management capabilities, domestic banks might be forced out of the market (Li et al., 2015). Bonin and Huang (2001) find that the opposite tends to be true as foreign banks' entry acts as a form of foreign direct investment (FDI) aiming at a long-term benefit. At the same time, foreign bank entry is subject to restrictions under regulation and supervision of CBRC, making it difficult to expand. There are a number of studies focusing on the impact of foreign bank entry on the competition and performance of Chinese domestic banks. Shen et al. (2009) solely focus on the impact of foreign bank presence on Chinese bank performance. MacroFP is calculated as the total

[11] Due to limitations of the database used in this thesis, only aggregate data is available for foreign banks and there is a lack of branch-level data. Therefore, we will focus on how the increase in foreign banks presence in China impact performance of Chinese domestic banks over the study period.

number of Chinese banks with foreign presence divided by the total number of Chinese banks, and MicroFP is calculated as the ratio of shares held by FSI in a bank to total shares in a bank. A wholly foreign-funded bank, in this case, would have 100% MicroFP. They find that MacroFP improves the profitability of Chinese banks but does not reduce cost. They attribute the improved bank performance to an enhanced competitive environment or the introduction of new technology. However, when foreign bank presence is measured on the micro level using MicroFP, no impact on profit or cost has been found. They ascribe it to many reasons, including the fact that due to relatively short study periods, the short-term cost of staffing and investment in software to accommodate new technologies might outweigh the benefits. Xu (2011) criticises the inaccuracy of using aggregate data for foreign bank presence in the previous literature and proposes a spatially disaggregated measurement for foreign bank presence on the local level. They find that disaggregated foreign exposure reduces interest margins, increases costs but increases non-interest margins in domestic banks. They conclude this result is in line with the theory that increased foreign presence does bring in healthy competition, improved efficiency with spill-over effects. It also increases the costs for domestic banks in the short run due to new investment in capital to adapt to the sudden competitive pressure from foreign entry. Berger et al. (2009) find that foreign bank presence increases efficiency in state-owned banks. Leung and Chan (2006) also find a positive relationship between foreign bank entry and local bank efficiency.

In our study, we focus on the presence of foreign-owned banks rather than the foreign investment in domestic banks as a measure for foreign presence, and it is measured on the national level. A negative relationship is expected between foreign bank presence and bank cost efficiency.

The correlation table below shows the correlation coefficients and their significance among microeconomic and macroeconomic environmental variables. The majority of the exogenous variables are of weak correlations, with a few indicating stronger or moderate correlations. ETA and ROAA are moderately correlated with a correlation coefficient of 0.6, which may indicate that higher capital requirements can be associated with higher profitability in the

banking sector. There is also a moderately negative correlation between NPL and FAS. This relationship is expected by observing the upward trend of NPL and declining of FAS in the diagrams. GDP per capita has a relatively higher correlation with the urbanization of 0.8133, as both are measurements of regional development.

Table 4.14. Correlation Table of Environmental Variables

	ETA	ROAA	NPL	GDPGR	LnGDPPC	URBAN	INF	HHI	FAS
ETA	1.0000								
ROAA	0.6000*	1.0000							
	(0.0000)								
NPL	-0.0897*	-0.2223*	1.0000						
	(0.0036)	(0.0000)							
GDPGR	-0.2680*	-0.0978*	0.2256*	1.0000					
	(0.0000)	(0.0015)	(0.0000)						
LnGDPPC	0.4351*	0.1289*	-0.3706*	-0.5935*	1.0000				
	(0.0000)	(0.0000)	(0.0000)	(0.0000)					
URBAN	0.4614*	-0.0517	-0.1329*	-0.3651*	0.8133*	1.0000			
	(0.0000)	(0.0943)	(0.0000)	(0.0000)	(0.0000)				
INF	-0.0617	0.1232*	-0.1020*	0.2507*	-0.1157*	-0.0794	1.0000		
	(0.0455)	(0.0001)	(0.0009)	(0.0000)	(0.0002)	(0.0100)			
HHI	0.2195*	-0.2139*	0.0456	0.0183	0.2424*	0.6025*	-0.0358	1.0000	
	(0.0000)	(0.0000)	(0.1400)	(0.5539)	(0.0000)	(0.0000)	(0.2470)		
FAS	0.1778*	0.3025*	-0.5227*	-0.2487*	0.4089*	0.1513*	0.3597*	-0.1339*	1.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	

This table shows the Pearson pairwise correlation coefficients between environmental variables with p-values in parentheses. Statistical significance is marked with *, corresponding to p-value of less than 0.01. ETA is the equity to total assets ratio, ROAA is the return of average assets, NPL is the non-performing loan to gross loan ratio, GDPGR is the GDP growth rate, LnGDPPC is the logarithm of real GDP per capita, URBAN is the urbanisation rate, INF is the inflation rate, HHI is the Herfindahl index, and FAS is the foreign asset share.

4.9.2 Pooled results

Three versions of the pooled model are presented in the table below. The differences in the three models are in the inefficiency model. Pooled 1 only includes microeconomic variables, pooled 2 only includes macroeconomic variables and pooled 3 contains both microeconomic and macroeconomic variables. The majority of the variables in the frontier are significant, suggesting a well-specified model. Coefficients and significance are very similar for all three models. An interesting phenomenon is that when including only microeconomic or macroeconomic variables in the pooled model, we could not observe any significance of the exogenous variables. The significance was only

revealed when both categories were included in the model.

The significant and negative coefficient of ETA suggests that increased capital levels in the banking sector tend to assist the improvement of bank cost efficiency.

The coefficient of inflation is found to be significantly positive, indicating that higher inflation contributes to higher cost and lower efficiency, as found by Kasman and Yildirim (2006), Pasiouras et al. (2009) and Demirguc-Kunt et al. (2003) in their European and cross-country studies. It is an indication that Chinese banks might exercise excessive expansion of bank branch networks in response to higher inflation. It might also be because inflation has increased bank operating costs, such as higher labour costs. As Staikouras and Wood (2004) have mentioned, during a high inflation period, borrowers may have difficulty with cash flow which may result in premature termination of loan agreements or an increase in loan losses, thus causing reduced loan volume and increased loan losses, reflected as reduced net loans in our model.

The sign of HHI is significant and positive, indicating that higher concentration reduces the cost efficiency of Chinese commercial banks. As suggested by Kasman and Yildirim (2006), this could be an indication that the concentration in the Chinese banking industry is more likely to come from market power rather than superior management. Banks in the areas with higher concentration and market power have less pressure for cost control (Maudos et al., 2002), thus resulting in lower cost efficiency. This result is in line with Kasman and Yildirim (2006), Maudos et al. (2002) and Goddard et al. (2014).

Foreign bank presence shows strong significance in its positive impact on cost efficiency in the pooled model. As measured by foreign asset shares in the Chinese banking sector, FAS has a significant and positive sign. Although Berger et al. (2009) and Xu (2011) find some evidence that foreign bank presence increases bank efficiency, their results are either partial or indirect. Berger et al. (2009) found that foreign presence increases efficiency in state-owned banks. Our study complements their results by finding that foreign bank presence also increases cost efficiency in the city and rural commercial banks. Xu (2011) conclude that foreign bank presence increases healthy competition

and increases domestic bank efficiency by finding that foreign bank presence is linked with increased costs, reduced interest margin and increased non-interest margin in domestic banks. They attribute the increased cost to initial investment in new technology in domestic banks in response to suddenly increased competition brought in by the foreign banks, and this increase in cost should be present in the short run. As our study covers the period 2005 to 2015, which can be seen as a continuation of Xu (2011)'s study period 1999 to 2006, foreign bank presence has been fully established in our study timeframe, and the role of foreign banks has evolved, especially after deregulations. Hence, it would not be of surprise but of interest if the impact of foreign presence has evolved over time. Another point that needs to be noted is that the effect of foreign presence on cost is different from its effect on cost efficiency. Cost efficiency measures how banks utilise capital and deposits to produce loans and other earnings assets, which is more complex than cost alone. Therefore, our results cannot be directly compared with studies that did not examine cost efficiency using SFA.

Table 4.15. Pooled Frontier Estimates

	(1)		(2)		(3)	
	Pooled 1		**Pooled 2**		**Pooled 3**	
			Frontier			
lnp1	0.6996***	(0.2072)	0.6235***	(0.2114)	0.5732***	(0.2094)
lnp2	-0.4295***	(0.1202)	-0.4053***	(0.1195)	-0.4141***	(0.1188)
lnp1lnp1	-0.0291	(0.0364)	-0.0490	(0.0370)	-0.0474	(0.0365)
lnp2lnp2	-0.0250	(0.0186)	-0.0190	(0.0188)	-0.0208	(0.0185)
lnp1lnp2	-0.0399**	(0.0163)	-0.0367**	(0.0164)	-0.0384**	(0.0161)
lny1	0.7199***	(0.1082)	0.7171***	(0.1105)	0.7435***	(0.1095)
lny2	0.6122***	(0.1088)	0.6297***	(0.1080)	0.6210***	(0.1075)
lny3	0.0423	(0.0610)	0.0591	(0.0606)	0.0332	(0.0603)
lny1lny1	0.2174***	(0.0126)	0.2237***	(0.0128)	0.2200***	(0.0127)
lny2lny2	0.1171***	(0.0129)	0.1147***	(0.0130)	0.1146***	(0.0127)
lny3lny3	0.0106***	(0.0038)	0.0101**	(0.0039)	0.0098**	(0.0039)
lny1lny2	-0.1823***	(0.0123)	-0.1854***	(0.0124)	-0.1833***	(0.0123)
lny1lny3	-0.0297***	(0.0064)	-0.0333***	(0.0064)	-0.0318***	(0.0062)
lny2lny3	0.0154**	(0.0071)	0.0201***	(0.0071)	0.0191***	(0.0070)
lny1lnp1	0.0387*	(0.0234)	0.0396*	(0.0237)	0.0449*	(0.0234)
lny1lnp2	0.0328***	(0.0106)	0.0321***	(0.0107)	0.0324***	(0.0105)
lny2lnp1	-0.0995***	(0.0211)	-0.1050***	(0.0212)	-0.1018***	(0.0206)
lny2lnp2	0.0403***	(0.0105)	0.0438***	(0.0106)	0.0419***	(0.0103)
lny3lnp1	-0.0004	(0.0124)	0.0065	(0.0123)	0.0005	(0.0122)
lny3lnp2	-0.0169***	(0.0062)	-0.0187***	(0.0063)	-0.0176***	(0.0061)
lneq	-0.2209***	(0.0469)	-0.2451***	(0.0497)	-0.2319***	(0.0490)
lneqlneq	0.0277***	(0.0062)	0.0289***	(0.0066)	0.0285***	(0.0064)
t	0.0725***	(0.0068)	0.0891***	(0.0099)	0.0893***	(0.0100)
t^2	-0.0081***	(0.0010)	-0.0099***	(0.0013)	-0.0100***	(0.0013)
Constant	-1.2272*	(0.6595)	-1.4335**	(0.6575)	-1.6409**	(0.6570)
			Inefficiency Model			
ETA_{pt}	-0.1588	(0.1458)			-0.0478**	(0.0240)
$ROAA_{pt}$	-0.4581	(0.4477)			-0.1331	(0.1180)
NPL_{pt}	0.0472	(0.0443)			0.0116	(0.0099)

$GDPGR_{pt}$			0.0203	(0.0212)	0.0057	(0.0121)
$GDPPCapita_{pt}$			-0.1877	(0.2081)	0.0414	(0.1441)
$Urban_{pt}$			-0.0045	(0.0100)	-0.0023	(0.0059)
INF_{pt}			0.1013	(0.0773)	0.0725*	(0.0391)
HHI_{ct}			0.0318	(0.0249)	0.0186*	(0.0106)
FAS_t			-0.5599	(0.3689)	-0.2983**	(0.1361)
Constant	0.0809	(0.3769)	-10.8559	(9.0800)	-8.0831	(4.9822)
σ_u						
Constant	-1.9704**	(0.9445)	-2.5134***	(0.7246)	-3.0282***	(0.4858)
σ_v						
Constant	-5.2876***	(0.0914)	-5.2914***	(0.1010)	-5.3186***	(0.0940)
N	1050		1050		1050	
ll	911.8143		913.1378		925.7662	

Notes: Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Pooled 1 contains only microeconomic variables, Pooled 2 contains only macroeconomic variables, and pooled 3 model contains both microeconomic and macroeconomic variables in the other two models. ETA_{pt} is the equity to total assets ratio, $ROAA_{pt}$ is the return of average assets, NPL_{pt} is the non-performing loan to gross loan ratio,$GDPGR_{pt}$ is the GDP growth rate, $GDPPCapita_{pt}$ is the real GDP per capita,$Urban_{pt}$ is the urbanisation rate, INF_{pt} is the inflation rate, $HHI_{ct}$is the Herfindahl index, and FAS_t is the foreign asset share. σ_u and σ_v are standard errors of two error components, namely the cost inefficiency component with distribution as $N^+(\mu, \sigma_u^2)$ and random error component v^2 with distribution as $N(0, \sigma_v)$.

4.9.3 Cluster-Specific Frontier Results

Table 4.16. Cluster Specific Frontier Estimates

	(1) Cluster 1		(2) Cluster 2		(3) Cluster 3	
			Inefficiency Model			
ETA_{pt}	-0.0859***	(0.0275)	-0.0979	(0.1730)	-0.2094	(0.1643)
$ROAA_{pt}$	-0.0813	(0.0878)	-1.4984	(2.3936)	0.3475	(0.3906)
NPL_{pt}	0.0285*	(0.0151)	0.0262	(0.0981)	0.0632	(0.0488)
Constant	0.4120**	(0.1666)	1.1845	(1.8878)	0.3669*	(0.1916)
σ_u						
Constant	-4.2571***	(0.8485)	-2.7191	(1.7995)	-2.6141***	(0.8624)
σ_v						
Constant	-5.3050***	(0.2526)	-6.8753***	(0.7309)	-5.3717***	(0.1113)
N	311		84		655	
ll	315.5896		118.6171		606.9226	

This table shows the cluster specific frontier estimates. ETA_{pt} is the equity to total assets ratio, $ROAA_{pt}$ is the return of average assets, NPL_{pt} is the non-performing loan to gross loan ratio . σ_u and σ_v are standard errors of two error components, namely the cost inefficiency component with distribution as $N^{+}(\mu, \sigma_u^2)$ and random error component v^2 with distribution as $N(0, \sigma_v)$. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

One important reason to justify the meta-frontier method is due to the fact that there is group differentiation in the sample. That is, it is not appropriate to estimate the x-efficiencies of the whole sample using a single pooled frontier as individual groups within the sample follow different technologies. To test whether the sample can be estimated using a single frontier, we calculate the log-likelihood for both the pooled model and cluster-specific frontiers. From the tables above, we can see that the log-likelihood of the single frontier with only microeconomic variables is 911.8143. The sum of the log-likelihood of all cluster-specific frontiers is 1041.1293. The likelihood ratio test suggests that

cluster-specific frontiers are preferred over a pooled single frontier. Therefore, we can confirm that our clusters indeed operate under different technologies[12]. The fact that the significance of the input prices, output variables and their interactions differ across clusters gives a hint that different clusters might follow different technologies.

In terms of technological change denoted by variable t and t^2, apart from the coefficient of t^2 in cluster 4, the signs of the coefficients of t are significantly positive and significantly negative for t^2, which indicate a non-linear technological change effect that reduces with time.

$$\frac{\partial lnTC}{\partial t} = b + a * t \qquad t = 1,2, \dots \dots 11 \tag{4.28}$$

In the above equation, a represents the coefficient of t and b the coefficient of t^2. If a is negative, as indicated in our results, the shape of the curve describing the relationship between t and the total cost would be a reversed U-shape curve. That is, the cost increases with time until a turning point then decreases with time. The turning point is calculated as –b/2a. The turning points for the 3 clusters are therefore 4.32, 8.41 and 4.57, respectively. It shows that Chinese banks in clusters 1 and 3 tend to become inefficient in the beginning half of the years; however, they generally increase their efficiency after the year 2010. This might be due to the banking reforms that happened at the beginning of the study period, as well as the cost spent on recovering and dumping the non-performing loans. After 2010, banks have generally completed the majority of the NPL recovering. Cluster 2 shows an increase in efficiency at a much later stage. This might be because banks headquartered in Shanghai and Beijing are the usually largest banks with the highest level of complexity in their products and activities. Therefore, the reform tends to be more difficult and takes longer than for the banks headquartered in the rest of the country.

In the meantime, the risk control variable equity also shows a nonlinear effect on banks' total costs. The relationship of the total cost with equity is the opposite

[12] The log likelihood ratio test is calculated as -2*(ll_OLS - ll_SFA) and the statistics for the 3 clusters are all larger than the critical value of the mixed chi-square distribution. The table of critical values can be found in Kodde and Palm (1986).

of the relationship with t. That is, the relationship is U-shaped. It indicates that total cost reduces when the equity level increases until a turning point where the total cost starts to increase when equity continues to increase. This finding is in line with the theory that a higher level of equity indicates less risk in the loan portfolio, therefore, less cost on monitoring and recovering bad loans (Lozano-Vivas et al., 2002). However, there is also a trade-off between capital level and profitability. Therefore, it is possible for banks to find an optimal level of equity to increase their efficiency.

The significance of the microeconomic variables can only be found in cluster 1, while microeconomic variables do not seem to affect cost efficiency in clusters 2 and 3. This is also an indication that banks in different clusters can increase their efficiency in different ways. In cluster 1, a higher level of capital ratio ETA tends to reduce inefficiency, and a higher level of non-performing loans NPLLOAN tends to increase inefficiency. Berger and DeYoung (1997) argue that a higher level of bad loans is likely to prompt banks to consume large amounts of resources to deal with potential loan losses, which will inevitably reduce cost-efficiency. This result is found to be consistent with the conclusions in the literature. The positive effects of higher capital can also be seen as a positive effect from the increased capital requirement from Basel II and Basel III implementation in China. Although the same signs of these two variables can be found in clusters 2 and 3, they are not significant. The profitability of banks is measured by ROAA. The sign of ROAA is positive for both cluster 1 and cluster 2 but negative for cluster 3. However, ROAAs in all three cluster-specific frontiers are not found to be significant. We can conclude that higher profitability does not seem to be linked with higher cost-efficiency in Chinese regional banks.

4.9.4 Meta-frontier Results

Table 4.17 below shows the frontier estimates from the second-stage meta-frontier estimation.

Table 4.17. Meta-frontier Estimates

Metafrontier Model					
	Frontier Estimates			***Cost Gap Model***	
lnp1	0.5780***	(0.1515)	$GDPGR_{pt}$	0.0001	(0.0019)
lnp2	-0.4141***	(0.0789)	$GDPPCapita_{pt}$	0.0141	(0.0160)
lnp1lnp1	-0.0085	(0.0423)	$Urban_{pt}$	-0.0019***	(0.0007)
lnp2lnp2	-0.0211	(0.0155)	INF_{pt}	0.0011	(0.0009)
lnp1lnp2	-0.0486**	(0.0229)	HHI_{ct}	0.0040**	(0.0016)
lny1	0.7670***	(0.0767)	FAS_{nt}	-0.0155**	(0.0070)
lny2	0.5372***	(0.0946)	Constant	0.4211**	(0.2041)
lny3	0.0332	(0.0459)	σ_u		
lny1lny1	0.2060***	(0.0145)	Constant	-6.6385***	(0.1557)
lny2lny2	0.1161***	(0.0134)	σ_v		
lny3lny3	0.0131***	(0.0031)	Constant	-6.9118***	(0.1550)
lny1lny2	-0.1712***	(0.0151)	*N*	1050	
lny1lny3	-0.0302***	(0.0047)	ll	1698.2943	
lny2lny3	0.0139*	(0.0071)			
lny1lnp1	0.0526***	(0.0177)			
lny1lnp2	0.0289***	(0.0096)			
lny2lnp1	-0.0927***	(0.0212)			
lny2lnp2	0.0387***	(0.0115)			
lny3lnp1	-0.0056	(0.0108)			
lny3lnp2	-0.0149***	(0.0057)			
lneq	-0.1607***	(0.0571)			
lneqlneq	0.0222***	(0.0076)			
t	0.0900***	(0.0052)			
t^2	-0.0101***	(0.0006)			
Constant	-2.2213***	(0.4992)			

This table shows the frontier estimates from the second-stage meta-frontier estimation. Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $GDPGR_{pt}$ is the GDP growth rate, $GDPPC_{pt}$ is the real GDP per capita, $Urban_{pt}$ is the urbanisation rate, INF_{pt} is the inflation rate, HHI_{ct} is the Herfindahl index, and is the foreign asset share. σ_u and σ_v are standard errors of two error components, namely the cost inefficiency component with distribution as $N^+(\mu, \sigma_u^2)$ and random error component v^2 with distribution as $N(0, \sigma_v)$.

Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

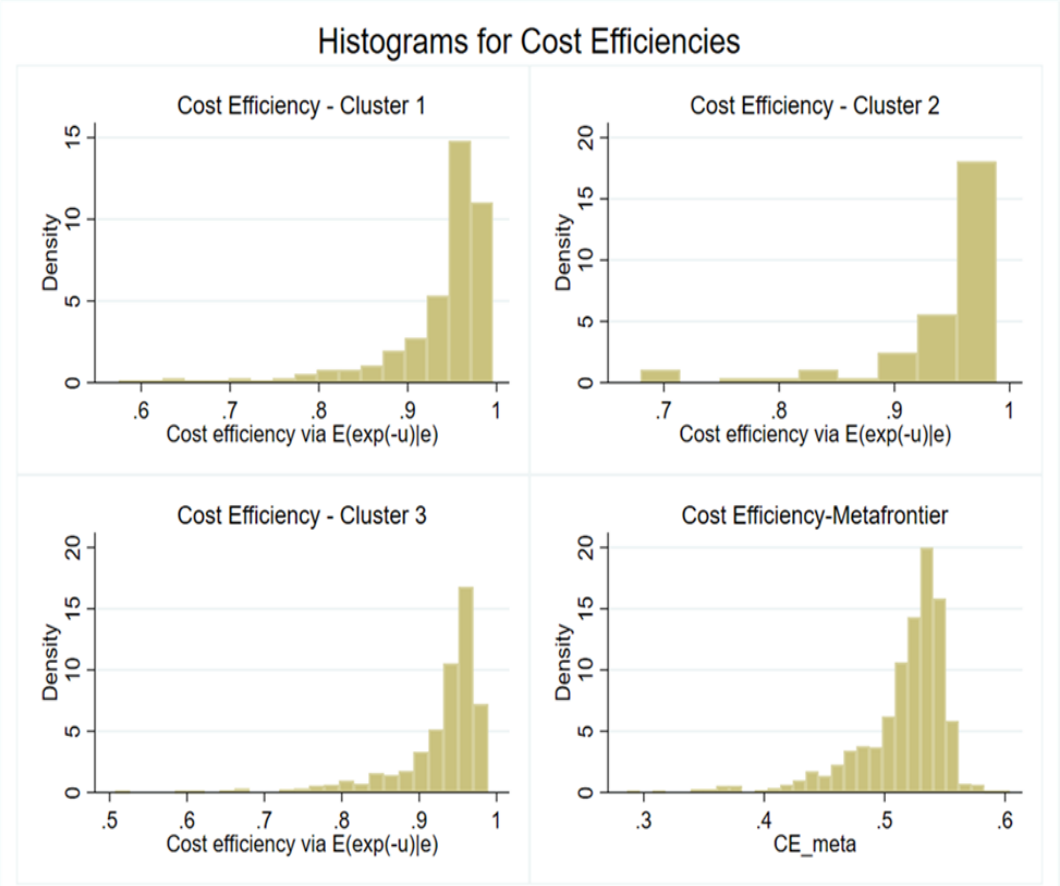
The majority of the coefficients of frontier variables are significant, indicating a well-specified meta-frontier model. The loglikelihood is also higher than both the pooled model and the total loglikelihood of three individual frontiers, further

confirming the preference of meta-frontier over the pooled model as well as individual frontier models. Again, we find a similar relationship between total cost and time as well as total cost and equity. That is, the relationship between equity and the total cost is U-shaped, and between time and the total cost is reverse U-shaped. In terms of environmental variables, the results from the meta-frontier are largely consistent with results from the pooled 3 model where all environmental variables are included. As is shown in the pooled models, the significance of macroeconomic variables is only revealed when microeconomic variables are included in the model. Similarly, the meta-frontier model reveals macroeconomic impact when enveloping the first stage individual frontiers where microeconomic variables were taken into account. The impact of HHI and FAS in the meta-frontier, which represent concentration and foreign presence, exhibit exactly the same directions of effects as pooled models. However, the interpretation is different as these macroeconomic factors directly impact the technological gap between meta-frontier and individual frontiers rather than inefficiency itself. As is witnessed in descriptive statistics, the phenomenon of new bank entry increases local competition during the period. These entry of newly established banks may possess superior technology and business model that may have an spill-over effect on the incumbent banks to improve their cost structure and efficiency. That is, higher concentration leads to a larger technological cost gap which may be a result of market power. Banks with higher market power in their specific cluster may be under less pressure to adopt the most advanced technology to compete with other banks. Foreign presence narrows down the technological gap between individual frontiers and meta-frontier, indicating banks that are operating in regions with larger foreign presence tend to be closer to the best available technology and highest cost-efficiency. Inflation in the meta-frontier model is unfortunately insignificant, although the sign of the INF is in the same direction as the pooled model indicating inflation tends to cause excessive branch expansion and loan losses leading to extra cost in the Chinese banking system. The variable Urban, which represents urbanization, is significant in the meta-frontier model but not in the pooled model. It is also a measurement for economic development and is found to help reduce the gap between individual frontiers and the meta-frontier.

Urbanised areas tend to be more populated with more concentrated productive and entrepreneurial activities, which could have a positive impact on bank deposits, loans and investments. This result is in line with Bikker and Bos (2008)'s findings in European banks. Unfortunately, both GDP growth rate and real GDP per capita were found to be insignificant in our model.

4.9.5 Cost Efficiencies and Cost Gap Ratios

Figure 4.8. Histogram of Cost Efficiencies



This histogram shows the frequency distribution of cost efficiency in frontiers obtained from the three clusters and the meta-frontier. The y-axis shows density, and the x-axis shows the cost efficiency. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the North-eastern and Coastal Region.

The histograms above display the frequency distribution of cost efficiency obtained from individual frontiers from three clusters and the meta-frontier. The

efficiencies in clusters 1, 2 and 3 display skewness towards unity, while meta-frontier efficiency scores seem to be less skewed.

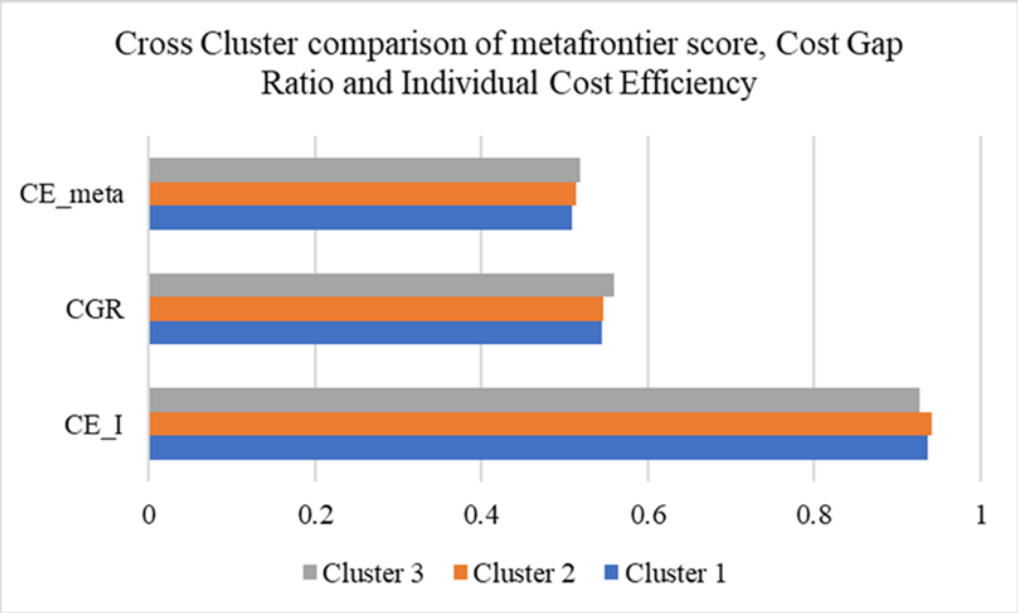
The following table presents the pooled cost x-efficiency scores (CE_pooled), cost efficiency scores estimated using individual frontiers (CE_I), technological cost gap ratio (CGR) and cost efficiency scores of individual banks with respect to meta-frontier.

Table 4.18. Cost Efficiencies and Cost Gap Ratios

Variables	No.of Obs.	Mean	S.D.	Max.	Min.
		Cluster 1- the Inland Region			
CE_pooled 1	311	0.9154	0.0670	0.9849	0.6240
CE_pooled 2	311	0.9136	0.0671	0.9799	0.6163
CE_pooled 3	311	0.9119	0.0698	0.9836	0.6205
CE_I	311	0.9358	0.0655	0.9958	0.5746
CGR_meta	311	0.5446	0.0187	0.6115	0.4686
CE_meta	311	0.5094	0.0375	0.5574	0.3424
		Cluster 2- the Eastern Municipalities			
CE_pooled 1	84	0.9245	0.0664	0.9795	0.7051
CE_pooled 2	84	0.9054	0.0817	0.9774	0.6413
CE_pooled 3	84	0.9062	0.0809	0.9793	0.6516
CE_I	84	0.9406	0.0660	0.9884	0.6797
CGR_meta	84	0.5457	0.0288	0.6047	0.4571
CE_meta	84	0.5135	0.0467	0.5724	0.3598
		Cluster 3- the North-eastern and Coastal Region			
CE_pooled 1	655	0.9292	0.0577	0.9890	0.4988
CE_pooled 2	655	0.9284	0.0598	0.9890	0.4800
CE_pooled 3	655	0.9256	0.0615	0.9894	0.5018
CE_I	655	0.9273	0.0608	0.9892	0.5070
CGR_meta	655	0.5594	0.0114	0.6194	0.5120
CE_meta	655	0.5189	0.0373	0.6038	0.2862
		All Regions			
CE_pooled 1	1050	0.9247	0.0616	0.9890	0.4988
CE_pooled 2	1050	0.9222	0.0645	0.9890	0.4800
CE_pooled 3	1050	0.9200	0.0661	0.9894	0.5018
CE_I	1050	0.9309	0.0628	0.9958	0.5070
CGR_meta	1050	0.5539	0.0173	0.6194	0.4571
CE_meta	1050	0.5157	0.0384	0.6038	0.2862

This table shows, for each cluster, the pooled cost x-efficiency scores (CE_pooled), cost efficiency scores estimated using individual frontiers (CE_I), technological cost gap ratio (CGR_meta) and cost efficiency scores of individual banks with respect to meta-frontier (CE_meta).

Figure 4.9. Bar Chart of Meta-frontier Efficiency Scores, Cost Gap Ratio and Individual Frontier Efficiency Scores



Green, orange and blue bars correspond to Cluster 3, Cluster 2 and Cluster 1 respectively. The x-axis shows cost efficiency scores estimated using individual frontiers (CE_I), technological cost gap ratio (CGR) and the meta-frontier score (CE_meta). CE_meta= CE_I* CGR. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the North-eastern and Coastal Region.

We can see that, in general, individual frontiers always generate higher efficiency scores than the meta-frontier. It is expected since the meta-frontier efficiency is the product of the individual frontier efficiency and the cost gap ratio, and the meta-frontier always envelopes individual frontiers.

From the conventional stochastic frontier model (pooled models) in our results, we can observe that the efficiency scores are more than 90% for all three clusters. That indicates that banks from each cluster do not operate very far from their own technology. That is, an average bank can become fully cost-efficient under regional technology by reducing its cost by slightly less than 10%. Compared to individual frontiers, pooled models always seem to underestimate efficiency

scores, which is also found in the cross-country study by Bos and Schmiedel (2007) in European banks. They conclude that the assumption of a single frontier technology leads to substantial biases in cross-country studies. Similarly, as a cross-region study, our results also show support that biases are observed when assuming banks in different regions within a country operate under the same frontier technology. We also find a similar downward bias as Bos and Schmiedel (2007) that, overall, the pooled model seems to have lower minimum cost efficiency scores than the individual frontier model if cost gap ratios are lower than one. However, this is not the case for each cluster, as this downward bias is only found in cluster 3 while clusters 1 and 2 exhibit upward bias. In our sample, all the cost gap ratios are noticeably lower than one and the maximum cost gap ratio is only around 0.6194, meaning that there might not be tangent point of individual frontier on the meta-frontier and individual frontiers all seem to deviate significantly far from the meta-frontier. In other words, cost efficiency decreases when the comparisons move from pooled frontier to meta-frontier.

As cost efficiencies of different groups operating under different technologies are only comparable when comparing to the meta-frontier, that is, only meta-frontier cost efficiencies should be compared across clusters, we can find that in this case, banks in cluster 3 (the Inland Region) have the highest level of cost-efficiency compared to its fellow clusters. If we use a pooled model assuming all banks operate under the same technology, which has been confirmed by the likelihood ratio test, we can still find cluster 3 to be the most efficient region. Regarding the worst-performing cluster in terms of cost efficiency, meta-frontier also points out cluster 1 to be the worst-performing region while pooled model points to either cluster 1 or cluster 2 depending on the choice of exogenous variables.

In terms of the Cost Gap ratio (CGR), the higher the CGR, the narrower the gap between individual frontier and meta-frontier. We find that cluster 3, which is the cluster with the highest meta-frontier efficiency, has the highest cost gap

ratio. It means that the technology adopted by banks in cluster 3 is closer to the best available technology in the country than in cluster 1 and cluster 2, although banks in all three clusters operate quite far from the meta-frontier technology. This may be explained by the fact that our sample excludes the biggest five state-owned banks, joint-stock banks and especially foreign banks, which could potentially be operating closer to the meta-frontier as literature has indicated that foreign banks in China perform significantly better than domestic banks in terms of x-efficiency. Similarly, the cluster with the lowest meta-frontier cost efficiency also has the lowest cost gap ratio.

According to Lee and Huang (2017), the meta-frontier inefficiency consists of two elements, the deviation of the individual bank from its cluster and the deviation of the individual frontier from the meta-frontier. The deviation of a bank from its cluster frontier could be interpreted as managerial inefficiency, which is inside banks' control and could be improved by better management. On the other hand, the deviation of the individual frontier from the meta-frontier can be regarded as the region's ability to operate under the best available technology, which is out of banks' control and is mainly affected by the macroeconomic conditions we have included in the model. One of the purposes and advantages of this stochastic meta-frontier model is to be able to separate internal causes of inefficiency from external causes, giving a clearer picture for policymakers in terms of region-tailored policies. Cluster 3 has the highest level of CGRs, meaning that banks in cluster 3 operate closer to meta-frontier compared to the other two clusters. From the previous analysis of macroeconomic variables and meta-frontier estimates, we find that this is most likely due to having the lowest level of HHIs among all clusters. As it is found that HHI has a significant and negative impact on meta-frontier cost efficiency, low HHI, which means a low level of concentration and a higher level of competition in the market, prompt a higher level of cost efficiency as banks tend to have more pressure to adopt the most advanced technology. Therefore, in order to increase cost-efficiency in cluster 3, policymakers could aim to increase regional competition and reduce the market concentration to narrow down the

technology gap and, in turn, increase inefficiency. Cluster 1 appears to be the least efficient bank compared to the meta-frontier. Again, from the descriptive data, we suspect that this is mostly due to cluster 1 being the least urbanised cluster, as it is found that urbanization has a significant and positive impact on the meta-frontier cost efficiency. Thus, to shrink the technology gap, we can recommend cluster 1 to advance its process in urbanization. Urbanization is likely to bring in technology through better infrastructure, more diverse investment and access to technological information. Foreign banks are also more likely to establish in urban areas than in rural areas due to their different product mix and goal in profit-maximization, and it has been found that foreign presence brings down the technology gap. Therefore, urbanization will mostly help rural commercial banks in cluster 1 to narrow down the technology gap.

From the results, we can infer that the banks in the most efficient cluster, which is cluster 3, obtain their efficiency from more favourable economic conditions and more advanced adoption of available technology compared to the rest of the country. While for the cluster with the lowest efficiency, the technology in their banks seems to be less efficiently adopted compared to the other clusters. This result gives us a clear picture in terms of what policy should be adopted in each region's banking sector.

While it is evident that in order to improve the cost-efficiency of banks in cluster 1, regulations should gear towards a more favourable banking environment through improving macroeconomic and financial sector conditions such as increasing foreign presence, supportive competition and urbanization rate and reducing market power, this policy could also be encouraged by banks in the other two clusters. This is because although the other two clusters seem to be operating closer to the meta-frontier than cluster 1, they are all somewhat distant from the best available technology. Banks in cluster 3, which have the highest meta-frontier efficiencies and highest cost gap ratios, also have the lowest individual frontier efficiency. As the equation in the methodology section indicates, meta-frontier efficiency is the product of individual frontier efficiency

CE and cost gap ratio CGR. Since it is clear that CGR is always much lower than cluster-specific CE, we can therefore conclude that in Chinese regional banks, out of the two components, CGR contributes more towards the determination of the meta-frontier cost-efficiency CE_meta. This is an indication that the problem of inferior technology is a more severe issue than managerial inability. Therefore, it seems that regional banks in China should focus more on adopting superior technology. One thing that needs to be emphasized here is that our sample only includes city and rural commercial banks, which means that state-owned banks, joint-stock banks and foreign banks are excluded from our study. It is known that joint-stock banks and foreign banks in China tend to be the groups that adopt the most superior technology. Therefore, confirmed in our results, the lagged-behind technology adoption seems to be the main reason that prevents city and rural commercial banks from being more efficient. With regards to regional differentiation in terms of the determination of meta-frontier cost efficiency, we can observe that cluster 3 appears to be the one region with the lowest cluster-specific CE. We can therefore recommend that banks in cluster 3 could give more attention to managerial efficiency as the managerial problem seems to be more severe than in the other two clusters. However, all three clusters should still focus on adopting better technology as their primary goal.

From the above results, we can summarize a few conclusions. Firstly, core regions, i.e. the Eastern Municipalities and the Northeastern and Coastal Region, perform better than peripherical regions, largely in line with the theories and findings by Alessandrini et al. (2009), Sun et al. (2013) and Yang et al. (2015). Secondly, from the relatively low value of cost gap ratios, we can conclude that cluster-specific frontiers deviate pretty far from the country meta-frontier. Therefore, in order to increase regional banking efficiency, city and rural commercial banks in China should focus on regional and economic and financial development to narrow the technological gaps, especially for peripheral regions. Thirdly, in terms of bank- and industry-specific factors, regional banks can attempt to improve managerial efficiency by reducing the amount of non-

performing loans in their books and increase their capital level.

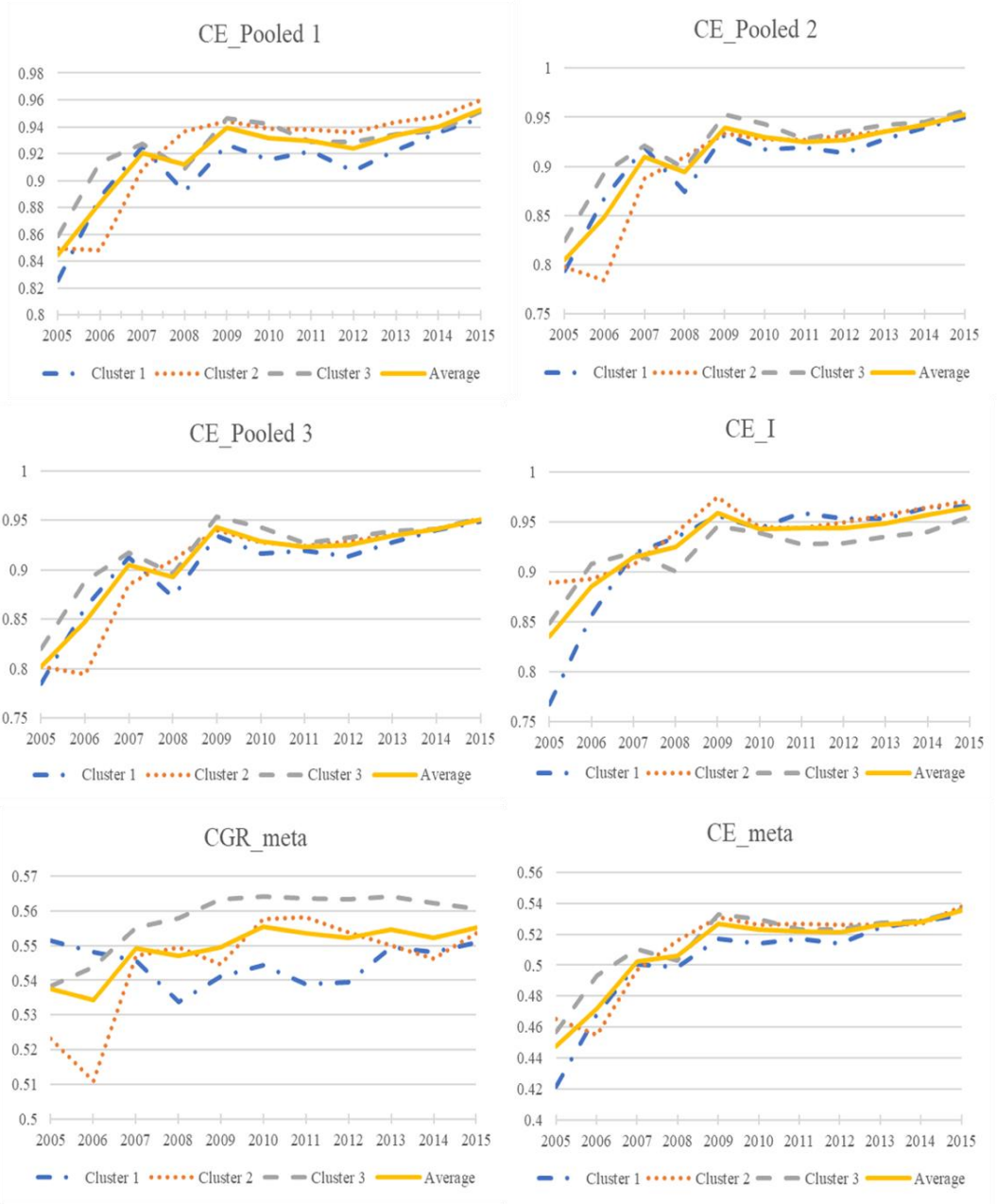
4.9.6 Trends of CE_pooled, CE_meta, CGR and CE_I

The following graphs plot the CGR, meta-frontier cost efficiency and pooled frontier cost efficiency against year to illustrate the trend. A salient feature is that for CE_pooled, CE_meta and CE_I, there is a clear trend of upward convergence towards the end of the period. That is, cost efficiencies increased, and gaps between the cluster-specific frontier and meta-frontier grow smaller over time. This is a clear indication of the success of policies and regulations in narrowing down regional technological gaps through digitalization and training.

There are relatively small variations in terms of the relationship among the three clusters in the pooled model, which is mainly caused by different combinations of exogenous variables. Although the trends for meta-frontier efficiency for the three clusters are more similar to the pooled model than the individual frontier model, meta-frontier efficiency is evidently much lower, owing to technology gaps. The cost gap ratio trend also has a gentle upward trend but is relatively flat compared to the trend of meta-frontier cost efficiency. In the CGR_meta graph, prior to 2007, cluster 2 has the lowest cost gap ratio, cluster 1 has the highest cost gap ratio, and cluster 3 lies in the middle. From 2007 to 2013, cluster 1 has the lowest cost gap ratio, and cluster 2 remains in the middle. Cluster 3 continues to be the region with the highest cost gap ratio since 2007. As we know, the higher the cost gap ratio, the closer the cluster frontier to meta-frontier. Therefore, since 2007 cluster 3 operates closest to the best available technology while cluster 1 deviates furthest from the best available technology. The distinct character of cluster 2 is that it has the sharpest increase in cost gap ratio, especially from 2006 to 2010[13].

[13] Please note that as our sample is an unbalanced panel dataset, the number of observations changes over the sample period. Thus, for example, increasing efficiency from one year to another does not necessarily mean that the same banks become more efficient over the years. Rather it indicates that the overall efficiency of banks in a region improves. For example, the entry of new banks may indirectly improve regional bank efficiency by bringing in new

Figure 4.10. Trends of Cost efficiency from Pooled models, Individual Frontiers and Meta-frontier with Cost Gap Ratio



These figures show, for each cluster, the trend of the pooled cost x-efficiency scores (CE_pooled), cost efficiency scores estimated using individual frontiers (CE_I), technological cost gap ratio (CGR_meta) and cost efficiency scores of individual banks with respect to meta-frontier (CE_meta). The x-axis represents the year.

technology and business know-how and healthy competition into the local banking market, indirectly narrowing the region-country technological gap and improves regional bank efficiency.

4.9.7 Rankings

To further test the stability of meta-frontier efficiency, we rank cost efficiency scores using different models and then divide rankings into four quartiles. We want to see if the best performing banks determined by cluster-specific frontiers would also stay top-ranked with meta-frontier cost efficiency. The results of similarity are listed in table 4.19. The similarity scores represent the percentage of the same banks that belong to each quartile for both CE_pooled and CE_meta. Quartile 4 represent the highest-ranked banks, and quartile 1 represents the lowest-ranked banks. For example, in cluster 1, we find that 81% of the banks which belong to quartile 4 in cluster-specific frontier also belong to quartile 4 in meta-frontier rankings. This shows that only 19% of the banks move to a different quartile from individual frontier CE_pooled to CE_meta. Although meta-frontier CE does not show the greatest overall stability across all quartiles, meta-frontier seems to be the most stable and consistent one in terms of predicting the top-performing banks. We can see that for the top-performing quartile, meta-frontier cost-efficiency CE_meta shows 81%, 71% and 91% similarities with individual frontier with clusters 1, 2 and 3, respectively, which seems to be more stable than pooled models. So, we can conclude that although the average stability of meta-frontier prediction is not better than pooled models, meta-frontier shows superior stability in terms of consistently predicting the best performing banks. This result is again in line with Bikker and Bos (2008)'s findings in meta-frontier's greater stability in predicting the highest-ranked banks in European countries.

Table 4.19. Rankings Similarity

Quartile of CE_I	Similarity with Pooled Frontier 1	Similarity with Pooled Frontier 2	Similarity with Pooled Frontier 3	Similarity with Meta-frontier
		Cluster 1- the Inland Region		
1	72%	58%	71%	55%
2	42%	38%	47%	32%
3	49%	45%	50%	47%
4	69%	71%	75%	81%
All quartiles	58%	53%	61%	54%
		Cluster 2- the Eastern Municipalities		
1	48%	57%	52%	33%
2	38%	33%	38%	33%
3	38%	33%	33%	33%
4	67%	67%	71%	71%
All quartiles	48%	48%	49%	43%
		Cluster 3- the North-eastern and Coastal Region		
1	88%	79%	82%	68%
2	68%	55%	63%	52%
3	69%	59%	68%	63%
4	88%	85%	87%	91%
All quartiles	78%	69%	75%	69%
All Cluster	61%	57%	62%	55%

This table shows the similarity of cost efficiencies scores between cost efficiencies estimated from cluster-specific frontiers with three different models of pooled frontiers and meta-frontier in 4 efficiency score quartiles. For example, left top corner of the table indicates that in the lowest ranked quartile in terms of individual frontier scores, the similarity between CE_I and CE_pooled 1 is 70%.

4.10 Determinants of Economies of Scale

In this section, we explore the determinants of returns to scale derived from the meta-frontier estimation in Section 4.8. Our empirical design follows a two-step approach. In the first stage, economies of scale (ES) are obtained both at the mean of the data and at each data point from region-specific frontiers and the meta-frontier. In the second step, we use the observations specific ES estimates as the dependent variable in a panel data regression, examining the determinants of cost economies. The sample consists of 136 regional Chinese commercial banks from 2005 to 2011 with 1049 observations.

4.10.1 Literature review and hypotheses development

Estimation of economies of scale received wide attention in the empirical banking literature. However, the majority of the studies focus on the US and the EU. Hughes et al. (2001) find evidence for economies of scale in U.S. banks in 1994 by incorporating capital structure and risk-taking into models of production. Hughes and Mester (2013) study U.S. banks during the years 2003,2007 and 2010 and find large scale economies that are not explained by too-big-to-fail considerations. Davies and Tracey (2014) investigate large international banks over the period 2001-2010 and find no evidence of scale economies when controlling for too-big-to-fail factors with two different methods. Restrepo-Tobon and Kumbhakar (2015) study banks in the U.S. from 2000 to 2010 with a new method based on input distance functions and find only a small degree of economies of scale. Recent studies on economies of scale in the US find a U-shaped curve where the smallest banks and largest banks enjoy the most economies of scale (Wheelock and Wilson, 2018, Wheelock and Wilson, 2012). Beccalli and Rossi (2020) examine economies of scope in EU banks between 2005 and 2018 and find cost economies of scope, revenue diseconomies of scope and profit diseconomies in scope.

In terms of literature on determinants of economies of scale, Bertay et al. (2013) investigate listed banks from 90 countries and find that banks with large absolute size tend to be more profitable, as they rely more on non-interest operations as a source of income and more on wholesale funding than customer deposits for debt financing. Beccalli et al. (2015) extend the study by Bertay et al. (2013) and investigate the impact of investment banking activities, liquidity risk, profit from traditional lending activities, capital strength, systemic risk and too big to fail status on the realization of economies of scale. They apply both univariate quantile analysis and multivariate regression analysis in a sample of EU banks, and confirm that investment banking activities and non-interest income have a positive impact on realizing economies of scale. They also find that banks with higher liquidity, lower core regulatory capital, less contribution to systemic risk and too-big-to-fail status realize higher economies of scale. Credit risk does not seem to have an impact on economies of scale in EU banking system.

In terms of estimating economies of scale, to the best of our knowledge, the literature concerning estimation of economies of scale in China is somewhat limited, and the majority of the studies focus on state-owned and joint-stock banks. Especially, little attention has been paid to city and rural commercial banks. In terms of economies of scale in China, Fu and Heffernan (2008) select a total of 187 observations of 5 state-owned and 12 joint-stock commercial banks from 1985 to 2002 and study the impact of three phases of banking reform on economies of scale and scope. They find an overall constant return to scale during the period. More specifically, they find constant returns to scale for joint-stock banks throughout the period and a change from decreasing returns to scale to constant return to scale during the second phases of reform for state banks. The authors also find that stochastic frontier approaches produce less biased results compared to other non-frontier approaches in terms of economies of scale estimation. In terms of determinants of economies of scale, a more recent study by Hou et al. (2015) finds that overall scale economies are not driven by loans or other earnings assets but mostly from off-balance-sheet operations using a

stochastic frontier cost model. They also find banks with lower profitability, smaller size and non-state ownership benefit most in terms of realizing economies of scale via increases in off-balance-sheet activities.

The aim of this section is threefold. Firstly, to fill the gap in the literature on scale economies studies in China for more recent years. Secondly, to test the hypotheses developed by Bertay et al. (2013) and Beccalli et al. (2015) for city and rural commercial banks in China - relatively small in size and neither a systemically important nor too big to fail part of the banking system. The operation of city and rural commercial banks in China is largely within their own provinces or regions with a trend of expansion to nearby regions. Thirdly, this section provides insights for banks in the process of intra-provincial and inter-provincial expansion and addresses the question of whether they should merge to achieve more significant scale economies. To the best of our knowledge, this is the first study in the literature to study determinants of economies of scale derived from a stochastic meta-frontier, with the assumption that different regions operating under different technologies.

4.10.2 Methodology

Estimating Economies of scale

In the first step of our empirical strategy, we estimate economies of scale as the total of cost elasticities with respect to output from equation (4.19) in section 4.26, that is:

$$\ln(\widehat{TC})^J{}_{it} = \alpha_0 + \sum_{m=1}^{3} \beta_m lny_{mit} + \sum_{n=1}^{2} \rho_n lnp_{nit} + \frac{1}{2}[\sum_{m=1}^{3}\sum_{n=1}^{3} \tau_{mn} lny_{mit} lny_{nit} + \sum_{m=1}^{2}\sum_{n=1}^{2} \varphi_{mn} lnp_{mit} lnp_{nit}] + \sum_{m=1}^{3}\sum_{n=1}^{2} \omega_{mn} lny_{mit} lnp_{nit} + \epsilon_1 t + \frac{1}{2}\epsilon_2 t^2 + \epsilon_3 lneq_{it} + \frac{1}{2}\epsilon_4 eq_{it}^2 + lnU_{it}^M + lnV_{it}^M \quad (4.19)$$

where $\ln(\widehat{TC})^J{}_{it}$ are predicted values of lnTC from group-specific frontiers. Thus, ES can be defined as:

$$ES_{it} = (\sum_{k=1}^{3} \frac{\partial TC_{it}}{\partial lnQ_k}) \quad (4.29)$$

More specifically,

$$ES_{it} = \sum_{i=1}^{3} \alpha_i + \sum_{i=1}^{3}\sum_{j=1}^{3} \gamma_{ij} \ln Q_j + \sum_{i=1}^{3}\sum_{j=1}^{3} \varphi_{ij} \ln P_j + \sum_{i=1}^{3} \omega_i \ln Eq \quad (4.30)$$

We obtain the economies of scale (ES) using the same bank-level output variables, input prices and equity as the cluster-specific model in equation (4.17) and meta-frontier model in equation (4.19). ES is estimated at the mean of data for each cluster using cluster-specific frontier and at the mean of data using the

metafrontier. An ES=1 represents constant returns to scale, with ES <1 indicating economies of scale and ES >1 diseconomies of scale. In other words, if a 1% simultaneous increase in output increases costs by 1%, we can conclude that the sample exhibits constant returns to scale (CRS). If cost increases more than 1%, decreasing returns to scale (DRS) is found and increasing returns to scale (IRS) is found if cost increases less than 1%. CRS suggest that a bank is operating at its optimal size with its available technology, DRS that the bank can benefit from size controlling or downsizing, and IRS indicates that the bank may choose expansion or mergers and acquisitions to increase its scale productivity (Badunenko and Kumbhakar 2017).

Estimating determinants of economies of scale

In the second stage, we follow Beccalli et al. (2015), and we estimate the following panel data regressions:

$$ES_{i,t} = \alpha_0 + \beta_1 SEC_TA_{i,t} + \beta_2 NIM_{i,t} + \beta_3 LR_{i,t} + \beta_4 LR_{i,t}^2 + \beta_5 LLPLOAN_{i,t} + \beta_6 Tier1_{i,t} + \beta_7 Tier1_{i,t}^2 + \beta_8 Zscore_{i,t} + \varepsilon_{i,t} \quad (4.31a)$$

$$ES_{i,t} = \alpha_0 + \beta_1 SEC_TA_{i,t} + \beta_2 NIM_{i,t} + \beta_3 LR_{i,t} + \beta_4 LR_{i,t}^2 + \beta_5 LLPLOAN_{i,t} + \beta_6 ETA_{i,t} + \beta_7 ETA_{i,t}^2 + \beta_8 Zscore_{i,t} + \varepsilon_{i,t} \quad (4.31b)$$

The difference between (4.31a) and (4.31b) is the different measure of capital adequacy. In equation (4.31a), we use the $Tier1_{i,t}$ ratio computed as Tier 1 capital scaled by risk-weighted assets, whereas in equation (4.31b), we include $ETA_{i,t}$ which is total equity to total assets. The reason for using two different measures for the capital ratio is that Tier 1 ratio data is not available for all banks,

thus reducing the sample size in the estimation process. Therefore, we use $ETA_{i,t}$ as well to maximise the number of observations in our model. Their respective squared terms are also included in the model to capture any non-linear effects. $SEC_TA_{i,t}$ are total securities divided by total assets as a proxy for investment banking activities. $NIM_{i,t}$ is the net interest margin, calculated as net interest income divided by total loans, representing the income from traditional lending operations. The liquidity ratio $LR_{i,t}$ is calculated as total liquid assets to total customer deposit and $LR_{i,t}^2$ is used to capture the non-linear relationship between liquidity and economies of scale. The loan loss provisions to gross loans $LLPLOAN_{i,t}$ is to measure the credit risk in each bank and $Zscore_{i,t}$, which is the sum of return on average asset and equity divided by the standard deviation of return on average asset, formulated as $Zscore_{it} = \frac{ROAA_{it}+Equity_{it}}{\sigma_{ROAA_i}}$. The fixed effects panel data model was used to estimate the above equations to capture individual unobserved heterogeneities.

4.10.3 Result Analysis

Economies of Scale

We begin the discussion of the empirical results by focusing on the estimated economies of scale. We calculate economies of scale using the parameters both from the individual frontiers and the meta-frontier estimated in Section 4.8.3. We use the data for each region, and we compare the individual frontier ES with the meta-frontier ES to understand how the movement towards a common technology would affect banks' strategy on size control or mergers for different areas in China. The result estimated at the mean of the data is shown in table 4.20. In column 'ES: 2005-2015', we can observe that with current technology, banks in the Inland Region (cluster 1) and the North-eastern and Coastal Region (cluster 3) exhibit diseconomies of scale while banks in the Eastern Municipalities (cluster 2) enjoy economies of scale. The ES estimated from the meta-frontier shows that banks experience increasing returns to scale, with ES less than 1 for all regions. The aforementioned findings suggest that the Inland region and the North-eastern region will start to enjoy economies of scale once

they adopt the meta-frontier technology. This result would suggest that for these two regions, banks can achieve lower unit costs by expanding production scale if they adopt better technology. Banks in the Eastern Municipalities, which are currently enjoying economies of scale, will achieve a greater cost reduction from expansion once they adopt better technology.

Table 4.20. Estimates of Scale Economies

Frontier Type	Cluster	ES: 2005-2015
Individual Frontiers	the Inland Region (1)	1.0035
	the Eastern Municipalities (2)	0.9636
	the North-eastern and Coastal Region (3)	1.0091
Stochastic Meta-Frontier (SMF)	the Inland Region (1)	0.9818
	the Eastern Municipalities (2)	0.9491
	the North-eastern and Coastal Region (3)	0.9869

This table presents the economies of scale estimated at the mean of data.

Determinants of Economies of Scale

In order to estimate the determinants of economies of scale, we also need scale economies estimated at the point of the data, i.e., bank-specific ES. Table 4.21 presents a summary of bank-specific economies of scale at each data point estimated from cluster-specific frontiers. The average scale economies are calculated for each province and each cluster [14] .

[14] To recall, the clusters present in this section have been defined in Section 4.8 Cluster Analysis Result using the hybrid Hierarchical K-means clustering methodology.

Table 4.21. Bank-Specific Economies of Scale for each Province and each Cluster

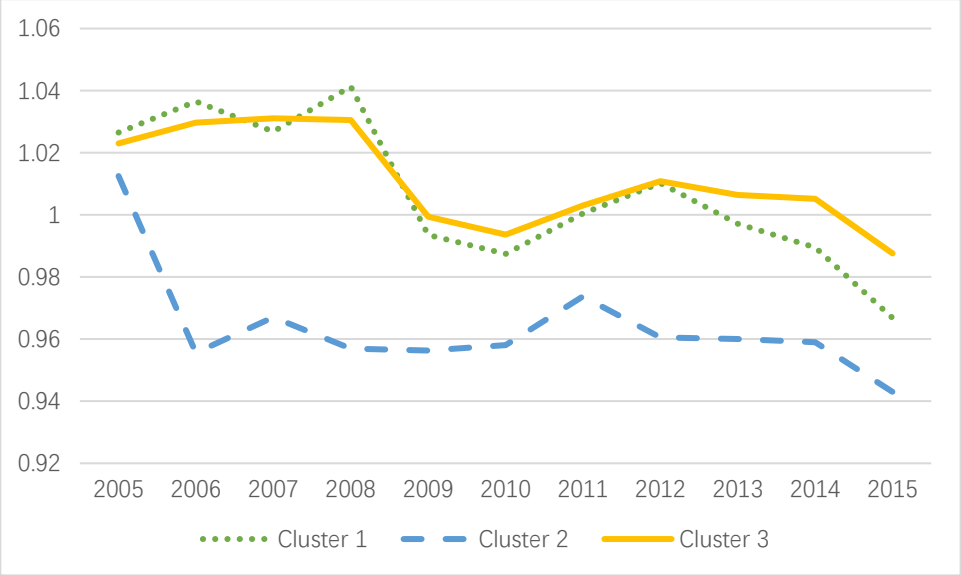
Province	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2005-2015
Panel A: Mean of Economies of Scale by Province												
Anhui	0.9393	0.9414	0.9377	0.9321	0.918	0.9305	0.9541	0.9531	0.9736	0.95	0.9333	0.9421
Beijing	0.9031	0.8999	0.9139	0.9111	0.8864	0.8811	0.899	0.9095	0.9104	0.9149	0.8965	0.9023
Chongqing	0.9851	0.9811	0.9842	0.9984	0.9394	0.9408	0.9861	0.9746	0.9564	0.9451	0.9356	0.9610
Fujian	0.9946	1.0044	1.0141	1.0285	0.9896	0.996	0.996	1.0127	1.0114	1.0096	0.9941	1.0049
Gansu	1.054	1.0505	1.0601					1.0202				1.0462
Guangdong	0.9331	0.9714	0.9999	1.0156	0.9843	0.9715	0.9742	1.0013	1.0045	1.009	0.9871	0.9920
Guangxi	1.0614	1.0447	1.0296	1.0517	1.0063	0.9936	1.006	1.0067	0.9964	1.0051	0.9976	1.0153
Guizhou		1.0377	1.0191		0.9631	0.9585	0.9641	0.9751		0.9467		0.9806
Hebei		1.0886	1.0383	1.0422	1.0096	1.0029	1.0185	1.0126	0.9991	0.972	0.9612	1.0087
Heilongjiang			1.0044	0.9824	0.9564	0.9402	0.9381	0.9447	0.9353	0.9401	0.9219	0.9485
Henan	1.0693	1.0598	1.0626	1.137	1.0536	1.0321	1.0377	1.0297	0.9824	0.9944	0.9524	1.0424
Hubei	1.0536	1.0241	0.9747	0.9929	0.9775	0.9385	0.9724	0.9779	0.9845	0.9885	0.9677	0.9833
Hunan	0.9551	0.968	0.944			0.9469	0.9606	0.9731	0.955	0.9479		0.9573
Inner Mongolia	0.9666	0.9889	1.0068	0.9957	0.9532	0.9278	0.944	0.9485		0.9599	0.9592	0.9611
Jiangsu	0.9677	0.9868	1.0155	1.0106	0.9759	0.987	1.0012	1.006	0.9941	0.9892	0.9703	0.9926
Jiangxi					1.0201	0.9832	0.9947	1.0379	0.9704	0.9833	0.9676	0.9939
Jilin			0.9677	0.9763	0.9459	0.9606	0.9797		0.9859	0.9892	0.9297	0.9709
Liaoning	1.059	1.0473	1.0439	1.0367	1.0101	0.996	1.0076	1.0089	0.9999	1.0009	0.9906	1.0129
Ningxia		1.0158	1.0021	1.0017	1.0204	0.9922	0.9896	0.9869	0.9883	0.9836	0.9813	0.9962
Shaanxi		1.0888			0.9705		0.9711	1.1033	1.0895	1.0706	0.9643	1.0535
Shandong	1.0828	1.0801	1.0666	1.0633	1.0357	1.0357	1.0595	1.0411	1.0258	1.016	1.0018	1.0445
Shanghai	1.069	0.9897	0.9786	0.9786	0.9954	1.0075	1.008	0.9782	0.9807	0.9678	0.9508	0.9944
Shanxi		1.0239	1.0649		0.9929	1.0389	1.0461	1.037	1.044	1.0269	0.9864	1.0358
Sichuan	1.0239	1.029	1.0283	1.0166	0.9916	0.9745	1.0011	1.0026	0.9995	0.9985	0.9781	1.0014
Tianjin	1.0051	0.9601	1.0029	0.9703	0.9394	0.9361	0.9776	0.9768	0.9723	0.9798	0.9662	0.9721
Xinjiang		1.0504	1.0611		1.0045	0.9813	0.9827					1.0063
Yunnan	0.9871	0.9869	0.9895	0.9875	0.9544	0.9801	0.9917	0.9856	0.9785	0.9733	0.9809	0.9814
Zhejiang	1.0227	1.0432	1.0431	1.0404	1.0034	1.0019	1.0146	1.0196	1.0148	1.0142	0.9967	1.0188
Panel B: Mean of Economies of Scale by Clustered Regions												
1	1.0265	1.0365	1.0268	1.0413	0.9935	0.9874	1.0004	1.0102	0.9971	0.9895	0.9668	1.0035
2	1.0125	0.9556	0.967	0.9569	0.9563	0.9581	0.9737	0.9605	0.96	0.959	0.943	0.9636
3	1.023	1.0297	1.0311	1.0305	0.9994	0.9936	1.003	1.0109	1.0064	1.0052	0.9876	1.0091

This table present mean value of economies of scale estimated at each data point for each province and each cluster. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

We can observe from Table 4.21 that the three provinces with the highest level of economies of scale are Beijing, Heilongjiang and Anhui, and most inland regions are showing diseconomies of scale. Figure 4.11 gives a visual illustration of the trends of ES for the three clusters over 2005-2015. We can see

that the ES value for all three clusters is decreasing over the period, following a similar trend of a steady increase in the level of economies of scale over the years. However, it is also clear that cluster 2, which are the Eastern Municipalities, seems to enjoy much greater economies of scale than the other two regions for the entire period. This is in line with the trend figure for cost efficiencies shown in Figure 4.10 that the core financial and economic centers represented by cluster 2 are also performing better in terms of meta-frontier cost efficiencies as they operate closer to the meta-frontier technology. One potential policy implication is that regional banks in China would benefit from mergers if they upgrade to the meta-frontier technology.

Figure 4.11. The trend of Economies of scale



This figure presents economies of scale for three clusters from 2005 to 2015. Economies of scale are estimated for each bank using cluster-specific frontiers. Cluster 1= the Inland Region, Cluster 2= the Eastern Municipalities and Cluster 3= the Northeastern and Coastal Region.

We now turn our attention to the second stage of our empirical estimation. We begin by presenting the summary statistics for the variables used in the regression (4.31a) and (4.31b) in table 4.22. In terms of security to total asset ratio $SEC_TA_{i,t}$, our sample range from banks with no investment banking activities to banks with 68% of the assets as securities. We also note that the sample includes an extensive range of bank liquidities, with the lowest liquid asset to customer deposit as 3% to the most liquid bank with the liquidity ratio of more than 340%.

Table 4.22. Summary Statistics for Determinants of Economies of Scale

Variable	Obs	Mean	Std. Dev.	Min	Max
$ES_{i,t}$	1,050	1.0038	0.0592	0.8391	1.3302
$SEC_TA_{i,t}$	1,050	0.2366	0.1188	0.0000	0.6808
$NIM_{i,t}$	1,050	0.0338	0.0115	0.0053	0.1005
$LR_{i,t}$	1,050	0.3933	1.1616	0.0318	34.5674
$LLP_{i,t}$	1,050	0.0089	0.0077	-0.0745	0.0662
$ETA_{i,t}$	1,050	0.0734	0.0539	0.0126	0.9471
$Tier1_{i,t}$	788	0.1327	0.2379	0.0278	4.2258
$Zscore_{i,t}$	1,050	29.0126	34.9322	1.7009	513.3940

$ES_{i,t}$ are the economies of scale calculated from cluster-specific frontiers. $SEC_TA_{i,t}$ are total securities divided by total asset as a proxy for investment banking activities. $NIM_{i,t}$ is the net interest margin, calculated as net interest income divided by total loans, representing the income from traditional lending operations. The liquidity ratio $LR_{i,t}$ is calculated as total liquid assets to total customer deposit and $LR_{i,t}^2$ is used to capture the non-linear relationship between liquidity and economies of scale. The loan loss provisions to gross loans $LLPLOAN_{i,t}$ is to measure the credit risk in each bank and $Zscore_{i,t}$, which is the sum of return on average asset and equity divided by standard deviation of return on average asset, formulated as $Zscore_{it} = (ROAA_{it} + Equity_{it})/\sigma_{ROAA_i}$. $Tier1_{i,t}$ ratio is computed as Tier 1 capital scaled by risk weighted assets and $ETA_{i,t}$ is total equity to total assets.

Table 4.23 presents the results from fixed effect panel data regression on determinants of economies of scale in the Chinese regional banking sector. Model (1) presents the estimates from (4.31a) without Zscore. Model (2) is the reduced model, including only variables with significance from Model (1) with Zscore. Model (3) presents the estimates from (4.31b) without Zscore. Model (4) is the reduced model, including only variables with significance from Model (2) with Zscore. We find that investment banking activities, higher liquidity ratio, and lower credit risk would contribute to economies of scale while banks' profitability from traditional lending activities does not have any impact on scale economies. We also find that a higher level of capital, measured by both overall capital ratio and Tier 1 capital, would have a positive impact on economies of scale. This funding supports China's compliance with Basel III capital and liquidity regulations, as more stringent regulation is likely to increase economies of scale in Chinese regional banks. Due to a lack of data on systemic risks, we

use Zscore as a measure for insolvency risk, and we do not find any evidence of a link between bankruptcy risk and economies of scale. By and large, these results align with findings by Beccalli et al. (2015) for EU banks and Bertay et al. (2013) for US banks.

Table 4.23. Determinants of Scale Economies

ES	(1)	(2)	(3)	(4)
$SEC_TA_{i,t}$	-0.124***	-0.124***	-0.134***	-0.129***
	(0.010)	(0.010)	(0.012)	(0.011)
$NIM_{i,t}$	-0.000		-0.001	
	(0.001)		(0.001)	
$LR_{i,t}$	0.035***	0.034***	-0.095***	-0.095***
	(0.003)	(0.003)	(0.013)	(0.013)
$LR_{i,t}^2$	-0.001***	-0.001***	0.075***	0.072***
	(0.000)	(0.000)	(0.010)	(0.010)
$LLPLOAN_{i,t}$	0.415***	0.416***	0.201	
	(0.126)	(0.125)	(0.139)	
$ETA_{i,t}$	-0.001***	-0.002***		
	(0.001)	(0.001)		
$ETA_{i,t}^2$	0.000***	0.000***		
	(0.000)	(0.000)		
$Zscore_{i,t}$		0.000		-0.000
		(0.000)		(0.000)
$Tier1_{i,t}$			-0.038**	-0.028
			(0.018)	(0.022)
$Tier1_{i,t}^2$			0.004	0.002
			(0.004)	(0.005)
$Constant$	1.026***	1.024***	1.062***	1.060***
	(0.006)	(0.004)	(0.007)	(0.005)
N	1050	1050	788	788
F	47.442	47.786	28.169	32.502
Fixed effects	Yes	Yes	Yes	Yes
r2	0.268	0.269	0.232	0.229

Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. Hausman test rejects the null hypothesis for all four equations, favouring fixed-effects over random-effects model. Model (1) present the estimates from (4.31a) without Zscore. Model (2) is the reduced model including only variables with significance from Model (1) with Zscore. . Model (3) present the estimates from (4.31b) without Zscore. Model (4) is the reduced model including only variables with significance from Model (2) with Zscore. $SEC_TA_{i,t}$ are total securities divided by total asset as a proxy for investment banking activities. $NIM_{i,t}$ is the net interest margin, calculated as net interest income divided by total loans, representing the income from traditional lending operations. The liquidity ratio $LR_{i,t}$ is calculated as total liquid assets to total customer deposit and $LR_{i,t}^2$ is used to capture the non-linear relationship between liquidity and economies of scale. The loan loss provisions to gross loans $LLPLOAN_{i,t}$ is to measure the credit risk in each bank and $Zscore_{i,t}$, which is the sum of return on average asset and equity divided by standard deviation of return on average asset, formulated as $Zscore_{it} = (ROAA_{it} + Equity_{it})/\sigma_{ROAA_i}$ . . $Tier1_{i,t}$ ratio is computed as Tier 1 capital scaled by risk weighted assets and $ETA_{i,t}$ is total equity to total assets.

4.11 Conclusion and Policy Implications

In this Chapter, we evaluate the performance of Chinese regional banks using a stochastic meta-frontier (SMF) estimation methodology. To the best of our

knowledge, this is the first study to explore this aspect of the Chinese banking system. Firstly, Chinese provinces are clustered by macroeconomic variables using HAC and k-means methods to reveal a natural classification of regions in China. We propose a new algorithm that conducts hierarchical and k-means clustering in one setting for k groups before deciding on the optimal number of clusters. This method allows the starting centres of k-means clustering to be determined by the results of hierarchical clustering without any arbitrary initial point. Three clusters were identified to be the most statistically viable solution, cross-validated by silhouette width.

We continue to estimate both cluster-specific cost frontiers and a meta-frontier in order to test the assumption of common banking technology. A series of log-likelihood tests confirm that banks in these clusters do not follow the same technology and a pooled model is not appropriate. Both equity and time show a significant and non-linear relationship with cost-efficiency. From the individual frontier and meta-frontier estimates, we can find that different clusters are affected by different microeconomic variables. For the clusters which are affected by microeconomic variables, we find a higher level of equity and a lower level of non-performing loans are associated with a higher level of cost-efficiency. Both the pooled models and meta-frontier consistently show that lower concentration, higher competition and foreign bank presence tend to increase bank cost efficiency and narrow down the gap between the group-specific technology and the best available technology portrayed by the meta-frontier. Individually, pooled models indicate the negative impact of higher inflation on cost efficiency, which could indicate potential excessive expansion of branch networks of Chinese banks in response to fluctuating and high inflation rates. The meta-frontier model further finds that urbanisation has a positive impact on shrinking the technology gap between individual clusters and meta-frontier.

From a methodological standpoint, our findings provide support to Bos and Schmiedel (2007) which also find that a meta-frontier gives a more complete set

of information than pooled frontiers. Furthermore, we concur that a meta-frontier is free of the biases which are present in the pooled models. Contrary to Bos and Schmiedel (2007) and Lee and Huang (2017), who find small technology gaps in the European meta-frontier, we find significant technology gaps in the Chinese banking industry. We link this wide technology gap to the nature of the banking sector, the choice of sample and the choice of macroeconomic variables. Sun et al. (2013) utilise a very similar homogeneous sample by only including city and rural commercial banks. However, since they estimate profit efficiency, the efficiency scores are not directly comparable with our cost efficiency scores. Nevertheless, their profit efficiency scores are generally on the same scale as our cost efficiency scores using pooled models. However, our meta-frontier models tell a very different story, and we observe an obvious difference in technology gaps by looking at the diagram of CGR trends in section 4.9.6.

In terms of cost efficiency scores, since only meta-frontier scores can be compared across different regions, we find banks in the North-eastern and Coastal Region (cluster 3) to be the most cost-efficient and banks in the Inland Region (cluster 1) to be the least cost-efficient. We find that for all clusters, the main contributor to cost inefficiency is inferior technology adopted compared to the best practice bank in the country, although there are regional variations in terms of the severity of the problem. This discovery gives policy makers a clear idea in terms of the source of the inefficiency for each region. For instance, the worst-performing cluster (the Inland Region) also deviates furthest from meta-frontier compared to the other two clusters, indicating that banks in the Inland Region (cluster 1) need to focus more on adopting a superior technology than the other two slightly better-performing clusters. For the best performing cluster, i.e. North-eastern and Coastal Region (cluster 3), it seems that its individual frontier cost-efficiency remains the lowest among the three clusters. This indicates that these clusters are lagged behind more in terms of management compared to other clusters. Therefore cluster 3 can also pay attention to management training apart from technology advancements. The overarching

theme of our findings is that all three clusters need to exercise more focus on adopting a superior technology.

After pointing out the sources of cost inefficiency, our results also inform policy makers in terms of the main approaches to reduce technology gaps and, in turn, reduce cost inefficiency. This information can be derived from the significance and sign of microeconomic and macroeconomic variables in the inefficiency model estimated as part of the meta-frontier, presented in table 4.17. Since the technology gap acts as the primary source of inefficiency, we find that speeding up urbanization, reducing banking concentration, increasing banking competition and increasing foreign bank presence could directly help to reduce technology gaps. In terms of managerial inefficiency implied by bank cluster-specific frontiers, we find that banks should mainly focus on increasing their capital and increasing the quality of loans and reducing their bad loans to achieve better bank management. Profitability does not seem to prompt cost efficiency in Chinese regional banks.

In terms of economies of scale, under current technology, banks in the Inland Region (cluster 1) and the North-eastern and Coastal Region (cluster 3) exhibit diseconomies of scale while banks in the Eastern Municipalities (cluster 2) enjoy economies of scale. If they moved towards the meta-frontier by adopting a better technology, Inland Region (cluster 1) and the North-eastern and Coastal Region (cluster 3) would start to enjoy economies of scale, and Eastern Municipalities (cluster 2) would continue to enjoy an even greater degree of economies of scale.

Regarding determinants of economies of scale for Chinese regional banks, we find that investment banking activities, higher liquidity ratio, lower credit risk would contribute to economies of scale, while banks' profitability from traditional lending activities does not have any impact on scale economies. We also find that a higher level of capital would have a positive impact on economies of scale.

The results from estimations of economies of scale and their determinants seem to suggest that it may be beneficial for Chinese regional commercial banks to continue adopting a prudent capital and liquidity regulation. A lower level of credit risk also promotes economies of scale. Furthermore, regional banks in China can look to expand and diversify the scope of their business such as growing investment banking and other non-traditional banking businesses.

Chapter 5. Loan Loss Provisioning Behaviour of Chinese Regional Banks

5.1 Introduction

This chapter investigates the determinants of loan loss provisioning behaviour, the relationship between loan risk and cost efficiency, and the impact of bank funding structure on earnings quality for Chinese regional banks. The contribution of this study is mainly three-fold. Firstly, it investigates various hypotheses of loan loss provisioning in the context of China's compliance with Basel regulations. It also examines the impact of the global financial crisis on the discretional use of LLPs. Six main hypotheses are tested under this context: procyclicality or countercyclicality, earnings management, capital management, signalling, and the 'bad management' hypothesis. To ensure the robustness of our inferences, we utilize two commonly used methods in the literature to test these hypotheses. The first method is to test all the hypotheses in a one-stage regression. The second method is to follow Beatty and Liao (2014) and Jin et al. (2018) to first extract the discretionary element of LLPs using a first stage regression on non-discretional components of LLPs. In the second stage, we test the hypotheses by regressing a series of test variables on the discretional component of LLPs, represented by the residuals from first-stage regressions.

The second contribution of this study is that we utilize the cost efficiency scores obtained from the meta-frontier estimations, and we examine the relationship between LLPs and cost efficiency to test Berger and DeYoung(1997)'s 'bad management' hypothesis. As mentioned in chapter 4, cost efficiency scores estimated via a meta-frontier framework allow comparisons of efficiency scores across different heterogeneous groups. In our case, we chose regional banks (city and rural commercial banks) where heterogeneity among banks primarily come from different macroeconomic conditions in different regions in China. We find low-efficiency banks tend to have a higher level of loan loss provisions, thus supporting the 'bad management' hypothesis.

Thirdly, this study investigates the relationship between bank funding structure and earnings quality in the context of China. Specifically, it examines whether reliance on retail deposit or wholesale funding makes a bank more likely to manipulate its earnings in the financial statement. We find a negative relationship between retail deposits and discretional LLPs, suggesting that banks with higher reliance on retail deposits as a funding source tend to be less likely to manipulate their earnings.

To the best of our knowledge, this study is the first to test all six hypotheses in this new context - China. This study provides investors and other stakeholders of Chinese commercial banks a simplified model to acknowledge the sensitivities of loan performance to macroeconomic conditions as well as evidence for any income-smoothing or signalling behaviour. This study also provides regulators insights on the indicators for banks' earnings quality.

5.2 Literature Review and Hypotheses Development

LLPs can be split into two components, general provisions and specific provisions. Specific provisions in their nature are procyclical, as they are related to specific loans that are highly probable to default. General provisions are provided against possible future losses, which involves a higher degree of judgement by bank managers (Laeven and Majnoni, 2003, Soedarmono et al., 2012). General provisions can be further divided into a discretionary component and a non-discretionary component. The non-discretionary component covers expected credit losses and is backwards-looking, while the discretionary component covers unexpected credit losses and is forward-looking (Beaver and Engel, 1996, Wahlen, 1994). Thus, the discretionary component of the LLPs is believed to be subject to management discretion for the purpose of creating a countercyclical buffer, smoothing income, meeting the regulatory capital requirements or signalling financial resilience and cash flow potential to stock markets and investors.

In this section, we will present a literature review on loan loss provisioning behaviours relating to theories of Procyclicality/ Countercyclicality, income-smoothing, capital management, signalling, funding quality and 'bad management'. Hypotheses are also developed for each theory at the end of each section.

5.2.1 Procyclicality/ Countercyclicality

The procyclicality hypothesis concerns the timing of LLP with respect to the business cycle, i.e. how LLPs respond to the business cycle and in turn exacerbate it. LLPs can be either procyclical or countercyclical with the economic cycle. LLPs are considered procyclical when they increase during an economic downturn, reflecting the increased loan default rates and increased bank cost, and deemed countercyclical when LLPs increase during an economic upturn, which creates a buffer against the economic cycle. That is, to increase the buffer against future losses and to smooth income, especially when LLPs are believed to build-up due to excessive loan growth during economic expansion.

The introduction of Basel Accords has raised concerns over the procyclical nature of the increased capital requirements. On the one hand, the new higher capital adequacy ratio would limit the credit supply in the economy in general. On the other hand, the risk-based capital ratio calculations inevitably lead to higher capital requirements during recession due to higher default risk, further amplifying capital shortages during the economic downturn, exactly when credit and loans are needed the most (Bikker and Metzemakers, 2005). There is a consensus in the literature that capital should cover unexpected losses while LLPs should cover the expected losses. If LLPs become inadequate to cover the expected losses, the losses can eat into capital, further discourages banks from granting loans, therefore worsening an economic recession (Soedarmono et al., 2012).

The rationale for the procyclical hypothesis is that, during economic recessions, incomes are likely to decrease, and the probability of default on loans in society

will increase consequently. Therefore, LLPs tend to increase during an economic upturn and decrease during an economic downturn, reflecting the likelihood of defaults in the loan portfolio following the business cycle (Bikker and Metzemakers, 2005).

In terms of the rationale for counter-cyclical provisioning, there are mainly two explanations. One explanation is by Borio et al. (2001) and Lowe (2003) that problematic loans are more likely to be generated during an economic upswing when there is substantial loan growth, and the credit losses tend to be materialised during a downturn. This is because the positive outlook held by bank managers during economic expansion can result in a lax lending behaviour as they are more willing to take higher risks in lending during good times, thus leading to increased riskiness of loan portfolio. This relaxed lending criteria during an economic boom can also be explained by the institutional memory hypothesis Berger and Udell (2004), which states that as time passes from the last economic bust, fewer and fewer bad loans will be associated with the current loan managers, especially when the staff turnover is high. This means that institutional memory on the previous problem loans during the last economic downturn tend to fade as the economy recovers. Consequently, as the economy climbs up from the bottom of the cycle, current bank managers tend to ease their criteria for lending, thus leading to a gradual increase in problem loans and a decrease in loan quality. Therefore, because of the abovementioned cyclical lending behaviour, Berger and Udell (2004) propose that LLPs should be increased during economic upturn to reflect the risk accumulation from lax lending as well as to build up loan loss reserves as a buffer which can then be drawn later during a potential downturn. The countercyclical behaviour may also result from banks managers using LLPs as a tool to smooth earnings, i.e., increasing provisions during upturn when earnings are high and reducing provisions during a downturn when earnings are low. (See section 5.2.2 income smoothing for more details). Fernandez de Lis et al. (2000) propose a dynamic provisioning framework which is generally adopted in Spain since 2000. A difference between expected credit losses and specific provisions using either banks' internal model or a standard model set by the regulators is calculated, in

addition to traditionally computed specific provisions and general provisions. The dynamic or statistical provisioning is set to reduce the correlation between the specific provisions and business cycle by anticipating an economic downturn in the future. Therefore, the aforementioned dynamic provisioning approach tends to increase LLPs during an economic upturn.

The cyclical effect of LLPs is of importance in China, especially due to bank loans being the primary financial source for private businesses in the country. Within the relevant literature, Bikker and Metzemakers (2005) find procyclicality of loan loss provisioning behaviour in 29 OECD countries. Another important study by Glen and Mondragón-Vélez (2011) finds that economic growth is the main driver for improved loan performance. Particularly, they find that loan performance only deteriorates when there is extreme economic distress in their sample of 22 major developing economies from 1996 to 2008. Laeven and Majnoni (2003) investigate the cyclical effect of regulatory capital requirements on bank lending using a large global dataset from BankScope. Their findings suggest that banks largely delay provisioning during a cyclical upturn with favourable income conditions until an economic downturn sets in. The authors propose a separate regulation on loan loss reserves from capital regulations, which would allow loan loss reserves to fluctuate during the economic cycle above a minimum level. A study on 12 countries in emerging Asia, including China, India and Southeast Asia, by Packer and Zhu (2012) find that, contrary to European countries, countercyclicality is the predominant behaviour of LLPs in emerging Asian economies during 2000 to 2009. However, they do not find any significant evidence to support either procyclical or countercyclical behaviour in China. We develop our hypothesis as below:

H1: Chinese Banks are procyclical in terms of loan loss provisioning practice.

5.2.2 Income Smoothing

The second hypothesis concerns the use of LLP to smooth earning streams. According to the income smoothing hypothesis, bank managers have

discretionary power towards LLPs, and they tend to use LLPs as an instrument to manipulate profit on the bank income statements. This is because LLPs, as a relatively large accrual for banks, are deducted in a bank's income statement to arrive at the profit for the year. The incentive for income smoothing is commonly explained by managerial self-interest, where managers' compensation schemes are related to bank performance, usually indicated by reduced volatility of reported profit and the stability of stock prices for listed banks (Lambert, 1984). Laeven and Majnoni (2003) provide an additional rationale that LLPs can prevent the possibility of credit loss eating into a bank's capital if perfect income smoothing is achieved. That is, banks draw from loan loss reserves when actual credit losses exceed the expected credit losses in bad times and increase loan loss reserves by adding in additional LLPs when actual losses are less than expected losses during good times. Therefore, expected losses would no longer affect capital.

There has been an inconclusive debate about whether income smoothing has a positive effect on the economy. Several studies argue that income smoothing can mitigate the procyclical effect of LLPs on bank lending (Borio et al., 2001, Lowe, 2003, Beatty and Liao, 2011). On the other hand, other studies argue that LLPs suffering from income smoothing cannot truly recognize expected loan losses. Therefore, this behaviour will reduce the transparency of bank financial statements, harm market discipline and increase difficulties for monitoring (Bushman and Williams, 2012).

There have been several studies on the income smoothing behaviour of LLPs. Early studies in the area provide mixed findings. For example, Greenawalt and Sinkey (1988) find significant evidence for 106 large bank holding companies in the US between 1976 to 1984. They also find that regional banking companies have more aggressive income smoothing behaviour than money-centre banks. Additionally, regulatory policy, risk management and compensation schemes all impact income smoothing through LLPs. Contrary to the study mentioned above, Beatty et al. (1995) and Ahmed et al. (1999) find no evidence of income smoothing in US banks. Anandarajan et al. (2007) find evidence of income

smoothing in Australian banks, while Bouvatier and Lepetit (2008) find the opposite that European Banks increase LLPs when earnings are low during 1992-2004, amplifying the economic cycle. Later research by Fonseca and González (2008) tests this hypothesis in a large sample of 40 countries and finds that when there is stronger investor protection, more effective supervision, more accounting disclosure and banking restrictions, banks are less likely to use income smoothing as a tool for risk management. A more recent study by Bushman and Williams (2012) further investigates whether income-smoothing harms market discipline. The results show that if the discretion in accounting is purely for the purpose of earnings management, then the income smoothing behaviour is likely to reduce transparency in banking systems and cause trouble for monitoring. However, if exercising discretionary power towards LLPs is a form of forward-looking accounting to recognise future expected losses in a timely fashion, then this behaviour can, in fact, enhance market discipline.

Due to a higher level of scrutiny and market discipline for listed banks, listed banks are found to engage more in income-smoothing than non-listed banks. Moreover, banks are also found to practice earnings management more intensively during a financial crisis period (Anandarajan et al., 2007, Curcio et al., 2017)[15].

Therefore, we develop our hypotheses as below:

[15] In addition to listed status and financial crisis, the degree of earnings management may also be impacted by corporate governance. The literature concerning the impact of corporate governance on earnings management mainly focuses on non-financial firms. Within the limited literature on the impact of corporate governance on banking firms, Sullivan and Spong (2007)find that bank managers are more likely to manipulate earnings when they have a higher level of wealth concentrated in the bank's stock holdings. Cornett et al. (2009) study the impact of corporate governance mechanism on earnings management using a sample of large publicly traded bank holding companies. They use pay-for-performance sensitivity, board independence and capital ratio as main corporate governance variables. They find that increased board independence, reduced CEO pay-performance sensitivity and a higher level of capital are linked to reduced earnings management. Leventis and Dimitropoulos (2012) investigate corporate governance and earnings management on US-listed banks during the Sarbanes-Oxley Act era (2003-2008) and find strong evidence that well-governed banks engage less in aggressive earnings management using LLPs, in addition to realised security gains and losses. We do not investigate how corporate governance, such as board independence and CEO compensation, impact earnings management, as it is outside the scope of this study.

H2: Chinese Banks engage in earnings management in terms of loan loss provisioning practice.

H2a: Listed Chinese Banks engage more aggressively in earnings management in terms of loan loss provisioning practice than non-listed banks.

H2b: Listed Chinese Banks engage more aggressively in earnings management in terms of loan loss provisioning practice during a financial crisis period.

5.2.3 Capital Management

According to the capital management hypothesis, banks use LLPs to reduce the cost of violating capital requirements. That is, banks with low capital levels tend to increase general provisions to meet the regulatory capital requirement and ensure their capital adequacy (Kim and Kross, 1998, Cortavarria et al., 2000). However, after implementing Basel Accords II and III, a less significant relationship between bank capital and LLPs is expected. The main reason is that under Basel I and II standardized approaches, LLPs do not count for Tier I capital and can only be counted as Tier II capital up to 1.25% of risk-weighted assets. Under Basel II Internal Rating Based (IRB) approach, the amount of provisions used to complement capital to cover unexpected losses is subject to 0.6% of credit risk-adjusted assets (BCBS, 2005). We will review studies in the US and cross-country studies below.

In terms of studies in the US, Collins et al. (1995) do not find any evidence of capital management. Kim and Kross (1998) find that the substantial exclusion of LLPs for bank capital calculation after the regulatory reform in 1989 led to reduced incentives for capital management through LLPs. However, no significant change in LLPs has been found for banks with large capital ratios after the reform. Ahmed et al. (1999) find evidence of capital management from 1986 to 1995.

Capital management is also researched in cross-country studies. Bikker and Metzemakers (2005) find capital management both on the OECD and EU levels from 1991 to 2001. However, they find no significant evidence of capital management in the UK. Anandarajan et al. (2007) find a negative relationship between LLPs and Tier 1 capital for Australian banks over 1991-2001, supporting the capital management hypothesis. However, they do not find any change in this relationship after the onset of Basel accords. Pérez et al. (2008) find no evidence for capital management in Spanish banks from 1986 until 2002. However, Bushman and Williams (2012) find capital management in their sample of 27 countries from 1995 to 2006. Curcio et al. (2017) find no evidence for capital management for the Euro Area for the period of 2005-2011.

H3: Chinese Banks use LLPs as a mechanism to meet capital requirements.

5.2.4 Signalling

In this section, we will discuss the theories and hypotheses relating to using LLPs as a method for signalling future cash flows and financial resilience.

LLPs are the largest accruals for banks. LLPs can be used as a device to signal bank investors and bank clients about their earnings ability and future cash flows. According to Akerlof (1970), information asymmetry between the firm's managers and investors may lead to stock mispricing. If managers' remuneration is tied to stock performance, managers may be incentivised to use LLPs as a tool to signal the firm's ability to generate positive cash flows to avoid under-pricing of the firm.

The other way to ascertain the incentive for firms to signalling is that the cost of false signalling must be significantly high (Spence, 1973). According to Kanagagretnam et al. (2004), poorly performing banks are unlikely to increase LLPs because increased LLPs will further reduce already poor retained earnings,

thus increase the risk of being audited by regulatory agencies and reduce managers' compensation that is based on reported earnings.

LLPs are be considered either as "good news" or "bad news". An increase in LLPs is often regarded as "good news" in the market when it is accompanied by other timely information about loan portfolios and loan defaults. Especially, the discretionary or unexpected element of LLPs is considered to signal good news as this is a signal that management believes the earnings ability of the bank to be strong and that it can withstand 'a hit in earnings' by increasing LLPs (Beaver et al., 1989). It is also believed that additional LLPs could be a sign of prudent accounting policies and better monitoring of loan portfolios. On the contrary, LLPs can also be seen as "bad news" when there is no other information available on loan defaults; thus, LLPs become the only sources of information on loan defaults, thus signalling risk in the loan portfolio (Anandarajan et al., 2007).

Two variables are commonly used to measure this signalling behaviour: stock returns or future earnings. In this study, we will use future earnings as the proxy since a large proportion of banks in our sample are not listed. Mixed results have been found in the literature. The earliest study by Wahlen (1994) finds a positive relationship between LLPs and pre-provision earnings as well as current stock returns. Beaver and Engel (1996) decompose LLPs into discretionary and non-discretionary components and find that they are positively and negatively related to stock price, respectively. On the contrary, Ahmed et al. (1999) find a negative relationship between LLPs and future earnings as well as contemporaneous stock returns. Eng and Nabar (2007) support the signalling hypothesis and find a positive relationship between unexpected LLPs and stock returns as well as future cash flows in a sample of banks in Hong Kong, Singapore and Malaysia from 1993 to 2000. They also find that signalling was less effective during a crisis period. However, a study on Australian banks by Anandarajan et al. (2007) finds no evidence of using LLPs as a tool for signalling from 1991 to 2001. Leventis et al. (2012) find that find insufficient evidence for signalling good news for financially healthy banks but find signalling is more pronounced for

financially distressed banks and more aggressively after the implementation of IFRS.

H4: Chinese Banks use discretionary LLPs to signal good news regarding future cash flows and financial resilience.

5.2.5 The impact of funding structure on earnings quality

In addition to the four hypotheses mentioned above, we also want to investigate the impact of the funding structure, i.e. the extent bank relies on retail or wholesale funding, on banks' manipulation of earnings figures on their financial statements. As opposed to earnings management, where lower LLPs are expected for periods with higher earnings, in order to investigate earnings quality, we do not focus on the direction of the earnings, rather the magnitude of the discretionary LLPs. That is, we only focus on the absolute values of the discretionary component of LLPs.

Jin et al. (2018) define earnings quality as the extent banks manipulate earnings information disclosed on their financial statement. A lower earnings quality is associated with the heavier use of discretionary LLPs as a tool for earnings management. The incentives to manipulate earnings, thus decreasing earnings quality, can be linked to a bank's funding structure (Jin et al., 2018).

There are two opposing views in the literature regarding the effect of the bank's funding structure, specifically, banks' reliance on retail funding or wholesale funding, on earnings quality. The first view that retail deposit is negatively related to banks earnings quality is proposed by Calomiris and Kahn (1991). They suggest that depositors have fewer resources and incentives to monitor bank earnings management, more so in the presence of deposit insurance schemes. This negative relationship may exist in the context of the Chinese banking industry. According to He (2016), while China's explicit deposit insurance is still at an early stage of establishment, there has been a long history of implicit deposit insurance in China, provided as a form of government

guarantee. With public ownership laying the foundation of the Chinese banking system, significant portions of ownership of Chinese banks come from local government or state-owned enterprises, who provide injections in case of potential liquidity issues and government bailouts in the event of potential bank failures. This is also evidenced in the multiple capital injections into Chinese commercial banks and the establishment of asset management companies to dispose of non-performing loans over the course of Chinese banking reform since 1998 (See details in section 2.2 in Chapter 2). This improves the security of deposits and reduces the need to scrutinize the banks' financial information for retail depositors. Li and Han (2008)'s panel data study finds that the implicit deposit insurance reduces banks' prudence in risk-taking and banks' self-discipline mechanisms are unsuccessful for both state-owned and non-state-owned banks due to it. Therefore, banks that rely more on retail deposit funding may disclose lower quality earnings information than banks that rely more on wholesale funding in China.

However, Jin et al. (2018) support the view that the opposite is more likely to be true. That is, earnings quality may, in fact, be lower for banks that rely more on wholesale funding. Firstly, wholesale funding lenders might not frequently use bank financial statements as their source of information. Iyer et al. (2016) find evidence that wholesale depositors tend to continue providing funding to troubled banks due to the difficulty to infer potential problems from the financial statement. This view is also supported by Gorton and Metrick (2012), who claim that some wholesale funding is relatively safe, sometimes even safer than retail deposits. They give the example of a repurchase agreement (repo), where lenders can terminate the agreement in the event of banks' default and obtain the collateral, which is typically of more value than the repo itself. Thus wholesale funders have little incentive to monitor bank earnings information due to the low risk of lending. Secondly, several studies in the literature suggest that depositors react to banks' non-financial information. For example, Homanen (2018) finds that US bank retail depositors react negatively to bank scandals. Thirdly, Kim et al. (2003) find that retail deposits, as a more stable source of funding, are stickier due to high cost for retail depositors to switch banks for transaction

services, as opposed to higher volatility of wholesale funding, which is short term and on a rollover basis, subject to market volatilities and economic shocks. Thus, banks relying on retail deposit are less incentivised to alter financial information due to their reduced sensitivity to market volatilities. Jin et al. (2018) indeed find evidence that earnings quality to be higher among banks relying more on retail funding than wholesale funding for US banks. Therefore, we develop the hypothesis as below:

H5: Chinese Banks which rely more on retail deposits as a source of funding will have less discretionary LLPs, thus higher earnings quality compared to banks that rely on wholesale funding.

5.2.6 The Relationship between Loan Quality and Bank Efficiency

To the best of our knowledge, there are limited studies in the literature that focus on examining the relationship between LLPs and bank efficiency (Williams, 2004)

Berger and DeYoung (1997) propose a 'bad management' hypothesis that inefficient banks tend to have higher credit risks within their loan portfolio. The rationale is that insufficient and inefficient screening, underwriting, and monitoring of loan portfolios result in higher risks in a bank's loan portfolio. This may be due to inefficient use of resources by the bank loan managers due to incompetence or lack of skills. Other studies also support the 'bad management' hypothesis that inefficient banks impose higher loan default risk, which may lead to bank failure. Some studies find that failing banks tend to be the most inefficient banks (Berger and Humphrey, 1990, Barr and Siems, 1997). Other studies such as Kwan and Eisenbeis (1997) and Resti (1996) find that problem loans are related to inefficiency even in banks that are unlikely to fail. Therefore, we develop our hypothesis as:

H6: More efficient banks incur less loan loss provisions.

5.3 Methodology

5.3.1 Baseline Model specification:

Following Anandarajan et al. (2007), the baseline model specification to test hypotheses 1-4 can be indicated as below:

Model 1a:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GDPGR_t + \alpha_4 EBTPTA_{it} + \alpha_5 ETA_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 SIZE_{it} + \varepsilon_{it} \quad (5.1a)$$

The dependent variable is LLPs to total gross loans denoted as $LLPLOAN_{it}$, $LLPLOAN_{it-1}$ is the dependent variable with one year lag. It is logical to assume that the LLPs in the previous year will affect the LLPs in the current year, as managers tend to use the provision from last year as a reference and make adjustments to it according to the conditions in the current year. Thus, lagged LLPs are expected to be positively correlated with LLPs.

$NPLLOAN_{it}$ is NPLs scaled by gross loans to capture the expected loan losses; LLPs are designed to cover the expected loss in the loan book, both general and specific provisions are required to be related to the bad loans. Therefore, I expect a positive relation between LLP and NPL.

$GDPGR_t$ is real GDP growth rate; As discussed in section 5.2.1, the classic view by Borio et al. (2001) and the Institutional memory hypothesis (Berger and Udell, 2004) supports procyclicality, and there is strong evidence of procyclicality in the banking Loan Loss Provisioning literature (Bikker and Metzemakers, 2005,

Glen and Mondragón-Vélez, 2011). Thus, we expected a negative relationship between LLP and GDP growth rate.

$EBTPTA_{it}$ is earnings before provisions and tax scaled by total assets; As mentioned above in Section 5.2.2, we test for income smoothing by including earnings before tax and provision as a proxy. It is expected that Chinese commercial banks may smooth income in order to have a better performance in terms of stock price and dividend payment. Managers may also smooth income as their compensation scheme is often linked with earnings stability. Thus, we expect LLPs to be positively related to EBTPTA.

ETA_{it} is total equity scaled total assets as a proxy for capital. As mentioned before in section 5.2.3, under Basel II and III, regulatory capital needs to cover expected loss if loan loss provisions cannot cover it entirely. Thus, banks have the incentive to increase LLPs if they have difficulty meeting CAR requirements.

We use equity over total assets as a proxy for total regulatory capital to test the capital management hypothesis. The direction of the sign of coefficient can be either positive or negative. This is because total regulatory capital contains Tier 1 and Tier 2 capital, and the impact of LLPs on Tier 1 and Tier 2 are in opposite directions. Under Basel accords, Tier 1 capital represents core capital containing common stocks plus retained earnings, fewer intangibles and regulatory adjustment. The regulatory adjustment is stricter under Basel III than in previous versions. Tier 2 capital contains cumulative preferred stocks, perpetual subordinate debt, revaluation reserves, unrealised gains on securities and excess loan loss reserves. Therefore, bank managers can manipulate LLPs to meet the total regulatory capital requirement by either increasing retained earnings through decreasing LLPs, in order to increase tier 1, or by increasing LLPs to achieve increased loan loss reserves, in order to increase Tier 2 capital, although loan loss reserves are subject to a maximum of 1.25% of risk-weighted assets.

$LOANTA_{it}$ are gross loans to total assets. Loan to asset measures credit risk

exposure. Banks with more credit risk exposures provide more provisions. Thus, the sign of the loan to asset ratio is expected to be positive.

$CFEER_{it}$ represents fee and commission income scaled by average assets as a proxy for trading activities; According to Anandarajan et al. (2007), banks with high commission fee income are the banks that engage in a high volume of different trading activities and services other than traditional lending. Thus, they tend to provide more LLPs to provide an image of a safer institution. Therefore, I expect the sign of commission fee income to be positive.

$CUSTF_{it}$ is the customer deposit ratio, calculated as total customer deposit to total funding. Customer deposit to total funding measures the stability of funding for lending. Thus, the lower the proportion of deposit intake, the less stable funding they have, the probability of insolvency increases. Thus, banks have the incentive to provision more to cover more of future expected losses. Thus, the sign of customer deposit to total funding is likely to be negative.

$SIZE_{it}$ is calculated as the logarithm of total assets. To control for size so that our explanatory variables do not just capture size effects, we use the logarithm of total assets. The bigger banks are expected to have a higher level of LLPs than smaller banks.

ε_{it} is the error term representing all the unobserved impacts on LLPs.

As the literature shows a variety of proxies for the variables above, we will include alternatives to see how they will impact our model.

Model 1b:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GDPGR_t$$

$$+\alpha_4 EBTPTA_{it} + \alpha_5 Tier1_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 SIZE_{it} + \varepsilon_{it} \quad (5.1b)$$

Model 1c:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GRPGR_{pt} + \alpha_4 EBTPTA_{it} + \alpha_5 ETA_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 SIZE_{it} + \varepsilon_{it} \quad (5.1c)$$

Model (1b) substitutes ETA with Tier1 ratio, as both proxies are present in the literature, although the signs of the coefficient are expected differently. For Tier 1, we expect a positive coefficient as a low Tier 1 ratio prompt bank managers to decrease LLPs in order to increase retained earnings as part of Tier 1 capital. In terms of capital proxied by total equity by total asset, the sign of coefficient can be positive or negative. Model 1c in equation $(5.1c)$ uses $GRPGR_{pt}$, the growth rate of gross regional product on provincial level instead of GDP growth $GDPGR_t$ to capture different regional economic environments for each province.

To further test the hypotheses that listed banks practice more income-smoothing than non-listed banks (H2a) and banks involve in more income-smoothing during the financial crisis period (H2b), we add in the interaction term between income-smoothing variable $EBTPTA_{it}$ and dummy variables *listed* and *crisis*, where *listed* takes the value of 1 for listed banks and the value of 0 for non-listed banks. *Crisis*'s value is 1 for the period 2007 to 2011 and 0 otherwise, in line with Curcio et al. (2017).

Model 1d:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GDPGR_t + \alpha_4 EBTPTA_{it} + \alpha_5 ETA_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 SIZE_{it} + \alpha_{10} EBTPTA * listed + \alpha_{11} EBTPTA * crisis + \alpha_{11} EBTPTA * listed * crisis + \varepsilon_{it} \quad (5.1d)$$

Listed:

Because listed banks use the stock market as a source of funding, we expect listed banks to practice more earnings management and signalling. Therefore, we include interactions terms Listed*EBTP.

As discussed in section 5.2.2, our hypothesis is: *H2a: Banks may engage in earnings management more aggressively if it is a publicly listed bank.*

Crisis:

Curcio et al. (2017) argue that during a financial crisis, deteriorated loan quality would lead to an increase in specific/non-discretionary provisions, which would, in turn, reduce the possibility for bank managers to use LLPs to manage income and capital.

Using the wide range of literature that investigates the impact of the global financial crisis on bank risks and performances, including studies such as (Curcio et al., 2017), it is generally considered that 2008-2011 is regarded as the crisis period. In order for our study to be comparable with current literature, we regard the same period as the crisis period and all other years as the non-crisis period.

As discussed in section 5.2.2, Our hypothesis is: *H2b: Banks may engage in earnings management more aggressively during financial crisis periods to smooth the income.*

To test H6, we add in cost efficiencies obtained from chapter 4 from the meta-frontier model for comparison in order to test if the 'bad management' hypothesis by Berger and DeYoung (1997) holds.

Model 2:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GDPGR_t + \alpha_4 EBTPTA_{it} + \alpha_5 Tier1_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 SIZE_{it} + \alpha_{10} CE_{Meta_{it}} + \alpha_{11} EBTPTA * listed + \alpha_{12} EBTPTA * crisis + \alpha_{13} EBTPTA * listed * crisis + \varepsilon_{it} \quad (5.2)$$

Cost Efficiency (CE_I, CE_meta):

Please see chapter 4 for a detailed calculation of these two variables. We expect the sign of CE_I, CE_meta to be negative. That is, the higher the efficiency score, the lower the loan risk, thus a lower level of LLPs.

5.3.2 Testing of Signalling

To test the signalling hypothesis, we follow Bouvatier and Lepetit (2008) and Soedarmono et al. (2012), and we include $SIGN_{it}$ into equation (5.3) - calculated as (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)].We expect a positive sign for $SIGN_{it}$, that is, higher loan loss provisions signal good news about higher future earnings.

Model 3:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 LLPLOAN_{it-1} + \alpha_2 NPLLOAN_{it} + \alpha_3 GDPGR_t + \alpha_4 EBTPTA_{it} + \alpha_5 Tier1_{it} + \alpha_6 LOANTA_{it} + \alpha_7 CFEER_{it} + \alpha_8 CUSTF_{it} + \alpha_9 CE_{Meta_{it}} + \alpha_{10} EBTPTA * listed + \alpha_{11} EBTPTA * crisis + \alpha_{12} EBTPTA * listed * crisis + SIGN_{it} + \varepsilon_{it} \quad (5.3)$$

An alternative measurement to test signalling is proposed by Soedarmono et al. (2012), where the one year ahead change in earnings before tax and provisions are divided by the average of current total assets and next year's total assets. The advantage of this measurement is that it can avoid the potential problem of multicollinearity if we use △EBTPTA.

5.3.3 Estimating discretionary LLPs

Before we present an alternative way to test the signalling hypothesis using the two-stage approach, we first need to obtain the discretionary component of LLPs.

The literature has presented several ways to measure the discretionary component of LLPs. Generally, it is measured by the residual of an OLS regression on the non-discretionary components.

We estimate discretionary LLPs using two different models in order to ensure the robustness of our results. Following Beatty and Liao (2014) and Bushman and Williams (2012), equation (5.4) models the non-discretionary component of LLPs as a function of past and future changes in NPLs, change in size, total size and economic environment. LLP_t is LLPs scaled by beginning gross loans. △

NPL_{it} is the change in non-performing loans scaled by the gross loans in their respective years; $Size_{it-1}$ is the logarithm of total asset. The other variables are calculated in the same way as the baseline model (Model 1).

Model 4:

$$LLP_{it} = \alpha_0 + \alpha_1 \triangle NPL_{it+1} + \alpha_2 \triangle NPL_{it} + \alpha_3 \triangle NPL_{it-1} + \alpha_4 \triangle NPL_{it-2} + \alpha_5 Size_{it-1} + \alpha_6 LOANTA_{it} + \alpha_7 GRP + Province\ Dummies + Year\ Dummies + \varepsilon_{it} \quad (5.4)$$

Furthermore, in order to confirm that our findings are not driven by the above specification, we follow Kanagaretnam et al. (2010), and we estimate equation (5.5):

Model 5:

$$LLP_{it} = \alpha_0 + \alpha_1 LLA_{t-1} + \alpha_2 NPL_{t-1} + \alpha_3 CO_t + \alpha_4 \Delta Loan_t + \alpha_5 Loan_t + \varepsilon_t \quad (5.5)$$

Where LLA_{it} is loan loss allowance scaled by the beginning of year gross loans, CO_t is loan charge-offs scaled by beginning gross loans and NPL_{t-1} is prior year's non-performing loans scaled by the beginning of year gross loans. $Loan_t$ and $\Delta Loan_t$ capture the loan volume and its change in the current year.

As both above models capture the non-discretional components of the LLPs, the residuals of the above equation $DLLP1_{it}$ $DLLP2_{it}$ are used as proxies for discretional LLPs. Taking absolute values of these residuals, we obtain $ABDLLP1_{it}$ and $ABDLLP2_{it}$.

5.3.4 Alternative test of Signalling

Model 3 tests signalling by including variable $SIGN_{it}$ as an explanatory variable. Two alternative models proposed by Ahmed et al. (1999) are tested here to validate our result further. Equation (5.6) uses one year ahead change in EBPT as a proxy for signalling while controlling for NPLs, economic cycle, earnings management, capital management and listing status. The second set of models (5.7a) and (5.7b) predict the one-year ahead change in earnings $EBTPA_{t+1}$ as a function of discretionary LLPs, where the discretionary components are obtained from equations (5.4) and (5.5) and are scaled by the beginning of year value of equity.

Model 6:

$$LLPLOAN_{it} = \alpha_0 + \alpha_1 \Delta LLA_{it} + \alpha_2 \Delta GDP_t + \alpha_3 Tier1_{it} + \alpha_4 EBTPTA_{it} + \alpha_5 Listed + \alpha_6 \triangle EBTPTA_{it} \quad (5.6)$$

Model 7a:

$$EBTPA_{t+1} = \alpha_0 + \alpha_1 EBTPA_t + \alpha_2 ABDLLP1E_{it} + year\ effect + \varepsilon_{it} \quad (5.7a)$$

Model 7b:

$$EBTPA_{t+1} = \alpha_0 + \alpha_1 EBTPA_t + \alpha_2 ABDLLP2E_{it} + year\ effect + \varepsilon_{it} \quad (5.7b)$$

Where $ABDLLP1E_{it}$ and $ABDLLP2E_{it}$ are the absolute values of discretional loan loss provisions, which are obtained as the residual from equations (5.4) and (5.5), and then divided by the beginning of year market value of total equity. ΔLLA is the change in loan loss allowance, and all other variables are defined the same way as above sections.

5.3.5 Testing the Impact of Retail Deposit on Earnings Quality

Following Jin et al. (2018), we can test the impact of customer deposits on earnings quality using the equation below:

Model 8a[16]:

$$ABDLLP1_{it} = \alpha_0 + \alpha_1 CLD_{lt-1} + \alpha_2 SIZE_{lt} + \alpha_3 SIZE_{lt}^2 + \alpha_4 LOAN_{lt} + \alpha_5 CAPR_{lt} + \alpha_6 EBTPTA_{lt} + \alpha_7 LLP_{lt-1} + \alpha_8 LIQUID_{lt} + \alpha_9 VOLROA_{lt} + \alpha_{10} \Delta EBTPTA_{lt+1} + \alpha_{11} Listed + \varepsilon_{lt} \quad (5.8a)$$

Model 8b:

$$ABDLLP1_{it} = \alpha_0 + \alpha_1 CLD_{lt-1} + \alpha_2 SIZE_{lt} + \alpha_3 SIZE_{lt}^2 + \alpha_4 LOAN_{lt} + \alpha_5 CAPR_{lt} + \alpha_6 EBTPTA_{lt} + \alpha_7 LLP_{lt-1} + \alpha_8 LIQUID_{lt} + \alpha_9 VOLROA_{lt} + \alpha_{10} \Delta EBTPTA_{lt+1} + \alpha_{11} Listed + \varepsilon_{lt} \quad (5.8b)$$

Model 8c:

$$ABDLLP2_{it} = \alpha_0 + \alpha_1 CLD_{lt-1} + \alpha_2 SIZE_{lt} + \alpha_3 SIZE_{lt}^2 + \alpha_4 LOAN_{lt} + \alpha_5 CAPR_{lt} + \alpha_6 EBTPTA_{lt} + \alpha_7 LLP_{lt-1} + \alpha_8 LIQUID_{lt} + \alpha_9 VOLROA_{lt} + \alpha_{10} \Delta EBTPTA_{lt+1} + \alpha_{11} Listed + \varepsilon_{lt} \quad (5.8c)$$

Model 8d:

$$ABDLLP2_{it} = \alpha_0 + \alpha_1 CLD_{lt-1} + \alpha_2 SIZE_{lt} + \alpha_3 SIZE_{lt}^2 + \alpha_4 LOAN_{lt} + \alpha_5 CAPR_{lt} + \alpha_6 EBTPTA_{lt} + \alpha_7 LLP_{lt-1} + \alpha_8 LIQUID_{lt} + \alpha_9 VOLROA_{lt} + \alpha_{10} \Delta EBTPTA_{lt+1} + \alpha_{11} Listed + \varepsilon_{lt} \quad (5.8d)$$

[16] As discussed in footnote 15, we do not investigate the effects of corporate governance variables such as executive compensation, internal controls and board independence on discretionary LLPs as they are out of the scope of this study. The only variables in our model that could potentially capture corporate governance of banks are capital ratio and listed status in model 8 (a,b,c,d). Cornett et al. (2009) argue that shareholders can limit bank managers' ability to invest in negative-NPV projects and force bank managers to distribute free cash flow to shareholders when banks have lower capital levels, i.e., higher leverage. Moreover, banks with lower capital adequacy levels are subject to more scrutiny by regulators, thus having less chance to manipulate earnings. However, our results from model 8 (a,b,c,d) (See result analysis in section 5.4) indicate that there is no evidence suggesting capital adequacy affect earnings management for Chinese banks.

Where $ABDLLP1_{it}$ and $ABDLLP2_{it}$ are the absolute values of discretional loan loss provisions, obtained as the residual from equations (5.4) and (5.5). CLD_{lt-1} is core deposit to total liabilities in the previous year. Moreover, it is our core variable to determine the relationship between earnings quality and funding structure. As discussed in Section 5.25 in this chapter, earnings quality is considered high if there is a lower level of discretionary LLPs, suggesting fewer manipulations on the earnings figure. $CAPR_{lt}$ controls for capital management, measured by total liability scaled by total equity, and $EBTPTA_{lt}$ controls for income-smoothing. $LIQUID_{lt}$ is calculated as liquid assets divided by total assets. Banks with a low level of liquid assets are expected to engage more in LLPs manipulation. $VOLROA_{lt}$ is the volatility of return on assets, which is expected to have a negative relationship with earnings quality thus a positive relationship with discretionary LLPs. The rest of the variables in these equations are calculated in the same way as the baseline equations.

5.3.6 Variable descriptions

Table 5.1 presents a summary of all the variables and their expected signs used in this study. As these variables are explained with the model specification, we will not explain further the variables.

Table 5.1. List of variables and expected sign

Variable	Description	Expected sign
$LLPLOAN_{it}$	Loan loss provisions/ Average total loans	N/A
$LLPLOAN_{it-1}$	Lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year)	+
$NPLLOAN_{it}$	Non-performing loans/ Average total loans	+
$GDPGR_t$	Gross domestic product growth rate	–
$GRPGR_{pt}$	Gross regional product growth rate	
$EBTPTA_{it}$	Earnings before tax and provisions/average total asset	+
$SIZE_{it}$	LnTA, i.e., the logarithm of total assets	+
$\triangle EBTPTA_{it}$	One year ahead EBTPTA- current year EBTPTA	+
$Tier1_{it}$	Tier 1 capital/ risk-adjusted weighted assets	+
$LOANTA_{it}$	Gross Loans/Total assets	+
$DLOAN_{it}$	Loan growth rate	+
$CFEER_{it}$	Commission and fee income/ Average total asset	+
$CUSTF_{it}$	Customer deposits to total funding	–
CE_I_{it}	Individual frontier cost-efficiency	–
CE_meta_{it}	Cost efficiencies estimated from the metafrontier in chapter 4	–
$SIGN_{it}$	(one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total assets +next year's total assets)]	+
CDL_{lt-1}	(Core deposits)/(total liabilities) in year t-1	+/-
$CAPR_{lt}$	Total liability/(total equity) in year t	+/-
$LIQUID_{lt}$	(liquid assets)/(total assets) in year t	–

$VOLROA_{lt}$	(standard deviation of net income)/(total asset) in year t	+
ETA_{it}	Equity / Total assets	+/-
listed	Takes the value of 1 for listed banks and value of 0 for non-listed banks	+
Crisis	Value is 1 for the period 2007 to 2009 and 0 otherwise	+
UNR_{it}	Unemployment rate	+
NPL_{it}	Non-performing loans / gross loans in the beginning of the year	+
LLA_{it}	Loan Loss Allowance / gross loans in the beginning of the year	+
CO_{it}	Net Charge Offs / gross loans in the beginning of the year	+
$\Delta Loan_{it}$	(Gross loans for current year – gross loans for previous year) / Total Assets in the previous year	+
$ABDLLP1_{it}$	The absolute value of discretionary LLPs estimated as the residual of equation (5.4)	N/A
$ABDLLP2_{it}$	The absolute value of discretionary LLPs estimated as the residual of equation (5.5)	N/A
$ABDLLP1E_{it}$	ABDLLP1 / beginning of year market value of total equity	N/A
$ABDLLP2E_{it}$	ABDLLP2 / beginning of year market value of total equity	N/A
$LOANDEP_{it}$	Loan divided by customer deposits	+
$EBTPTE_{it}$	Earnings before tax and provisions/average total equity	+

5.4 Empirical Results

This section firstly describes data and provides descriptive statistics, and then we present the test results on hypotheses 1-6. We then offer an alternative way for testing signalling using discretionary LLPs. Finally, we present the results for the test on the impact of retail deposits on earnings quality.

5.4.1 Data Description

The dataset contains a unique panel dataset of 136 regional Chinese commercial banks during the 11-year period from 2005 to 2015 with a total of 928 observations, as well as a set of province-level macroeconomic variables for all the provinces these banks operate in. Our 11 years of data is sufficiently long enough to capture both economic upswings and downturns in order to investigate cyclicality, especially when it covers a 2008-2009 global financial crisis. Bank-level balance sheet and income statement information are collected from the Fitch BankScope database (now Bureau van Dijk Orbis Bank Focus). All accounts comply with IFRS standards, and all bank data are deflated by the GDP deflator using 2005 as the base year. We collect province-level macroeconomic variables from the official website of the National Bureau of Statistics of China, and we deflate them using again 2005 as the reference year. We only include regional commercial banks to ensure the homogeneity of the objectives of the banks in the dataset. We also excluded banks that have less than three years of observations to accommodate for the panel nature of our analysis.

We do not further truncate the data in the upper or lower bounds of the distributions as banks at the end of the spectrums may be more prone to engage in manipulating LLPs. Removing such observations may remove meaningful patterns of the data (Balboa et al., 2013) 1-3[17].

[17] It is worth noting that examining nonlinearities in income smoothing is beyond the scope of this thesis, though an interesting avenue for future research.

5.4.2 Descriptive Statistics

In Table 5.2, we present mean values for each variable used in baseline models 1-3 in equations (5.1) to (5.3)). Furthermore, we compare listed with non-listed banks for both the all bank sample and the regional bank sample. In addition, we perform a Wlocoxon_Mann-Whitney test of mean equality between regional banks and national banks.

Firstly, for all banks, the mean value of $LLPLOAN_{it}$ is 0.87%, validating the fact that LLPs are a relatively large accrual in banks. National banks report a statistically lower level of LLPs, which could be linked to the lower relative level of NPLs, although the difference of $NPLLOAN_{it}$ between national banks and regional banks are not significant. The average value of $NPLLOAN_{it}$ is 1.95%, with a generally decreasing trend over the period partly because of the practice of dumping NPLs to government-owned asset management companies(Kwan, 2009). The fee and commission income ratio
$CFEER_{it}$ is 0.1% for all banks and statistically higher for national banks compared to regional banks. This can be due to national banks being large universal banks offering a wider range of services while regional banks focus more on traditional lending activities as the primary source of income. National banks are also found to have higher loan to asset ratio levels as expected, as national banks lend to larger businesses and state-owned projects while local banks focus more on consumer lending. The average Tier 1 ratio is 10%, which is more than what is required by Basel II and III. National banks also have statistically significant higher earnings and higher capital levels, with the latter possibly due to market discipline. The average customer deposit as a proportion of total funding is 82%, with regional banks relying on customer deposits more than national banks. This indicates that Chinese banks still largely rely on customer deposits as their primary source of funding.

In terms of comparing listed and unlisted banks, listed banks, on average, have lower levels of LLPs than non-listed banks, are larger in size and have a higher

level of investment banking trading activities. Listed banks also, on average, have higher earnings to asset ratio. Table 5.3 presents the mean values in crisis and non-crisis periods. The mean value of $LOANTA_{it}$ is statistically higher for the crisis period compared to the non-crisis period, for both the all banks sample and the regional banks sample, indicating that there is a higher demand for credit during the crisis period. The tier 1 ratio is also found to be higher during the crisis period for both samples, in line with increased credit risk in the loan portfolio and higher market disciplining activities during the crisis period. Banks also seem to rely more on customer deposits as bank funding during the crisis period for both samples to lower the risk. For the regional bank sample, banks seem to be more cost-efficient during the crisis period, as banks seem to attempt to minimise costs during the financial crisis to mitigate further damage caused by GFC.

Table 5.4 presents the correlation coefficients of variables in the baseline models 1-3 for all banks, listed banks and regional banks. We can observe that $LLPLOAN_{it}$ has statistically significant correlation coefficients with most of the variables in the table. It has positive correlation coefficients with $NPLLOAN_{it}, LOANTA_{it}$ and $CUSTF_{it}$ and is negatively correlated with $CFEERit,\ SIZE_{it}$ and $Tier1_{it}$. The correlations between explanatory variables are relatively low, meaning that there is little risk of the problem of multicollinearity.

Table 5.2. Descriptive Statistics – Listed vs non-listed and National vs Regional Banks

	All Banks	Listed	Non-listed	Listed vs unlised banks	National Banks	Regional Banks	National vs Regional Banks	Regional Banks		
	Mean (sd)	Mean (sd)	Mean (sd)	Difference in means (p-value)	Mean (sd)	Mean (sd)	Difference in means (p-value)	Listed Mean (sd)	Unlisted Mean (sd)	Listed vs Unlisted Difference in means (p-value)
$LLPLOAN_{it}$	0.0087	0.0073	0.0095	-0.0023***	0.0070	0.0090	-0.0021***	0.0076	0.0095	-0.0020***
	(0.0063)	(0.0046)	(0.0070)	(p<0.0001)	(0.0041)	(0.0066)	(0.0010)	(0.0052)	(0.0070)	(0.00040)
$NPLLOAN_{it}$	0.0195	0.0177	0.0204	-0.0028	0.0182	0.0197	-0.0015	0.0172	0.0205	-0.0033
	(0.0269)	(0.0244)	(0.0282)	(0.3010)	(0.0284)	(0.0266)	(0.6710)	(0.0200)	(0.0285)	(0.2016)
$CFEER_{it}$	0.0019	0.0027	0.0014	0.0013***	0.0040	0.0014	0.0026***	0.0017	0.0013	0.0003***
	(0.0017)	(0.0021)	(0.0013)	(p<0.0001)	(0.0023)	(0.0012)	(p<0.0001)	(0.0014)	(0.0012)	(p<0.0001)
$SIZE_{it}$	11.4537	12.9583	10.6155	2.3428***	14.4666	10.8682	3.5984***	11.7844	10.5612	1.2232***
	(1.7888)	(1.7683)	(1.1308)	(p<0.0001)	(1.2162)	(1.1945)	(p<0.0001)	(1.0484)	(1.0782)	(p<0.0001)
$LOANTA_{it}$	0.4950	0.4947	0.4952	-0.0005	0.5220	0.4898	0.0322***	0.4752	0.4947	-0.0195***
	(0.1061)	(0.0823)	(0.1174)	(0.2134)	(0.0622)	(0.1120)	(0.0035)	(0.0891)	(0.1183)	(0.0023)
$EBTPTA_{it}$	0.0190	0.0181	0.0195	-0.0014***	0.0175	0.0193	-0.0018***	0.0183	0.0196	-0.0013**
	(0.0059)	(0.0036)	(0.0068)	(0.0087)	(0.0037)	(0.0062)	(0.0057)	(0.0037)	(0.0068)	(0.0265)

Table 5.2. Descriptive Statistics Listed vs Unlisted -continued

	All Banks				All Banks			Regional Banks		
	All Banks Mean (sd)	Listed Mean (sd)	Unlisted Mean (sd)	Listed vs unlised banks Difference in means (p-value)	National Banks Mean (sd)	Regional Banks Mean (sd)	National vs Regional Banks Difference in means (p-value)	Listed Mean (sd)	Unlisted Mean (sd)	Listed vs Unlisted Difference in means (p-value)
ETA_{it}	0.0651	0.0598	0.0680	-0.0082***	0.0546	0.0671	-0.0125***	0.0624	0.0687	-0.0063***
	(0.0235)	(0.0161)	(0.0263)	(p<0.0001)	(0.0152)	(0.0243)	(p<0.0001)	(0.0167)	(0.0261)	(0.0021)
$Tier1_{it}$	0.1068	0.0991	0.1110	-0.0119***	0.0885	0.1106	-0.0220***	0.1063	0.1119	-0.0056
	(0.0394)	(0.0246)	(0.0451)	(p<0.0001)	(0.0205)	(0.0413)	(p<0.0001)	(0.0251)	(0.0453)	(0.4340)
$CUSTF_{it}$	0.8222	0.8054	0.8316	-0.0262***	0.7949	0.8275	-0.0326***	0.8095	0.8336	-0.0241***
	(0.1231)	(0.0999)	(0.1335)	(p<0.0001)	(0.0942)	(0.1274)	(p<0.0001)	(0.1039)	(0.1339)	(0.0003)
$SIGN_{it}$	0.0039	0.0036	0.0041	-0.0004	0.0033	0.0040	-0.0008**	0.0039	0.0041	-0.0002
	(0.0041)	(0.0025)	(0.0048)	(0.1957)	(0.0022)	(0.0044)	(0.0402)	(0.0029)	(0.0048)	(0.7988)
CE_meta_{it}	-	-	-	-	-	0.7792	-	0.7831	0.7779	0.0052*
	-	-	-	-	-	(0.0568)	-	(0.0566)	(0.0568)	(0.0665)
Observations	928	332	596	-	151	777		195	582	

$LLPLOAN_{it}$= Loan loss provisions/ Average total loans, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$= Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4

Table 5.3. Descriptive Statistics- Crisis vs Non-crisis period

	All Banks			Regional Banks		
	Crisis Mean (sd)	Non-crisis Mean (sd)	Crisis vs . Non-crisis Difference in means (p-value)	Crisis Mean (sd)	Non-crisis Mean (sd)	Crisis vs . Non-crisis Difference in means (p-value)
$LLPLOAN_{it}$	0.0077	0.0093	-0.0016***	0.0081	0.0097	-0.0016***
	(0.0055)	(0.0067)	(p<0.0001)	(0.0058)	(0.0070)	(0.0008)
$NPLLOAN_{it}$	0.0157	0.0218	-0.0061***	0.0164	0.0218	-0.0054***
	(0.0269)	(0.0267)	(p<0.0001)	(0.0291)	(0.0247)	(p<0.0001)
$CFEER_{it}$	0.0017	0.0019	-0.0002	0.0014	0.0015	-0.0001
	(0.0014)	(0.0019)	(0.4475)	(0.0010)	(0.0014)	(0.2688)
$SIZE_{it}$	11.3413	11.5243	-0.1829**	10.766	10.9324	-0.1665**
	(1.6811)	(1.8512)	(0.0264)	(1.0132)	(1.2925)	(0.0101)
$LOANTA_{it}$	0.5107	0.4852	0.0254***	0.5074	0.4788	0.0286***
	(0.0933)	(0.1124)	(0.0003)	(0.0994)	(0.1180)	(0.0004)
$EBTPTA_{it}$	0.0193	0.0188	0.0005	0.0198	0.0190	0.0008
	(0.0062)	(0.0057)	(0.9677)	(0.0065)	(0.0059)	(0.4437)
ETA_{it}	0.0670	0.0639	0.0032	0.0696	0.0656	0.0040**
	(0.0251)	(0.0224)	(0.1927)	(0.0261)	(0.0229)	(0.0473)
$Tier1_{it}$	0.1131	0.1025	0.0106***	0.1183	0.1053	0.0129***
	(0.0421)	(0.0370)	(0.0003)	(0.0437)	(0.0388)	(p<0.0001)
$CUSTF_{it}$	0.8564	0.8008	0.0557***	0.8654	0.8037	0.0618***
	(0.1113)	(0.1254)	(p<0.0001)	(0.1137)	(0.1298)	(p<0.0001)
$SIGN_{it}$	0.0039	0.0039	0.0000	0.0041	0.0040	0.0001*
	(0.0041)	(0.0041)	(0.1337)	(0.0043)	(0.0045)	(0.0776)
CE_meta_{it}	-	-	-	0.7868	0.7744	0.0124**
	-	-	-	(0.0495)	(0.0605)	(0.0101)
Observations	358	570		300	477	

$LLPLOAN_{it}$= Loan loss provisions/ Average total loans, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$= Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA-current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4

Table 5.4. Correlation Coefficient

	$LLPLOAN_{it}$	$GRPGR_{it}$	$GDPGR_{it}$	$NPLLOAN_{it}$	$CFEER_{it}$	$SIZE_{it}$	$LOANTA_{it}$	$EBTPTA_{it}$	ETA_{it}	$Tier1_{it}$	$CUSTF_{it}$	$SIGN_{it}$
	Panel A: All Banks											
$LLPLOAN_{it}$	1											
$GRPGR_{pt}$	0.026	1										
$GDPGR_t$	0.063	0.707***	1									
$NPLLOAN_{it}$	0.262***	0.169***	0.308***	1								
$CFEERit$	-0.140***	-0.425***	-0.316***	-0.206***	1							
$SIZE_{it}$	-0.238***	-0.347***	-0.179***	-0.087**	0.671***	1						
$LOANTA_{it}$	0.097**	0.064	0.320***	0.126***	-0.131***	-0.166***	1					
$EBTPTA_{it}$	0.243***	-0.098**	-0.090**	-0.096**	0.037	-0.194***	0.341***	1				
ETA_{it}	-0.036	-0.207***	-0.289***	-0.203***	0.016	-0.207***	0.086**	0.312***	1			
$Tier1_{it}$	-0.104***	-0.087**	-0.224***	-0.240***	-0.042	-0.208***	-0.144***	0.107***	0.840***	1		
$CUSTF_{it}$	0.126***	0.177***	0.273***	0.189***	-0.244***	-0.155***	0.495***	0.358***	0.007	-0.101**	1	
$SIGN_{it}$	-0.111***	0.267***	0.264***	-0.038	-0.143***	-0.196***	0.024	-0.218***	-0.047	-0.046	0.031	1
	Panel B: Listed Banks											
$LLPLOAN_{it}$	1											
$GRPGR_{pt}$	-0.040	1										
$GDPGR_t$	0.034	0.693***	1									
$NPLLOAN_{it}$	0.386***	0.301***	0.501***	1								
$CFEER_{it}$	-0.040	-0.579***	-0.372***	-0.230***	1							
$SIZE_{it}$	-0.122*	-0.482***	-0.128**	-0.029	0.742***	1						
$LOANTA_{it}$	0.008	0.045	0.320***	0.201***	-0.057	0.079	1					
$EBTPTA_{it}$	0.206***	-0.156**	-0.277***	-0.175***	0.316***	-0.026	0.079	1				
ETA_{it}	-0.089	-0.106*	-0.297***	-0.279***	0.138**	-0.149**	-0.183***	0.492***	1			
$Tier1_{it}$	-0.105	0.048	-0.196***	-0.263***	-0.010	-0.247***	-0.379***	0.408***	0.866***	1		
$CUSTF_{it}$	0.038	0.270***	0.364***	0.339***	-0.180***	-0.076	0.526***	0.187***	0.062	0.049	1	
$SIGN_{it}$	-0.187***	0.206***	0.234***	-0.041	-0.136**	-0.193***	-0.042	-0.358***	-0.100	-0.085	-0.054	1

Table 5.4. Correlation Coefficient - Continued

	$LLPLOAN_{it}$	$GRPGR_{it}$	$GDPGR_{it}$	$NPLLOAN_{it}$	$CFEER_{it}$	$SIZE_{it}$	$LOANTA_{it}$	$EBTPTA_{it}$	ETA_{it}	$Tier1_{it}$	$CUSTF_{it}$	$SIGN_{it}$	CE_meta_{it}
	Panel C: Regional Banks												
$LLPLOAN_{it}$	1												
$GRPGR_{pt}$	0.001	1											
$GDPGR_t$	0.086*	0.691***	1										
$NPLLOAN_{it}$	0.260***	0.139***	0.301***	1									
$CFEER_{it}$	-0.100**	-0.325***	-0.371***	-0.253***	1								
$SIZE_{it}$	-0.218***	-0.262***	-0.329***	-0.100**	0.373***	1							
$LOANTA_{it}$	0.138***	0.081*	0.339***	0.137***	-0.285***	-0.435***	1						
$EBTPTA_{it}$	0.236***	-0.095**	-0.047	-0.095**	0.001	-0.260***	0.394***	1					
ETA_{it}	-0.055	-0.252***	-0.282***	-0.217***	0.075*	-0.168***	0.135***	0.272***	1				
$Tier1_{it}$	-0.144***	-0.137***	-0.221***	-0.266***	0.060	-0.106**	-0.104**	0.050	0.826***	1			
$CUSTF_{it}$	0.142***	0.165***	0.283***	0.171***	-0.348***	-0.241***	0.512***	0.382***	-0.021	-0.142***	1		
$SIGN_{it}$	-0.117***	0.264***	0.281***	-0.040	-0.130***	-0.195***	0.028	-0.225***	-0.067	-0.069	0.024	1	
CE_meta_{it}	-0.348***	-0.301***	-0.343***	-0.324***	0.233***	0.240***	-0.009	0.058	0.293***	0.253***	-0.296***	-0.054	1

$LLPLOAN_{it}$= Loan loss provisions/ Average total loans, $GRPGR_{pt}$ = Gross Regional Product growth rate, $GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$= Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4

5.4.3 A test of the Loan loss provisioning hypotheses

Tables 5.6 and 5.7 present the tests of six hypotheses on behaviour of LLPs using a system GMM dynamic panel data model, with Table 5.5 presenting the result on a sample of all banks and Table 5.6 reporting the results of the regression results on only regional banks. Columns (1) to (4) correspond to model 1a, 1b, 1c and 1d, with different measures for capital and alternation between gross domestic product and gross regional product. Column (5) offers interaction between crisis and $EBTPTA_{it}$, column (6) offers the interaction between Listed and $EBTPTA_{it}$. column (7) includes three-way interaction between listed, crisis and $EBTPTA_{it}$.

A system GMM model is used to estimate the baseline specification. It is designed to mainly address the below econometric issues: (i) to control for the unobservable bank effect; (ii) dynamic nature of LLPs; and (iii) potential endogeneity of the explanatory variables. As our model contains lagged dependent variable, using a static model such as OLS would cause an upward bias in the coefficient of the lagged dependent variables as the latter is positively correlated with the residuals, and using the within estimator (fixed effects) would cause a negative bias (Nickel bias) in the coefficient of the lagged dependent variable (Baum, 2006).

The generalized method of moments (GMM) estimator is used for the parameters of this model, which was originally developed by Arellano and Bond (1991). This estimator is designed for datasets with many panels (N) and few periods (T). The data after cleaning contains 136 banks with 11 years, which fit the small T large N data structure. As mentioned before, the reason for using the Arellano-Bond technique is that linear dynamic panel data models, which include endogenous variables (such as lagged dependent variables), contain unobserved panel level effects. The unobserved panel-level effects are correlated with the endogenous variables, making standard estimators

inconsistent (Baum, 2006). The Arellano-Bond dynamic GMM estimator requires that there be no autocorrelation in the idiosyncratic errors. The GMM model's consistency also depends on the validity of instruments (Fonseca and González, 2008).

(i) The Sargan test of over-identifying restrictions will be used to test the validity of instruments as a whole. The rejection of the Sargan test means that the model is satisfactorily identified and the instruments are valid overall. That is, there is no correlation between instruments and the error term.

(ii) The Arellano-Bond test for second-order autocorrelation in first differences will be used. The null hypothesis for second-order autocorrelation should always be rejected. The first-order autocorrelation test in the differentiated residuals does not need to be rejected because this autocorrelation is due to the first differences in the model(Mileva, 2007). In this study, I will use the *xtabond2* command in STATA written by (Roodman, 2006) to implement the Arellano-Bond estimator as it is the latest version and is more advanced than *xtabond*.

The lagged dependent variable is treated as a pre-determined variable, meaning that it is independent of current disturbances but influenced by past disturbances; therefore, it is not strictly exogenous. GDPGR is treated as strictly exogenous, independent of the error term. Because it is a macroeconomic variable, thus it is unlikely that LLPs would explain the growth rate of the gross regional product. We treat the rest of the explanatory variables as potentially endogenous since they all have the potential to be determined within the system and reverse causality with the dependent variable. The SGMM model automatically creates the instrumental variables as SGMM uses one equation in levels (using lagged difference as IVs) and one in first differences (using lagged levels as IVs).

Therefore, before we go into the result analysis, we will first validate that our

results do not come from models that are misspecified. In table 5.5, AR (1) is the first-order autocorrelation, and AR (2) is the second-order serial correlation. As mentioned above, autocorrelation indicates that lags of the dependent variable are endogenous and thus are bad instruments. Therefore, we should not accept second-order autocorrelation. However, first-order correlations are expected in differences. Table 5.5 indicates that AR(2) are all insignificant for all seven models, indicating that there is no second-order serial correlation. In terms of the Hansen J-Statistic, all of the p values are between 0.25 to 0.7, and therefore we do not reject the joint null hypothesis that the instruments are valid. Therefore, we can say that the instruments used in our SGMM model are valid instruments.

Table 5.5 presents the SGMM estimation results for all banks. Firstly, column (2) shows statistically significant and negative coefficients for $GDPGR_{it}$. This indicates that there is some limited evidence that when GDP increases, banks provision less, following the economic cycle. This indicates that banks in China could be procyclical in terms of provisioning practice. This is consistent with the finding that provisioning in China is strongly associated with non-performing loans, seen as the strongly significant coefficient for NPLs. We do not find any evidence of capital management. This may be because Chinese banks are generally well-capitalised, therefore reducing the need to use LLPs to meet regulatory capital adequacy requirements. We find consistent evidence for income-smoothing across all samples and all models during the study period in terms of earnings management. The interaction term $EBTPTA_{it} * crisis$ is significant and positive across all samples, indicating that banks engage in even more income-smoothing during the crisis period, in line with the findings by Curcio et al. (2017) for EU banks and El Sood (2012). They believe that this is because banks managers tend to shift risk and signal financial resilience to various stakeholders during a crisis by reducing earnings volatility. The customer deposit variable is significant and negative across all models, indicating a possible association between retail deposits and earnings quality proposed by Jin et al. (2018), which will be examined in detail in section 5.4.6.

The lagged dependent variable is also significant and positive across all samples and all models, indicating that banks adjust provisions gradually over time to match the change in the loan portfolio, in line with findings by Bouvatier and Lepetit (2008) and Soedarmono et al. (2012).

Table 5.5. Determinants of Loan Loss Provisions-All Banks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$LLPLOAN_{it}$	0.207***	0.207***	0.229***	0.227***	0.244***	0.244***	0.250***
	(0.060)	(0.057)	(0.058)	(0.058)	(0.060)	(0.063)	(0.061)
$GDPGR_t$	-0.022	-0.031*		-0.006	-0.015	-0.017	-0.021
	(0.018)	(0.019)		(0.019)	(0.019)	(0.018)	(0.020)
$NPLLOAN_{it}$	0.184***	0.209***	0.180***	0.198***	0.200***	0.201***	0.204***
	(0.047)	(0.047)	(0.049)	(0.052)	(0.049)	(0.049)	(0.049)
$CFEER_{it}$	0.270	0.422	0.366	0.208	0.222	0.260	0.283
	(0.395)	(0.375)	(0.428)	(0.409)	(0.486)	(0.445)	(0.451)
$SIZE_{it}$	0.001	-0.000	0.000	0.001*	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$LOANTA_{it}$	0.005	0.000	0.002	0.008	0.005	0.006	0.004
	(0.010)	(0.008)	(0.010)	(0.009)	(0.009)	(0.007)	(0.008)
$EBTPTA_{it}$	0.560***	0.444***	0.583***	0.632***	0.481***	0.420**	0.418***
	(0.128)	(0.115)	(0.135)	(0.133)	(0.153)	(0.168)	(0.155)
ETA_{it}	0.010		0.007	0.026	0.015	0.021	0.019
	(0.033)		(0.035)	(0.028)	(0.032)	(0.031)	(0.030)
$CUSTF_{it}$	-0.016***	-0.012**	-0.016**	-0.021***	-0.020***	-0.021***	-0.021***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
$Tier1_{it}$		-0.002					
		(0.017)					
$GRPGR_{pt}$			-0.005				
			(0.014)				
$crisis$				0.001**	-0.004*	-0.005*	-0.005**
				(0.001)	(0.002)	(0.003)	(0.003)
$Listed$				-0.006	-0.004	-0.002	-0.000
				(0.004)	(0.004)	(0.004)	(0.004)
$EBTPTA_{it} * crisis$					0.283**	0.313**	0.324**
					(0.126)	(0.136)	(0.135)
$EBTPTA_{it} * listed$						-0.045	-0.083
						(0.181)	(0.169)
$EBTPTA_{it} * listed * crisis$							-0.005
							(0.052)
$Constant$	-0.001	0.007	-0.001	-0.012	-0.003	0.000	0.004
	(0.011)	(0.010)	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)
No. of Observations	745	652	745	745	745	745	745
F-statistics	151.89***	99.83***	147.14***	139.55***	141.01***	147.67***	163.39***
AR(1)	-3.59***	-3.46***	-3.66***	-3.53***	-3.46***	-3.51***	-3.50***
AR(2)	0.33(0.741)	0.53(0.593)	0.39 (0.695)	0.53(0.593)	0.59(0.553)	0.62(0.535)	0.64(0.525)
Hansen J-Stat	90.61(0.403)	89.61(0.432)	89.98(0.421)	106.23(0.368)	107.73(0.489)	96.41(0.555)	98.81(0.402)
No. Instruments	98	98	98	114	121	113	111
No.Banks	123	117	123	123	123	123	123

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, $LLPLOAN_{it}$= Loan loss provisions/ Average total loans, $GRPGR_{pt}$ = Gross Regional Product growth rate, $GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$= Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4, $*$ = Interaction term, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, *Crisis* = value is 1 for period 2007 to 2009 and 0 otherwise.

Similar to the estimation results for the all-bank sample in table 5.5, in table 5.6, we find some limited evidence of procyclicality in column (3) and strong evidence for earnings management. We can see that most models in table 5.6 show a significant and negative coefficient for CE_meta_{it}, indicating that more cost-efficient banks tend to have a lower level of credit risk within their loan portfolio, supporting Berger and Deyoung (1997)'s 'bad management' hypothesis.

Table 5.6. Determinants of Loan Loss Provisions- Regional Banks Panel A

	(1) $LLPLOAN_{it}$	(2) $LLPLOAN_{it}$	(3) $LLPLOAN_{it}$	(4) $LLPLOAN_{it}$
$LLPLOAN_{it-1}$	0.191***	0.199***	0.208***	0.209***
	(0.061)	(0.064)	(0.065)	(0.057)
$GDPGR_t$	-0.019	-0.025	-0.041*	
	(0.024)	(0.029)	(0.021)	
$NPLLOAN_{it}$	0.185***	0.168***	0.186***	0.166***
	(0.059)	(0.053)	(0.047)	(0.053)
$CFEER_{it}$	0.672	0.576	0.719*	0.586
	(0.451)	(0.484)	(0.373)	(0.409)
$SIZE_{it}$	0.001	0.001*	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
$LOANTA_{it}$	0.009	0.011	0.009	0.006
	(0.012)	(0.017)	(0.009)	(0.012)
$EBTPTA_{it}$	0.593***	0.575***	0.448***	0.581***
	(0.133)	(0.159)	(0.122)	(0.152)
ETA_{it}	0.033	0.032		0.034
	(0.035)	(0.032)		(0.039)
$CUSTF_{it}$	-0.011*	-0.016*	-0.012**	-0.014*
	(0.006)	(0.008)	(0.005)	(0.008)
CE_meta_{it}		-0.027	-0.030**	-0.017
		(0.022)	(0.014)	(0.018)
$Tier1_{it}$			0.007	
			(0.022)	
$GRPGR_{pt}$				-0.007
				(0.017)
$Constant$	-0.017	0.006	0.013	-0.001
	(0.012)	(0.023)	(0.014)	(0.023)
No. of Observations	611	611	530	611
F-statistics	131.52***	182.21***	122.52***	149.56***
AR(1)	-3.45***	-3.29***	-3.36***	-3.51***
AR(2)	0.19(0.853)	0.19(0.847)	0.48(0.631)	0.25(0.801)
Hansen J-Stat	77.47(0.590)	87.16(0.535)	80.72(0.488)	83.60(0.400)
No.of Instruments	91	100	92	92
No.of Banks	108	108	102	108

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, $LLPLOAN_{it-1}$= lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year), $GRPGR_{pt}$ = Gross Regional Product growth rate, $GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$= Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it} = cost efficiencies estimated from metafrontier in chapter 4

Table 5.6 (Panel B) introduces interaction terms into the model. Variables such as $LLPLOAN_{it-1}$ and $NPLLOAN_{it}$ are showing consistent signs and significance. $CUSTF_{it}$ is also found to be consistently significant and negative, indicating that funding stability impacts loan risk. That is, the more stable the finding is, the less likely bank will provision to cover future expected losses. Again, we find strong evidence of earnings management and listed banks tend

to engage more aggressively to smooth income than non-listed banks, in line with Anandarajan et al. (2007).

Table 5.6. Determinants of Loan Loss Provisions- Regional Banks Panel B

	(1) $LLPLOAN_{it}$	(2) $LLPLOAN_{it}$	(3) $LLPLOAN_{it}$	(4) $LLPLOAN_{it}$
$LLPLOAN_{it-1}$	0.245***	0.239***	0.265***	0.263***
	(0.063)	(0.065)	(0.058)	(0.063)
$GRPGR_{pt}$	-0.006	-0.004	-0.012	-0.007
	(0.017)	(0.018)	(0.013)	(0.019)
$NPLLOAN_{it}$	0.171***	0.174***	0.156***	0.163***
	(0.056)	(0.056)	(0.046)	(0.055)
$CFEER_{it}$	0.614*	0.629	0.585	0.531
	(0.352)	(0.402)	(0.428)	(0.378)
$SIZE_{it}$	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
$LOANTA_{it}$	0.011	0.012	0.011	0.011
	(0.011)	(0.010)	(0.009)	(0.010)
$EBTPTA_{it}$	0.597***	0.483***	0.436***	0.412***
	(0.134)	(0.139)	(0.141)	(0.142)
ETA_{it}	0.040	0.041	0.034	0.040
	(0.034)	(0.033)	(0.033)	(0.035)
$CUSTF_{it}$	-0.021***	-0.022***	-0.022***	-0.023***
	(0.007)	(0.007)	(0.006)	(0.008)
$crisis$	0.002***	-0.001	-0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.003)
$Listed$	-0.001	-0.001	-0.009*	-0.010**
	(0.005)	(0.005)	(0.005)	(0.004)
CE_meta_{it}	-0.042**	-0.037**	-0.040***	-0.042**
	(0.020)	(0.018)	(0.015)	(0.017)
$EBTPTA_{it} * Crisis$		0.142	0.117	0.135
		(0.114)	(0.107)	(0.116)
$EBTPTA_{it} * Listed$			0.425*	0.451**
			(0.228)	(0.207)
$EBTPTA_{it} * Listed * Crisis$				0.019
				(0.073)
$Constant$	0.012	0.009	0.020	0.019
	(0.020)	(0.020)	(0.017)	(0.020)
No. of Observations	611	611	611	611
F-statistics	153.03***	117.96***	143.55***	121.28***
AR(1)	-3.37***	-3.41***	-3.63***	-3.57***
AR(2)	0.47(0.635)	0.55(0.583)	0.64(0.522)	0.60(0.546)
Hansen J-Stat	85.52(0.433)	88.91(0.513)	76.34(0.712)	85.50(0.585)
No.of Instruments	97	104	99	105
No.of Banks	108	108	108	108

Standard errors in parentheses. $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$. $LLPLOAN_{it-1}$ = lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year), $GRPGR_{pt}$ = Gross Regional Product growth rate, $NPLLOAN_{it}$=Non performing loans/ Average total loans, $CFEER_{it}$= commission and fee income/ Average total asset, $SIZE_{it}$ = LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$ = Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $CUSTF_{it}$= customer deposits to total funding, CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4, $*$ = Interaction term, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, *Crisis* = value is 1 for period 2007 to 2009 and 0 otherwise.

5.4.4 Test on Signalling theory

Table 5.7 presents the results of the test of signalling theory, following Bouvatier and Lepetit (2008) and Soedarmono et al. (2012). We use $SIGN_{it}$, which is the average of current and future earnings, as our proxy for signalling of future earnings and find that the variable is only significant for the regression using the all banks sample, while insignificant when including only the regional sample (Table 5.7, Columns 2 and 3). This shows that for the Chinese commercial banking sector as a whole, banks provision less when there is a higher level of future earnings and provision more when there is a lower level of future earnings. This indicates that LLPs in China mainly indicate 'bad news', that is, risks in the loan portfolio and future expenses, rather than 'good news' as a sign of financial resilience.

Table 5.7. Test on signalling

	All banks (1) $LLPLOAN_{it}$	Regional (2) $LLPLOAN_{it}$	Regional (3) $LLPLOAN_{it}$
$LLPLOAN_{it-1}$	0.110**	0.109*	0.117**
	(0.053)	(0.057)	(0.058)
$GDPGR_t$	0.003	-0.001	-0.014
	(0.016)	(0.025)	(0.027)
$NPLLOAN_{it}$	0.177***	0.165***	0.143***
	(0.046)	(0.063)	(0.050)
$CFEER_{it}$	0.210	0.406	0.441
	(0.558)	(0.758)	(0.514)
$SIZE_{it}$	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
$LOANTA_{it}$	0.003	0.005	0.009
	(0.009)	(0.011)	(0.015)
$EBTPTA_{it}$	0.184	0.269	0.307*
	(0.182)	(0.218)	(0.174)
ETA_{it}	0.060*	0.045	0.044
	(0.034)	(0.038)	(0.045)
$CUSTF_{it}$	-0.020***	-0.015**	-0.024***
	(0.006)	(0.007)	(0.007)
$crisis$	-0.005**	-0.005	-0.002
	(0.002)	(0.003)	(0.003)
$Listed$	0.005	0.002	0.003
	(0.006)	(0.004)	(0.006)
$EBTPTA_{it} * Crisis$	0.329***	0.269*	0.178
	(0.120)	(0.149)	(0.132)
$EBTPTA_{it} * Listed$	-0.219	-0.165	-0.161
	(0.228)	(0.218)	(0.258)
$EBTPTA_{it} * Listed * Crisis$	-0.055	-0.025	-0.024
	(0.059)	(0.070)	(0.071)
CE_meta_{it}			-0.029
			(0.020)
$SIGN_{it}$	-0.365*	-0.106	-0.095
	(0.211)	(0.195)	(0.177)
$Constant$	0.017	0.016	0.043**
	(0.014)	(0.019)	(0.020)
No. of Observations	671	500	500
F-statistics	78.94***	59.67***	103.86***
AR(1)	-2.50**	-3.09***	-2.87***
AR(2)	0.50(0.617)	0.74(0.457)	0.78(0.438)
Hansen J-Stat	80.71(0.251)	81.09(0.630)	73.73(0.585)
No.of Instruments	89	102	94
No.of Banks	131	106	106

Standard errors in parentheses. $* p < 0.10$,** $p < 0.05$,*** $p < 0.01$, $LLPLOAN_{it-1}$ = lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year), $GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$ =Non performing loans/ Average total loans, $CFEER_{it}$ = commission and fee income/ Average total asset, $SIZE_{it}$ = LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$ = Gross Loans/Total assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $CUSTF_{it}$= customer deposits to total funding, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4, $*$ = Interaction term, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, *Crisis* = value is 1 for period 2007 to 2009 and 0 otherwise.

5.4.5 Estimation of Discretionary LLPs

Table 5.8 presents the coefficients of the equations that estimate discretionary components of LLPs using equations (5.4) and (5.5) in panel A and panel B, respectively. Recall, the residuals from these equations are the discretionary component of LLPs. We first obtain the residuals $DLLP1_{it}$ and $DLLP2_{it}$ from estimating the two alternative models proposed by Beatty and Liao (2014). The discretional component is derived from modelling LLPs as a function of a set of variables that are considered to be non-discretionary. The equations for obtaining discretionary LLPs are model 4 and model 5 in section 5.3.3. Then we take the absolute values of them, and we call them $ABDLLP1_{it}$ and $ABDLLP2_{it}$.

These are the discretionary LLPs that we will be using for the alternative test of signalling in section 5.4.5. These discretionary LLPs will also be used to test the impact of retail deposits on earnings quality in section 5.4.6. Again, discretionary LLPs are estimated using both an all-bank sample and a regional bank sample.

Table 5.8. Estimation of discretionary Loan loss provisions (DLLP)

Panel A		
	All Banks	Regional Banks
	LLP_{it}	LLP_{it}
ΔNPL_{it+1}	0.008	0.005
	(0.026)	(0.030)
ΔNPL_{it}	0.057**	0.057*
	(0.027)	(0.031)
ΔNPL_{it-1}	0.027	0.025
	(0.020)	(0.022)
ΔNPL_{it-2}	0.000	-0.001
	(0.011)	(0.013)
$SIZE_{it-1}$	-0.001***	-0.001***
	(0.000)	(0.000)
$\Delta Loan_{it}$	-0.008	-0.007
	(0.007)	(0.008)
$GDPGR_t$	0.013	0.029
	(0.050)	(0.066)
UNR_{it}	-1.130	-1.339
	(0.689)	(0.868)
$Constant$	0.066**	0.077**
	(0.027)	(0.034)
N	364.000	276.000
F	6.019	3.811
r2	0.194	0.170

Panel B		
	All Banks	Regional Banks
	LLP_{it}	LLP_{it}
LLA_{it-1}	-0.029	-0.027
	(0.043)	(0.052)
$NPLLOAN_{it-1}$	0.052***	0.069***
	(0.014)	(0.017)
CO_{it}	0.510***	0.000***
	(0.040)	(0.000)
$\Delta LOANTA_{it}$	0.021***	0.029***
	(0.005)	(0.007)
$LOANTA_{it}$	-0.004	-0.001
	(0.003)	(0.004)
$Constant$	0.008***	0.008***
	(0.002)	(0.002)
N	523.000	401.000
F	24.661	8.520
r2	0.422	0.249

Standard errors in parentheses* p < 0.10, ** p < 0.05, *** p < 0.01. $\Delta NPL_{it+1,}\Delta NPL_{it}, \Delta NPL_{it-1}, \Delta NPL_{it-2}$ = Change in nonperforming loans / gross loans in the beginning of the year (for one year ahead, current year, lagged, and lagged two years, respectively), $SIZE_{it-1}$= lagged LnTA, i.e., the logarithm of total assets, $\Delta Loan_{it}$ = (Gross loans for current year – gross loans for previous year) / Total Assets in the previous year, $GDPGR_t$ = Gross domestic product growth rate, UNR_{it} = Unemployment rate, LLA_{it} = Loan Loss Allowance / gross loans in the beginning of the year, $NPLLOAN_{it-1}$= lagged Non performing loans/ Average total loans, CO_{it} = Net Charge Offs / gross loans in the beginning of the year, $LOANTA_{it}$= Gross Loans/Total assets, $\Delta LOANTA_{it}$= Current $LOANTA$ − lagged $LOANTA$

5.4.6 Alternative tests on signalling theory

We follow Ahmed et al. (1999) and Anandarajan et al. (2007) to test signalling using two alternative models using Model 6 (equation (5.6)) and Model 7a and 7b ((5.7a, b)). Results from Model 6 are presented in column (1) and column (2), and results from Model 7a and 7b are presented in column (3) and column (4). In terms of columns (1) and (2), we can see that our key variable $\Delta EBTPTA_{it}$ is negative and significant for both the model with and without listed dummy. This is also true for both all bank sample and regional bank sample.

Then we use Model 7a and 7b (equation (5.7a, b)) by regressing one year ahead change of earnings before tax and provision on the current year's EBTPTA. We can see from table 5.9 that our core variables ABDLLP1E and ABDLLP2E, calculated as Discretionary LLP scaled by the beginning of year equity, are significant and negative, indicating that the higher the current year discretional LLPs, the lower the one year head earning. This result is consistent with the finding in section 5.4.3 that discretional LLPs for Chinese commercial banks mainly signal an expense rather than an indication of positive future cash flows. This result is consistent with the result from Anandarajan et al. (2007)'s results on Australian banks. Thus, we can conclude that the traditional good-news signalling hypothesis is not supported in the Chinese banking system.

Table 5.9. Alternative test on Signalling

Panel A All Banks

	(1) $LLPLOAN_{it}$	(2) $LLPLOAN_{it}$	(3) $\Delta EBTPE_{it+1}$	(4) $\Delta EBTPE_{it+1}$
$GDPGR_t$	0.040***	0.040***		
	(0.013)	(0.013)		
$Tier1_{it}$	-0.015**	-0.013**		
	(0.006)	(0.006)		
$EBTPTA_{it}$	0.153***	0.171***		
	(0.044)	(0.045)		
$Listed$	-0.002***			
	(0.001)			
$\Delta EBTPTA_{it}$	-0.184***	-0.169***		
	(0.060)	(0.061)		
$\Delta EBTPE_{it}$			-0.256***	0.004*
			(0.048)	(0.002)
$ABDLLP1E_{it}$			-3186.540*	
			(1818.841)	
$ABDLLP2E_{it}$				-1790.541
				(2042.889)
$Constant$	0.004**	0.003	-0.012*	-0.015*
	(0.002)	(0.002)	(0.007)	(0.008)
N	550	550	275	328
F	12.941	12.801	15.288	2.191
r2	0.106	0.086	0.101	0.013

Panel B- Regional Banks

	(1) $LLPLOAN_{it}$	(2) $LLPLOAN_{it}$	(3) $\Delta EBTPE_{it+1}$	(4) $\Delta EBTPE_{it+1}$
$GDPGR_t$	0.040***	0.040***		
	(0.011)	(0.011)		
$Tier1_{it}$	-0.013**	-0.008		
	(0.006)	(0.006)		
$EBTPTA_{it}$	0.198***	0.226***		
	(0.040)	(0.041)		
$Listed$	-0.002***			
	(0.000)			
$\Delta EBTPTA_{it}$	-0.264***	-0.242***		
	(0.055)	(0.056)		
$\Delta EBTPE_{it}$			-0.244***	0.003*
			(0.041)	(0.002)
$ABDLLP1E_{it}$			-3766.748**	
			(1646.790)	
$ABDLLP2E_{it}$				-909.270
				(1901.258)
$Constant$	0.003*	0.001	-0.011*	-0.017***
	(0.002)	(0.002)	(0.006)	(0.006)
N	687.000	687.000	363.000	432.000
F	24.964	23.803	19.399	1.944
r2	0.155	0.123	0.097	0.009

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, $GDPGR_t$ = Gross domestic product growth rate, $Tier1_{it}$ = Tier 1 capital/ risk adjusted weighted assets, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, $\triangle EBTPTA_{it}$= one year ahead EBTPTA- current year EBTPTA, $\triangle EBTPTE_{it}$= one year ahead EBTPTE- current year EBTPTE, $EBTPTE_{it}$= Earnings before tax and provisions/average total equity, $ABDLLP1_{it}$ is the absolute value of discretionay LLPs estimated as the residual of equation (5.4), $ABDLLP2_{it}$ is the absolute value of discretionay LLPs estimated as the residual of equation (5.5), $ABDLLP1E_{it}$ = ABDLLP1E / average total equity, $ABDLLP2E_{it}$ = ABDLLP2E / average total equity

5.4.7 Test of Impact of Retail Deposits on Earnings Quality

Table 5.10 presents the descriptive statistics for variables used in equation (5.8, a-d). Table 5.11 presents results of Model 8(a-d) using fixed effects within estimators to examine the impact of customer deposit on discretionary loan loss provisions for both regional bank and all-bank samples. Our core variable CDL and CUSTF are significant and negative, indicating that a higher reliance on retail deposit for funding decrease banks' incentive to increase discretionary use of LLPs and thus increase the quality of banks' reported earnings. That is, the earnings quality represented by a lower level of loan loss provisions improves when there is a higher percentage of retail deposit as a source for bank funding. Our result is consistent for both samples across all four model specification variations, in line with Iin et al. (2018) that retail deposit is a stable source of funding for banks, and banks have fewer incentives to manipulate accounting information in order to attract and obtain customer deposits.

Consistent with Jin et al. (2018), we also find that lagged loan loss provisions have a positive impact on discretionary LLPs, and bank size has a negative impact on LLP manipulations. This indicates that a higher level of LLPs in the previous period may prompt bank managers to manipulate LLPs more in order to reduce the earnings volatility. $SIZE_{it}$ is found to be significant and negative using an all-bank sample, indicating that bigger banks have less discretional LLPs, which may be due to closer monitoring by regulatory authorities and more market disciplining. Again, in terms of signalling variable $\Delta EBTPTA_{it}$, we find a consistent negative relationship with discretionary LLPs. This indicates bank managers manipulate less for LLPs when future earnings are expected to increase.

Table 5.10. Descriptive Statistics for Test of Impact of Retail Deposit on Earnings Quality

All Banks					
Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A					
$ABDLLP1_{it}$	312	0.0033	0.0030	0.0000	0.0204
$ABDLLP2_{it}$	432	0.0034	0.0033	0.0000	0.0227
CDL_{it-1}	432	0.8115	0.1036	0.4270	0.9822
$CUSTF_{it-1}$	432	0.8308	0.1064	0.4313	0.9999
$SIZE_{it}$	432	12.0905	1.7469	8.9163	16.4543
$LOANDEP_{it}$	432	0.6607	0.0977	0.3181	1.0798
$VOLROA_{it}$	432	0.0042	0.0007	0.0035	0.0053
$CAPR_{it}$	432	15.5026	5.5609	5.8356	78.1553
$LIQUID_{it}$	432	0.2249	0.0924	0.0453	0.6117
$EBTPTA_{it}$	432	0.0194	0.0053	0.0071	0.0412
$LLPLOAN_{it-1}$	432	0.0068	0.0048	-0.0065	0.0246
$\Delta EBTPTA_{it}$	432	0.0001	0.0039	-0.0274	0.0182
Panel B					
$ABDLLP1_{it}$	227	0.0038	0.0032	0.0000	0.0205
$ABDLLP2_{it}$	328	0.0041	0.0042	0.0000	0.0470
CDL_{it-1}	328	0.8199	0.1083	0.4270	0.9822
$CUSTF_{it-1}$	328	0.8396	0.1112	0.4313	0.9999
$SIZE_{it}$	328	11.3096	1.0263	8.9163	14.2108
$LOANDEP_{it}$	328	0.6431	0.1004	0.3181	1.0798
$VOLROA_{it}$	328	0.0044	0.0008	0.0036	0.0057
$CAPR_{it}$	328	14.8706	5.2390	5.8356	78.1553
$LIQUID_{it}$	328	0.2323	0.0961	0.0453	0.6117
$EBTPTA_{it}$	328	0.0198	0.0057	0.0073	0.0412
$LLPLOAN_{it-1}$	328	0.0071	0.0052	-0.0065	0.0246
$\Delta EBTPTA_{it}$	328	0.0000	0.0042	-0.0274	0.0182

$ABDLLP1_{it}$ is the absolute value of discretionay LLPs estimated as the residual of equation (5.4), $ABDLLP2_{it}$ is the absolute value of discretionay LLPs estimated as the residual of equation (5.5), CDL_{lt-1} = (Core deposits)/(total liabilities) in year t-1, $CUSTF_{it-1}$= lagged customer deposits to total funding, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANDEP_{it}$ = Loan devided by customer deposits, $VOLROA_{lt}$ = (standard deviation of net income)/(total asset) in year t, $CAPR_{lt}$ =Total liability/(total equity) in year t, $LIQUID_{lt}$ =(liquid assets)/(total assets) in year t, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, $LLPLOAN_{it-1}$= lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year), $\triangle EBTPTA_{it}$= one year ahead EBTPTA- current year EBTPTA

Table 5.11. Test of Impact of Retail Deposit on Earnings Quality

Regional Banks

	(1) $ABDLLP1_{it}$	(2) $ABDLLP1_{it}$	(3) $ABDLLP2_{it}$	(4) $ABDLLP2_{it}$
$\boldsymbol{CDL_{it-1}}$	**-0.009****		**-0.007***	
	(0.004)		**(0.004)**	
$\boldsymbol{CUSTF_{it-1}}$		**-0.008****		**-0.007***
		(0.004)		**(0.004)**
$SIZE_{it}$	-0.007	-0.007	-0.003	-0.003
	(0.007)	(0.007)	(0.007)	(0.007)
$SIZE_{it}^2$	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$LOANDEP_{it}$	-0.004	-0.004	-0.002	-0.002
	(0.006)	(0.006)	(0.005)	(0.005)
$CAPR_{it}$	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$VOLROA_{it}$	0.108	0.097	-0.497	-0.494
	(0.366)	(0.366)	(0.400)	(0.399)
$LIQUID_{it}$	0.003	0.003	-0.004	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)
$EBTPTA_{it}$	-0.121	-0.115	0.323***	0.319***
	(0.115)	(0.115)	(0.093)	(0.092)
$LLPLOAN_{it-1}$	0.110**	0.114**	0.098*	0.100*
	(0.053)	(0.053)	(0.052)	(0.052)
$\Delta EBTPTA_{it}$	-0.198**	-0.192**	0.021	0.019
	(0.084)	(0.084)	(0.075)	(0.075)
$Constant$	0.060	0.061	0.039	0.040
	(0.043)	(0.043)	(0.040)	(0.040)
N	275.000	275.000	328.000	328.000
F	2.119	2.080	4.594	4.679
r2_w	0.099	0.098	0.165	0.167
r2_b	0.117	0.132	0.023	0.023
r2_o	0.116	0.120	0.058	0.059

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All Banks

	(1) $ABDLLP1_{it}$	(2) $ABDLLP1_{it}$	(3) $ABDLLP2_{it}$	(4) $ABDLLP2_{it}$
$\boldsymbol{CDL_{it-1}}$	**-0.007****		**-0.004***	
	(0.003)		**(0.003)**	
$\boldsymbol{CUSTF_{it-1}}$		**-0.007****		**-0.005***
		(0.003)		**(0.003)**
$SIZE_{it}$	-0.008*	-0.008*	-0.003	-0.003
	(0.004)	(0.004)	(0.003)	(0.003)
$SIZE_{it}^2$	0.000*	0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$LOANDEP_{it}$	-0.003	-0.003	0.000	0.000
	(0.004)	(0.004)	(0.003)	(0.003)
$CAPR_{it}$	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$VOLROA_{it}$	0.096	0.087	-0.274	-0.273
	(0.297)	(0.297)	(0.270)	(0.269)
$LIQUID_{it}$	0.002	0.002	-0.000	-0.000
	(0.003)	(0.003)	(0.002)	(0.002)
$EBTPTA_{it}$	-0.137	-0.134	0.080	0.078
	(0.095)	(0.095)	(0.064)	(0.064)
$LLPLOAN_{it-1}$	0.102**	0.105**	0.057	0.058
	(0.045)	(0.045)	(0.036)	(0.036)
$\Delta EBTPTA_{it}$	-0.209***	-0.205***	-0.135***	-0.136***
	(0.070)	(0.070)	(0.051)	(0.051)
$Constant$	0.062**	0.062**	0.031	0.032
	(0.029)	(0.029)	(0.022)	(0.022)
N	363.000	363.000	432.000	432.000
F	2.759	2.745	4.317	4.377
r2_w	0.094	0.094	0.119	0.120
r2_b	0.074	0.087	0.179	0.173
r2_o	0.073	0.080	0.126	0.124

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, CDL_{lt-1} = (Core deposit)/ (total liabilities) in year t-1, $CUSTF_{it-1}$= lagged customer deposits to total funding, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $SIZE_{it}^2$= SIZE squared, $LOANDEP_{it}$ = Loan devided by customer deposits, $VOLROA_{lt}$ = (standard deviation of net income)/(total asset) in year t, $CAPR_{lt}$ =Total liability/(total equity) in year t, $LIQUID_{lt}$ =(liquid assets)/(total assets) in year t, $EBTPTA_{it}$= Earnings before tax and provisions/average total asset, $LLPLOAN_{it-1}$= lagged $LLPLOAN$ (loan loss provisions/ Total loans in the previous year), $\triangle EBTPTA_{it}$= one year ahead EBTPTA- current year EBTPTA

5.5 Robustness Check

In order to ensure the robustness of our results, we re-estimate models 1-3 and re-test our main hypotheses using the fixed effects within estimator instead of SGMM. Tables 5.12 and 5.13 present the results using fixed effects instead of GMM as the estimation technique to re-test our hypotheses 1-4 and 6.

We use an alternative measurement for signalling, that is, one year-ahead change in earnings before provisions and tax, to test the signalling hypothesis rather than SIGN. It can be denoted as △EBTPTA, calculated as next year's EBTPTA ratio minus the current year's EBTPTA ratio, following Anandarajan et al. (2007) and Leventis et al. (2012). In our previous estimations, the results have consistently pointed out to reject the 'good new' signalling hypothesis. Therefore, we also expect a negative relationship between △EBTPTA and LLPs in the robustness check.

We can see from tables 5.12 and 5.13 that the fixed effects model also gives strong evidence of 'expense' or 'cost' signalling. It indicates that the choice of measure for signalling does not have an impact on the conclusion. That is, for both $SIGN_{it}$ and △EBTPTA, the coefficients are negative, showing that LLPs in Chinese commercial banks signal an expense rather good news in terms of future earnings.

We continue to find strong evidence for income-smoothing for the all-bank sample and national-bank and regional-bank sub-samples. although the coefficients for the interaction between $EBTPTA_{it}$ and crisis are not significant. Please note that here we exclude lagged dependent variables, as lagged dependent variables would not be suited to be included in a static model due to the potential problems of endogeneity and reverse causation. Therefore, because we exclude a relevant variable here, the results from fixed effects within estimator modelling could potentially have some form of model misspecification.

Table 5.12. Robustness check- all banks

	(1) $LLPLOAN_{it}$	(2) $LLPLOAN_{it}$	(3) $LLPLOAN_{it}$
$GDPGR_t$	-0.035**	-0.027*	-0.028*
	(0.014)	(0.015)	(0.016)
$NPLLOAN_{it}$	0.069***	0.057***	0.059***
	(0.008)	(0.009)	(0.009)
$CFEER_{it}$	0.498**	0.276	0.241
	(0.211)	(0.249)	(0.250)
$SIZE_{it}$	-0.003***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.001)
$LOANTA_{it}$	-0.008**	-0.012***	-0.012***
	(0.004)	(0.004)	(0.004)
$EBTPTA_{it}$	0.474***	0.514***	0.552***
	(0.063)	(0.083)	(0.088)
ETA_{it}	-0.007	-0.004	-0.005
	(0.011)	(0.011)	(0.011)
$CUSTF_{it}$	-0.011***	-0.009***	-0.009***
	(0.003)	(0.003)	(0.003)
$crisis$	-0.002*	-0.002	-0.002
	(0.001)	(0.001)	(0.001)
$EBTPTA_{it} * crisis$	0.066	0.063	0.061
	(0.065)	(0.068)	(0.069)
$EBTPTA_{it} * listed$	-0.026	-0.126	-0.148
	(0.123)	(0.133)	(0.133)
$EBTPTA_{it} * listed * crisis$	0.026	0.026	0.029
	(0.041)	(0.043)	(0.043)
$SIGN_{it}$		-0.112*	
		(0.060)	
$\Delta EBTPTA_{it}$			-0.049
			(0.061)
$Constant$	0.046***	0.059***	0.058***
	(0.011)	(0.012)	(0.012)
N	957.000	809.000	809.000
F	18.274	15.154	14.874
r2_w	0.214	0.230	0.227
r2_b	0.152	0.172	0.170
r2_o	0.142	0.125	0.122

Standard errors in parentheses,* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$,$GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$ =Nonperforming loans/ Average total loans, $CFEER_{it}$ = commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$ = Gross Loans/Total assets, $EBTPTA_{it}$ = Earnings before tax and provisions/average total asset, ETA_{it} = Equity / Total assets, $CUSTF_{it}$= customer deposits to total funding, $*$ = Interaction term, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, *Crisis* = value is 1 for period 2007 to 2009 and 0 otherwise, $\triangle EBTPTA_{it}$= one year ahead EBTPTA- current year EBTPTA, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)]

Table 5.13. Robustness check- National vs Regional Banks

	(1) National Banks	(2) National Banks	(3) National Banks	(1) Regional Banks	(2) Regional Banks	(3) Regional Banks	(4) Regional Banks
	LLPLOAN	LLPLOAN	LLPLOAN	LLPLOAN	LLPLOAN	LLPLOAN	LLPLOAN
$GDPGR_t$	-0.030	-0.027	-0.003	-0.041**	-0.031**	-0.034*	-0.037**
	(0.022)	(0.021)	(0.023)	(0.016)	(0.013)	(0.018)	(0.019)
$NPLLOAN_{it}$	0.014	0.006	-0.002	0.079***	0.070***	0.065***	0.067***
	(0.015)	(0.015)	(0.015)	(0.009)	(0.009)	(0.010)	(0.010)
$CFEER_{it}$	0.505	0.062	0.168	0.655**	0.397*	0.468	0.444
	(0.333)	(0.334)	(0.323)	(0.273)	(0.224)	(0.324)	(0.325)
$SIZE_{it}$	-0.005***	-0.005***	-0.003*	-0.002***	-0.000	-0.004***	-0.004***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)
$LOANTA_{it}$	-0.021**	-0.024***	-0.022***	-0.006	0.001	-0.010**	-0.010**
	(0.008)	(0.007)	(0.007)	(0.004)	(0.003)	(0.005)	(0.005)
$EBTPTA_{it}$	1.009***	0.589*	0.314	0.412***	0.410***	0.465***	0.506***
	(0.368)	(0.352)	(0.366)	(0.067)	(0.054)	(0.088)	(0.094)
ETA_{it}	-0.090**	-0.088***	-0.084***	-0.003	0.004	-0.001	-0.001
	(0.035)	(0.030)	(0.030)	(0.011)	(0.010)	(0.012)	(0.012)
$CUSTF_{it}$	-0.022**	-0.010	-0.007	-0.009***	-0.009***	-0.007**	-0.008**
	(0.009)	(0.008)	(0.008)	(0.003)	(0.002)	(0.003)	(0.003)
$crisis$	0.001	-0.001	-0.000	-0.001	0.001	-0.001	-0.001
	(0.004)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
$EBTPTA_{it} * crisis$	0.017	0.201	0.142	-0.001	-0.071	-0.004	-0.008
	(0.283)	(0.253)	(0.248)	(0.072)	(0.068)	(0.077)	(0.077)
$EBTPTA_{it} * listed$	-0.453	-0.014	0.041	-0.017	-0.056	-0.150	-0.170
	(0.307)	(0.290)	(0.286)	(0.153)	(0.048)	(0.169)	(0.169)
$EBTPTA_{it} * listed * crisis$	-0.104	-0.183	-0.181	0.033	0.026	0.038	0.041
	(0.147)	(0.131)	(0.128)	(0.051)	(0.049)	(0.054)	(0.054)
$SIGN_{it}$		-0.039				-0.087	
		(0.163)				(0.064)	
$\Delta EBTPTA_{it}$			-0.303*				-0.023
			(0.162)				(0.066)
CE_meta_{it}					-0.033***		
					(0.005)		
$Constant$	0.106***	0.091***	0.065**	0.040***	0.038***	0.054***	0.054***
	(0.029)	(0.027)	(0.029)	(0.011)	(0.006)	(0.012)	(0.013)
N	151	136	136	777	777	655	655
F	8.980	6.037	6.493	13.647		11.064	10.899
r2_w	0.465	0.421	0.439	0.200	0.236	0.212	0.210
r2_b	0.247	0.118	0.125	0.108	0.182	0.154	0.155
r2_o	0.164	0.119	0.182	0.144	0.225	0.128	0.123

Standard errors in parentheses,* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, $GDPGR_t$ = Gross domestic product growth rate, $NPLLOAN_{it}$ =Non performing loans/ Average total loans, $CFEER_{it}$ = commission and fee income/ Average total asset, $SIZE_{it}$= LnTA, i.e., the logarithm of total assets, $LOANTA_{it}$ = Gross Loans/Total assets, $EBTPTA_{it}$ = Earnings before tax and provisions/average total assets, ETA_{it} = Equity / Total assets, $CUSTF_{it}$= customer deposits to total funding, $*$ = Interaction term, *listed* = takes value of 1 for listed banks and value of 0 for non-listed banks, *Crisis* = value is 1 for period 2007 to 2009 and 0 otherwise, $SIGN_{it}$= (one year ahead EBTPTA- current year EBTPTA)/ [0.5*(current year's total asset +next year's total asset)], $\triangle EBTPTA_{it}$= one year ahead EBTPTA- current year EBTPTA, CE_meta_{it}= cost efficiencies estimated from metafrontier in chapter 4.

5.6 Conclusion and Policy Implications

From our test of the six hypotheses of the behaviour of loan loss provisions, we find strong evidence of income smoothing behaviour among Chinese banks. Moreover, Chinese banks engaged more in income smoothing practices during the global financial crisis period, in line with Curcio et al. (2017). We do not find sufficient evidence to support the traditional signalling theory that LLPs signal good news of higher future earnings and financial strength. Instead, we find LLPs signal more of an expense rather than good news to the investors. We also find limited evidence of the potential procyclicality of LLPs in China.

Moreover, we find that reliance on retail deposits has a strong impact on banks' engagement in earnings management, reducing the incentive for banks to use discretionary LLPs to manipulate accruals. Thus, banks with higher levels of retail deposits seem to report higher-quality earnings figures in their financial statements.

These hypotheses are tested over different model specifications and different estimation techniques, i.e., SGMM and fixed effected within estimators. We also tested these hypotheses over different samples. The results remain consistent.

In terms of policy implications, firstly, to resolve the procyclicality of LLPs, a separate regulation on Loan loss reserves that allows it to fluctuate over a period of time above a minimum requirement, rather than meeting capital requirements at all points in time, may be beneficial. Secondly, regulators should pay more attention to monitor the accounting information of banks that rely more on wholesale funding rather than retail deposits in China as part of market disciplinary measures.

Chapter 6. Conclusion

6.1 Thesis Summary

This thesis investigates performance and risk in the Chinese regional banking sector. We first provided the background for our research in Chapter 2: the evolution of the Chinese banking system and why city and rural commercial banks are the best candidates to fit our research purpose to investigate regional disparities.

Then we provided a justification for using cost efficiency as the performance measure for Chinese Banks in Chapter 3. This is not only due to the problems of non-proportionality and inter-temporal instability of accounting-based ratios but also because it can better reflect the cost-minimizing behaviour of Chinese banks. Then we reviewed the literature on bank efficiency and bank theories and chose the Battese and Coelli (1995) model to be the foundation that we build upon due to its favourable properties, such as its ability to incorporate environmental variables to affect the in-efficiency term directly. We also follow the intermediation theory and regard Chinese banks as intermediaries to use deposits and other inputs to produce loans and other earning assets, which forms the basis for selecting inputs and outputs for frontier models in chapter 4.

In chapter 4, we first gave our rationale regarding why regional disparities is an under-investigated issue in the Chinese banking sector, building upon Sun et al. (2013). We offered a new way of classifying Chinese provinces into regions using a hybrid hierarchical k-means clustering method based on macroeconomic characteristics. The results of cluster analysis gave us three new regions, which we define as: The Inland Region (1), The Eastern Municipalities (2) and the North-eastern and Coastal Region (3). We then adopted a stochastic meta-frontier (SMF) method so that we do not assume banks in different regions of China operate under the same technology. The likelihood ratio test confirmed this hypothesis. We find that for banks in each region to move towards the meta-frontier, which presents a simulated best practice bank in the country, changes in the economic and market conditions need to be made as well as improvements to managerial practice.

We finally linked the two aspects of the thesis: performance and risk, and

connected chapter 4 and chapter 5 by incorporating cost efficiency scores derived from the meta-frontier into the modelling of loan loss provisions. We find the more efficient banks indeed incur less credit risk, which may result from more effective monitoring and management.

6.2 Answers to the research questions

In this section, we provide answers to each of the research questions listed in the introduction in chapter 1:

(1) Regional commercial banks in China indeed operate under different technological frontiers due to large disparities in regional economic environments.

(2) Cost inefficiency for Chinese commercial banks primarily comes from the technological gaps between the region's best practice and the country's best practice, with a relatively smaller proportion coming from inefficient management.

(3) The deviation between each bank and its regional frontier is mainly affected by capital levels and NPLs, although this is only found for the Inland Region (1). The technological gap is mainly influenced by macroeconomic conditions and the banking market competitive environment, the urbanisation rate, market concentration, and foreign bank presence in each region.

(4) Regarding regional differences in cost efficiency, the North-eastern and Coastal Region (3) is the most cost-efficient, and The Inland Region (1) is the least cost-efficient. The inefficiencies for all three regions are primarily driven by the deviation of the region's best practice from the country's best practice. However, relatively speaking, managerial inefficiency seems to be a relatively bigger issue in the North-eastern and

Coastal Region (3) than in other regions. This can be mainly because the North-eastern and Coastal Region (3) is already more advanced at adopting a better technology than the other two regions.

(5) Under current technology, banks in the Inland Region (cluster 1) and the North-eastern and Coastal Region (cluster 3) exhibit diseconomies of scale while banks in the Eastern Municipalities (cluster 2) enjoy economies of scale. If they moved towards the meta-frontier by adopting a better technology, Inland Region (cluster 1) and the North-eastern and Coastal Region (cluster 3) would start to enjoy economies of scale, and Eastern Municipalities (cluster 2) would continue to enjoy an even greater degree of economies of scale.

(6) In terms of determinants of economies of scale for Chinese regional banks, we find that investment banking activities, higher liquidity ratio, and lower credit risk would contribute to economies of scale, while banks' profitability from traditional lending activities does not have any impact on scale economies. We also find that higher levels of capital would have a positive impact on economies of scale.

(7) Regarding loan loss provisions in China, we find strong evidence of income-smoothing, with some weak evidence of procyclicality. We do not find evidence of using LLPs as a mechanism for capital management. We also find a strong negative relationship between customer deposits and LLPs. This finding is in line with Curcio et al. (2017). In terms of the relationship between meta-frontier cost efficiency and bank risk, we find that banks that are more cost-efficient provision less for LLPs.

(8) Contrary to the traditional beliefs, LLPs in China do not signal good news of financial resilience. Instead, they are merely a signal of expenses. This is also found by Anandarajan et al. (2007) for Australian banks.

(9) Retail deposit is found to have a positive impact on earnings quality. That is, banks with a higher level of retail deposits report earnings in their financial statement with fewer manipulations.

6.3 Policy and Practical Implications

The main policy implications that can be derived from this Thesis are as below:

6.3.1 Chinese Regional Division

This Thesis offers a new way of regional division based on macroeconomic conditions, including GDP growth rate, GDP per capita, urbanisation, inflation rate and Herfindahl index, which are found to influence bank performance in the literature (Kasman and Yildirim, 2006, Shen et al., 2009, Bos and Schmiedel, 2007).

The results from our hybrid hierarchical k-means clustering method suggest that mainland Chinese provinces can be divided into the Eastern Municipalities (Beijing, Shanghai and Tianjin), the Northeastern and Coastal Region (Fujian, Guangdong, Heilongjiang, Inner Mongolia, Jiangsu, Jilin, Liaoning, Shandong, Zhejiang) and the Inland region (the rest of the provinces in mainland China), based on macroeconomic conditions and financial market development from 2005 to 2015. This statistical division offers a more dynamic way to define regions compared to the traditional eastern-central-western division, which is geography and policy-driven. This new classification method offers policymakers a more updated view of Chinese regional banking markets based on China's recent economic development.

6.3.2 Bank Efficiency Policies

The meta-frontier estimation method allows for a region-tailored policy for

improving bank performance in the Chinese regional banking sector, with regional disparities in economic conditions, financial development, and banking market conditions.

Firstly, our results suggest that estimating a country-frontier by pooling all banks in China, assuming that they operate under the same technology, is inappropriate. Instead, they are more appropriately modelled by assuming banks in each region operate under a different available technology, deviating from a simulated country's best available technology. This results from disparities in regional economic and financial market development. The SMF method gives a more accurate insight into the performance of the regional banking market and allows us to compare bank efficiency across regions without assuming a common frontier. By incorporating microeconomic and macroeconomic variables into the technological gap model, our results also inform the determinants of the deviation from country best practice, thus offering region-specific strategies to improve bank efficiency.

For all regional banks in China, their inefficiency mainly comes from the inaccessibility of country technology due to regional constraints, i.e., the technological gap. However, managerial inefficiency, which is the distance between individual banks and regional frontiers, still exist. Moreover, the extent to which a bank's inefficiency comes from the technological gap or managerial inefficiency differs across regions. These findings indicate that policy implications may differ across regions, despite an overarching result of a significant technological gap for all regions.

Among all three regions, the Inland Region has a slightly wider technological gap compared to the other two regions. In contrast, the North-eastern and Coastal Region has the smallest technological gap from the country meta-frontier. In terms of managerial inefficiency, banks in the North-eastern and Coastal Region deviate the furthest from their own cluster-specific frontier, indicating a relatively lower level of managerial efficiency compared to other regions. Meanwhile, the Eastern Municipalities have relatively the highest level of

managerial efficiency among all clusters. This indicates that banks in the North-eastern and Coastal Region may benefit from adopting a better managerial practice to catch up with the other two regions. Our findings suggest that better managerial efficiency can be achieved by improving bank profitability and reducing the level of riskiness in their loan portfolios, such as better pricing and loan monitoring.

Our findings suggest that banks in all three regions would benefit significantly from better technology. This can be firstly achieved through the acceleration of urbanisation. Since the Inland region still consists of a relatively large proportion of rural areas, progressing on urbanisation and developing local infrastructure can help banks improve their ability to transform inputs into outputs more efficiently. Secondly, our model indicates that the regional Chinese banking sector is highly concentrated. Thus its efficiency can be improved through a higher level of competition. Policymakers should welcome new entrants and challengers into the local banking market. In the meantime, policymakers may encourage increased foreign banks presence as there may exit a spillover effect from foreign assets, bringing managerial know-how and healthy competition into the domestic market. The methods to encourage foreign bank entry may include lowering the barrier of entry, further liberating restrictions and allowing a broader scope of banking services to be provided by foreign banks to create more competitive banking markets.

6.3.3 Economies of Scale Policies

Narrowing the gap between the regional and the national technological frontiers may also increase the degree of economies of scale for banks in all three regions. Notably, for banks in the Inland region and North-eastern and Coastal Region, which are currently operating with diseconomies of scale, adopting a superior technology can allow them to start enjoying economies of scale. This indicates that if regional banks move towards the national technology meta-frontier, mergers and acquisitions will reduce average costs. Banks in Eastern Municipalities are currently operating with economies of scale. Our findings

indicate that by adopting a better technology, banks can enjoy an even greater degree of economies of scale.

Regarding determinants of economies of scale, it is beneficial for Chinese regional commercial banks to continue complying with Basel III or even more stringent capital and liquidity regulations, as a higher level of capital and liquidity ratio is likely to increase economies of scale for Chinese regional banks. In order to achieve a higher degree of economies of scale, banks are also encouraged to diversify their business model and engage in more non-traditional bank activities such as investment banking businesses. It is also found that lowering the credit risk in their loan portfolio would help banks achieve a higher level of economies of scale.

6.3.4 Loan Loss Provisioning Behaviour

In terms of loan loss provisioning, regulators should be aware of banks' earnings management practice, which is prevalent across all regions and bank types. That is, it is practised by state-owned banks, joint-stock banks as well as regional city and rural commercial banks. Chinese banks exercise income smoothing more aggressively during economically distressing periods, and listed banks engage more in income smoothing than non-listed banks. Furthermore, a bank's funding structure, i.e., whether it relies more on retail deposits or wholesale funding, affects income manipulations and earnings quality. Policymakers should pay closer attention to banks that rely more on wholesale funding, as earnings figures for these banks are found to have a higher degree of discretionary component. Therefore, regulating and monitoring earnings should be more stringent for banks with listed status, relying more on wholesale funding and during economic downturns. This finding should prompt regulators to favour policies that can enhance the market discipline of the Chinese commercial banking sector and support a sounder financial system. In terms of practical implications, investors should pursue other supplementary information when making investment decisions on Chinese banks, particularly for banks with listed status and higher reliance on wholesale funding.

In terms of banks' loan loss provisioning behaviours, Chinese banks are largely procyclical to economic conditions. This behaviour can amplify the economic cycle, making credit shortages more severe during an economic downturn. To avoid such problems, regulators should encourage a more forward-looking dynamic provisioning method rather than the current backwards-looking practice.

Furthermore, loan loss provisions in Chinese banks also do not signal good news. Instead, it is a signal of extra expense. Therefore, investors should be warned that a higher provisioning level is not a sign of financial resilience against future losses. Instead, it is an expense that reflects a higher level of credit risk in the bank loan portfolio. Hence, the loan loss provisioning behaviour should be continuously monitored to reflect any change in its behaviour in real-time in order to provide an up-to-date picture of the current provisioning practice. This may also help enhance market transparency and disclosure of accounting information, helping investors make better-informed decisions.

References

ABDUL-MAJID, M., FALAHATY, M. & JUSOH, M. 2017. Performance of Islamic and conventional banks: A meta-frontier approach. *Research in International Business and Finance,* 42**,** 1327-1335.

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,* 43**,** 2033-2054.

ADELOPO, I., LLOYDKING, R. & TAURINGANA, V. 2018. Determinants of bank profitability before, during, and after the financial crisis. 14**,** 378-398.

AHMED, A. S., TAKEDA, C. & THOMAS, S. 1999. Bank loan loss provisions: a reexamination of capital management, earnings management and signaling effects. *Journal of Accounting and Economics,* 28**,** 1-25.

AIGNER, D., LOVELL, C. A. K. & SCHMIDT, P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics,* 6**,** 21-37.

AKERLOF, G. A. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics,* 84**,** 488-500.

ALESSANDRINI, P., PRESBITERO, A. F. & ZAZZARO, A. 2009. Global banking and local markets: a national perspective. *Cambridge Journal of Regions, Economy and Society,* 2**,** 173-192.

ANANDARAJAN, A., HASAN, I. & MCCARTHY, C. 2007. Use of loan loss provisions for capital, earnings management and signalling by Australian banks. *Accounting and Finance,* 47**,** 357-379.

ANBAR, A. & ALPER, D. 2011. Bank Specific and Macroeconomic Determinants of Commercial Bank Profitability: Empirical Evidence from Turkey. *Business and Economics Research Journal,* 2**,** 139-152.

ANDRIES, A. M. 2011. The Determinants of Bank Efficiency and Productivity Growth in the Central and Eastern European Banking Systems. *Eastern European Economics,* 49**,** 38-59.

ARELLANO, M. & BOND, S. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies,* 58**,** 277-297.

ASHTON, J. 2009. Does the United Kingdom have regional banking markets? An assessment of UK deposit provision 1992-2006. *Applied Economics Letters,* 16**,** 1123-1128.

ASHTON, J. K. & GREGORIOU, A. 2014. The Influence of Banking Centralization on Depositors: Regional Heterogeneities in the Transmission of Monetary Policy. *Regional Studies,* 48**,** 1467-1482.

ATHANASOGLOU, P. P., BRISSIMIS, S. N. & DELIS, M. D. 2008. Bank-specific, industry-specific and macroeconomic determinants of bank profitability.

Journal of International Financial Markets, Institutions and Money, 18**,** 121-136.

BALBOA, M., LÓPEZ-ESPINOSA, G. & RUBIA, A. 2013. Nonlinear dynamics in discretionary accruals: An analysis of bank loan-loss provisions. *Journal of Banking & Finance,* 37**,** 5186-5207.

BALTAGI, B. H. & CHANG, Y.-J. 1994. Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model. *Journal of Econometrics,* 62**,** 67-89.

BARAJAS, A., STEINER, R. & SALAZAR, N. 2000. The impact of liberalization and foreign investment in Colombia's financial sector. *Journal of Development Economics,* 63**,** 157-196.

BARR, R. & SIEMS, T. 1997. Bank Failure Prediction Using DEA to Measure Management Quality. *Advances in Metaheuristics, Optimization, and Stochastic Modeling Techniques,* 7.

BATTESE, G. E. & COELLI, T. J. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics,* 38**,** 387-399.

BATTESE, G. E. & COELLI, T. J. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics,* 20**,** 325-332.

BATTESE, G. E., RAO, D. S. P. & O'DONNELL, C. J. 2004. A Metafrontier Production Function for Estimation of Technical Efficiencies and Technology Gaps for Firms Operating Under Different Technologies. *Journal of Productivity Analysis,* 21**,** 91-103.

BAUM, C. F. 2006. *An introduction to modern econometrics using Stata*, Stata press.

BCBS 2005. An explanatory note on the Basel II IRB risk weight functions.

BEATTY, A., CHAMBERLAIN, S. L. & MAGLIOLO, J. 1995. Managing Financial Reports of Commercial Banks: The Influence of Taxes, Regulatory Capital, and Earnings. *Journal of Accounting Research,* 33**,** 231-261.

BEATTY, A. & LIAO, S. 2011. Do delays in expected loss recognition affect banks' willingness to lend? *Journal of Accounting and Economics,* 52**,** 1-20.

BEATTY, A. & LIAO, S. 2014. Financial accounting in the banking industry: A review of the empirical literature. *Journal of Accounting and Economics,* 58**,** 339-383.

BEAVER, W., EGER, C., RYAN, S. & WOLFSON, M. 1989. Financial Reporting, Supplemental Disclosures, and Bank Share Prices. *Journal of Accounting Research,* 27**,** 157-178.

BEAVER, W. H. & ENGEL, E. E. 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics,* 22**,** 177-206.

BECCALLI, E., ANOLLI, M. & BORELLO, G. 2015. Are European banks too big? Evidence on economies of scale. *Journal of Banking & Finance,* 58**,** 232-246.

BECCALLI, E. & ROSSI, L. 2020. Economies or diseconomies of scope in the EU banking industry? *European Financial Management,* 26**,** 1261-1293.

BENSTON, G. J. 1972. Economies of Scale of Financial Institutions. *Journal of Money, Credit and Banking,* 4**,** 312-341.

BENSTON, G. J. & SMITH, C. W. 1976. A Transactions Cost Approach to the Theory of Financial Intermediation. *The Journal of Finance,* 31**,** 215-231.

BERGER, A. & HUMPHREY, D. 1990. Measurement and Efficiency Issues in Commercial Banking.

BERGER, A. N. 1995. The Profit-Structure Relationship in Banking--Tests of Market-Power and Efficient-Structure Hypotheses. *Journal of Money, Credit and Banking,* 27**,** 404-431.

BERGER, A. N. & DEYOUNG, R. 1997. Problem loans and cost efficiency in commercial banks. *Journal of Banking & Finance,* 21**,** 849-870.

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,* 33**,** 113-130.

BERGER, A. N. & HUMPHREY, D. B. 1991. The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics,* 28**,** 117-148.

BERGER, A. N. & MESTER, L. J. 1997. Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of Banking & Finance,* 21**,** 895-947.

BERGER, A. N. & MESTER, L. J. 2001. Explaining the Dramatic Changes in Performance of U.S. Banks: Technological Change, Deregulation and Dynamic Changes in Competition. Wharton School Center for Financial Institutions, University of Pennsylvania.

BERGER, A. N. & MESTER, L. J. 2003. Explaining the dramatic changes in performance of US banks: technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation,* 12**,** 57-95.

BERGER, A. N. & UDELL, G. F. 2004. The institutional memory hypothesis and the procyclicality of bank lending behavior. *Journal of Financial Intermediation,* 13**,** 458-495.

BERTAY, A. C., DEMIRGÜÇ-KUNT, A. & HUIZINGA, H. 2013. Do we need big banks? Evidence on performance, strategy and market discipline. *Journal of Financial Intermediation,* 22**,** 532-558.

BIKKER, J. & BOS, J. W. 2008. *Bank Performance: A theoretical and empirical framework for the analysis of profitability, competition and efficiency*, Routledge.

BIKKER, J. A. & METZEMAKERS, P. A. J. 2005. Bank provisioning behaviour and procyclicality. *Journal of International Financial Markets, Institutions and Money,* 15**,** 141-157.

BOLLEN, K. A. & WARD, S. 1979. Ratio variables in aggregate data analysis their uses, problems, and alternatives. *Sociological Methods & Research,* 7**,** 431-450.

BONIN, J. P. & HUANG, Y. 2001. Dealing with the bad loans of the Chinese banks. *Journal of Asian Economics,* 12**,** 197-214.

BORIO, C., FURFINE, C. & LOWE, P. 2001. Procyclicality of the financial system and financial stability: issues and policy options. *In:* SETTLEMENTS, B. F. I. (ed.) *Marrying the macro- and micro-prudential dimensions of financial stability.* Bank for International Settlements.

BOS, J. & SCHMIEDEL, H. 2007. Is there a single frontier in a single European banking market? *Journal of Banking & Finance,* 31**,** 2081-2102.

BOUVATIER, V. & LEPETIT, L. 2008. Banks' procyclical behavior: Does provisioning matter? *Journal of International Financial Markets, Institutions and Money,* 18**,** 513-526.

BOUVATIER, V., LEPETIT, L. & STROBEL, F. 2014. Bank income smoothing, ownership concentration and the regulatory environment. *Journal of Banking & Finance,* 41**,** 253-270.

BRYCE, C., DADOUKIS, A., HALL, M., NGUYEN, L. & SIMPER, R. 2015. An analysis of loan loss provisioning behaviour in Vietnamese banking. *Finance Research Letters,* 14**,** 69-75.

BUSHMAN, R. M. & WILLIAMS, C. D. 2012. Accounting discretion, loan loss provisioning, and discipline of Banks' risk-taking. *Journal of Accounting and Economics,* 54**,** 1-18.

CALIŃSKI, T. & HARABASZ, J. 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods,* 3**,** 1-27.

CALOMIRIS, C. W. & KAHN, C. M. 1991. The Role of Demandable Debt in Structuring Optimal Banking Arrangements. *The American Economic Review,* 81**,** 497-513.

CERUTTI, E., KOCH, C. & PRADHAN, S. 2020. Banking Across Borders: Are Chinese Banks Different? *IMF Working Papers,* 20.

CHEN, B., W. HARRISON, R., PAN, Y. & TAI, P. 2005. *Novel Hybrid Hierarchical-K-means Clustering Method (H-K-means) for Microarray Analysis*.

CLAESSENS, S., DEMIRGÜÇ-KUNT, A. & HUIZINGA, H. 2001. How does foreign entry affect domestic banking markets? *Journal of Banking & Finance,* 25**,** 891-911.

COLLINS, J. H., SHACKELFORD, D. A. & WAHLEN, J. M. 1995. Bank Differences in the Coordination of Regulatory Capital, Earnings, and Taxes. *Journal of Accounting Research,* 33**,** 263-291.

CORNETT, M. M., MCNUTT, J. J. & TEHRANIAN, H. 2009. Corporate governance and earnings management at large U.S. bank holding companies. *Journal of Corporate Finance,* 15**,** 412-430.

CORNWELL, C., SCHMIDT, P. & SICKLES, R. C. 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics,* 46**,** 185-200.

CORTAVARRIA, L., DZIOBEK, C., KANAYA, A. & SONG, I. 2000. Loan Review, Provisioning, and Macroeconomic Linkages. *IMF Working Papers no.00/195.*

CROCCO, M., SANTOS, F. & AMARAL, P. V. 2010. The Spatial Structure of Financial Development in Brazil. *Spatial Economic Analysis,* 5**,** 181-203.

CURCIO, D., DE SIMONE, A. & GALLO, A. 2017. Financial crisis and international supervision: New evidence on the discretionary use of loan loss provisions at Euro Area commercial banks. *The British Accounting Review,* 49**,** 181-193.

DALY, M., KRAINER, J. & LOPEZ, J. 2003. Does regional economic performance affect bank health? New analysis of an old question. *IDEAS Working Paper Series from RePEc*.

DAVIES, R. & TRACEY, B. 2014. Too Big to Be Efficient? The Impact of Implicit Subsidies on Estimates of Scale Economies for Banks. *Journal of Money, Credit and Banking,* 46**,** 219-253.

DE ANDRÉS, J., LANDAJO, M. & LORCA, P. 2009. Flexible quantile-based modeling of bivariate financial relationships: The case of ROA ratio. *Expert Systems with Applications,* 36**,** 8955-8966.

DEMIRGUC-KUNT, A., LAEVEN, L. & LEVINE, R. 2003. Regulations, market structure, institutions, and the cost of financial intermediation. National Bureau of Economic Research.

DEMSETZ, H. 1973. Industry Structure, Market Rivalry, and Public Policy. *The Journal of Law & Economics,* 16**,** 1-9.

DEYOUNG, R. 1997. Measuring bank cost efficiency: Don't count on accounting. *Financial Practice & Education,* 7**,** 20-31.

DIONNE, G., GAGNÉ, R. & VANASSE, C. 1998. Inferring technological parameters from incomplete panel data. *Journal of Econometrics,* 87**,** 303-327.

DOW, S. C. & MONTAGNOLI, A. 2007. The Regional Transmission of UK Monetary Policy. *Regional Studies,* 41**,** 797-808.

DU, J. & GIRMA, S. 2011. Cost economies, efficiency and productivity growth in the Chinese banking industry: evidence from a quarterly panel dataset. *Empirical Economics,* 41**,** 199-226.

DUDA, R. O., HART, P. E. & STORK, D. G. 2012. *Pattern classification*, John Wiley & Sons.

EL SOOD, H. A. 2012. Loan loss provisioning and income smoothing in US banks pre and post the financial crisis. *International Review of Financial Analysis,* 25**,** 64-72.

ELLIOTT, D. J. & YAN, K. 2013. *The Chinese financial system: An introduction and overview*, Brookings Institution.

EMMONS, W., GILBERT, R. & YEAGER, T. 2004. Reducing the Risk at Small Community Banks: Is it Size or Geographic Diversification that Matters?*. *Journal of Financial Services Research,* 25**,** 259-281.

ENG, L. L. & NABAR, S. 2007. Loan Loss Provisions by Banks in Hong Kong, Malaysia and Singapore. *Journal of International Financial Management & Accounting,* 18**,** 18-38.

FEINBERG, W. E. & TROTTA, J. R. 1984. Definitional dependency is dead. Does chaos still reign? A possible resolution of the ratio-correlation dilemma. *Social Forces,* 62**,** 1068-1075.

FERNANDEZ DE LIS, S., PAGÉS, J. & SAURINA, J. 2000. Credit Growth, Problem Loans and Credit Risk Provisioning in Spain.

FLAMINI, V., MCDONALD, C. A. & SCHUMACHER, L. T. 2009. The Determinants of Commercial Bank Profitability in Sub-Saharan Africa. IMF Working Paper**,** 1-30.

FONSECA, A. R. & GONZÁLEZ, F. 2008. Cross-country determinants of bank income smoothing by managing loan-loss provisions. *Journal of Banking & Finance,* 32**,** 217-228.

FREEMAN, J. H. & KRONENFELD, J. E. 1973. Problems of definitional dependency: the case of administrative intensity. *Social Forces,* 52**,** 108-121.

FU, X. & HEFFERNAN, S. 2007. Cost X-efficiency in China's banking sector. *China Economic Review,* 18**,** 35-53.

FU, X. & HEFFERNAN, S. 2008. Economies of scale and scope in China's banking sector. *Applied Financial Economics,* 18**,** 345-356.

FURLONG, F. & KRAINER, J. 2007. Regional economic conditions and the variability of rates of return in commercial banking. *IDEAS Working Paper Series from RePEc*.

GLEN, J. & MONDRAGÓN-VÉLEZ, C. 2011. Business cycle effects on commercial bank loan portfolio performance in developing economies. *Review of Development Finance,* 1**,** 150-165.

GODDARD, J., MOLYNEUX, P. & WILLIAMS, J. 2014. Dealing with cross-firm heterogeneity in bank efficiency estimates: Some evidence from Latin America. *Journal of Banking & Finance,* 40**,** 130-142.

GOOD, D. H., NADIRI, M. I., RÖLLER, L.-H. & SICKLES, R. C. 2013. Efficiency and productivity growth comparisons of European and U.S. Air carriers: A first look at the data. *Journal of Productivity Analysis,* 4**,** 115-125.

GORTON, G. & METRICK, A. 2012. Securitized banking and the run on repo. *Journal of Financial Economics,* 104**,** 425-451.

GOYAL, J., SINGH, M., SINGH, R. & AGGARWAL, A. 2019. Efficiency and technology gaps in Indian banking sector: Application of meta-frontier directional distance function DEA approach. *The Journal of Finance and Data Science,* 5**,** 156-172.

GREENAWALT, M. & SINKEY, J., JR. 1988. Bank loan-loss provisions and the income-smoothing hypothesis: An empirical analysis, 1976–1984. *Journal of Financial Services Research,* 1**,** 301-318.

GREENE, W. 2005. Fixed and random effects in stochastic frontier models. *Journal of productivity analysis,* 23**,** 7-32.

GREENE, W. H. 2008. The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth***,** 92-250.

HAYAMI, Y. 1969. Sources of Agricultural Productivity Gap among Selected Countries. *American Journal of Agricultural Economics,* 51**,** 564-575.

HAYAMI, Y. & RUTTAN, V. W. 1970. Agricultural Productivity Differences among Countries. *The American Economic Review,* 60**,** 895-911.

HAYAMI, Y. & RUTTAN, V. W. 1971. *Agricultural development: an international perspective*, Baltimore, Md/London: The Johns Hopkins Press.

HE, D. 2016. *Financial Security in China: Situation Analysis and System Design*, Springer.
HEFFERNAN, S. 2005. *Modern Banking*, Wiley.
HICKS, J. R. 1935. Annual Survey of Economic Theory: The Theory of Monopoly. *Econometrica,* 3**,** 1-20.
HOMANEN, M. 2018. Depositors Disciplining Banks: The Impact of Scandals *Chicago Booth Research Paper* 28.
HOPWOOD, W. & MCKEOWN, J. C. 2003. Market Effects, Size Contingency and Financial Ratios. *Review of Accounting and Finance,* 2**,** 3-15.
HOU, X., WANG, Q. & LI, C. 2015. Role of off-balance sheet operations on bank scale economies: Evidence from China's banking sector. *Emerging Markets Review,* 22**,** 140-153.
HUANG, C. J., HUANG, T.-H. & LIU, N.-H. 2014. A new approach to estimating the metafrontier production function based on a stochastic frontier framework. *Journal of Productivity Analysis,* 42**,** 241-254.
HUANG, M.-Y. & FU, T.-T. 2013. An examination of the cost efficiency of banks in Taiwan and China using the metafrontier cost function. *Journal of Productivity Analysis,* 40**,** 387-406.
HUGHES, J., MESTER, L. & MOON, C.-G. 2001. Are scale economies in banking elusive or illusive?: Evidence obtained by incorporating capital structure and risk-taking into models of bank production. *Journal of Banking & Finance,* 25**,** 2169-2208.
HUGHES, J. & MESTER, L. J. 2013. Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation,* 22**,** 559-585.
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,* 4**,** 293-315.
HUMPHREY, D. 1992. Flow versus stock indicators of banking output: Effects on productivity and scale economy measurement. *Journal of Financial Services Research,* 6**,** 115-135.
HUMPHREY, D. B. 1990. Why do estimates of bank scale economies differ? *Economic Review***,** 38-50.
HUSSON, F., LÊ, S. & PAGÈS, J. 2017. *Exploratory multivariate analysis by example using R*, CRC press.
IYER, R., PURI, M. & RYAN, N. 2016. A Tale of Two Runs: Depositor Responses to Bank Solvency Risk. *The Journal of Finance,* 71**,** 2687-2725.
JACKSON, W. E. & EISENBEIS, R. A. 1997. Geographic integration of bank deposit markets and restrictions on interstate banking: A cointegration approach. *Journal of Economics and Business,* 49**,** 335-346.
JIANG, C., YAO, S. & ZHANG, Z. 2009. The effects of governance changes on bank efficiency in China: A stochastic distance function approach. *China Economic Review,* 20**,** 717-731.
JIN, J. Y., KANAGARETNAM, K. & LIU, Y. 2018. Banks' funding structure and earnings quality. *International Review of Financial Analysis,* 59**,** 163-178.

KANAGAGRETNAM, K., LOBO, G. J. & YANG, D.-H. 2004. Joint Tests of Signaling and Income Smoothing through Bank Loan Loss Provisions*. *Contemporary Accounting Research,* 21**,** 843-884.

KANAGARETNAM, K., KRISHNAN, G. V. & LOBO, G. J. 2010. An Empirical Analysis of Auditor Independence in the Banking Industry. *The Accounting Review,* 85**,** 2011-2046.

KASMAN, A. & YILDIRIM, C. 2006. Cost and profit efficiencies in transition banking: the case of new EU members. *Applied Economics,* 38**,** 1079-1090.

KAUFMAN, L. & ROUSSEEUW, P. 1990. *Finding Groups in Data: An Introduction To Cluster Analysis*.

KIM, M.-S. & KROSS, W. 1998. The impact of the 1989 change in bank capital standards on loan loss provisions and loan write-offs. *Journal of Accounting and Economics,* 25**,** 69-99.

KIM, M., KLIGER, D. & VALE, B. 2003. Estimating switching costs: the case of banking. *Journal of Financial Intermediation,* 12**,** 25-56.

KLAGGE, B. & MARTIN, R. 2005. Decentralized versus centralized financial systems: is there a case for local capital markets? *J. Econ. Geogr.,* 5**,** 387-421.

KODDE, D. A. & PALM, F. C. 1986. Wald Criteria for Jointly Testing Equality and Inequality Restrictions. *Econometrica,* 54**,** 1243-1248.

KONTOLAIMOU, A. & TSEKOURAS, K. 2010. Are cooperatives the weakest link in European banking? A non-parametric metafrontier approach. *Journal of Banking & Finance,* 34**,** 1946-1957.

KOOPMANS, T. C. 1951. *Activity Analysis of Production and Allocation. Proceedings of a conference. Edited by T. C. Koopmans [and others], etc*, John Wiley & Sons.

KOSMIDOU, K. 2008. The determinants of banks' profits in Greece during the period of EU financial integrationnull. *Managerial Finance,* 34**,** 146-159.

KPMG. 2013. Global Debt Sales, China, Third Edition. Available: http://www.kpmg.com/global/en/issuesandinsights/articlespublications/global-debt-sales/pages/default.aspx.

KUMBHAKAR, S. C. & LOVELL, C. A. K. 2003. *Stochastic Frontier Analysis*, Cambridge University Press.

KUMBHAKAR, S. C. & WANG, D. 2007. Economic reforms, efficiency and productivity in Chinese banking. *Journal of Regulatory Economics,* 32**,** 105-129.

KWAN, C. H. 2009. Reform of China's State-Owned Banks a Success. *Nomura Journal of Capital Markets,* 1.

KWAN, S. & EISENBEIS, R. 1997. Bank Risk, Capitalization, and Operating Efficiency. *Journal of Financial Services Research,* 12**,** 117-131.

KWAST, M. L., STARR-MCCLUER, M. & WOLKEN, J. D. 1997. Market definition and the analysis of antitrust in banking. *The Antitrust Bulletin,* 42**,** 973-995.

LAEVEN, L. & MAJNONI, G. 2003. Loan loss provisioning and economic slowdowns: too much, too late? *Journal of Financial Intermediation,* 12**,** 178-197.

LAMBERT, R. A. 1984. Income Smoothing as Rational Equilibrium Behavior. *The Accounting Review,* 59**,** 604–618.

LEE, C.-C. & HUANG, T.-H. 2017. Cost efficiency and technological gap in Western European banks: A stochastic metafrontier analysis. *International Review of Economics & Finance,* 48**,** 161-178.

LEE, C.-C. & HUANG, T.-H. 2019. WHAT CAUSES THE EFFICIENCY AND THE TECHNOLOGY GAP UNDER DIFFERENT OWNERSHIP STRUCTURES IN THE CHINESE BANKING INDUSTRY? *Contemporary Economic Policy,* 37**,** 332-348.

LEIBENSTEIN, H. 1966. Allocative Efficiency vs. "X-Efficiency". *The American Economic Review,* 56**,** 392-415.

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,* 28**,** 553-568.

LEUNG, M. K. & CHAN, R. Y. K. 2006. Are foreign banks sure winners in post-WTO China? *Business Horizons,* 49**,** 221-234.

LEV, B. & SUNDER, S. 1979. Methodological issues in the use of financial ratios. *Journal of Accounting and Economics,* 1**,** 187-210.

LEVENTIS, S. & DIMITROPOULOS, P. 2012. The role of corporate governance in earnings management: experience from US banks. *Journal of Applied Accounting Research,* 13**,** 161-177.

LEVENTIS, S., DIMITROPOULOS, P. E. & ANANDARAJAN, A. 2012. Signalling by banks using loan loss provisions: the case of the European Union. *Journal of Economic Studies,* 39**,** 604-618.

LEVINE, R. 1997. Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature,* 35**,** 688-726.

LI, R., LI, X., LEI, W. & HUANG, Y. 2015. Consequences of China's Opening to Foreign Banks. *In:* LIGANG, S., GARNAUT, R., CAI, F. & JOHNSTON, L. (eds.) *China's Domestic Transformation in a Global Context.* ANU Press.

LI, Y. P. & HAN, L. Y. 2008. Charter value, official implicit insurance and risk-taking behaviors of Chinese banks. *Journal of Beijing Institute of Technology (English Edition),* 17**,** 132-135.

LITWACK, J. M. & QIAN, Y. 1998. Balanced or Unbalanced Development: Special Economic Zones as Catalysts for Transition. *Journal of Comparative Economics,* 26**,** 117-141.

LIU, H. & WILSON, J. O. S. 2010. The profitability of banks in Japan. *Applied Financial Economics,* 20**,** 1851-1866.

LIU, W. & WU, W. 2008. Development of Local Financial Systems in Mainland China. *Eurasian Geography and Economics,* 49**,** 160-179.

LOWE, P. 2003. Credit Risk Measurement and Procyclicality. *In:* KUPPERS, J. G., PRAST, H. N. & WESSEING, A. A. T. (eds.) *Supervision on a Cross-road.* Cheltenham, UK and Northampton, MA., USA: Edward Elgar.

LOZANO-VIVAS, A. & PASIOURAS, F. 2014. Bank Productivity Change and Off-Balance-Sheet Activities Across Different Levels of Economic Development. *Journal of Financial Services Research,* 46**,** 271-294.

LOZANO-VIVAS, A., PASTOR, J. T. & PASTOR, J. M. 2002. An Efficiency Comparison of European Banking Systems Operating under Different Environmental Conditions. *Journal of Productivity Analysis,* 18**,** 59-77.

LUO, D., DONG, Y., ARMITAGE, S. & HOU, W. 2017. The impact of foreign bank penetration on the domestic banking sector: new evidence from China. *The European Journal of Finance,* 23**,** 752-780.

MACFADDEN, D. 1966. *Cost, Revenue, and Profit Functions: A Cursory Review*.

MAKLES, A. 2012. Stata tip 110: How to get the optimal k-means cluster solution. *Stata Journal,* 12**,** 347-351.

MAUDOS, J. N., PASTOR, J. M., PÉREZ, F. & QUESADA, J. 2002. Cost and profit efficiency in European banks. *Journal of International Financial Markets, Institutions and Money,* 12**,** 33-58.

MCALLISTER, P. H. & MCMANUS, D. 1993. Resolving the scale efficiency puzzle in banking. *Journal of Banking & Finance,* 17**,** 389-405.

MCLEAY, S. & TRIGUEIROS, D. 2002. Proportionate Growth and the Theoretical Foundations of Financial Ratios. *Abacus,* 28**,** 297-316.

MEEUSEN, W. & VAN DEN BROECK, J. 1977. Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review,* 18**,** 435-444.

MERRYMAN, S. 2008. CHINA_MAP: Stata module to provide map of China's provinces. S456958 ed.: Boston College Department of Economics.

MEYER, A. & YEAGER, T. 2001. Are small rural banks vulnerable to local economic downturns? *Review - Federal Reserve Bank of St. Louis,* 83**,** 25-38.

MILEVA, E. 2007. Using Arellano – Bond Dynamic Panel GMM Estimators in Stata. *Tutorial with Examples using Stata 9.0* [Online]. Available: http://www.fordham.edu/economics/mcleod/Elitz-UsingArellano%E2%80%93BondGMMEstimators.pdf [Accessed 20 Aug 2014].

MUKHERJEE, A., NATH, P. & NATH PAL, M. 2002. Performance benchmarking and strategic homogeneity of Indian banks. *International Journal of Bank Marketing,* 20**,** 122-139.

MULLINEAUX, D. J. 1978. ECONOMIES OF SCALE AND ORGANIZATIONAL EFFICIENCY IN BANKING: A PROFIT-FUNCTION APPROACH. *The Journal of Finance,* 33**,** 259-280.

NAUGHTON, B. 2006. *The Chinese economy : transitions and growth / Barry Naughton,* Cambridge, Mass. ; London, England, MIT Press.

O'DONNELL, C. J., RAO, D. S. P. & BATTESE, G. E. 2008. Metafrontier frameworks for the study of firm-level efficiencies and technology ratios. *Empirical Economics,* 34**,** 231-255.

PACKER, F. & ZHU, H. 2012. Loan loss provisioning practices of Asian banks. Bank for International Settlements.

PARK, K. H. & WEBER, W. L. 2006. A note on efficiency and productivity growth in the Korean Banking Industry, 1992–2002. *Journal of Banking & Finance,* 30**,** 2371-2386.

PASIOURAS, F., TANNA, S. & ZOPOUNIDIS, C. 2009. The impact of banking regulations on banks' cost and profit efficiency: Cross-country evidence. *International Review of Financial Analysis,* 18**,** 294-302.

PENG, Q. & HAN, W. 2021. China Allows Baoshang Bank to Go Bankrupt in Final Cleanup. Available: https://www.caixinglobal.com/2020-08-07/china-allows-baoshang-bank-to-go-bankrupt-in-final-cleanup-101589758.html [Accessed 1 July 2021].

PÉREZ, D., SALAS-FUMÁS, V. & SAURINA, J. 2008. Earnings and Capital Management in Alternative Loan Loss Provision Regulatory Regimes. *European Accounting Review,* 17**,** 423-445.

RESTI, A. Linear Programming and Econometric Methods for Bank Efficiency Evaluation: an Empirical Comparison Based on a Panel of Italian Banks. *In:* BERTOCCHI, M., CAVALLI, E. & KOMLÓSI, S., eds. Modelling Techniques for Financial Markets and Bank Management, 1996// 1996 Heidelberg. Physica-Verlag HD, 111-139.

RESTREPO-TOBON, D. & KUMBHAKAR, S. 2015. Nonparametric estimation of returns to scale using input distance functions: an application to large U.S. banks. *Empirical Economics,* 48**,** 143-168.

REUTERS. 2021. China's Liaoning province plans merging local banks as early as June -sources. Available: https://www.reuters.com/article/china-banks-idUSL4N2N118Z [Accessed 1 July 2021].

REVELL, J. 1979. Inflation and financial institutions. *Financial Times*.

ROODMAN, D. 2006. How to do xtabond2: an introduction to "Difference" and "System" GMM in Stata. *Center for Global Development Working Paper.*

ROSE, P. S. & HUDGINS, S. C. 2012. Bank Management & Financial Services. London: McGraw-Hill.

ROUSSEEUW, P. J. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics,* 20**,** 53-65.

SAUNDERS, A. & CORNETT, M. M. 2018. *Financial institutions management : a risk management approach*.

SCHMIDT, P. & SICKLES, R. C. 1984. Production Frontiers and Panel Data. *Journal of Business & Economic Statistics,* 2**,** 367-374.

SEALEY, C. W. & LINDLEY, J. T. 1977. Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions. *The Journal of Finance,* 32**,** 1251-1266.

SHEN, C.-H., LU, C.-H. & WU, M.-W. 2009. Impact of Foreign Bank Entry on the Performance of Chinese Banks. 17**,** 102-121.

SHERMAN, H. D. & GOLD, F. 1985. Bank branch operating efficiency: Evaluation with data envelopment analysis. *Journal of banking & finance,* 9**,** 297-315.

SOEDARMONO, W., TARAZI, A., AGUSMAN, A., MONROE, G. S. & GASBARRO, D. 2012. Loan loss provisions and lending behaviour of banks: Asian evidence during 1992-2009. *Available at SSRN 2049217*.

SPENCE, M. 1973. Job Market Signaling. *The Quarterly Journal of Economics,* 87**,** 355-374.

STAIKOURAS, C. & WOOD, G. 2004. *The Determinants Of European Bank Profitability*.

STEVENSON, R. E. 1980. Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics,* 13**,** 57-66.

SUDARSANAM, P. S. & TAFFLER, R. J. 1995. Financial ratio proportionality and inter-temporal stability: An empirical analysis. *Journal of Banking & Finance,* 19**,** 45-60.

SUFIAN, F. & HABIBULLAH, M. S. 2009. Bank specific and macroeconomic determinants of bank profitability: Empirical evidence from the China banking sector. *Frontiers of Economics in China,* 4**,** 274-291.

SULLIVAN, R. J. & SPONG, K. R. 2007. Manager wealth concentration, ownership structure, and risk in commercial banks. *Journal of Financial Intermediation,* 16**,** 229-248.

SUN, J., HARIMAYA, K. & YAMORI, N. 2013. Regional economic development, strategic investors, and efficiency of Chinese city commercial banks. *Journal of Banking & Finance,* 37**,** 1602-1611.

SUN, J. & YAMORI, N. 2009. REGIONAL DISPARITIES AND INVESTMENT–CASH FLOW SENSITIVITY: EVIDENCE FROM CHINESE LISTED FIRMS. *Pacific Economic Review,* 14**,** 657-667.

TAN, Y. & FLOROS, C. 2012. Bank profitability and inflation: the case of Chinanull. *Journal of Economic Studies,* 39**,** 675-696.

TANDON, A., MURRAY, C. J., LAUER, J. A. & EVANS, D. B. 2000. Measuring overall health system performance for 191 countries. *Geneva: World Health Organization*.

TRUJILLO-PONCE, A. 2013. What determines the profitability of banks? Evidence from Spain. *Accounting & Finance,* 53**,** 561-586.

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,* 27**,** 2323-2345.

USCC 2013. Chapter 1; Section 3 Governance and Accountability in China's Financial System. *U.S.-China Economic and Security Review Commission 2013 ANNUAL REPORT TO CONGRESS.*

WAHLEN, J. M. 1994. The nature of information in commercial bank loan loss disclosures. *Accounting Review***,** 455-478.

WANG, H.-J. & SCHMIDT, P. 2002. One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels. *Journal of Productivity Analysis,* 18**,** 129-144.

WHEELOCK, D. C. & WILSON, P. W. 2012. Do Large Banks Have Lower Costs? New Estimates of Returns to Scale for U. S. Banks. *Journal of Money, Credit and Banking,* 44**,** 171-199.

WHEELOCK, D. C. & WILSON, P. W. 2018. The evolution of scale economies in US banking. *Journal of Applied Econometrics,* 33**,** 16-28.

WHITE, H. 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica,* 50**,** 1-25.

WHITTINGTON, G. 1980. SOME BASIC PROPERTIES OF ACCOUNTING RATIOS. *Journal of Business Finance & Accounting,* 7**,** 219-232.

WILLIAMS, J. 2004. Determining management behaviour in European banking. *Journal of Banking & Finance,* 28**,** 2427-2460.

WINSTEN, C. 1957. Discussion on Mr. Farrell's Paper. *Journal of the Royal Statistical Society. Series A (General)* 120**,** 282-284.

WISEMAN, R. M. 2009. On the use and misuse of ratios in strategic management research. *Research methodology in strategy and management,* 5**,** 75-110.

WU, W. 2018. Asia Bank Primer 2018. Bank of America Merrill Lynch.

XIANG, D., SHAMSUDDIN, A. & WORTHINGTON, A. C. 2013. The differing efficiency experiences of banks leading up to the global financial crisis: A comparative empirical analysis from Australia, Canada and the UK. *Journal of Economics and Finance,* 39**,** 327-346.

XU, Y. 2011. Towards a more accurate measure of foreign bank entry and its impact on domestic banking performance: The case of China. *Journal of Banking & Finance,* 35**,** 886-901.

YANG, W., LIU, Y. C., MAI, S. Y. & MAI, C. C. 2015. Regional Performance of China's Banks: Evidence from Industrial and Commercial Bank of China. *Review of Pacific Basin Financial Markets and Policies,* 18.

YAO, S., JIANG, C., FENG, G. & WILLENBOCKEL, D. 2007. WTO challenges and efficiency of Chinese banks. *Applied Economics,* 39**,** 629-643.

YIN, Z., YU, Y. & HUANG, J. 2018. Evaluation and evolution of bank efficiency considering heterogeneity technology: An empirical study from China. *PLOS ONE,* 13**,** e0204559.

ZHAO, Y. & XU, Z. 2021. The Impact of Cross-Border Capital Flows on the Chinese Banking System. *SAGE Open,* 11**,** 21582440211021410.

ZHU, C. & CHEN, L. 2016. An Empirical Study on the Capital Buffer of Rural Commercial Banks in China. *Journal of Finance and Economics,* 4**,** 97-102.

ZHU, N., BALEŽENTIS, T., YU, Z. & WU, W. 2020. Multi-Directional Meta-Frontier DEA Model for Total Factor Productivity Growth in the Chinese Banking Sector: A Disaggregation Approach. *Informatica,* 31**,** 185-204.

Appendix 1

Table A.1. Bank Performance Studies using Profitability Ratios

Authors	Country	Purpose	Sample	Methodology	Variables	Reasons for choosing studied dependent variable	Results
Kosmidou (2008)	Greece	Examine the determinants of performance of Greek banks during the period of EU financial integration	Unbalanced pooled time series dataset of 23 banks during 1990-2002	fixed effects regression	Dependent variable: ROAA Independent Variables: COST, EQAS,LODEP,LORES, SIZE (bank-specific) GDPGR,INF,MSG, MACPASS, ASSGDP, COC(macroeconomic factors)	ROAA indicates the returns generated from the assets financed by the bank	ROAA was found to be associated with well-capitalized banks and lower cost to income ratios. GDP has a significantly positive effect while inflation has a significantly negative effect on ROAA
Liu and Wilson (2010)	Japan	Profitability of Japanese banks following the major financial crisis that	Accounts data for all banks operating in Japan over the period 2000 to 2007	dynamic specification model by including a	Dependent variables: ROA, ROE, NIM Independent Variables: DIV, LA, KA,CI,IMP,MS(bank-specific)	Commonly used profitability measures	Well-capitalized, efficient banks, with lower credit risks tend to outperform less capitalized, less efficient counterparts with higher credit

		affected the country's economy in the mid-1990s.		lagged dependent variable (System GMM)	HH, GDPG, MC (macroeconomic factors)	(They also find interest income is less volatile than noninterest income)	risks; industry concentration, Gross Domestic Product (GDP) growth and the extent of stock market development play an important role in determining the profitability of Japanese banks.
Anbar and Alper (2011)	Turkey	Examine bank-specific and macroeconomic determinants of the bank's profitability	10 commercial banks observed over the period 2002 – 2010 consisting of 90 observations in Turkey	bank-specific and macroeconomic determinants of the banks profitability	Dependent variables: ROA ROE, Independent Variables: logA, CA, LA, LQD, DP, NIM, NII(bank-specific) GDP, INF, RI (macroeconomic factors)	ROA is a general measure for bank profitability that reflects bank's ability to achieve returns on its sources of funds to generate profits. ROE is defined as net profit divided by shareholders' equity and is expressed in percent.	Asset size and non-interest income have a positive and significant effect on bank profitability; size of credit portfolio and loans under follow-up have a negative and significant impact on bank profitability.only the real interest rate affects the performance of banks positively
Tan and Floros (2012)	China	Evaluate the determinants of bank profitability in China	101 banks (five state-owned banks, 12 joint-stock commercial banks and 84 city	two step generalized methods of moments	Dependent variables: ROA, NIM Independent Variables: lagged dependent variables, LTA, LLPTA, LA, TOPBT, ETA, CE, NTA, LP (bank-specific) C3,	ROA is widely used to compare the efficiency and operational performance of banks as it looks at the returns	Positive relationship between bank profitability, cost efficiency, banking sector development, stock market development and inflation in China;

			commercial banks) over 2003-2009.	(GMM) estimators	C5(, BSD, SMD, IR(macroeconomic factors)	generated from the assets financed by the bank; Providing convenience in comparing our results to other findings reported in the literature; NIM focuses on the profit earned on lending, investing and funding activities. ROE is not considered due to the fact that ROA and NIM are better representatives of bank profitability in China.	low profitability is also associated with higher volume of non-traditional activity and higher taxation
Trujillo-Ponce (2013)	Spain	Analyses the factors that determine the profitability of	All Spanish commercial banks, savings banks and credit cooperatives in the Bankscope	system-GMM	Dependent variables: ROA, ROE Independent Variables: Loan/TA , NPL/G , LLP/NL , Eq/TA , Dep/TL , DepGR , CIR , Size , HHIRD ,(bank-specific)	ROA- the single most important ratio for comparing the efficiency and operational performance of banks.	High bank profitability during these years is associated with a large percentage of loans in

		Spanish banks for the period of 1999–2009.	database during the period 1999–2009		HHIIC, GDP , Inflation , Interest (macroeconomic factors)	This ratio considers the returns generated from the assets that the bank finances; it is primarily an indicator of managerial efficiency, although it may be misleading because of off-balance-sheet activities. (ROE), which is a measure of the return on shareholder funds. .	total assets, a high proportion of customer deposits, good efficiency and a low doubtful assets ratio; higher capital ratios increase the bank's return only when return on assets (ROA) is used as the profitability measure.

Appendix 2

Table A.2. Summary of literature on SFA studies in Chinese Banking Sector

Authors	study period	Methodology	Inputs and Outputs	Environmental Variables	Findings
Fu and Heffernan (2007)	1985–2002	Two stage regressions with cost function	Inputs: costs of funds, fixed assets, and labour Outputs: Total deposits, Total loans, non-interest income and total investment. Input prices: The cost of funds, cost of employees and cost of fixed assets	Ownership dummy, reform dummy, purchased funds over total assets, total loans over total assets; total investment over total assets non-interest income over pre-tax profits.	Banks are operating 40–60% below the frontier. joint-stock banks are more x-efficient than state-owned banks. privatisation, greater foreign bank participation, and liberalised interest rates help with increasing efficiency
Kumbhakar and Wang (2007)	1993–2002	input distance function BC(95)	Outputs: y1 is total loans net; y2 is other earning assets. Inputs: x1 is labor; x2 is total fixed assets;	Bank characteristics and environmental variables (the Z variables): z1 is ratio of	EQ/TA ratio to be negatively related to inefficiency; private ownership increases efficiency;

			x3 is total deposits plus total borrowed funds.	equity to total assets; z2 is ownership dummy; z3 is deregulation dummy; and z4 includes four dummy variables for different bank sizes	diminishing negative marginal effects of bank size on efficiency; deregulation did not improve efficiency
Berger et al. (2009)	1994–2003	Standard Cost and profit efficiency with year dummies	Input Prices: Unit interest cost of deposits; Unit interest cost of deposits Output Quantities: Total loans; Total deposits; Liquid Assers;Other earning assets	N.A.	Big Four banks the least efficient; foreign banks most efficient; minority foreign ownership improves efficiency
Jiang et al. (2009)	1995–2005	Output distance function with BC(95)	Model 1:2 Inputs: total interest expense and non-interest expense 2 outputs:net interest income and non-interest income; Model 2: interest expense and labour and physical capital and three outputs, total loans total deposits and non-interest income	governance indicators (a set of dummy variables such as listed, foreign acquisition and going public, risk-taking characteristics (e,g. E/A ratio, LLR/Loans) GDP and time trend	Increase in efficiency over the study period; Joint-stock ownership increases efficiency; Foreign acquisition increases efficiency in the long run while IPOs only have short-term benefit
(Du and Girma, 2011)	1995–2001	Cost function with time trend and a share equation,	Intermediation Approach: 4 outputs-short-term loans, long-term loans, investments	N.A.	Joint-stock commercial banks outperform state-owned

	Quarterly data	labour is modelled as a quasi-input, where u having a half-normal distribution	and securities and other earning assets 4 inputs: short-term deposits, long-term deposits, other borrowing and deposits and inter-bank deposits; TC=interest expenses and operating costs		commercial banks in productivity growth and cost-efficiency