May 2020

Essays on Benchmarking and Shadow Pricing

Shirong Zhao
Clemson University, shironz@g.clemson.edu

Essays on Benchmarking and Shadow Pricing

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Economics

---

by
Shirong Zhao
May 2020

Accepted by:
Dr. Paul Wilson, Committee Chair
Dr. Matthew Lewis
Dr. Babur De los Santos
Dr. Howard Bodenhorn

Abstract

In the first chapter, we examine the performance of Chinese commercial banks before, during, and after the 2008 global financial crisis and the 2008–2010 China's 4 trillion Renminbi stimulus plan. Fully nonparametric methods are used to estimate technical efficiencies. Recently-developed statistical results are used to test for changes in efficiencies as well as productivity over time, and to test for changes in technology over time. We also test for differences in efficiency and productivity between big and small banks, and between domestic and foreign banks. We find evidence of the non-convexity of banks' production set. The data reveal that technical efficiency declined at the start of the global financial crisis (2007–2008) and after the China's stimulus plan (2010–2011), but recovered in the years later (2011–2013), and declined again from 2013 to 2014, ending lower in 2014 than in 2007. We find that productivity declined during and just after the China's stimulus plan (2009–2011), but recovered in the years later (2013–2014), ending lower in 2014 than in 2007. We also find that the technology shifted downward from 2012 to 2013, and then shifted upward from 2013 to 2014. Over the period 2007–2014, technology shifted upward. We provide evidence that in general big banks were more efficient and productive than small banks. Finally, domestic banks had higher efficiency and productivity than foreign banks over this period except in 2008.

In the second chapter, I estimate shadow price of equity for U.S. commercial banks over 2001–2018 using nonparametric estimators of the underlying cost frontier and tests the existence of "Too-Big-to-Fail" (TBTF) banks. Evidence on the existence of TBTF banks are found. Specifically, I find that a negative correlation exists between the shadow price of equity and the size of banks in each year, suggesting that big banks pay less in equity than small banks. In addition, in each year there are more banks with a negative shadow price of equity in the fourth quartile based on total assets than in the other three quartiles. The data also reveal that for each year, the estimated mean shadow price of equity for the top 100 largest banks is smaller than the mean price of deposits,

even though equity is commonly viewed as a riskier asset than deposits. Finally, I find that the top 10 largest banks are willing to pay much more at the start of the global financial crisis and after the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 than the other periods. These results imply that these regulations are effective in reducing the implicit subsidy, at least for the top 10 largest banks. However, it is also evident that the recapitalization has imposed significant equity funding costs for the top 10 largest banks.

In the third chapter, we examine the performance of 144 countries in the world before, during, and after the 2007–2012 global financial crisis. Fully nonparametric methods are used to estimate technical efficiencies. Recently-developed statistical results are used to test for changes in efficiencies as well as productivity over time, and to test for changes in technology over time. We also test for these differences between developing and developed countries. We find evidence of the non-convexity of countries' production set. The data revealed that technical efficiency declined at the start of the global financial crisis (2006–2008), but recovered in the years later (2008–2014), ending higher in 2014 than in 2004. We also find that mean productivity continued decreasing from 2004 to 2010. Moreover, productivity in 2004 stochastic dominants in the first order that in 2014. Statistical tests indicate that the frontier continued shifting downward from 2004 to 2010, and then continued shifting upward from 2010 to 2014. Overall, the technology has shifted downward from 2004 to 2014. Finally, we provide evidence that developing economies have lower technical efficiency but higher productivity than developed economies over this period.

Dedication

To my wife.

Acknowledgments

I would first like to thank Dr. Paul Wilson, the chairman of my committee. He has taught me far more than I could give him credit. He is my mentor in both research and life. He has shown me, not only what a good econometrician should be, but also what a good person should be. I am greatly indebted to him for his guidance, patience and encouragement. Specially I thank Dr. Paul Wilson and Lane for inviting me every year to their Thanksgiving Dinner, treating me with delicious food, making me feel at home, and giving me the best memory at Clemson.

I would also like to thank the guidance and support provided by Dr. Mathew Lewis, Dr. Babur De los Santos, and Dr. Howard Bodenhorn. I give thanks to Dr. Curtis Simon, who helps me make the right decision in some specific times. I would like to thank Dr. Cheng Chen for highly recommending me during job searching. I would like to thank Dr. Xin Liu for guidance in math project and career advice. I am also grateful to Dr. Taufiquar Khan for supporting me financially and admitting me as a graduate student in mathematics. I appreciate Dr. Robert Lund highly recommends me and invites me to his Thanksgiving Dinner. I give thanks to Dr. Shinichi Nishiyama, who gives me the chance to study in US and supports me over the last five years.

I would like to thank my family for being highly supportive of my life choices. I want to first thank my father Guangjun Zhao and my mother Guihua Gao since they give me freedom to pursue the dream. In particular, I'd like to thank my wife for her unconditional love, kindness and support. Without her continuous encouragement, this would never have been possible. I am deeply grateful and also blessed to have her in my life.

Finally, I would like to thank my friends at Clemson for accompanying me with help and laugh. In particular, I'd like to thank Dr. Yanli Wang for giving me numerous advice and support on job searching. I would also like to thank Dr. Ben Posmanick, Fun Choi John Chan, Dr. Haobin Fan, Dr. Guanghua Wang, Jaejeong Shin, Haibin Jiang, Dr. Caitlin O'Loughlin for the happy studying, research and lunch time. Thanks also give to Yanchao Li, Dr. Guolin Yao, Dr. Wenxin Pu, Sha Wan, Jun Li, Guangshun Qiao, Yinlin Dai, Guncha Babajanova, Gangtong Huang for bringing me a lot of fun during the difficult times in the last year of PhD life. The basketball team, tennis team and football team at Clemson also greatly enrich my life at Clemson. Finally I am thankful to Dr. Sarah Wilson and Shannon Graham for editing my papers.

Table of Contents

List of Tables

List of Figures

Chapter 1

Performance of Chinese Banks over 2007–2014

1.1 Introduction

Over the last 40 years, China's GDP growth rate averaged 10 percent per year. China is currently the second largest economy in the world and is projected to become the largest economy in the coming decades. China's banks have played an important role in this transition because China's economy has always relied heavily on domestic investment.

The period 2007–2014 was especially disruptive to the Chinese banking industry. In 2008, the global financial crisis had an influential negative effect on China's exports. To minimize the effect of the crisis on China's economy, China's central government announced a plan in November 2008 to stimulate domestic demand. This plan invested 4 trillion Renminbi (RMB) (about 586 billion U.S. dollars) in rural infrastructure, transportation, disaster rebuilding, health and education, housing, and other areas by the end of 2010.[1] State-owned enterprises (SOEs) mainly undertook the investment plans. To help China's banks make loans to SOEs, the central bank lowered the reserve ratio requirement, the borrowing and lending benchmark rate, and even canceled the credit limit on commercial banks.[2] With the help of the China's stimulus, China's GDP growth averaged 10 percent while the GDP of North America and Europe was slowing.[3]

China's banking industry was undoubtedly heavily influenced by the crisis and the stimulus. According to *International Monetary Institute at Renmin University of China*, by the end of 2007, the total assets of the banking industries in China, U.S., and Germany were 7.48 trillion, 13.56 trillion, and 11.70 trillion U.S. dollars, respectively. By the end of 2016, the total assets of the banking industries in China, U.S., and Germany were 30.32 trillion, 15.22 trillion, and 7.49 trillion U.S. dollars, respectively.[4] China's banking industry expanded far more quickly than the other countries over the period 2007–2016. The latest 2018 *S&P Global Market Intelligence report* lists the 100 largest banks in the world, of which 18 banks were from China and 11 banks were from the U.S. Industrial and Commercial Bank of China (ICBC), China Construction Bank (CCB), Agricultural Bank of China (ABC), and Bank of China (BOC) are the four largest banks in the world in terms of total assets.[5] Also, the latest 2018 *16th Forbes Global 2000* list shows that China and the U.S.

[1] Renminbi is also known as Chinese yuan. All dollar amounts in the following are given in 2010 U.S. dollars and all yuan amounts in the following are given in 2010 Chinese yuan.

[2] For example, the reserve ratio requirement was lowered from 17 percent on 08 October 2008 to 15.5 percent on 18 January 2010.

[3] GDP growth rate for China is 9.65 percent, 9.4 percent, 10.64 percent for 2008, 2009 and 2010 respectively, while GDP growth rate for U.S. is -0.1 percent, -2.5 percent, 2.6 percent for 2008, 2009 and 2010 respectively.

[4] See http://bank.jrj.com.cn/2018/05/16152224548488.shtml.

[5] See https://www.spglobal.com/marketintelligence/en/news-insights/research/the-world-s-100-largest-banks.

split the top 10 of the world's largest public companies evenly this year, in which there were four big Chinese banks and three U.S. banks.

Starting from 1978, Chinese banking industry underwent significant reforms and are now functioning more like western banks. Nonetheless, Chinese banking industry is still very different from U.S. and European banking industry. China's domestic banks has remained in the government's hands even though they have gained more autonomy. Banks may focus on different types of loans and hence have different business plans. Given the importance of banks in the economy and the complexity of the world's largest banking industry, it is reasonable to investigate the performance of China's banks before, during and after the 2008 global financial crisis and the 2008–2010 China's stimulus.

There exists a vast literature on efficiency and productivity of China's commercial banks. Berger et al. (2009) analyze the cost and profit efficiency of Chinese banks over 1994–2003 by specifying translog functional form for cost and profit functions. They find that foreign banks are most efficient and big four banks are by far the least efficient. However, Chinese banking industry is heavily right-skewed even after taking log s, and hence the translog specification could be easily rejected. Chen et al. (2005) examines the cost, technical and allocative efficiency of 43 Chinese banks over the period 1993–2000 using variable returns to scale (VRS) Data Envelopment Analysis (DEA). They specify loans, deposits, and non-interest income as outputs, and interest expenses, non-interest expenses, the price of deposits, and the price of capital as inputs. They map prices to outputs, which does not represent the production process of banks. They find that big four banks and smaller banks are more efficient than medium-sized banks. Yao et al. (2008) use constant returns to scale (CRS) Data Envelopment Analysis to estimate efficiency for the 15 largest Chinese national commercial banks over the period 1998-2005. They specify interest income and non-interest income as outputs, and interest expense, non-interest expense, the ratio of non-performing loans to gross loans as inputs. They map cost to revenue, which again does not represent the typical production function. They find that the three large state-owned banks (CCB, BOC, and ICBC) have high technical efficiency and profitability. These two results are not surprising because the research that studies the efficiency of China's banks using nonparametric methods (either free-disposal hull (FDH) or DEA) has few observations with many inputs and outputs. Thus many of the estimated efficiencies are equal to 1. They naturally encounter the "curse of dimensionality", which is a serious problem in nonparametric efficiency estimation.[6] For example, Chen et al. (2005) specify 3 outputs and 4 inputs for only 43 observations. Yao et al. (2008) specify 2 outputs and 3 inputs for only 15 observations. The effective parametric sample size defined by Wilson (2018) is then, $43^{\frac{4}{8}} \approx 7$ for VRS estimators in Chen et al. (2005), and $15^{\frac{4}{5}} \approx 9$ for CRS estimators in Yao et al. (2008). Hence, dimension reduction is needed in the context of nonparametric efficiency estimation.[7]

Among the nonparametric methods, DEA estimators which impose convexity assumption, have been widely applied to examine technical, cost and profit efficiency in the Chinese banking sector. Recently published examples include Chen et al. (2005), Ariff and Can (2008), Laurenceson and Yong (2008), Yao et al. (2008), Sufian and Habibullah (2009), Luo and Yao (2010), Avkiran (2011), Barros et al. (2011), Gu and Yue (2011), Sufian and Habibullah (2011), Ji et al. (2012), Lee and Chih (2013), Tan and Floros (2013), Dong et al. (2014), Wang et al. (2014), Dong et al. (2014), Wang et al. (2014), An et al. (2015), Liu et al. (2015), Zha et al. (2016), Du et al. (2018). However, they did not test the convexity of the production set, nor do they test CRS versus VRS. Some of these studies have used a two-stage approach, in which in the first stage, efficiency is estimated, and then the estimated efficiencies are regressed on covariates which are usually environmental variables. Published examples are Sufian and Habibullah (2009), Luo and Yao (2010), Sufian and Habibullah (2011), Lee and Chih (2013), Tan and Floros (2013), Du et al. (2018). As mentioned by Simar and Wilson (2007), second-stage regression requires separability condition. However, none of these papers test whether the separability condition holds. Moreover, some of these studies simply report efficiency estimates without any inference and compare the mean efficiency of two groups without correcting the bias of the estimated efficiency. Of course, point estimates without inference are largely meaningless. Hence, the results of these studies are dubious. Recently, Kneip et al. (2016), using the central limit theorem results from Kneip et al. (2015), develop hypotheses testing the model structure. They provide tests of the convexity of the production set, returns to scale and differences in mean efficiency across groups of producers. They reject the convexity assumption of the production set using U.S. commercial banks, which casts doubt on the results of many banking studies that have imposed convexity assumptions.

This paper provides evidence on the performance of China's commercial banks just before,

[6] Curse of dimensionality means the convergence rates of nonparametric estimators decrease with increasing dimensions (number of inputs and outputs).

[7] Recently Wilson (2018) proposes a new dimension reduction technique that is advantageous in terms of reducing estimation error. Results also suggest that FDH estimator is a viable, attractive alternative to the VRS-DEA in many cases when dimension reduction is used.

during and after the 2008 global financial crisis and the 2008–2010 China stimulus. The approach is fully nonparametric and exploits recently developed theoretical results. Estimates of technical efficiency and productivity at one-year intervals from 2007 to 2014 are examined in a statistical paradigm permitting inference and hypothesis testing. Therefore, this paper both (i) contributes to the banking literature by shedding light on the reaction of Chinese commercial banks to the recent crisis and the stimulus, and (ii) fills the gap between point estimates and inference in the empirical research on China's commercial banks' technical efficiency and productivity.

The rest of the paper is organized as follows. Section 1.2 provides background on the Chinese banking industry. Estimators of technical efficiency and their properties are discussed in Section 1.3. Section 1.4 discusses various statistical results needed for testing hypotheses about model features. Section 1.5 describes the Chinese commercial banks data, specially the input and output variables. Section 1.6 discusses the empirical results of the tests. Major conclusions and directions for future works are presented in Section 1.7. Additional details on model assumptions, data, and results are provided in separate Appendices A–C, which are available online.

1.2 Background on the Chinese Banking Industry

Before 1978, the only bank in China was the People's Bank of China (PBC). The PBC took on the responsibilities of central and commercial banking. After the reforms in 1978, the banking system was expanded by establishing four big state owned commercial banks: ICBC, CCB, ABC, and BOC. These four banks took over the role of commercial banking from the PBC and the PBC only undertakes the role of implementing monetary policy. However, the four big state owned commercial banks at that time mainly served as the government's policy lending institutions. They did not have much flexibility and there was little competition among them.[8]

Starting in 1986, 13 joint stock commercial banks were created. They are partially owned by local governments and SOEs, and sometimes by the private sector. They finance small SOEs and firms with partial private ownership, including small and medium-sized enterprises (SMEs). They maintain much smaller branch networks than four big state owned commercial banks, typically confined to the region of origin or to the fast-growing coastal area. However they are generally allowed to operate at the national level.

[8]For example, they could not set the deposits and lending interest rates without authorization from the central government. Even until now, the interest rates are still not totally determined by the market.

In 1994, China's government established three policy lending non-commercial banks: the Export-Import Bank of China, Agricultural Development Bank of China and China Development Bank. These three policy banks took over the policy lending roles from the four big state owned commercial banks.

Since the mid-1990s, city commercial banks have been created by restructuring and consolidating urban credit cooperatives. Their capital is in the hands of urban enterprises and local governments. They mainly lend to SMEs, collective and local residents in their municipalities.

In 1999, 1.4 trillion RMB of nonperforming loans of the four big state owned commercial banks were sold to four newly created asset management companies. At this time there were a lot of strict policies on the internal management of the four big state owned commercial banks. The evaluation of their performance and the risk management have significantly improved since then. After China joined the World Trade Organization in 2001, there was more pressure to reform of China's banks. After 5 years, China's banking industry would open to foreign banks and China promised that there would be fewer restrictions on ownership takeovers and fewer regulations on interest rates.

In 2003, the China Banking Regulation Commission was created to achieve better monitoring of China's banking industry and to oversee reforms and regulations.[9] At the same time, aimed to improve the efficiency and competitiveness of the domestic banks, China government started a new reform on the ownership of domestic banks (especially the four big state owned commercial banks) and hope that they could all be listed in the market.

In addition, rural commercial banks are also one important part of China banking sector. They regard SMEs as their key clients to provide them with business operations aimed at serving the agricultural sector and other rural industries. Historically, bank lending to rural areas has not performed on par with lending to urban areas. In order to encourage lending to rural areas, the China Banking Regulation Commission and central government have considered and initialized some new incentives, such as tax cuts, a lowered capital requirement for rural banks, and subsidy programs that include infrastructure development.

Since 2006, all of the foreign banks were permitted to conduct RMB business and were treated theoretically the same as the domestic banks. In 2010, the ABC became the last bank listed

[9] Starting in April 2018, the China Banking Regulation Commission merges with China Insurance Regulatory Commission as China Banking and Insurance Regulatory Commission.

on the market among the big four banks.

In 2014, in order to enact even more reforms on banks, reduce the financial risk, and provide better banking services, three private banks solely owned by private companies were allowed by the China Banking Regulation Commission to open. As of September 2018, there are already 17 private banks, which have greatly enriched China's banking sector.

1.3 The Statistical Model

To establish notation, let $X \in \mathbb{R}_+^p$ and $Y \in \mathbb{R}_+^q$ denote (random) vectors of input and output quantities, respectively. Similarly, let $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ denote fixed, nonstochastic vectors of input and output quantities. The production set

$$\Psi := \{(x, y) \mid x \text{ can produce } y\} \tag{3.1}$$

gives the set of feasible combinations of inputs and outputs. Several assumptions on Ψ are common in the literature. The assumptions of Shephard (1970) and Färe (1988) are typical in microeconomic theory of the firm and are used here.

Assumption 1.3.1. Ψ *is closed.*

Assumption 1.3.2. $(x, y) \notin \Psi$ *if* $x = 0,\ y \geq 0,\ y \neq 0$*; i.e., all production requires use of some inputs.*

Assumption 1.3.3. *Both inputs and outputs are strongly disposable, i.e.,* $\forall\,(x, y) \in \Psi$*, (i)* $\widetilde{x} \geq x \Rightarrow (\widetilde{x}, y) \in \Psi$ *and (ii)* $\widetilde{y} \leq y \Rightarrow (x, \widetilde{y}) \in \Psi$*.*

Here and throughout, inequalities involving vectors are defined on an element-by-element basis, as is standard. Assumption 1.3.1 ensures that the *efficient frontier* (or *technology*) Ψ^∂

$$\Psi^\partial := \big\{(x, y) \mid (x, y) \in \Psi,\ (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for any } \alpha \in (1, \infty)\big\} \tag{3.2}$$

is the set of extreme points of Ψ and is contained in Ψ. Assumption 1.3.2 means that production of any output quantities greater than 0 requires use of some inputs so that there can be no free lunches. Assumption 1.3.3 imposes weak monotonicity on the frontier.

The Farrell (1957) input efficiency measure

$$\theta(x, y \mid \Psi) := \inf \{\theta \mid (\theta x, y) \in \Psi\} \tag{3.3}$$

gives the amount by which input levels can feasibly be scaled downward, proportionately by the same factor, without reducing output levels. The Farrell (1957) output efficiency measure gives the feasible, proportionate expansion of output quantities and is defined by

$$\lambda(x, y \mid \Psi) := \sup \{\lambda \mid (x, \lambda y) \in \Psi\}. \tag{3.4}$$

Both (3.3) and (3.4) provide *radial* measures of efficiency since all input or output quantities are scaled by the same factor θ or λ, holding output or input quantities fixed (respectively). Clearly, $\theta(x, y \mid \Psi) \leq 1$ and $\lambda(x, y \mid \Psi) \geq 1$ for all $(x, y) \in \Psi$.

Alternatively, Färe et al. (1985) provide a hyperbolic, graph measure of efficiency defined by

$$\gamma(x, y \mid \Psi) := \inf \left\{\gamma > 0 \mid (\gamma x, \gamma^{-1} y) \in \Psi\right\}. \tag{3.5}$$

By construction, $\gamma(x, y \mid \Psi) \leq 1$ for $(x, y) \in \Psi$. Just as the measures $\theta(x, y \mid \Psi)$ and $\lambda(x, y \mid \Psi)$ provide measures of the *technical efficiency* of a firm operating at a point $(x, y) \in \Psi$, so does $\gamma(x, y \mid \Psi)$, but along the hyperbolic path from (x, y) to the frontier of Ψ. The measure $\gamma(x, y \mid \Psi)$ gives the amount by which input levels can be feasibly, proportionately scaled downward while simultaneously scaling output levels upward by the same proportion.

All of the quantities and model features defined so far are unobservable, and therefore must be estimated. The set Ψ can be estimated using the free-disposal hull (FDH) estimator introduced by Deprins et al. (1984) or either the variable returns to scale (VRS) or constant returns to scale (CRS) versions of the data envelopment analysis (DEA) estimator proposed by Farrell (1957). But, inference is needed in order to know what might be learned from data, and inference requires a well-defined statistical model.

1.4 Estimation and Inference

Let $\mathcal{S}_n = \{(X_i, Y_i)\}_{i=1}^n$ be a random input-output pairs sample, where $X_i \in \mathbb{R}_+^p$ and $Y_i \in \mathbb{R}_+^q$. Given a random sample $\mathcal{S}_n = \{(X_i, Y_i)\}$, the production set Ψ can be estimated by the free disposal hull of the sample observations in $\mathcal{S}_n$,

$$\widehat{\Psi}_{\text{FDH},n} := \bigcup_{(X_i,Y_i)\in\mathcal{S}_n} \left\{(x,y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i,\ y \leq Y_i\right\}, \tag{4.1}$$

proposed by Deprins et al. (1984). Alternatively, Ψ can be estimated by the convex hull of $\widehat{\Psi}_{\text{FDH},n}$, i.e., by

$$\widehat{\Psi}_{\text{VRS},n} := \left\{(x,y) \in \mathbb{R}^{p+q} \mid y \leq \boldsymbol{Y\omega},\ x \geq \boldsymbol{X\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{4.2}$$

where $\boldsymbol{X} = \begin{pmatrix} X_1, & \ldots, & X_n \end{pmatrix}$ and $\boldsymbol{Y} = \begin{pmatrix} Y_1, & \ldots, & Y_n \end{pmatrix}$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors, respectively; $\boldsymbol{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights. The estimator $\widehat{\Psi}_{\text{VRS},n}$ imposes convexity, but allows for VRS. This is the VRS (DEA) estimator of Ψ proposed by Farrell (1957) and popularized by Banker et al. (1984). The CRS (DEA) estimator $\widehat{\Psi}_{\text{CRS},n}$ of Ψ is obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (4.2). FDH, VRS or CRS estimators of $\theta(x,y \mid \Psi)$, $\lambda(x,y \mid \Psi)$ and $\gamma(x,y \mid \Psi)$ defined in Section 1.3 are obtained by substituting $\widehat{\Psi}_{\text{FDH},n}$, $\widehat{\Psi}_{\text{VRS},n}$ or $\widehat{\Psi}_{\text{CRS},n}$ for Ψ in (3.3)–(3.5) (respectively). In the case of VRS estimators, this results in

$$\widehat{\theta}_{\text{VRS}}(x,y \mid \mathcal{S}_n) = \min_{\theta,\boldsymbol{\omega}} \left\{\theta \mid y \leq \boldsymbol{Y\omega},\ \theta x \geq \boldsymbol{X\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{4.3}$$

$$\widehat{\lambda}_{\text{VRS}}(x,y \mid \mathcal{S}_n) = \max_{\lambda,\boldsymbol{\omega}} \left\{\lambda \mid \lambda y \leq \boldsymbol{Y\omega},\ x \geq \boldsymbol{X\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\} \tag{4.4}$$

and

$$\widehat{\gamma}_{\text{VRS}}(x,y \mid \mathcal{S}_n) = \min_{\gamma,\boldsymbol{\omega}} \left\{\gamma \mid \gamma^{-1} y \leq \boldsymbol{Y\omega},\ \gamma x \geq \boldsymbol{X\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}. \tag{4.5}$$

The corresponding CRS estimators $\widehat{\theta}_{\text{CRS}}(x,y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{CRS}}(x,y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{CRS}}(x,y \mid \mathcal{S}_n)$ are obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (4.3)–(4.5). The estimators in (4.3)–(4.4) can be computed using linear programming methods, but the hyperbolic estimator in (4.5) is a non-linear program. Nonetheless, estimates can be computed easily using the numerical algorithm developed by Wilson (2011). Substituting $\widehat{\Psi}_{\text{FDH},n}$ into (3.3)–(3.5) (respectively) will yield FDH estimators

$\widehat{\theta}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$. However, this leads to integer programming problems, but the estimators can be computed using simple numerical methods.[10]

The statistical properties of these efficiency estimators are well-developed. Kneip et al. (1998) derive the rate of convergence of the input-oriented VRS estimator, while Kneip et al. (2008) derive its limiting distribution. Park et al. (2010) derive the rate of convergence of the input-oriented CRS estimator and establish its limiting distribution. Park et al. (2000) and Daouia et al. (2017) derive both the rate of convergence and limiting distribution of the input-oriented FDH estimator. These results extend trivially to the output orientation after straightforward (but perhaps tedious) changes in notation. Wheelock and Wilson (2008) extend these results to the hyperbolic FDH estimator, and Wilson (2011) extends the results to the hyperbolic DEA estimator.

Kneip et al. (2015) derive moment properties of both the input-oriented FDH, VRS and CRS estimators and also establish new central limit theorem (CLT) results for mean input-oriented efficiency after showing that the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold unless $(p + q) < 4$ in the CRS case, $(p + q) < 3$ in the VRS case, or unless $p + q < 2$ in the FDH case.[11] Kneip et al. (2016) use these CLT results to establish asymptotically normal test statistics for testing differences in mean efficiency across two groups, convexity versus non-convexity of Ψ, and CRS versus VRS (provided Ψ is weakly convex).[12] All of these results extend trivially to the output-oriented FDH, VRS and CRS estimators. These results could also be extended to the hyperbolic VRS and CRS estimators following Wilson (2011). The hyperbolic FDH estimator can be viewed as an input-oriented FDH estimator in a transformed space, hence moment results for the hyperbolic FDH estimator could also be extended trivially (but again, tediously) from the input-oriented FDH estimator. The new CLT results of Kneip et al. (2015) as well as the results from Kneip et al. (2016) on tests of differences in means, convexity versus non-convexity of Ψ, and CRS versus VRS carry over to the hyperbolic FDH estimator.

To summarize, in all cases, the FDH, VRS and CRS estimators are consistent, converge at rate n^{κ} (where $\kappa = 1/(p+q)$, $2/(p+q+1)$ or $2/(p+q)$ for the FDH, VRS and CRS estimators) and possess non-degenerate limiting distributions under the appropriate set of assumptions. In addition, the bias of each of the three estimators is of order $O\left(n^{-\kappa}\right)$. Bootstrap methods proposed by Kneip

[10]For details, see Kneip et al. (2015) and Wilson (2011).

[11]In other words, standard CLT results hold in the FDH case if and only if $p = 1$ and output is fixed and constant, or $q = 1$ and input is fixed and constant.

[12]If Ψ^{∂} is globally CRS, then the VRS estimator attains the faster convergence rate of the CRS estimator due to the Theorem 3.1 of Kneip et al. (2016).

et al. (2008, 2011) and Simar and Wilson (2011) provide consistent inference about $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ for a fixed point $(x, y) \in \Psi$. Kneip et al. (2015) provide new CLT results enabling inference about the expected values of these measures over the random variables (X, Y), and they show that the sample mean of these measures is a consistent estimator of population mean, with a bias term of order $O\left(n^{-\kappa}\right)$. In addition, if $\kappa \leq 1/2$, the bias term will "kill" the variance and the bias term need to be estimated using a jackknife method. Kneip et al. (2016) develop additional theoretical results permitting consistent tests of differences in mean efficiency across groups of producers, convexity of the production set and returns to scale.

Additional technical assumptions required for moment properties and central limit theorem results of means of FDH, VRS and CRS estimates, established by Kneip et al. (2015) and used below are given in the separate Appendix A.

1.5 Data and Variable Specification

The sample is an unbalanced panel including data from the balance sheets and income statements of commercial banks in China from 2007 to 2014. We have one year of data (2007) before the crisis, 3 years of data (2008–2010) during the crisis and the stimulus, and 4 years of data (2011–2014) after the stimulus. The main data source is BankScope database maintained by Bureau Van Dijk.

According to China Banking Regulation Committee, in 2014, China has 4 big state owned commercial banks, 12 joint stock commercial banks, 133 city commercial banks, 665 rural commercial banks, and 41 foreign banks. The total assets in 2014 were 150.95 trillion RMB. In 2014, our sample includes 4 big state owned commercial banks, 12 joint stock commercial banks, 58 city commercial banks, 18 rural commercial banks, and 32 foreign banks. The total assets of the sample are 108.90 trillion RMB, accounting for 72 percent of the total assets of the China's commercial banks in population. Therefore the sample is a good representation of commercial banks in population.[13]

Following Wheelock and Wilson (2018), $p = 3$ inputs and $q = 5$ outputs are defined. Specifically, the five output variables are defined as: consumer loans (Y_1), real estate loans (Y_2), business and other loans (Y_3), securities (Y_4), and off-balance sheet items (Y_5) consisting of net non-interest income. The three input variables are defined as: total funding (X_1), consisting of total

[13]The numbers of China's commercial banks across years in population and sample are provided in Tables C1–C2 of the separate Appendix C.

customer deposits, deposits from banks, repos and cash collateral, other deposits and short-term borrowings, senior debt maturing after 1 year, subordinated borrowing, other funding, total long-term funding, derivatives and trading liabilities; labor services, measured by the personnel expenses (X_2); and fixed asset (X_3). The first input quantity X_1 captures non-equity sources of investment funds for the bank. All RMB amounts are measured in constant 2010 RMB. The input-output specification is typical and standard, reflecting the basic production process of banks.

We assume that all commercial banks operate in the same production set Ψ defined by (3.1), and therefore they face the same frontier in the eight-dimensional input-output space. Banks may have different business plans and hence may operate in different areas of the production set Ψ. The model described in Section 1.3 is fully non-parametric, and hence quite flexible. The assumptions listed in Section 1.3 impose only minimal restrictions involving free-disposability, continuity, and some smoothness of the frontier, etc. Note that there is no assumption of convexity of Ψ, which is tested below in Section 1.6.

The flexibility of the non-parametric model specified in Section 1.3 comes with a price, however, in terms of the well-known "curse of dimensionality." The convergence rate of non-parametric efficiency estimators decreases with increasing inputs and outputs. The number of observations in each period that we consider ranges from 24 to 124. The effective parametric sample size defined by Wilson (2018) is then, in the worst case, $24^{\frac{2}{8}} \approx 2$ for FDH estimators, $24^{\frac{4}{9}} \approx 4$ for VRS estimators, and $24^{\frac{4}{8}} \approx 5$ for CRS estimators; and in the best case, $124^{\frac{2}{8}} \approx 3$ for FDH estimators, $124^{\frac{4}{9}} \approx 9$ for VRS estimators and $124^{\frac{4}{8}} \approx 11$ for CRS estimators. With the maximum sample size of 124 and the highest converge rate of $n^{\frac{2}{8}}$, nonparametric estimators should be expected to result in estimation error of order no better than that one would achieve with only 11 observations in a typical parametric estimator. Given the relatively small sample size and the high dimensions, it is not surprising that the estimated efficiency for many banks is equal to 1 and hence is not reliable.

To address this, the dimension reduction technique proposed by Wilson (2018) is applied. Considering the $(n \times p)$ and $(n \times q)$ matrices $\mathbf{X}$ and $\mathbf{Y}$ of observed non-negative inputs and outputs, we compute the $(n \times 1)$ vectors of principle components $X^* = \boldsymbol{X}\Lambda_x$ and $Y^* = \boldsymbol{Y}\Lambda_y$, where Λ_x and Λ_y are the $(p \times 1)$ and $(q \times 1)$ eigenvectors corresponding to the largest eigenvalues of $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$, respectively. The dimensions of both $\mathbf{X}$ and $\mathbf{Y}$ are then reduced to only one dimension. However, we need to examine R_x and R_y, which are the ratios of the largest eigenvalue of the moment matrices $\mathbf{X}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{Y}$ to the corresponding sums of the eigenvalues for these moment matrices. Wilson

(2018) mentions that R_x and R_y provide measures of how close the corresponding moment matrices are to rank-one, regardless of the joint distributions of inputs and outputs.

The eigensystem analysis on the full data yields $R_x \geq 98.25\%$ and $R_y \geq 95.87\%$ for all years.[14] It is clear that X^* and Y^* contain most of the independent information of $\boldsymbol{X}$ and $\boldsymbol{Y}$. Wilson (2018) shows that in many cases, but not in general, this dimension reduction method is advantageous in terms of reducing efficiency estimation error. In addition, dimension reduction could significantly increase the convergence rate of non-parametric efficiency estimators and lead to a more accurate estimation of efficiency. Now the convergence rates for FDH, VRS, and CRS are $n^{\frac{1}{2}}$, $n^{\frac{2}{3}}$ and n respectively.[15] The tradeoff is that a small amount of information may be lost, but the mean squared error is reduced. All estimation in the following is done using X^* and Y^*.

Figure 1.1 shows kernel density estimates of the log of total assets of Chinese banks over the years 2007, 2011, and 2014. The estimates displayed in Figure 1.1 illustrate the evolution of bank sizes over the period covered by our sample. The distribution of bank sizes has shifted rightward over time, suggesting that the Chinese banking sector is expanding over time. This density distribution is right-skewed, reflecting some banks have very large sizes. Similar phenomenon has been observed in the U.S. banking industry.[16]

Table 1.1 shows the summary statistics for year 2014. After removing the data with 0 values in any of the three inputs defined above, we have 124 observations in 2014, of which 32 observations are foreign banks. Comparing differences between the median and Q1 and between Q3 and the median for the input and output variables reveals that the marginal distributions are heavily skewed to the right, again reflecting the skewness of the distribution of bank sizes.

1.6 Empirical Results

1.6.1 Efficiency and Productivity Evolution

As a robustness check to the need for dimension reduction, we estimate the hyperbolic efficiency for each year first using full data with eight dimensions, and then using reduced data with only two dimensions. The FDH, VRS and CRS estimators are applied for both cases. For

[14]The details about eigensystem analysis of input and output moment matrices are shown in Table C3 of the separate Appendix C.

[15]The slowest rate is the root-n parametric rate after dimension reduction.

[16]See Wheelock and Wilson (2018).

each year, the FDH estimator produces more estimates equal to 1 than the VRS estimator, which produces more estimates equal to 1 than the CRS estimator.[17] This is expected since there are more restrictions for the CRS estimator than the VRS estimator, which has more restrictions than the FDH estimator. More importantly, when using the full data, the FDH estimator results in all the observations with estimates equal to one in any given year. The proportions for the VRS are between 61 percent and 92 percent, and for the CRS are between 31 percent and 54 percent. This is clear evidence of too many dimensions for the given sample size. With dimension reduction, when using either estimator for any given year, the number of observations with estimates equal to 1 is much smaller than that without dimension reduction. In addition, the numbers using the FDH estimator are at least 5 times those using the VRS estimator, suggesting that the production set Ψ may be non-convex. In addition to large values of R_x and R_y discussed in Section 1.5, we provide another piece of evidence that dimension reduction likely reduces estimation error relative to what would be obtained when using the full data without dimension reduction. Therefore, the principal components X^* and Y^* described in Section 1.5 are used for obtaining all the following results.

The next question is to determine which estimator we should use. As discussed in Section 1.3, in decreasing order of restrictions and rates of convergence lies the CRS, VRS, and FDH estimators. Kneip et al. (2016) and Daraio et al. (2018) develop a test to test the null hypothesis of convexity of the production set Ψ versus the alternative hypothesis that Ψ is non-convex. Two randomly split subsamples for a given year are needed for this test. The first subsample of size $n_1 = \lfloor n/2 \rfloor$ is used for computing VRS estimates, and the second subsample of size $n_2 = n - n_1$ is used for computing FDH estimates for a given sample size n. The test statistic given in equation (50) of Kneip et al. (2016) involves the difference of the means of these two sets of estimates, with generalized jackknife estimates of biases and corresponding sample variances, and is asymptotically normally distributed with mean zero and unit variance. The test is a one-sided test since under the null the two means should be roughly similar, but should diverge with increasing departures from the null. The statistic given in equation (50) of Kneip et al. (2016) is defined in terms of input-oriented estimators but extends trivially to output-oriented and hyperbolic estimators. The tests are one-sided and we define the statistics so that "large" positive values indicate rejection of the null hypothesis. While the results of Kneip et al. (2016) and Daraio et al. (2018) hold for a single split of the original sample,

[17] Additional details about the number of observations with estimates equal to one are given in the Table C5 of the separate Appendix C.

some have noticed that p-values resulting from the tests vary across different random splits of the original sample. Simar and Wilson (2020) develop a method that eliminates much of this ambiguity by repeating the random splits a large number of times and then use a bootstrap algorithm to exploit the information from the multiple sample-splits and enable inference-making from multiple sample-splits. Using the test developed by Simar and Wilson (2020), we use the principal components of X^* and Y^* to test the null hypothesis of convexity of the production set Ψ versus the alternative hypothesis that Ψ is non-convex. [18]

The results of the convexity tests for each year are shown in Table 1.2. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. Over 2007–2008, none of the six statistics reject convexity. It could be due to the fact that we have a small sample size for the first two years and it is hard to find the evidence of non-convexity. However, it is evident that over 2009–2014 convexity is rejected for all cases at the 10 percent significance level except two cases (output-oriented, 2010 and output-orientation, 2011), and rejected at the 1 percent significance level for most cases. Hence, the results in Table 1.2 provide strong evidence of the non-convexity of the production set Ψ.[19] When the production set is convex, both FDH and DEA estimators remain consistent. However, when the production set is non-convex, FDH estimators remain consistent, whereas DEA estimators do not. Consequently, the FDH estimators are applied for the remainder of the analysis. [20]

Table 1.3 presents summary statistics of the FDH technical efficiency estimates in the input, output, and hyperbolic orientations. To compare with the input-oriented and hyperbolic-oriented estimates, we report the statistics of the reciprocals of the output-oriented estimates. For each orientation, the closer the estimates are to 1, the more technically efficient the banks and the closer to the true frontier the banks. As might be expected, the hyperbolic estimates are more conservative on average, with mean efficiencies ranging from 0.9595 to 0.9951. By contrast, the means of the input-oriented estimates range from 0.9215 to 0.9905, while the means of the output-oriented estimates range from 0.9174 to 0.9890. These differences are due to the geometry of the efficiency measures as discussed by Wilson (2011). The mean efficiency in hyperbolic orientation first decreased from 2007 to 2008, then slightly increased from 2008 to 2009, and then continued declining until 2012, after

[18] We randomly split the samples for a given year by 1000 times and we bootstrap 1000 times.

[19] As a robustness check, we also consider convexity tests with unevenly split subsamples of the sample. The results of these convexity tests are shown in Table C6 of the separate Appendix C. The results in Table C6 are consistent with that in Table 1.2.

[20] We also use the KS-statistics developed in Simar and Wilson (2020). Even though KS-statistics are less likely to reject convexity, it does not mean the null of convexity is true. As mentioned previously, when the production set is non-convex, FDH estimators still remain consistent, whereas DEA estimators do not. Therefore, the FDH estimators are applied for the remainder of the analysis.

which the means rose again until 2014. The pattern of mean efficiency in the output orientation appears to be the same, while the pattern in the input orientation is a little mixed. Mean efficiency in the input orientation first declined from 2007 to 2009, then increased from 2009 to 2010, and then declined until 2012, after which the mean rose and declined alternately from 2013 to 2014.

We use the test described by Kneip et al. (2016, Section 3.1.1) to test for significant differences between the means reported in Table 1.3 from one year to the next, as well as from the first year to the last year. As discussed in Kneip et al. (2015, 2016), even with the reduced dimensionality so that $p + q = 2$, the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold for means of FDH efficiency estimates. As with the convexity test discussed above, the test statistic given by equation (18) of Kneip et al. (2016) involves not only the difference in sample means of efficiency estimates in a pair of years, but also the corresponding difference in generalized jackknife estimates of bias. The test extends trivially to the output-orientation, and the hyperbolic orientation. In each case, the statistic used here is defined so that a positive value indicates that efficiency increases from year 1 to year 2, while a negative value indicates that efficiency decreases from year 1 to year 2.[21] As shown by Kneip et al. (2016), the test statistics are asymptotically normal with zero mean and unit variance. Since our data is unbalanced panel, there may exist time correlation, which violates the independent assumption of the test for differences of mean efficiency. The technical details dealing with time correlations are given in the separate Appendix B Section B.1.

Table 1.4 gives the results of the tests of significant differences in mean efficiency over time. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. The tests provide clear evidence that the mean efficiency decreased from 2007 to 2008. As Table 1.3 shows that mean efficiency in output orientation decreased from 0.9890 to 0.9581 over 2007–2008, it suggests that given the same input, Chinese commercial banks averagely produced about 3 percent less output in 2008 compared to that in 2007. This possibly reflects the negative effect of the crisis. However, mean efficiency increased from 2008 to 2009 without any significant statistic. The result from 2009 to 2010 is mixed, where one statistic is positive but insignificant and the other two statistics are negative with only one significant. This is not surprising since the crisis and stimulus happened at the same time, which disrupted the Chinese banking system. Mean efficiency declined significantly from 2010 to 2011. This decline could be the reversal effect of the stimulus. Mean efficiency then

[21] Consequently, the statistic we use for the output orientation is the negative of the statistic appearing in equation (18) of Kneip et al. (2016).

increased from 2011 to 2013 with only two insignificant statistics and then decreased from 2013 to 2014 with only one insignificant statistics. Overall, from 2007 to 2014, mean efficiency declined significantly.

Taken together, we find strong evidence that mean efficiency decreased at the start of the crisis. This suggests that banks were on average farther away from the frontier in 2008 than in 2007. Moreover, the crisis and the stimulus heavily disrupted Chinese banking industry, making the change of mean efficiency from 2008 to 2010 unclear. However, it is apparent that there was a reversal effect from the stimulus. Banks moved far away from the frontier from 2010 to 2011. Even though banks seemed to recover from 2011 to 2013, overall, banks actually lied much farther away from the frontier in 2014 compared with 2007.

In order to measure productivity, note that with the dimension reduction to $(p + q) = 2$ dimensions using the principal components X_i^*, Y_i^* as described in Section 1.5, productivity can be measured by Y_i^*/X_i^* for bank i. Summary statistics for this measure is displayed in Table 1.5. Mean productivity first increased from 2007 to 2009, then decreased continuously from 2009 to 2013, after which, it rose again from 2013 to 2014. Since productivity is measured by a simple ratio that does not involve estimators of efficiency, standard CLT results can be used to test for significant changes in means over time. However, we need to deal with time correlation, see the separate Appendix B Section B.2 for technical details. The results of these tests are shown in Table 1.6. Cells in columns 7 are shaded whenever p-value is less than 0.10. Note that there are only three one-year intervals in which the change of mean productivity is significant at the 10 percent level. Mean productivity significantly declined from 2009 to 2010 and from 2010 to 2011, and significantly increased from 2013 to 2014. These results show that the productivity did not change at the start of the crisis. However, it is evident that during or just after the stimulus, the mean productivity declined from 2009 to 2011 and banks finally recovered from 2013 to 2014. Overall, it appears that there was a significant decrease in mean productivity from 2007 to 2014.

The results presented so far provide clear evidence of changes in mean technical efficiency and productivity over the years represented in the sample. To gain further insight, we test whether the frontiers change over time. This involves the test of "separability" developed by Daraio et al. (2018), in which time is treated as a binary "environmental" variable. We examine it using pairs of years 2007–2008, . . . , 2013–2014 as well as 2007–2014.

Implementation of the separability test of Daraio et al. (2018) involves pooling the data for two periods and then randomly shuffling the observations using the randomization algorithm presented by Daraio et al.. Then the pooled, randomly shuffled observations are split into two subsamples of equal size (or, if the combined number of observations is odd, one subsample will have one more observation than the other). Using the first subsample, efficiency is estimated as usual for each observation, ignoring which period a particular observation comes from, and the sample mean of the efficiency estimate is computed. The second subsample is split into the set of observations from period 1 and the set of observations from period 2. Efficiency is estimated for the period 1 observations using only the observations from period 1, while efficiency for the period 2 observations is estimated using only those observations from period 2. Then the sample mean of these two sets of efficiency estimates from the two sub-subsamples (of the second subsample) is computed. The resulting test statistic involves differences in the two subsample means as well as differences in the corresponding generalized jackknife estimates of bias. See Daraio et al. (2018) for discussion and details.

Results of the separability tests are shown in Table 1.7. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. From 2007 to 2008 and from 2009 to 2010, none of the six statistics are significant, showing that technology did not change over these two periods. Two statistics from the period 2008–2009 are significant at the 10 percent level, while the remaining statistic is significant at the 1 percent level. Therefore, the technology changed from 2008 to 2009, but, the evidence is not strong. It is evident that technology changed for the remaining five periods from 2010 to 2014, even though during this period one statistic (hyperbolic-oriented, period 2010–2011) is only significant at the 10 percent level. Overall, for the entire period 2007–2014, separability is rejected with a p-value less than .05. The separability tests provide clear evidence of changes in the technology during the crisis and the stimulus (2008–2009). They also suggest changes in the technology after the stimulus (2010–2014), as well as over the full period 2007–2014.

In order to learn something about the *direction* in which technology may have shifted, we use new results from Simar and Wilson (2018) who provide CLT results for components of productivity changed measured by Malmquist indices. Simar and Wilson define the Malmquist index in terms of hyperbolic distances, and then consider various decompositions that can be used to identify components of productivity change. In particular, let Ψ^t represent the production set at time $t \in \{1,\ 2\}$ and let $Z_i^t = (X_i^t, Y_i^t)$ denote the i-th firm's observed input-output pair at time t.

Then technical change relative to firm i's position at times 1 and 2 is measured by

$$\mathcal{T}_i = \left[\frac{\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^2 \mid \Psi^2)} \times \frac{\gamma(Z_i^1 \mid \Psi^1)}{\gamma(Z_i^1 \mid \Psi^2)}\right]^{1/2}. \tag{6.1}$$

This is the hyperbolic analog of the output-oriented technical-change index that appears in the decompositions of Ray and Desli (1997), Gilbert and Wilson (1998), Simar and Wilson (1998) and Wheelock and Wilson (1999). The first ratio inside the brackets in (6.1) measures technical change relative to firm i's position at time 2, while the second ratio measures technical change relative to the firm's position at time 1. The measure $\mathcal{T}_i$ is the geometric mean of these two ratios. Values greater than 1 indicate an upward shift in the technology, while values less than 1 indicate a downward shift (a value of 1 indicates no change from time 1 to time 2).

Estimates $\widehat{\mathcal{T}}_i$ are obtained by substituting the hyperbolic FDH estimator for each term in (6.1). Simar and Wilson (2018) develop CLT results for geometric means $\widehat{T}^{1,2}$ of $\mathcal{T}_i$ over firms $i = 1, \ldots, n$, for periods 1 and 2, and these results can be used to test significant differences of the geometric means from 1. Table 1.8 shows the results of tests of technology change for each one-year interval as well as for the entire period 2007–2014. Cells in columns 7 are shaded whenever p-value is less than 0.10. The geometric mean $\widehat{T}^{1,2}$ is less than 1 for each one-year interval from 2009 to 2013. This suggests continuing downward shifts of the technology from 2009 to 2013 and upward shifts for the remaining periods. However, the p-value is well less than 0.01 only for 2012–2013 and 2013–2014. Consequently, the data only provide evidence that the technology shifted downward from 2012 to 2013 and then shifted upward from 2013 to 2014. Overall, the technology shifted upward over the full period 2007–2014.

1.6.2 Big Versus Small

In China, big banks (especially the four big state owned commercial banks) take some government orders explicitly and implicitly and thus they face more political pressures than small banks. Therefore, big banks could not be more efficient and productive than small banks, which are often considered to be more market-based. Our test could also be used to answer this question.

We split our sample into two subsamples in terms of the median total assets for each year. The big banks are then defined as those with total assets larger than the median total assets for each year, and the remaining are defined as small banks. Table 1.9 shows the results of tests on the difference of technical efficiency between big and small banks for each year. Cells in columns 5, 7 and 9 are shaded whenever p-value is less than 0.10. Note that all of the statistics except one (output-oriented, 2009) are negative, suggesting that small banks were more technically inefficient. Moreover, out of 24 total statistics (three orientations and eight years), only one statistic (output-oriented, 2009) is insignificant, two statistics (output-oriented, 2008; hyperbolic-oriented 2008) are only significant at the 10 percent level and all the remaining statistics are significant at the 5 percent level.

Table 1.10 shows the results of tests on the difference of productivity. Cells in columns 7 are shaded whenever p-value is less than 0.10. The statistics are positive only for each year over the period 2007–2010, suggesting that small banks were more productive than big banks at the early periods, while big banks performed better in terms of productivity after 2010. However, the p-values are significant at the 5 percent level only for 2012, 2013 and 2014. There is no evidence in our sample showing that big banks had lower productivity. Big banks actually had higher productivity than small banks in 2012, 2013, 2014.

Our results refute criticisms of the low efficiency and low productivity of big banks. As a robustness check, we also consider different definitions of big banks and small banks based on different quantiles of total assets for each year. The results of these tests are shown in Tables C7–C10 of the separate Appendix C. The results are quite consistent with our baseline estimates.

1.6.3 Domestic Versus Foreign

Foreign banks are typically considered as having more advanced technology and more experienced managers. Therefore they are usually more efficient and productive than domestic banks in China. Our tests could also be used to examine this outcome.

Table 1.11 shows the results of tests of the difference in mean technical efficiency between domestic and foreign banks. Cells in columns 5, 7 and 9 are shaded whenever p-value is less than 0.10. In the first three years 2007–2009, all p-values are less than 0.05. However, the sign of the statistics alternates, first negative for 2007, then positive for 2008, and negative again for 2009. This suggests that foreign banks only had higher technical efficiency than domestic banks in 2008. From 2010 to 2014, most statistics are negative and seven statistics are significant at the 10 percent level. In contrast, two statistics (input-oriented, 2011; input-oriented, 2013) are positive and insignificant. Combining together, the data show that in general domestic banks performed better in terms of technical efficiency than foreign banks over 2007–2014, while foreign banks only performed better in 2008.

Table 1.12 provides the results of tests of the difference in productivity between domestic banks and foreign banks. Cells in columns 7 are shaded whenever p-value is less than 0.10. The statistics are positive only for the first three years, of which only the one in 2008 is significant at the 5 percent level. From 2010 to 2014, all statistics are negative and most are also significant at the 5 percent level (except the one in 2010). This result suggests that foreign banks had higher productivity only in 2008. However, domestic banks were more productive than foreign banks over the period 2011–2014.

Our results refute criticisms of the low efficiency and low productivity of domestic banks. However, the low efficiency and productivity of foreign banks could be due to more regulations compared with domestic banks.

1.7 Summary and Conclusions

Among studies that use either FDH or DEA estimators to estimate efficiency and benchmark the performances of firms, the vast majority use VRS (DEA) estimators which impose convexity on the production set. The test of convexity versus non-convexity of Ψ developed by Kneip et al. (2016) allows researchers to let the data tell them whether DEA estimators are appropriate in a given setting. Here, in the context of Chinese commercial banks, convexity is strongly rejected. This is consistent with the results of Wheelock and Wilson (2012, 2018), who find evidence of increasing returns to scale among even the largest banks operating in the U.S.

Because we reject convexity of the production set, we use FDH estimators which remain consistent when Ψ is not convex, whereas DEA estimators do not. We exploit collinearity in the data to reduce inputs and outputs to their first principle components, resulting in a two-dimensional problem. Results from Wilson (2018) indicate that this substantially reduces mean square error of efficiency estimates. Moreover, the simulation evidence provided by Wilson (2018) suggests that when production sets are convex, FDH estimates often have less mean square error than DEA estimators after dimension reduction.

By rigorously comparing estimates and testing differences across the years represented in our data, we find that technical efficiency declined at the start of the global financial crisis (2007–2008), and also after the China stimulus (2010–2011). However, technical efficiency finally recovered from 2011 to 2013 but declined again from 2013 to 2014. Overall, banks lied much farther away from the frontier in 2014 compared to 2007. We find similar results for productivity. Productivity declined during or just after the stimulus (2009–2011), but recovered from 2013 to 2014. Overall, there was a decrease in mean productivity from 2007 to 2014. We also find that the frontier shifted downward from 2012 to 2013 and shifted upward from 2013 to 2014. Over the period 2007–2014, technology shifted upward. Our results show that in general big banks were more efficient and productive than small banks. Domestic banks had higher efficiency and productivity than foreign banks over this period except in 2008.

In terms of policy implications, recently the higher efficiency and productivity of big banks compared to small banks suggests that there is a benefit for big banks to become even larger since they could produce more output given the same input. If the government restricts the size of big banks in case of "Too Big To Fail", it will also restrict the total output of society given the same input, and hence reduce the total welfare of the society. The higher technical efficiency of domestic banks compared to foreign banks suggests that the domestic banks in China should be more confident about their efficiency. These banks could operate in the international market and compete with foreign banks in this market.

Table 1.1: Summary Statistics for Year 2014 in Billions of 2010 RMB

Variable	Min	Q1	Median	Mean	Q3	Max
Total Funding (X_1)	1.0480×10^{-01}	$4.7570 \times 10^{+01}$	$9.4980 \times 10^{+01}$	$7.9430 \times 10^{+02}$	$2.9780 \times 10^{+02}$	$1.6240 \times 10^{+04}$
Labor Services (X_2)	3.8550×10^{-03}	2.5850×10^{-01}	6.8050×10^{-01}	$4.7010 \times 10^{+00}$	$1.5160 \times 10^{+00}$	$9.8140 \times 10^{+01}$
Fixed Asset (X_3)	8.1210×10^{-04}	2.0960×10^{-01}	8.4810×10^{-01}	$7.3030 \times 10^{+00}$	$1.7780 \times 10^{+00}$	$1.7460 \times 10^{+02}$
Consumer Loans (Y_1)	$0.0000 \times 10^{+00}$	7.2740×10^{-01}	$5.0890 \times 10^{+00}$	$5.0610 \times 10^{+01}$	$1.3700 \times 10^{+01}$	$8.7010 \times 10^{+02}$
Real Estate Loans (Y_2)	$0.0000 \times 10^{+00}$	$0.0000 \times 10^{+00}$	$1.0500 \times 10^{+00}$	$7.2320 \times 10^{+01}$	$1.1960 \times 10^{+01}$	$1.9910 \times 10^{+03}$
Business Loans (Y_3)	$0.0000 \times 10^{+00}$	$2.0460 \times 10^{+01}$	$3.9970 \times 10^{+01}$	$3.1960 \times 10^{+02}$	$9.4740 \times 10^{+01}$	$6.9760 \times 10^{+03}$
Securities (Y_4)	$0.0000 \times 10^{+00}$	$8.4100 \times 10^{+00}$	$2.3790 \times 10^{+01}$	$2.0190 \times 10^{+02}$	$9.3590 \times 10^{+01}$	$3.9300 \times 10^{+03}$
Off-balance Sheet Items (Y_5)	$0.0000 \times 10^{+00}$	1.3790×10^{-01}	4.8440×10^{-01}	$6.0180 \times 10^{+00}$	$1.5930 \times 10^{+00}$	$1.2250 \times 10^{+02}$
First Principle Component of Inputs (X^*)	6.5380×10^{-02}	$2.8300 \times 10^{+01}$	$5.6240 \times 10^{+01}$	$4.6930 \times 10^{+02}$	$1.7430 \times 10^{+02}$	$9.6100 \times 10^{+03}$
First Principle Component of Outputs (Y^*)	2.2200×10^{-02}	$1.6450 \times 10^{+01}$	$3.4090 \times 10^{+01}$	$2.9300 \times 10^{+02}$	$1.1280 \times 10^{+02}$	$6.1750 \times 10^{+03}$

Table 1.2: Results of Convexity Tests, Average over 1000 splits, Bootstrap 1000 times (Even Split, with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	-4.5343	0.1370	3.8204	0.4500	-4.3650	0.3310
2008	-4.3240	0.2820	3.0105	0.4230	-4.8880	0.1920
2009	-6.1165	0.0130	6.6732	0.0030	-7.0909	0.0020
2010	-8.1269	0.0040	6.1395	0.1380	-7.1028	0.0310
2011	-5.5079	0.0070	1.6444	0.3830	-4.1084	0.0770
2012	-6.0398	0.0050	4.0242	0.0520	-5.7870	0.0060
2013	-7.2708	0.0000	5.4913	0.0060	-10.2836	0.0010
2014	-5.2326	0.0040	4.7402	0.0030	-6.1213	0.0010

Table 1.3: Summary Statistics for FDH Technical Efficiency Estimates (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
			— **Input Orientation** —			
2007	0.7727	1.0000	1.0000	0.9905	1.0000	1.0000
2008	0.3429	0.9876	1.0000	0.9616	1.0000	1.0000
2009	0.6075	0.9819	1.0000	0.9596	1.0000	1.0000
2010	0.7320	0.9873	1.0000	0.9743	1.0000	1.0000
2011	0.5644	0.9260	1.0000	0.9427	1.0000	1.0000
2012	0.3427	0.8879	1.0000	0.9232	1.0000	1.0000
2013	0.6089	0.8986	1.0000	0.9363	1.0000	1.0000
2014	0.2600	0.8716	0.9978	0.9215	1.0000	1.0000
			— **Output Orientation** —			
2007	0.7367	1.0000	1.0000	0.9890	1.0000	1.0000
2008	0.4934	0.9682	1.0000	0.9581	1.0000	1.0000
2009	0.7062	0.9810	1.0000	0.9756	1.0000	1.0000
2010	0.6951	0.9608	1.0000	0.9624	1.0000	1.0000
2011	0.1630	0.8848	1.0000	0.9219	1.0000	1.0000
2012	0.4068	0.8414	1.0000	0.9174	1.0000	1.0000
2013	0.5505	0.8725	1.0000	0.9285	1.0000	1.0000
2014	0.6376	0.8887	0.9951	0.9314	1.0000	1.0000
			— **Hyperbolic Orientation** —			
2007	0.8831	1.0000	1.0000	0.9951	1.0000	1.0000
2008	0.8205	0.9903	1.0000	0.9826	1.0000	1.0000
2009	0.8762	0.9897	1.0000	0.9865	1.0000	1.0000
2010	0.8071	0.9913	1.0000	0.9836	1.0000	1.0000
2011	0.7233	0.9435	1.0000	0.9601	1.0000	1.0000
2012	0.7343	0.9330	1.0000	0.9595	1.0000	1.0000
2013	0.6798	0.9429	1.0000	0.9622	1.0000	1.0000
2014	0.7266	0.9451	0.9994	0.9623	1.0000	1.0000

NOTE: Statistics for the reciprocals of the output efficiency estimates are given to facilitate comparison with the input-oriented and hyperbolic estimates.

Table 1.4: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007–2008	−3.2015	1.37×10^{-03}	−4.2474	2.16×10^{-05}	−4.0501	5.12×10^{-05}
2008–2009	1.0134	3.11×10^{-01}	1.3244	1.85×10^{-01}	0.8801	3.79×10^{-01}
2009–2010	0.6543	5.13×10^{-01}	−0.2108	8.33×10^{-01}	−2.3697	1.78×10^{-02}
2010–2011	−2.2646	2.35×10^{-02}	−2.6639	7.72×10^{-03}	−2.3636	1.81×10^{-02}
2011–2012	0.6169	5.37×10^{-01}	2.3769	1.75×10^{-02}	1.8283	6.75×10^{-02}
2012–2013	1.2657	2.06×10^{-01}	3.1364	1.71×10^{-03}	3.8639	1.12×10^{-04}
2013–2014	−1.8348	6.65×10^{-02}	−2.1342	3.28×10^{-02}	−0.5801	5.62×10^{-01}
2007–2014	−11.4955	1.39×10^{-30}	−13.4240	4.37×10^{-41}	−12.0046	3.36×10^{-33}

NOTE: The numerator of statistics for each period is the difference of estimated mean efficiency of the second year minus the first year.

Table 1.5: Summary Statistics for Productivity (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
2007	0.4586	0.5898	0.6352	0.6575	0.6781	1.3950
2008	0.3066	0.6225	0.6509	0.6640	0.6923	1.0507
2009	0.4831	0.6016	0.6410	0.6822	0.6943	2.0020
2010	0.4756	0.5808	0.6333	0.6259	0.6684	0.7827
2011	0.1085	0.5131	0.5884	0.5665	0.6320	0.7430
2012	0.1621	0.5089	0.5842	0.5622	0.6327	0.7637
2013	0.1502	0.5121	0.5788	0.5556	0.6212	0.7166
2014	0.2019	0.5292	0.5925	0.5759	0.6425	1.0687

NOTE: Productivity for bank i is defined as Y_i^*/X_i^*.

Table 1.6: Tests of Differences in Means for Productivity Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	Mean1	Mean2	Statistic	p-value
2007–2008	24	41	0.6575	0.6640	0.2201	8.26×10^{-01}
2008–2009	41	45	0.6640	0.6822	0.7273	4.67×10^{-01}
2009–2010	45	65	0.6822	0.6259	−1.7409	8.17×10^{-02}
2010–2011	65	82	0.6259	0.5665	−4.9801	6.36×10^{-07}
2011–2012	82	108	0.5665	0.5622	−0.4428	6.58×10^{-01}
2012–2013	108	123	0.5622	0.5556	−0.8445	3.98×10^{-01}
2013–2014	123	124	0.5556	0.5759	2.4747	1.33×10^{-02}
2007–2014	24	124	0.6575	0.5759	−2.2556	2.41×10^{-02}

NOTE: The numerator of statistics for each period is the difference of estimated mean productivity of the second year minus the first year.

Table 1.7: Tests for Separability with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007–2008	-0.9241	8.22×10^{-01}	-2.2895	9.89×10^{-01}	-0.2231	5.88×10^{-01}
2008–2009	2.7892	2.64×10^{-03}	1.2855	9.93×10^{-02}	1.4563	7.26×10^{-02}
2009–2010	0.3538	3.62×10^{-01}	-0.5795	7.19×10^{-01}	-0.2364	5.93×10^{-01}
2010–2011	2.8456	2.22×10^{-03}	1.8413	3.28×10^{-02}	1.6287	5.17×10^{-02}
2011–2012	5.9338	1.48×10^{-09}	4.1313	1.80×10^{-05}	5.3978	3.37×10^{-08}
2012–2013	3.7983	7.28×10^{-05}	4.9201	4.32×10^{-07}	4.3795	5.95×10^{-06}
2013–2014	5.0238	2.53×10^{-07}	5.1497	1.30×10^{-07}	4.4403	4.49×10^{-06}
2007–2014	3.3405	4.18×10^{-04}	4.0621	2.43×10^{-05}	3.6674	1.23×10^{-04}

NOTE: The numerator of the statistics is the difference of the conditional mean estimates minus the unconditional mean estimates.

Table 1.8: Tests for Technology Change with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	n	$\widehat{T}^{1,2}$	Var	p-value
2007–2008	24	41	23	1.0080	0.0033	7.55×10^{-01}
2008–2009	41	45	33	1.0105	0.0085	8.05×10^{-01}
2009–2010	45	65	42	0.9925	0.0041	3.87×10^{-01}
2010–2011	65	82	60	0.9771	0.0033	3.73×10^{-01}
2011–2012	82	108	79	0.9962	0.0047	1.99×10^{-01}
2012–2013	108	123	101	0.9910	0.0031	8.60×10^{-04}
2013–2014	123	124	111	1.0138	0.0048	6.78×10^{-03}
2007–2014	24	124	22	1.0437	0.0091	9.24×10^{-03}

NOTE: For each period, the number of banks in the first year is n_1, while the number of banks in the second year is n_2. The number of banks existing in both years is n. Mean of the technology ratio $\widehat{T}^{1,2}$ is greater than 1 if and only if the technology shifts upward.

Table 1.9: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	12	12	−2.1830	2.90×10^{-02}	−2.1830	2.90×10^{-02}	−2.1830	2.90×10^{-02}
2008	21	20	−2.8359	4.57×10^{-03}	−1.6860	9.18×10^{-02}	−1.8349	6.65×10^{-02}
2009	23	22	−4.9276	8.33×10^{-07}	0.5838	5.59×10^{-01}	−3.5620	3.68×10^{-04}
2010	33	32	−2.9834	2.85×10^{-03}	−2.2602	2.38×10^{-02}	−2.3524	1.87×10^{-02}
2011	41	41	−5.1218	3.03×10^{-07}	−2.9776	2.90×10^{-03}	−5.1261	2.96×10^{-07}
2012	54	54	−4.2521	2.12×10^{-05}	−4.3841	1.16×10^{-05}	−3.6539	2.58×10^{-04}
2013	62	61	−3.0905	2.00×10^{-03}	−2.2699	2.32×10^{-02}	−3.3622	7.73×10^{-04}
2014	62	62	−5.7251	1.03×10^{-08}	−5.6388	1.71×10^{-08}	−6.3183	2.64×10^{-10}

NOTE: We split the total observations of each year into two even subsamples by the median total assets in that year. The number of big banks is n_1, while the number of small banks is n_2. The numerator of statistics is the difference of estimated mean efficiency of small banks minus big banks.

Table 1.10: Tests of Differences in Means for Productivity Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2007	12	12	0.6301	0.6848	0.7758	4.38×10^{-01}
2008	21	20	0.6356	0.6938	1.5655	1.17×10^{-01}
2009	23	22	0.6394	0.7270	1.3274	1.84×10^{-01}
2010	33	32	0.6188	0.6333	0.8462	3.97×10^{-01}
2011	41	41	0.5838	0.5491	−1.5846	1.13×10^{-01}
2012	54	54	0.5839	0.5405	−2.3468	1.89×10^{-02}
2013	62	61	0.5819	0.5288	−3.0986	1.94×10^{-03}
2014	62	62	0.6025	0.5493	−2.8746	4.05×10^{-03}

NOTE: We split the total observations of each year into two even subsamples by the median total assets in that year. The number of big banks is n_1, while the number of small banks is n_2. The numerator of statistics is the difference of estimated mean productivity of small banks minus big banks.

Table 1.11: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	18	6	−2.5210	1.17×10^{-02}	−2.5210	1.17×10^{-02}	−2.5210	1.17×10^{-02}
2008	27	14	4.9225	8.54×10^{-07}	4.0976	4.17×10^{-05}	4.5152	6.33×10^{-06}
2009	29	16	−2.2062	2.74×10^{-02}	−2.0177	4.36×10^{-02}	−2.5641	1.03×10^{-02}
2010	46	19	−1.0914	2.75×10^{-01}	−1.7238	8.47×10^{-02}	−1.0063	3.14×10^{-01}
2011	62	20	0.6964	4.86×10^{-01}	−2.2187	2.65×10^{-02}	−0.4654	6.42×10^{-01}
2012	81	27	−2.0105	4.44×10^{-02}	−2.9061	3.66×10^{-03}	−1.5096	1.31×10^{-01}
2013	93	30	0.2699	7.87×10^{-01}	−0.7012	4.83×10^{-01}	−0.9955	3.20×10^{-01}
2014	92	32	−2.9846	2.84×10^{-03}	−1.8352	6.65×10^{-02}	−2.6650	7.70×10^{-03}

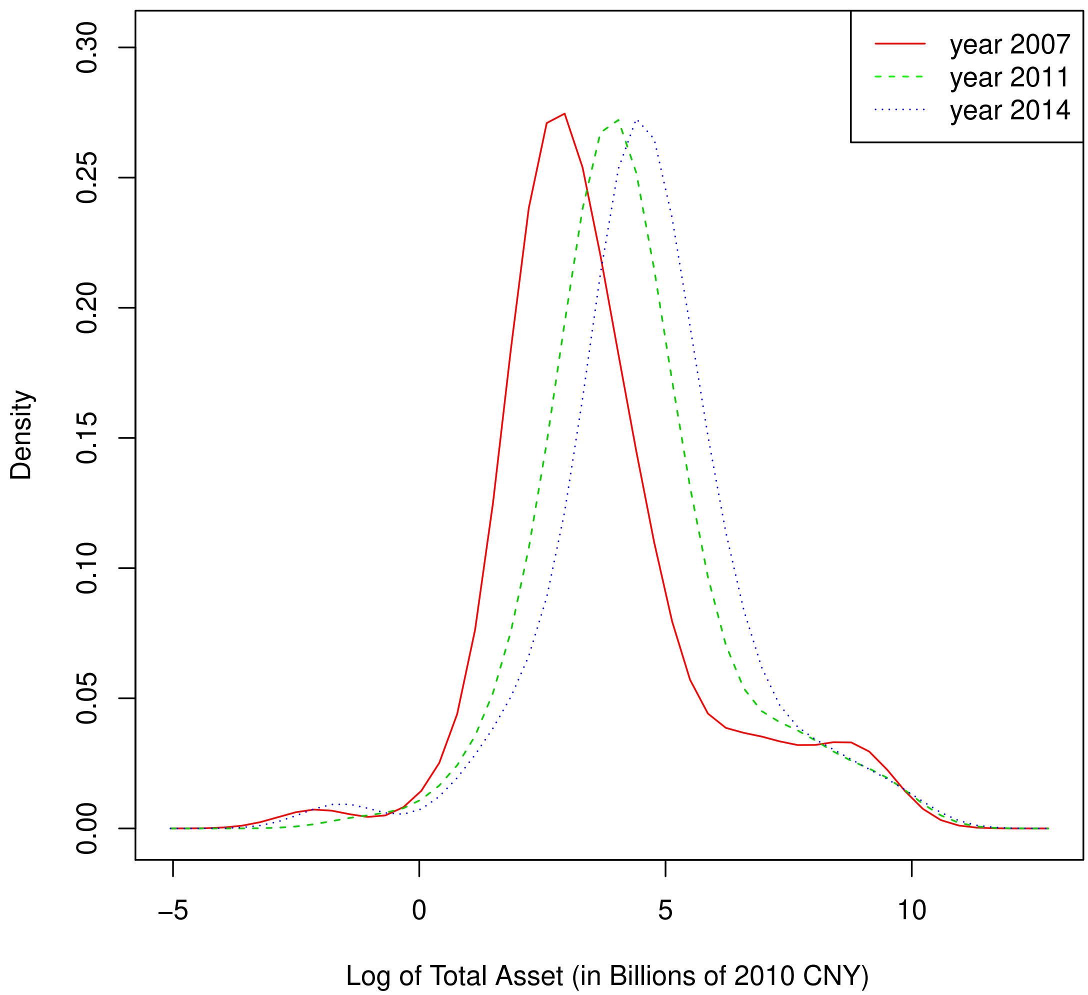
NOTE: The number of domestic banks is n_1, while the number of foreign banks is n_2. The numerator of statistics is the difference of estimated mean efficiency of foreign banks minus domestic banks.

Table 1.12: Tests of Differences in Means for Productivity Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2007	18	6	0.6220	0.7640	1.0614	2.89×10^{-01}
2008	27	14	0.6272	0.7349	2.5058	1.22×10^{-02}
2009	29	16	0.6374	0.7634	1.4312	1.52×10^{-01}
2010	46	19	0.6276	0.6219	-0.2598	7.95×10^{-01}
2011	62	20	0.5821	0.5180	-2.0464	4.07×10^{-02}
2012	81	27	0.5892	0.4814	-4.4668	7.94×10^{-06}
2013	93	30	0.5806	0.4782	-4.3776	1.20×10^{-05}
2014	92	32	0.6040	0.4951	-3.8912	9.98×10^{-05}

NOTE: The number of domestic banks is n_1, while the number of foreign banks is n_2. The numerator of statistics is the difference of estimated mean productivity of foreign banks minus domestic banks.

Figure 1.1: Density of (log) Total Assets of China's Commercial Banks in 2007, 2011 and 2014



NOTE: Solid red line shows density for 2007; dashed green line shows density for 2011; dotted blue line shows density for 2014.

Chapter 2

Evidence from Shadow Price of Equity on “Too-Big-to-Fail” Banks

2.1 Introduction

The global financial crisis of 2007–2012 may have started in the U.S. banking sector and was the worst U.S. economic disaster since the 1929 Great Depression. After Lehman’s failure, the U.S. Congress passed the Troubled Asset Relief Program (TARP) to funnel hundreds of billions of dollars to support banks in a period of extraordinary financial turbulence. In addition, the Federal Reserve Board lent hundreds of billions of dollars to the banks through a series of newly created special lending facilities.

As a result of this recent global financial crisis, “Too-Big-To-Fail” (TBTF) is now a virtually official “policy”. TBTF “policy” means that since some banks are so big and so important that their failure would be disastrous to the whole economic system, they must be protected by the government whenever they face potential failure. Access to the federal government’s safety net allows TBTF banks to operate with a lower funding cost relative to non-TBTF banks since the public believe that the government would protect the TBTF banks again whenever there is another crisis, hence their uninsured creditors (e.g., the equity investors) do not charge as high a price for the use of their funds as they would in the absence of this perception. Therefore, there may exist an implicit subsidy for TBTF banks. On the other hand, the Dodd-Frank Wall Street Reform and Consumer Protection

Act of 2010 were intended to remove TBTF "policy" by establishing a formal process for resolving failures of large financial institutions, as well as by imposing a tighter financial regulatory regime.

All in all, TBTF has become a heated topic after the recent global financial crisis. However, until now there lack enough evidence on the existence of TBTF banks. If TBTF banks do exist, then it is rational for the government to remove the implicit subsidy. Moreover, we need to know whether the implicit subsidy has decreased after the tighter regulations. In other words, we need to know whether the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 are effective in removing the implicit subsidy for TBTF.

There exists some literature providing evidence on the existence of TBTF banks. Santos and Santos (2014) use information from bonds issued to find that the largest banks have a relatively larger cost advantage over their smaller peers, compared with the largest nonbanks over their smaller peers. This difference supports investors' beliefs of TBTF banks. Brewer and Jagtiani (2013), using data from the merger boom of 1991–2004, find that banking organizations were willing to pay an added premium for mergers that would put them over the asset sizes that are commonly viewed as the thresholds for being TBTF. Ueda and Weder di Mauro (2013) use the level of government support embedded in the credit rating and its impact on the overall credit rating to provide estimates for the structural subsidy values. They found a significant funding cost advantage for Systemically Important Financial Institutions (SIFIs), about 60 basis points as of the end of 2007, before the crisis and 80 basis points by the end of 2009. Baker and Mcarthur (2009) use data from the Federal Deposit Insurance Corporation (FDIC) on the relative cost of funds for TBTF banks and other banks, before and after the crisis, to quantify the value of the government protection provided by the TBTF "policy". They find that the spread between the average cost of funds for smaller banks and the cost of funds for institutions with assets in excess of $100 billion averaged 0.29 percentage points in the period from the first quarter of 2000 through the fourth quarter of 2007, the last quarter before the collapse of Bear Steams. In the period from the fourth quarter of 2008 through the second quarter of 2009, after the government bailouts had largely established TBTF, the gap had widened to an average of 0.78 percentage points. If this gap is attributable to the TBTF "policy", it implies a substantial taxpayer subsidy for the TBTF banks. As a conclusion, previous research uses different methods to show that TBTF banks indeed do exist in U.S. banking sector and they enjoy the benefit of the implicit guarantee from government.

This paper estimates shadow price of equity for U.S. commercial banks over 2001–2018 using nonparametric estimators of the underlying cost function and then tests the existence of "Too-Big-to-Fail" (TBTF) banks. Since TBTF banks are believed to be implicitly protected by the government, they are considered safer and are willing to pay a lower price of equity than non-TBTF banks. If a bank is public, the price of equity can be derived from the bank's price of stock. However, most U.S. commercial banks are not public, and the price of equity is also not directly observed from the banks' balance sheets and income statement information. Hence the price of equity needs to be estimated for private banks. Following previous literature, the estimated price of equity using balance sheet and income statement information is called the "shadow price of equity" in this paper, to be differentiated from the price of equity measured using the price of stocks. An important advantage of this approach is that the shadow price of equity can be estimated for both listed and non-listed banks without using the market information. The shadow price of equity will equal the market price of equity when banks' cost is minimized at the current used amount of equity. Even if the current amount of equity does not minimize the cost, the shadow price of equity still provides a measure of opportunity cost of using the current amount of equity.

There exists some literature providing estimates of the shadow price of equity. Hughes (1999) explicitly derives the shadow price of equity capital, however, he does not show the estimates of shadow price of equity. Hughes et al. (2001) may be the first one to estimate the shadow price of equity capital using translog specification for the cost function. They find that there a positive relationship exists between asset size and the estimated shadow price of equity for the bank holding companies in 1994. Fethi et al. (2012) estimate the shadow price equity for 22 banks from Turkey over the period 2006–2009. They find that the shadow price on equity is negative in the post-financial crisis period, suggesting that the massive recapitalization of the banks during the recovery from the financial crisis drove them a long way from the equilibrium, and thus the involved deleveraging has imposed significant costs. Boucinha et al. (2013) estimate the shadow price of equity for Portuguese banks between 1992–2006 through the estimation of a translog cost frontier. The obtained measure of the shadow price of equity is in general higher than the short-term money market interest rate, however, it is lower than what is generally acknowledged to be a reasonable value for the actual price of equity. Restrepo et al. (2013) present new nonparametric measures of scale economies and total factor productivity growth for U.S. commercial banks over 2001–2010. Their results show that the sign of shadow price of equity depends on the models they used and also on the bank size. Duygun et al. (2015) use the same methods to estimate the shadow price of equity for 485 banks from emerging economies over period 2005–2008. They find a very consistent and statistically significant positive shadow price of equity capital of 4.1% to 4.9% on the capital constraint at the sample mean. Consequently the regulatory requirement to hold equity capital as a proportion of total assets is a strongly binding constraint at the sample mean. They also find that, at some sample points, the estimated shadow price of equity is negative, indicating that they have identified an "Excessive Capitalizer" operating in the uneconomic region of the banking production function because it is having to achieve a much higher equity capital to assets ratio. Radić (2015) estimates the shadow price of equity for the Japanese banking system over 1999–2011. Radić finds that at the sample mean, the shadow price on equity is between 2.8% and 6.1%. For the city banks, the cost of equity over time is significantly negative. Dong et al. (2016) estimate the shadow price of equity for Chinese commercial banks over the period 2002–2013. They find that there is a decreasing trend in the shadow price of equity over this period. They also find that the sign on the shadow price of equity is positive initially, but becomes negative by the end of the period. This may be because, during or after a severe recapitalization period, banks tend to deviate from their long-run equilibrium, which can cause the shadow price of equity to become negative. Fiordelisi et al. (2018) estimate the shadow price of equity using the data from commercial banks in Japan over the period 2000 to 2010. They find that at the sample mean the shadow price on equity is between 2.8 percent and 3.4 percent. They also find that for part of the period, the asset-weighted mean of the estimated shadow prices of equity capital did turn negative for both listed and unlisted banks indicating strong efforts at deleveraging and recapitalization. Hasannasab et al. (2019) use quadratic functional form of directional distance functions to obtain shadow prices of bank equity capital for listed and unlisted banks. They find that shadow prices for equity capital had reached abnormally high levels in the years leading up to the subprime crisis in the US indicative of excessive risk-taking behavior.

Instead of using the standard translog cost function approach, this paper uses nonparmatric methods initially developed by Simar et al. (2017) to estimate the cost frontier and thereby derive the shadow price of equity. The translog cost function is not flexible enough in estimating the cost function for U.S. commercial banks, where the size distribution is heavily right-skewed (Wheelock and Wilson, 2018). I also control for cost inefficiency when estimating the shadow price of equity. I use an almost fully-nonparametric specification of the noise and inefficiency processes, as opposed to estimating the more typical parametric stochastic frontier model where the noise and inefficiency distributions do not vary. The approach only requires symmetry of the two-sided noise process and that inefficiency be distributed half-normal. However, I allow the inefficiency to depend on the same covariates in the response function. The method is along the line of Simar et al. (2017) and Wheelock and Wilson (2019). Specifically, in the first regression, a nonparametric local-linear estimator is used to estimate the conditional mean cost function. In the second regression, I regress the cubed residuals from the first regression on the same covariates in the first regression. Using the information in the second regression, I can adjust the original estimates of the conditional mean cost function to estimate the cost frontier as well as the estimates of derivatives by exploiting the right skewness of the estimated residuals. Consequently, the approach is almost fully nonparametric. Although nonparametric estimators face the "curse of dimensionality", I take two steps to mitigate this problem.[1] Specially, I estimate my model using a large dataset consisting of over 119,000 observations on all U.S. commercial banks over the period 2001–2018. I also use an eigensystem decomposition of the correlation matrix of the right-hand-side variables to reduce the dimensions of the empirical model. My estimation methodology follows that of Wheelock and Wilson (2018, 2019). However, Wheelock and Wilson (2018) focus exclusively on the estimation of return to scale for U.S. banks for 1986–2015, and Wheelock and Wilson (2019) focus exclusively on the estimation of Lerner indices for U.S. bank holding companies for 2001–2018. Here, I focus on the estimation of the shadow price of equity for U.S. commercial banks for 2001–2018.

The nonparametric estimates of the shadow price of equity show that there indeed exist implicit subsidies for the TBTF banks. Specifically, I find that in each year, the estimated median value of shadow prices of equity for the banks in the fourth quartile based on total assets is much smaller than the banks in the other three quartiles. Moreover, for any given year, there exists a negative correlation between the shadow prices of equity and the sizes of banks, suggesting that big banks pay less in equity than small banks. In addition, there are more banks with a negative shadow price of equity in the fourth quartile than the other three quartiles in each year. The data reveal that for any given year in the sample, the estimated mean shadow price of equity for the top 100 largest banks is smaller than the mean price of deposits, even though equity is commonly viewed as a riskier asset than deposits. Finally, I find that the top 10 largest banks are willing to pay much more in equity at the start of the global financial crisis and after 2010. Therefore, the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 are effective in removing the

[1] "Curse of dimensionality" means the convergence rate of nonparametric estimators will decrease with the number of dimensions. In this paper, the number of dimensions are the number of independent variables included.

implicit subsidy, at least for the top 10 largest banks, and the deleveraging has imposed significant costs on the top 10 largest banks.

In the next section, a microeconomic model is presented for commercial banks, and the components needed to compute shadow price of equity are defined. In Section 2.3 the data used to define variables described in Section 2.2 are discussed. Section 2.4 presents the statistical model and gives details for estimation and inference. Empirical results are presented in Section 2.5. Summary and conclusions are given in Section 2.6.

2.2 The Economic Model

2.2.1 Deriving the Shadow Price of Equity

In this section, following Braeutigam and Daughety (1983), Duygun et al. (2015) and Weyman-Jones (2016), I employ a model of a representative bank's cost function that takes account of the requirement for the equity-asset ratio. Specifically, banks are required to hold the equity fixed in the short run, or to maintain a fixed equity-asset ratio to satisfy the government regulations. However, in the long run, the equity is allowed to be variable.

A representative bank's production function has p variable inputs $\boldsymbol{x} = (x_1, \ldots, x_p)$, q outputs $\boldsymbol{y} = (y_1, \ldots, y_q)$, input prices $\boldsymbol{w} = (w_1, \ldots, w_p)$, and an additional quasi-fixed input, equity q_1, i.e., an input which may be a fixed input in the short run but is variable in the long run. Assume the tranformation function $F(\boldsymbol{y}, \boldsymbol{x}, q_1, t) = 0$ for banks has the properties of convexity and weak disposability, where t is time. Weak disposability means $F_{x_i} = \frac{\partial F}{\partial x_i}$, $F_{y_j} = \frac{\partial F}{\partial y_j}$, $F_{q_1} = \frac{\partial F}{\partial q_1}$ are not restricted in sign. Therefore, banks are allowed to operate in the uneconomic region, and hence the shadow price of equity is not restricted in sign.

The long run cost function, with all inputs including q_1 treated as variables, takes the form

$$c^l(\boldsymbol{y}, \mathbf{w}, w_0, t) = \min_{\boldsymbol{x}, q_1} \{\mathbf{w}'\boldsymbol{x} + w_0 q_1 : F(\boldsymbol{y}, \boldsymbol{x}, q_1, t) = 0\}, \tag{2.1}$$

where $c^l(\boldsymbol{y}, \mathbf{w}, w_0, t)$ is the long run cost function and w_0 is the shadow price of equity. Following Duygun et al. (2015), the regulated short run cost function, modeled by specifying a fixed equity-asset ratio, $r_0 = q_1/q_2$, has the form

$$\begin{aligned} c^s(\boldsymbol{y}, \mathbf{w}, r_0, t) =& c(\boldsymbol{y}, \mathbf{w}, r_0, t) + w_0 q_1 \\ &= \min_{\boldsymbol{x}} \{\mathbf{w}'\boldsymbol{x} + w_0 q_1 : F(\boldsymbol{y}, \boldsymbol{x}, q_1, t) = 0, q_1 = r_0 q_2\}, \end{aligned} \tag{2.2}$$

where $c^s(\boldsymbol{y}, \mathbf{w}, r_0, t)$ is the short run cost function, q_2 is assets for the bank and $c(\boldsymbol{y}, \mathbf{w}, r_0, t)$ is the short run variable cost. The envelope theorem confirms that the long run cost function defines the envelope of the short run cost function

$$c^l(\boldsymbol{y}, \mathbf{w}, w_0, t) = \min_{r_0} \{c(\boldsymbol{y}, \mathbf{w}, r_0, t) + w_0 q_1, q_1 = r_0 q_2\}. \tag{2.3}$$

Consequently, the envelope theorem gives

$$\frac{\partial c^l(\boldsymbol{y}, \mathbf{w}, w_0, t)}{\partial r_0} = 0 = \frac{\partial c(\boldsymbol{y}, \mathbf{w}, r_0, t)}{\partial r_0} + w_0 q_2, \tag{2.4}$$

and rearranging this equation leads to

$$w_0 = -\frac{\partial c(\boldsymbol{y}, \mathbf{w}, r_0, t)}{\partial r_0} \frac{1}{q_2} = -\frac{\partial c(\boldsymbol{y}, \mathbf{w}, r_0, t)}{\partial q_1}, \tag{2.5}$$

where w_0 is the **shadow price** of equity, q_1. Rearranging equation (2.5) and expressing it in elasticity form gives

$$\frac{w_0 q_1}{c} = -\frac{\partial \log c(\boldsymbol{y}, \mathbf{w}, r_0, t)}{\partial \log r_0} = -\varepsilon_{q_1}, \tag{2.6}$$

where ε_{q_1} is the elasticity of short run variable cost with respect to equity-asset ratio. Therefore, the **shadow share** of equity expenses to total expenses could be estimated by the negative of the elasticity of short run variable cost with respect to equity-asset ratio.

2.2.2 Interpreting the Shadow Price of Equity

The shadow price of equity is derived in the previous subsection, and it is shown in equation (2.5). This equation is particularly relevant, since there is no explicit information on the price of equity, except that equity is an input fixed in the short run. Given fixed total outputs, when the bank has one more dollar increase in equity, some amount of deposits must be freed up since deposits and equity are substitutes. Expenses on deposits will surely decrease. Consequently, the variable cost, which is the sum of deposit expenses, labor expenses, and physical capital expenses, will also decrease. Therefore, the negative value of the derivative of the variable cost function with respect to equity is the shadow price for equity, as shown in equation (2.5). The rationale underlying the computation of the shadow price of equity is to provide a measure of how much banks are willing to pay for one more dollar increase in the level of equity, since it indicates the amount that they would save in the variable cost as a result of one more dollar increase in the level of equity. Consequently given the price of outputs, the shadow price of equity also indicates the amount that they would increase in the profit as a result of one more dollar increase in the level of equity. Even though equity is fixed, and hence "free" in the short run, there still exits a price for equity in the long run.

In this paper, I use year-end balance sheet and income statement information to estimate the cost function for banks. Therefore all dollar amounts are measured in book values rather than market values. Expenses on deposits are measured as the total deposits times the average annual interest rate on total deposits. Since all the other expenses are also measured annually, the estimated shadow price of equity could be interpreted as the "annual interest rate" on equity, if equity is treated the same as deposits. Therefore, the estimated shadow price of equity is directly comparable to the average price of deposits and the average price of loans and leases.

2.3 Data and Variable Specification

To obtain estimates of the shadow price of equity in equation (2.5), I must specify the variable cost function $c(\boldsymbol{y}, \mathbf{w}, r_0, t)$. My specification of right-hand-side (RHS) explanatory variables closely follows much of the banking literature. I use year-end data on U.S. commercial banks for 2001–2018 from the FFIEC (The Federal Financial Institutions Examination Council) call reports.[2] The widely used intermediation method of Sealey and Lindley (1977) is used to model a bank's technology as using deposits, labor, and physical capital (consisting of premises and fixed assets) to produce loans and leases, investments and off-balance items.

For the model, I specify three output quantities: total loans and leases (y_1), total securities (y_2), and off-balance sheet items consisting of non-interest income (y_3). Further, I specify three input prices: price of deposits (w_1), price of labor (w_2), and price of physical capital (w_3). The input

[2]See https://cdr.ffiec.gov/public/PWS/DownloadBulkData.aspx

price variables are measured by dividing expenditures on inputs by the corresponding quantities of inputs. Equity-asset ratio is also included to reflect that equity (q_1) is a quasi-fixed input. As an additional control for banks' differences in risk, I also include a measure of non-performing loans (npl), consisting of total loans and lease financing receivables past due 30 days or more and still accruing. As a final control variable, I index the years 2001–2018 by $t = 1, 2, \ldots, 18$. Although t is an ordered, categorical variable, it is treated as a continuous variable since its range is relatively large. Including t as an explanatory variable controls for changes in regulation, the global financial crisis, and all other changes by allowing the functional form of cost function to change over time.

The summary statistics for the variables over 2001–2018 are shown in Table 2.1. All monetary values are reported in constant 2018 U.S. dollars. Comparing differences between the first quartile and the median, and between the median and the third quartile for the input and output variables reveals that the marginal densities for both input and output variables are all skewed to the right. This implies that the translog specification for the cost function is not likely to be well specified. The translog specification for the cost function in U.S. banks is rejected in Wheelock and Wilson (2012, 2018, 2019).

Figure 2.1 shows the kernel density estimates of the log of total assets for U.S. commercial banks in 2001, 2009, and 2018. The estimates displayed in Figure 2.1 illustrate the evolution of commercial banks' sizes over the period covered by the sample. The distribution of U.S. commercial banks' sizes has shifted rightward over time, suggesting that U.S. commercial banks are expanding and some commercial banks have very large sizes. The nonparametric local estimator is more suitable than parametric estimators when the right skewness exists.

The medians and means for equity-asset ratio for each year are shown in Table 2.2. The mean and median values of the equity-asset ratio across years are around 10 percent, and the mean value is slightly larger than the median value for each year. The median value continuously increases from 2001 to 2007, and then continuously decreases from 2007 to 2009, after which, it continuously increases until 2012, where it maintains a much higher level than before. The pattern of the mean value of the equity-asset ratio appears to be the same. The decrease of the equity-asset ratio from 2007 to 2009 reflects the negative effect of the global financial crisis on U.S. commercial banks. In contrast, the increase of the equity-asset ratio after 2009 reflects the recapitalization process after the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010.

The medians of the equity-asset ratio for each size quartile based on total assets are reported in Table 2.3. In general, for each year the median value of the equity-asset ratio decreases when the size of banks increases. The correlations between the equity-asset ratio and size for each year in 2001–2018 are reported in Table 2.4. It is clear that there exists a negative correlation between the equity-asset ratio and the size of banks, suggesting that big banks tend to hold less equity, given total assets, compared with small banks. This indicates that small banks lack other resources, except increasing the equity-asset ratio to mitigate the potential market risk. However, big banks may get an implicit guarantee of bailout from the government when there is another crisis, and hence there is no need for them to hold too much equity than the level required by the government. Table 2.5 shows the weights for the assets of the top 10 largest commercial banks in the U.S. for each year. As shown in the table, the asset weight for the top 10 largest banks takes more than 50 percent after 2005, reflecting that the top 10 largest banks indeed have very large bank sizes and dominate the U.S. banking sector.

The relatively larger sizes of big banks may give them some market power in pricing the deposits and loans and leases. The results of the tests of differences in means for price of deposits between big and small banks for each year are shown in Table 2.6. For each year I split the sample into two subsamples based on the median value of total assets. The small banks are then defined as those with total assets smaller than the median value of total assets in each year, and the remaining are defined as big banks. Table 2.6 shows that the mean price of deposits for big banks is significantly smaller than that for small banks before 2003, while it is significantly larger for most cases after 2003. This reflects the demand effect since big banks usually need larger amount of deposits than small banks, and hence they are willing to pay a higher price for deposits than small banks. The results of the tests of differences in means for price of loans and leases between small and big banks for each year are shown in Table 2.7. Table 2.7 shows that the mean price of loans and leases for big banks is significantly smaller than that for small banks in each year. This reflects the supply effect since big banks usually provide a larger amount of loans and leases, and hence they will charge a lower price than small banks. As a robustness check, I also consider a different definition of small banks and big banks based on the bottom and top 25th percentile of total assets in each year. The results of the tests of the differences in means for price of deposits and price of loans and leases are shown in Tables 2.8 and 2.9, respectively. The results are quite consistent with my baseline results in Tables 2.6 and 2.7.

2.4 The Econometric Model

Simar et al. (2017) propose an almost fully-nonparametric framework for stochastic frontier models. This method involves less assumptions on the cost function and is more general than the translog or other parametric specifications of the cost function. A number of papers have rejected the translog specifications of the cost function.[3] Therefore, I use nonparametric least squares methods for stochastic frontier models to estimate the cost function for the U.S. commercial banks and thereby derive the shadow price of equity. The nonparametric estimation strategy avoids specification error that might be obtained when using a mis-specified model. A disadvantage of nonparametric estimators, however, is that they suffer from the "curse of dimensionality", i.e., the convergence rates fall as the number of dimensions in the model increases. However, the slow convergence of nonparametric estimators is mitigated by using a large dataset and an eigensystem decomposition of the correlation among the right-hand-side variables to reduce dimensions.

The variable cost function $c(\boldsymbol{y}, \mathbf{w}, r_0, t)$ must be homogeneous of degree one with respect to input prices $\mathbf{w}$ since the cost minimization problem implies that factor demand equations must be homogeneous of degree zero in input prices. Consequently I divide the input prices and the variable cost by the price of physical capital (w_3). Following Wheelock and Wilson (2012, 2018), I define the vector of covariates

$$z_i = [\frac{w_{i1}}{w_{i3}} \quad \frac{w_{i1}}{w_{i3}} \quad y_{i1} \quad y_{i2} \quad y_{i3} \quad r_{i0} \quad npl_i \quad \exp(t_i)]$$

for the right-hand-side variables (RHS) of the cost function. In order to estimate cost frontiers and to allow for inefficiency, I employ the moment-based method of Simar et al. (2017) and eigensystem decomposition of Wheelock and Wilson (2019) as described below.

I first take logs of each RHS variable, then standardize the logs by subtracting means and dividing by standard deviations of the logs. This will transform z_i to $\widetilde{z_i}$. The RHS variables are usually highly correlated. Following Wheelock and Wilson (2019), I use eigensystem decomposition of the correlation matrix of $\widetilde{z_i}$ to reduce dimensions. Let E denote the matrix of eigenvectors of the correlation matrix of $\widetilde{z_i}$. The eigenvectors in the columns of E are ordered so that the first column corresponds to the largest eigenvalue and the last column corresponds to the smallest eigenvalue.

[3] See Wheelock and Wilson (2012, 2018, 2019)

Then I compute the $(n \times 8)$ matrix

$$\boldsymbol{\Psi}_{full} = [\boldsymbol{\Psi} \quad \boldsymbol{\Psi}_{del}] = \widetilde{z}_i E \tag{4.7}$$

of principal components, where $\boldsymbol{\Psi}$ contains the desired principal components and $\boldsymbol{\Psi}_{del}$ contains the removed components. Let e_j denote the eigenvalues, sorted in decreasing order, and let $\tilde{e_j} = \sum_{k=1}^{j} e_k / \sum_{k=1}^{8} e_k$ for $j = 1, 2, \ldots, 8$. Then $\tilde{e_j}$ gives the proportion of the independent linear transformation in $\widetilde{z}_i$ contained in the first j principal components, i.e., the first j columns of $\boldsymbol{\Psi}_{full}$. These values are 0.3986, 0.6122, 0.7610, 0.8742, 0.9341, 0.9708, 0.9870, and 1.0000. Consequently, I define the partition in equation (4.7) so that $\boldsymbol{\Psi}$ is an $(n \times 5)$ matrix, and I use these first $d = 5$ principal components to estimate the cost function. By construction, $\boldsymbol{\Psi}$ contains more than 93 percent of the independent linear information in $\widetilde{z}_i$ for the period 2001–2018, and consequently the number of dimensions are reduced from 8 to 5.

Now let $\boldsymbol{\Psi}_i = (\Psi_{i1}, \Psi_{i2}, \ldots, \Psi_{id})$ denote the ith row of $\boldsymbol{\Psi}$. I use the local-linear estimator to estimate the following cost function

$$\log(\frac{c_i}{w_{i3}}) = m(\Psi_{i1}, \Psi_{i2}, \ldots, \Psi_{id}) + V_i + U_i, \tag{4.8}$$

where $m(\boldsymbol{\Psi}_i) = m(\Psi_{i1}, \Psi_{i2}, \ldots, \Psi_{id})$ is a conditional mean function measuring the cost function frontier, and V_i is the statistical noise term, for which I assume that $E(V_i|\boldsymbol{\Psi}_i) = 0$ and $\mathrm{Var}(V_i|\boldsymbol{\Psi}_i) \in (0, \infty)$ for all i and that U_i is a nonnegative random variable, capturing the individual cost inefficiency. Moreover, U_i is assumed to be independent from V_i. Conditionally on $\boldsymbol{\Psi}_i$, $U_i|\boldsymbol{\Psi}_i$ is assumed to follow half-normal distribution $|N(0, \sigma_U^2(\boldsymbol{\Psi}_i))|$, and hence $\mu_U(\boldsymbol{\Psi}_i) = \sqrt{\frac{2}{\pi}}\sigma_U(\boldsymbol{\Psi}_i)$. In addition, I make no functional form assumptions regarding $m(\boldsymbol{\Psi}_i)$ and only make the regular assumptions to ensure the consistency of nonparametric estimators.

Following Simar et al. (2017) and Wheelock and Wilson (2019), let $\varepsilon_i = V_i + U_i - \mu_U(\boldsymbol{\Psi}_i)$, and $r_1(\boldsymbol{\Psi}_i) = m(\boldsymbol{\Psi}_i) + \mu_U(\boldsymbol{\Psi}_i)$. Using the local-linear estimator I estimate the following equation

$$\log(\frac{c_i}{w_{i3}}) = r_1(\boldsymbol{\Psi}_i) + \varepsilon_i, \tag{4.9}$$

where $r_1(\boldsymbol{\Psi}_i)$ is the estimated individual cost function. Denote $r_3(\boldsymbol{\Psi}_i) = E(\varepsilon_i^3|\boldsymbol{\Psi}_i)$. It can be easily shown that

$$E(\varepsilon_i|\mathbf{\Psi}_i) = 0 \tag{4.10}$$

and

$$E(\varepsilon_i^3|\mathbf{\Psi}_i) = E[(U_i - \mu_U(\mathbf{\Psi}_i))^3|\mathbf{\Psi}_i], \tag{4.11}$$

where the distribution of inefficiency $U_i|\mathbf{\Psi}_i$ has a positive skewness and therefore $r_3(\mathbf{\Psi}_i) \geq 0$.

After estimating the cost function, I now have

$$\hat{\varepsilon}_i = \log(\frac{c_i}{w_{i3}}) - \hat{r}_1(\mathbf{\Psi}_i), \tag{4.12}$$

and I can get the local linear estimate of $r_3(\mathbf{\Psi}_i)$ from the data points $\{\hat{\varepsilon}_i^3, \mathbf{\Psi}_i|i = 1, \ldots, n\}$. After some algebra, it can be shown that the variance function for U_i can be consistently estimated by

$$\hat{\sigma}_U^2(\mathbf{\Psi}_i) = \max\left\{0, \left[\sqrt{\frac{\pi}{2}}\Big(\frac{\pi}{4-\pi}\Big)\hat{r}_3(\mathbf{\Psi}_i)\right]^{\frac{2}{3}}\right\}, \tag{4.13}$$

and

$$\hat{c}_i = w_{i3}\exp(\hat{m}(\mathbf{\Psi}_i)) = w_{i3}\exp(\hat{r}_1(\mathbf{\Psi}_i) - \hat{\mu}_U(\mathbf{\Psi}_i)). \tag{4.14}$$

Therefore according to (2.5) in Section 2.2, the shadow price of equity is equal to

$$w_{i0} = -\frac{\partial c_i}{\partial r_{i0}}\frac{1}{q_{i2}} = -c_i(\frac{\partial r_1(\mathbf{\Psi}_i)}{\partial r_{i0}} - \frac{\partial \mu_U(\mathbf{\Psi}_i)}{\partial r_{i0}})\frac{1}{q_{i2}}, \tag{4.15}$$

where for z_{il}, the l-th element of z_i, I have

$$\frac{\partial r_1(\mathbf{\Psi}_i)}{\partial z_{il}} = s_l^{-1} z_{il}^{-1} \sum_{j=1}^{d} \hat{\beta}_{1ij} E_{lj}, \tag{4.16}$$

and

$$\frac{\partial \mu_U(\mathbf{\Psi}_i)}{\partial z_{il}} = \begin{cases} \frac{2^{1/3}}{3}(4-\pi)^{-1/3}\hat{r}_3(\mathbf{\Psi}_i)^{-2/3}s_l^{-1}z_{il}^{-1}\sum_{j=1}^{d}\hat{\beta}_{3ij}E_{lj}, & \forall \hat{r}_3(\mathbf{\Psi}_i) > 0 \\ 0 \quad \text{otherwise}, \end{cases} \tag{4.17}$$

where E_{lj} is the (l, j)-th element of the matrix E of eigenvectors, and s_l is the standard deviation of the logged l-th variable, i.e., the standard deviation of the l-th column of $\widetilde{z}$. The slope terms are $\hat{\beta}_{1ij} = \frac{\partial r_1(\mathbf{\Psi}_i)}{\partial \Psi_{ij}}$ and $\hat{\beta}_{3ij} = \frac{\partial r_3(\mathbf{\Psi}_i)}{\partial \Psi_{ij}}$, $j = 1, 2, \ldots, 5$. Moreover, the $\hat{\beta}_{1ij}$s and $\hat{\beta}_{3ij}$s are computed at each observation i in each regression due to the local nature of the local-linear estimator. The estimation approach described here is almost fully nonparametric. Although I assume that inefficiency is distributed half-normal, the shape parameter is estimated locally and is allowed to vary continuously across observations. A fully nonparametric approach does not seem possible, as some structures are needed in order to identify expected inefficiency.

To implement the local-linear estimator a bandwidth parameter must be selected to control the smoothing over the continuous dimensions in the data. Following Wheelock and Wilson (2011, 2012, 2018, 2019), I use least-squares cross-validation to optimize an adaptive, κ-nearest-neighbor bandwidth. In addition, I employ a second-order Epanechnikov kernel function. I use the bandwidth inside the kernel function. This means that when estimating cost at any fixed point of interest in the space of the RHS variables, only the κ observations closet to that point can influence estimated cost. In addition, among these κ observations, the influence that a particular observation has on estimated cost diminishes with distance from the point at which the response is being estimated. The estimator here is thus a *local* estimator and is very different from typical, parametric, *global* estimation strategies (e.g. ordinary least squares, maximum likelihood, etc.) where all observations in the sample influence (with equal weights) estimation at any given point in the data space. Moreover, because I use adaptive nearest-neighbor bandwidths, the bandwidths automatically adapt to variation in the sparseness of data throughout the support of the RHS variables. This results in relatively larger values for the bandwidths where the data are sparse (and where more smoothing is required), and smaller values for the bandwidths where the data are relatively dense (where less smoothing is needed).

For making inference about the shadow prices of equity or differences in these across different years or groups based on the nonparametric estimates, I use the wild bootstrap introduced by Härdle (1990) and Härdle and Mammen (1993), which avoids making specific distributional assumptions. Although the estimators are asymptotically normal, the bootstrap avoids the needs to estimate those unknown parameters in an asymptotically normal distribution. I estimate confidence intervals using methods described in Wheelock and Wilson (2011, 2012, 2018, 2019).

First, I obtain bootstrap estimates $\{\hat{w}^*_{0b}\}_{b=1}^{B}$ (set $B = 1,000$), then sort the values in $\{\hat{w}^*_{0b} - \hat{w}_0\}_{b=1}^{B}$ by algebraic value, delete $(\frac{\alpha}{2} \times 100)\%$ of the elements at either end of this sorted array, and denote the lower and upper end points of the remaining, sorted array as $-b^*_\alpha$ and $-a^*_\alpha$, respectively. Then a bootstrap estimate of a $(1-\alpha)\%$ confidence interval for $\hat{w}_0$ is

$$\hat{w}_0 + a^*_\alpha \leq w_0 \leq \hat{w}_0 + b^*_\alpha. \tag{4.18}$$

The idea underlying equation (4.18) is that the empirical distribution of the bootstrap values $(\hat{w}^*_{0b} - \hat{w}_0)$ mimics the unknown distribution of $(\hat{w}_0 - w_0)$, with the approximation improving as $n \to \infty$. As $B \to \infty$, the choices of $-b^*_\alpha$ and $-a^*_\alpha$ become increasingly accruate estimates of the percentiles of the distribution of $(\hat{w}^*_{0b} - \hat{w}_0)$. Any bias in $\hat{w}_0$ relative to w_0 is reflected in the bias of $\hat{w}^*$ relative to $\hat{w}$. The estimated confidence interval may not contain the original estimates of $\hat{w}$ if the bias is large because the estimated confidence interval corrects for the bias in $\hat{w}$.

2.5 Empirical Results

I use nonparametric local linear methods and some moment conditions to estimate the frontier, from which I derive the shadow price of equity for each commercial bank over 2001–2018. However, there are some unreasonable estimates of the shadow prices of equity, therefore, I first remove the outliers in the estimates before analyzing the results. If the estimated shadow price of equity is smaller than $Q1 - 3 \times IQR$ or larger than $Q3 + 3 \times IQR$, then this estimate is defined as an outlier, where $Q1$ is the first quartile, $Q3$ is the third quartile, and IQR is the interquartile range (i.e., $IQR = Q3 - Q1$). In total there are 1443 outliers out of $119,028$ observations. After removing the outliers, I do not lose much information, and the estimates, after removing the outliers, still capture the changes in the trend of quartiles of the estimates.

The summary statistics for the estimated shadow prices of equity are reported in Table 2.10. The mean value of the estimated shadow prices of equity is 0.0875 in 2001. This means that for a typical U.S. commercial bank in the beginning of 2001, if it borrows one more dollar of equity from an investor for just one year, then it is estimated to pay back 1.0875 dollars by the end of 2001. In other words, the "interest rate" on equity is 0.0875 in 2001. The median value of the estimated shadow prices of equity continuously decreases from 2001 to 2005, and then increases from 2005 to 2006, after which, it decreases again from 2006 to 2008. The median value increases from 2008 to 2010, and then decreases from 2010 to 2014, after which it increases from 2014 to 2016. The median value then decreases from 2016 to 2017 and increases from 2017 to 2018. The mean values of the estimated shadow prices of equity have similar trend over the period 2001–2018. I find that the median values before 2008 are, in general, higher than that after 2008, thus implying that the recapitalization process in U.S. commercial banks leads to a decrease in the shadow price of equity. The lower median values after 2008 may also reflect the relatively low funding costs, low potential market risk, and the increased competition in the U.S. banking sector. The increase in shadow price of equity from 2008 to 2010 implies that banks are willing to pay more to increase their equity capital during the global financial crisis since the market risk during this time is perceived to be very high. Specifically, the larger median value (0.0324) in 2010 may be due to the tighter regulations of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 that require banks to hold a much higher level of equity than before. Banks are willing to pay more in 2010 to raise their equity capital to satisfy the government's regulations. This result is consistent with Dong et al. (2016), who find that there is an increase in the shadow price of equity for Chinese banks from 2008 to 2009. Surprisingly, some amount of banks with negative estimates of shadow prices of equity show up across years. This suggests that for any given year, these banks actually operate in an uneconomic region of the production function. These banks may hold a much higher equity level than their efficient level, causing their shadow prices of equity to be negative.

Stated previously, the estimated shadow price of equity is directly comparable to the price of deposits and the price of loans and leases. The comparisons among the median values of these three prices are reported in Table 2.11. Equity is commonly considered more risky than deposits because equity holders are the last to receive any distribution of assets as a result of bankruptcy proceedings. Therefore, equity holders expect greater returns from their investment in the firm's stock than depositors. Table 2.11 shows that for most years, the estimated shadow price of equity is larger than the price of deposits and smaller than the price of loans. However, over the period 2006–2008, the estimated shadow price of equity is smaller than the price of deposits. Moreover,

over the period 2006–2008, the price of loans and leases is very high. The potential profits for banks over 2006–2008 may lead banks to be willing to pay an unreasonable price for deposits.

Table 2.12 reports the summary statistics for the estimated shadow shares of equity costs to total variable expenses in each year. The median values of the estimated shadow shares of equity costs have a decreasing trend from 2001 to 2008, and then have an increasing trend from 2008 to 2018. The mean values have a similar pattern. Even though Table 2.10 shows that the median values of the estimated prices of equity before 2008 are in general higher than that after 2008, Table 2.12 shows that the median values of estimated shadow shares of equity costs increase in general after 2008, reflecting that banks use much more equity than before, and U.S. banking systems have been undergoing recapitalization since the global financial crisis.

Table 2.13 shows the summary statistics for the estimated cost inefficiency in each year. Even though the changes of median and mean values of the cost inefficiency over this period are mixed, it is evident that the cost inefficiency before 2008 is lower in general than that after 2008. This suggests that banks became less cost efficient after 2008, and on average they are much farther away from the cost frontier. The increase in inefficiency may be caused by the tighter regulations from the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, suggesting that the deleveraging has imposed significant costs on banks.

The difference in medians for the estimated shadow prices of equity by size quartile based on log of total assets are reported in Table 2.14. It is evident that the median values of estimated shadow prices of equity for the fourth quartile are much smaller than the other three quartiles for any given year. Moreover, the median values of shadow prices of equity for the fourth quartile are negative over 2005–2008 and 2015–2016. Specifically the median values for the banks in the fourth quartile have the lowest negative value in 2008, suggesting that these banks do not pay anything to equity investors at the start of the global financial crisis, and instead they get paid implicitly by the equity investors. This fact implies that big banks indeed get an implicit guarantee from the government during the upheaval global financial crisis. To further understand the relationship between the estimated shadow price of equity and the size of banks, I report their correlation in Table 2.15. For any given year, there is a negative correlation between the shadow price of equity and the size of banks, suggesting that big banks indeed consistently pay less in equity than small banks over the sample period. Our results are different from Hughes et al. (2001) who find that a positive relationship exists between the size and the estimated shadow price of equity for the bank holding companies in 1994. This is not surprising since I use nonparametric methods instead of translog cost function and also here I focus on commercial banks over 2001–2018 rather than bank holding companies in 1994.

The results of tests about whether the estimated shadow prices of equity are different from 0 are reported in Table 2.16. The table reports the number of banks for which I reject that the shadow price of equity is significantly different from 0 (at .05 significance) in favor of a positive shadow price of equity or a negative shadow price of equity, or for which I cannot reject that the shadow price of equity is equal to 0 in each quartile of total assets in each year. In each quartile, I find that a small number of banks have a negative shadow price of equity in each year, suggesting that only a small number of banks operate in an uneconomic region of production. However, among banks in the fourth quartile (the largest 25 percent of banks by assets), I find that there are more banks having negative shadow price of equity than the other three quartiles in each year. Therefore, the results suggest that more banks in the fourth quartile get paid implicitly by equity investors than the other three quartiles, providing evidence of TBTF banks.

Table 2.17 reports the correlations between the equity-asset ratios and the shadow prices of equity across years. Economic theory predicts that in a free market, if a bank uses more equity given fixed outputs (or assets), the price (or opportunity cost) of equity should be lower. Table 2.17 shows that the correlations are only negative for 2001, 2002, and 2006. This means that if a bank uses more equity given fixed assets, the shadow price of equity will be lower over these three years. However, for most years, the correlations are positive. This result is not surprising and supports the choice of treating equity capital as a quasi-fixed input rather than a variable input. In the short run, equity is not variable due to government's regulations and constraints.

Turning to differences in the means of estimated shadow price of equity between the top 100 largest banks and the other banks in Table 2.19, the data reveal that for any given year in the sample, the estimated mean shadow price of equity for the top 100 largest banks is smaller than the mean price of deposits, which is smaller than the mean price of loans and leases. This result thus provides evidence that, on average, the top 100 largest banks get the implicit guarantee from the government, so they are able to pay a much lower price on equity than on deposits, even though equity is commonly viewed as a much riskier asset than deposits. Also, for most years, the mean shadow price of equity for the top 100 largest banks is negative, again implying that they get paid implicitly by the equity investors. In addition, the mean shadow price of equity for the top 100 largest banks in 2007 has the largest positive value (0.0148), suggesting that on average they would like to pay more to get equity capital at the start of the global financial crisis, even though this value is still much smaller than the price of deposits. Comparing the difference in the mean shadow prices of equity between the top 100 largest banks and the other banks, I find that the mean value for the top 100 largest banks is much smaller than that for the other banks, again providing evidence that the top 100 largest banks get paid implicitly by the equity investors.

I further check differences in the means of the estimated shadow prices of equity for the top 10 largest banks and the other banks with results reported in Table 2.20. The data reveal that in 2001–2005, 2007, and 2009–2011, the estimated mean shadow prices of equity for the top 10 largest banks are smaller than the mean prices of deposits, which is smaller than the mean price of loans and leases. This is evidence that there exits an implicit subsidy for the top 10 largest banks. However, for the remaining periods, the estimated mean shadow price of equity for the top 10 largest banks is larger than the mean price of deposits and even larger than the mean price of loans and leases in some cases. In addition, the mean shadow price of equity for the top 10 largest banks in 2007 has a very large positive value (0.133), suggesting that on average they would like to pay much more to get equity capital to mitigate the market risk at the start of the global financial crisis. Moreover, the mean shadow price of equity for the top 10 largest banks maintains a very high level after 2010. Comparing the difference in mean shadow prices of equity between the top 10 largest banks and the other banks, I find that the mean values for the top 10 largest banks are larger than for the other banks at the start of the global financial crisis and after the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. Comparing these results with those in Table 2.19, I find that even though on average the top 100 largest banks pay less in equity for each year, the top 10 largest banks actually pay more for some years, especially at the start of the global financial crisis and after the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. These results imply that the regulations are effective in reducing the implicit subsidy at least for the top 10 largest banks. However, it is also evident that the recapitalization has imposed significant equity funding costs on the top 10 largest banks.

2.6 Summary and Conclusions

After the global financial crisis, TBTF has become a heated topic. However, until now there was little evidence of the existence of TBTF banks. This paper contributes to the literature on the existence of TBTF banks by providing a new piece of evidence.

By estimating the shadow price of equity, using nonparametric local-linear method for stochastic frontier models initially introduced by Simar et al. (2017), I find that there are indeed implicit subsidies for the TBTF banks. Specifically, I find that the estimated median values of shadow prices of equity for the banks in the fourth quartile based on total assets are much smaller than the banks in the other three quartiles. Moreover, for any given year there exists a negative correlation between the shadow prices of equity and the sizes of banks, suggesting that big banks pay less in equity than small banks. In addition, there are more banks with a negative shadow price of equity in the fourth quartile than the other three quartiles in each year. The data reveal that for any given year in the sample, the estimated mean shadow price of equity for the top 100 largest banks is smaller than the mean price of deposits, even though equity is commonly viewed as a riskier asset than deposit. Finally, I find that the top 10 largest banks are willing to pay much more at the start of the global financial crisis and after the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. Therefore, the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 is effective in removing the implicit subsidy for the top 10 largest banks. However, it is also evident that the recapitalization has imposed significant equity funding costs on the top 10 largest banks.

My results are consistent with the evidence of the existence of TBTF banks, provided by Baker and Mcarthur (2009), Brewer and Jagtiani (2013), Ueda and Weder di Mauro (2013), and Santos and Santos (2014). Given the importance of TBTF in the U.S., the policies trying to remove the implicit subsidy on TBTF have significant impacts on the U.S. banking market. Policy makers should be cautious of the fact that the increased safety of banks may be offset by the adjustment costs (decreased efficiency and increased shadow price of equity) imposed by the recapitalization process. This is a critical question for both policy makers and banking regulators.

Table 2.1: Summary Statistics for Years 2001–2018

Variable	N	Min	Q1	Median	Mean	Q3	Max
Total Deposits (x_1)	119028	$2.2470 \times 10^{+02}$	$7.3740 \times 10^{+04}$	$1.5320 \times 10^{+05}$	$1.7880 \times 10^{+06}$	$3.5260 \times 10^{+05}$	$1.9300 \times 10^{+09}$
No. of Full Time Employee (x_2)	119028	$1.0000 \times 10^{+00}$	$2.0000 \times 10^{+01}$	$4.0000 \times 10^{+01}$	$2.9210 \times 10^{+02}$	$8.7000 \times 10^{+01}$	$2.3520 \times 10^{+05}$
Physical Capital (x_3)	119028	$1.0000 \times 10^{+00}$	$9.4750 \times 10^{+02}$	$2.8440 \times 10^{+03}$	$1.8650 \times 10^{+04}$	$7.1960 \times 10^{+03}$	$1.3430 \times 10^{+07}$
Total Loans (y_1)	119028	$2.1820 \times 10^{+01}$	$4.8060 \times 10^{+04}$	$1.0840 \times 10^{+05}$	$1.1450 \times 10^{+06}$	$2.6070 \times 10^{+05}$	$9.7350 \times 10^{+08}$
Total Securities (y_2)	119028	$1.0230 \times 10^{+00}$	$1.3460 \times 10^{+04}$	$3.2700 \times 10^{+04}$	$3.9260 \times 10^{+05}$	$8.0820 \times 10^{+04}$	$4.3280 \times 10^{+08}$
Off-balance Sheet Itmes (y_3)	119028	$1.0000 \times 10^{+00}$	$3.7640 \times 10^{+02}$	$9.7780 \times 10^{+02}$	$3.6950 \times 10^{+04}$	$2.8490 \times 10^{+03}$	$4.5050 \times 10^{+07}$
Price of Deposit (w_1)	119028	1.4910×10^{-05}	6.9250×10^{-03}	1.7620×10^{-02}	2.0080×10^{-02}	2.9900×10^{-02}	$2.9720 \times 10^{+00}$
Price of Labor (w_2)	119028	2.4000×10^{-01}	$5.6740 \times 10^{+01}$	$6.6070 \times 10^{+01}$	$7.0360 \times 10^{+01}$	$7.8980 \times 10^{+01}$	$8.5300 \times 10^{+02}$
Price of Physical Capital (w_3)	119028	1.2660×10^{-04}	1.8460×10^{-01}	2.7560×10^{-01}	6.3330×10^{-01}	4.6640×10^{-01}	$4.2380 \times 10^{+03}$
Non-performing Loans (npl)	119028	$1.0000 \times 10^{+00}$	$7.4000 \times 10^{+02}$	$2.0980 \times 10^{+03}$	$3.7720 \times 10^{+04}$	$5.7650 \times 10^{+03}$	$8.5170 \times 10^{+07}$
Equity (q_1)	119028	$2.7890 \times 10^{+01}$	$8.8220 \times 10^{+03}$	$1.7850 \times 10^{+04}$	$2.1750 \times 10^{+05}$	$4.0300 \times 10^{+04}$	$2.1650 \times 10^{+08}$
Total Assets (q_2)	119028	$3.1080 \times 10^{+03}$	$8.3840 \times 10^{+04}$	$1.7300 \times 10^{+05}$	$2.0590 \times 10^{+06}$	$3.9560 \times 10^{+05}$	$2.2190 \times 10^{+09}$
Equity-asset Ratio (r_0)	119028	8.0210×10^{-05}	8.6600×10^{-02}	1.0070×10^{-01}	1.0860×10^{-01}	1.2070×10^{-01}	9.8790×10^{-01}
Time (t)	119028	$1.0000 \times 10^{+00}$	$4.0000 \times 10^{+00}$	$8.0000 \times 10^{+00}$	$8.7980 \times 10^{+00}$	$1.3000 \times 10^{+01}$	$1.8000 \times 10^{+01}$
Variable Cost (c)	119028	$2.2730 \times 10^{+01}$	$2.7050 \times 10^{+03}$	$5.6660 \times 10^{+03}$	$5.5390 \times 10^{+04}$	$1.2990 \times 10^{+04}$	$6.8930 \times 10^{+07}$
Price of Loans and Leases(p_1)	119028	1.8180×10^{-04}	5.4130×10^{-02}	6.4010×10^{-02}	6.5240×10^{-02}	7.4240×10^{-02}	$3.0900 \times 10^{+00}$

Note: All dollar amounts are given in 2018 U.S. thousand dollars

Table 2.2: Medians and Means for Equity-Asset Ratio

Year	N	Median	Mean
2001	8111	0.0932	0.1028
2002	7905	0.0955	0.1050
2003	7782	0.0954	0.1044
2004	7550	0.0961	0.1062
2005	7371	0.0961	0.1059
2006	7168	0.0977	0.1085
2007	7065	0.1002	0.1114
2008	6878	0.0977	0.1068
2009	6578	0.0975	0.1051
2010	6293	0.0987	0.1058
2011	6109	0.1033	0.1098
2012	6468	0.1046	0.1110
2013	6301	0.1024	0.1089
2014	6025	0.1064	0.1132
2015	5724	0.1069	0.1140
2016	5491	0.1061	0.1129
2017	5234	0.1075	0.1150
2018	4975	0.1092	0.1168

Table 2.3: Medians of Equity-Asset Ratio by Size Quartile Based

Year	1st Quartile	2nd Quartile	3nd Quartile	4th Quartile
2001	0.1049	0.0953	0.0909	0.0863
2002	0.1069	0.0974	0.0926	0.0890
2003	0.1058	0.0971	0.0922	0.0889
2004	0.1065	0.0980	0.0925	0.0906
2005	0.1070	0.0979	0.0929	0.0898
2006	0.1090	0.0998	0.0941	0.0919
2007	0.1117	0.1022	0.0961	0.0938
2008	0.1104	0.1002	0.0932	0.0906
2009	0.1059	0.0995	0.0939	0.0931
2010	0.1041	0.0994	0.0967	0.0967
2011	0.1067	0.1046	0.1010	0.1023
2012	0.1078	0.1045	0.1024	0.1045
2013	0.1043	0.1028	0.1004	0.1036
2014	0.1075	0.1070	0.1045	0.1068
2015	0.1101	0.1071	0.1050	0.1063
2016	0.1104	0.1067	0.1034	0.1054
2017	0.1118	0.1082	0.1045	0.1068
2018	0.1138	0.1093	0.1051	0.1091

Note: The size quartiles are defined in terms of total assets for each year.

Table 2.4: Correlation Between Equity-Asset Ratios and Sizes of Banks

Year	Correlation
2001	-0.1960
2002	-0.1856
2003	-0.1594
2004	-0.1394
2005	-0.1455
2006	-0.1512
2007	-0.1497
2008	-0.2007
2009	-0.1423
2010	-0.0803
2011	-0.0555
2012	-0.0411
2013	-0.0131
2014	-0.0446
2015	-0.0637
2016	-0.0658
2017	-0.0645
2018	-0.0687

Table 2.5: The Weight of Assets for Top 10 Largest Banks

Year	Weight
2001	0.3862
2002	0.4077
2003	0.4221
2004	0.4651
2005	0.4817
2006	0.5127
2007	0.5343
2008	0.5552
2009	0.5495
2010	0.5613
2011	0.5753
2012	0.5437
2013	0.5467
2014	0.5566
2015	0.5465
2016	0.5376
2017	0.5380
2018	0.5438

Table 2.6: Tests of Differences in Means for Price of Deposits Between Big and Small Banks

Year	$n1$	$n2$	Mean1	Mean2	Statistic	p-value
2001	4055	4056	0.0497	0.0485	-4.5192	6.21×10^{-06}
2002	3952	3953	0.0326	0.0313	-6.2582	3.90×10^{-10}
2003	3891	3891	0.0237	0.0231	-3.5160	4.38×10^{-04}
2004	3775	3775	0.0194	0.0197	1.9707	4.88×10^{-02}
2005	3685	3686	0.0231	0.0249	11.0570	2.03×10^{-28}
2006	3584	3584	0.0313	0.0346	12.3775	3.46×10^{-35}
2007	3532	3533	0.0357	0.0393	4.5629	5.04×10^{-06}
2008	3439	3439	0.0281	0.0294	7.3395	2.14×10^{-13}
2009	3289	3289	0.0208	0.0210	0.1990	8.42×10^{-01}
2010	3146	3147	0.0148	0.0152	3.0753	2.10×10^{-03}
2011	3054	3055	0.0107	0.0109	1.6835	9.23×10^{-02}
2012	3234	3234	0.0081	0.0082	1.2454	2.13×10^{-01}
2013	3150	3151	0.0062	0.0062	0.5783	5.63×10^{-01}
2014	3012	3013	0.0052	0.0052	0.5788	5.63×10^{-01}
2015	2862	2862	0.0048	0.0048	-0.1576	8.75×10^{-01}
2016	2745	2746	0.0047	0.0048	0.5972	5.50×10^{-01}
2017	2617	2617	0.0049	0.0052	3.4847	4.93×10^{-04}
2018	2487	2488	0.0061	0.0070	9.4219	4.43×10^{-21}

Note: I split the total observations of each year into two even subsamples by the median total assets in that year. The number of small banks is $n1$, while the number of big banks is $n2$.

Table 2.7: Tests of Differences in Means for Price of Loans and Leases Between Big and Small Banks

Year	$n1$	$n2$	Mean1	Mean2	Statistic	p-value
2001	4055	4056	0.0887	0.0843	-11.2364	2.70×10^{-29}
2002	3952	3953	0.0788	0.0733	-13.0674	5.05×10^{-39}
2003	3891	3891	0.0730	0.0665	-18.3902	1.58×10^{-75}
2004	3775	3775	0.0681	0.0626	-6.1035	1.04×10^{-09}
2005	3685	3686	0.0710	0.0666	-14.2099	7.96×10^{-46}
2006	3584	3584	0.0771	0.0743	-9.2144	3.13×10^{-20}
2007	3532	3533	0.0796	0.0759	-9.5987	8.09×10^{-22}
2008	3439	3439	0.0716	0.0667	-14.8066	1.33×10^{-49}
2009	3289	3289	0.0673	0.0624	-10.7416	6.49×10^{-27}
2010	3146	3147	0.0666	0.0618	-12.3055	8.46×10^{-35}
2011	3054	3055	0.0643	0.0596	-14.9504	1.55×10^{-50}
2012	3234	3234	0.0612	0.0566	-11.8717	1.66×10^{-32}
2013	3150	3151	0.0576	0.0526	-15.2305	2.22×10^{-52}
2014	3012	3013	0.0549	0.0500	-15.5815	9.73×10^{-55}
2015	2862	2862	0.0538	0.0486	-15.8820	8.44×10^{-57}
2016	2745	2746	0.0535	0.0479	-18.0996	3.21×10^{-73}
2017	2617	2617	0.0531	0.0485	-11.2615	2.03×10^{-29}
2018	2487	2488	0.0546	0.0506	-12.1416	6.36×10^{-34}

Note: I split the total observations of each year into two even subsamples by the median total assets in that year. The number of small banks is $n1$, while the number of big banks is $n2$.

Table 2.8: Tests of Differences in Means for Price of Deposits Between Big and Small Banks

Year	$n1$	$n2$	Mean1	Mean2	Statistic	p-value
2001	2028	2028	0.0494	0.0479	-3.5437	3.95×10^{-04}
2002	1976	1977	0.0326	0.0306	-6.1715	6.76×10^{-10}
2003	1946	1946	0.0235	0.0226	-3.5992	3.19×10^{-04}
2004	1888	1888	0.0189	0.0195	2.6699	7.59×10^{-03}
2005	1843	1843	0.0222	0.0253	13.1272	2.30×10^{-39}
2006	1792	1792	0.0300	0.0353	17.9959	2.10×10^{-72}
2007	1766	1767	0.0344	0.0391	8.8235	1.11×10^{-18}
2008	1720	1720	0.0272	0.0294	8.2329	1.83×10^{-16}
2009	1645	1645	0.0208	0.0209	0.0710	9.43×10^{-01}
2010	1573	1574	0.0141	0.0150	4.2883	1.80×10^{-05}
2011	1527	1528	0.0104	0.0108	2.4803	1.31×10^{-02}
2012	1617	1617	0.0078	0.0081	2.3653	1.80×10^{-02}
2013	1575	1576	0.0060	0.0061	1.0969	2.73×10^{-01}
2014	1506	1507	0.0050	0.0051	0.7422	4.58×10^{-01}
2015	1431	1431	0.0047	0.0047	0.2447	8.07×10^{-01}
2016	1373	1373	0.0046	0.0047	0.9334	3.51×10^{-01}
2017	1309	1309	0.0048	0.0052	3.9932	6.52×10^{-05}
2018	1244	1244	0.0059	0.0072	9.2318	2.66×10^{-20}

Note: I split the total observations of each year into the top 25% quantile group and the bottom 25% quantile group in terms of total assets. The number for small banks is $n1$, while the number of big banks is $n2$.

Table 2.9: Tests of Differences in Means for Price of Loans and Leases Between Big and Small Banks

Year	$n1$	$n2$	Mean1	Mean2	Statistic	p-value
2001	2028	2028	0.0897	0.0831	-11.3313	9.18×10^{-30}
2002	1976	1977	0.0807	0.0718	-13.9986	1.59×10^{-44}
2003	1946	1946	0.0752	0.0646	-19.7196	1.46×10^{-86}
2004	1888	1888	0.0699	0.0598	-15.9104	5.37×10^{-57}
2005	1843	1843	0.0721	0.0656	-13.2573	4.10×10^{-40}
2006	1792	1792	0.0777	0.0734	-9.5689	1.08×10^{-21}
2007	1766	1767	0.0804	0.0750	-8.3365	7.66×10^{-17}
2008	1720	1720	0.0733	0.0656	-14.0755	5.37×10^{-45}
2009	1645	1645	0.0688	0.0616	-8.5787	9.60×10^{-18}
2010	1573	1574	0.0678	0.0609	-11.3459	7.77×10^{-30}
2011	1527	1528	0.0655	0.0585	-13.1973	9.10×10^{-40}
2012	1617	1617	0.0624	0.0555	-10.7737	4.58×10^{-27}
2013	1575	1576	0.0589	0.0513	-13.6655	1.63×10^{-42}
2014	1506	1507	0.0563	0.0485	-14.9686	1.18×10^{-50}
2015	1431	1431	0.0552	0.0469	-15.3249	5.21×10^{-53}
2016	1373	1373	0.0551	0.0465	-16.2377	2.73×10^{-59}
2017	1309	1309	0.0545	0.0476	-9.2036	3.46×10^{-20}
2018	1244	1244	0.0557	0.0498	-10.3783	3.11×10^{-25}

Note: I split the total observations of each year into the top 25% quantile group and the bottom 25% quantile group in terms of total assets. The number for small banks is $n1$, while the number of big banks is $n2$.

Table 2.10: Summary Statistics for Estimated Shadow Prices of Equity

Year	N	Min	Q1	Median	Mean	Q3	Max
2001	7907	-0.7965	-0.0646	0.0830***	0.0875***	0.2437	0.8760
2002	7793	-0.7991	-0.0559	0.0647***	0.0759***	0.2048	0.8757
2003	7694	-0.7863	-0.0746	0.0466***	0.0526**	0.1840	0.8726
2004	7458	-0.7973	-0.0930	0.0366***	0.0402**	0.1740	0.8747
2005	7234	-0.7989	-0.1156	0.0304***	0.0299	0.1816	0.8711
2006	6945	-0.7974	-0.1267	0.0310***	0.0368**	0.2029	0.8768
2007	6843	-0.7894	-0.1287	0.0307***	0.0345*	0.2072	0.8754
2008	6761	-0.7937	-0.1148	0.0229**	0.0232	0.1640	0.8671
2009	6526	-0.7791	-0.0846	0.0276***	0.0299*	0.1460	0.8715
2010	6255	-0.7953	-0.0739	0.0324***	0.0328**	0.1415	0.8766
2011	6087	-0.7904	-0.0622	0.0286***	0.0333**	0.1292	0.8632
2012	6452	-0.7822	-0.0565	0.0275***	0.0333**	0.1222	0.8262
2013	6281	-0.7924	-0.0681	0.0226***	0.0198	0.1148	0.8236
2014	6014	-0.7804	-0.0599	0.0223***	0.0235*	0.1086	0.8463
2015	5700	-0.7972	-0.0668	0.0237***	0.0243*	0.1164	0.8600
2016	5468	-0.7808	-0.0666	0.0255***	0.0290**	0.1261	0.8760
2017	5215	-0.7787	-0.0681	0.0245***	0.0278**	0.1278	0.8739
2018	4952	-0.7992	-0.0788	0.0267***	0.0256*	0.1311	0.8695

Note: Statistical significance (difference from 0) for the median and mean values at the ten, five, or one percent levels is denoted by one, two, or three asterisks, respectively.

Table 2.11: Differences in Medians for Estimated Shadow Prices of Equity, Prices of Deposits and Prices of Loans and Leases

Year	Equity	Deposits	Loans
2001	0.0830	0.0496	0.0860
2002	0.0647	0.0321	0.0753
2003	0.0466	0.0234	0.0687
2004	0.0366	0.0195	0.0636
2005	0.0304	0.0242	0.0679
2006	0.0310	0.0331	0.0749
2007	0.0307	0.0374	0.0771
2008	0.0229	0.0289	0.0683
2009	0.0276	0.0204	0.0640
2010	0.0324	0.0148	0.0632
2011	0.0286	0.0106	0.0610
2012	0.0275	0.0077	0.0576
2013	0.0226	0.0057	0.0536
2014	0.0223	0.0048	0.0509
2015	0.0237	0.0044	0.0499
2016	0.0255	0.0043	0.0494
2017	0.0245	0.0046	0.0494
2018	0.0267	0.0062	0.0511

Table 2.12: Summary Statistics for Estimated Shadow Shares of Equity Costs to Total Expenses

Year	N	Min	Q1	Median	Mean	Q3	Max
2001	7907	-2.1198	-0.1431	0.2027	0.1955	0.5460	2.2585
2002	7793	-2.6517	-0.1762	0.2070	0.2161	0.6035	2.5181
2003	7694	-2.8934	-0.2569	0.1744	0.1816	0.6150	3.0163
2004	7458	-2.7709	-0.3392	0.1577	0.1534	0.6329	3.1989
2005	7234	-3.4110	-0.3800	0.1174	0.1142	0.6111	3.7346
2006	6945	-2.9897	-0.3886	0.1073	0.1205	0.6114	3.0056
2007	6843	-3.5348	-0.3682	0.1038	0.1085	0.5880	3.1385
2008	6761	-3.6798	-0.3398	0.0735	0.0931	0.5189	2.9500
2009	6526	-4.5818	-0.2869	0.1086	0.1252	0.5328	2.9416
2010	6255	-2.9489	-0.2910	0.1396	0.1444	0.5866	2.8939
2011	6087	-2.7813	-0.2939	0.1472	0.1617	0.6162	3.0730
2012	6452	-3.5784	-0.2961	0.1609	0.1748	0.6389	3.5406
2013	6281	-3.2565	-0.3653	0.1343	0.1357	0.6383	3.7433
2014	6014	-4.0487	-0.3614	0.1449	0.1551	0.6586	5.3166
2015	5700	-3.8380	-0.4031	0.1575	0.1678	0.7358	4.7627
2016	5468	-6.2755	-0.3878	0.1733	0.1806	0.7470	5.5381
2017	5215	-5.1794	-0.3943	0.1630	0.1797	0.7419	4.1570
2018	4952	-4.2895	-0.4272	0.1540	0.1550	0.7191	3.3417

Table 2.13: Summary Statistics for Estimated Mean Cost Inefficiency

Year	N	Min	Q1	Median	Mean	Q3	Max
2001	7907	1.0000	1.0603	1.1684	1.1641	1.2510	1.6734
2002	7793	1.0000	1.1194	1.1955	1.1896	1.2695	1.6526
2003	7694	1.0000	1.1346	1.2116	1.2014	1.2820	1.6786
2004	7458	1.0000	1.1444	1.2210	1.2093	1.2893	1.6842
2005	7234	1.0000	1.1451	1.2204	1.2086	1.2901	1.6692
2006	6945	1.0000	1.1400	1.2171	1.2059	1.2888	1.6837
2007	6843	1.0000	1.1404	1.2188	1.2087	1.2914	1.6498
2008	6761	1.0000	1.1493	1.2238	1.2190	1.3004	1.6558
2009	6526	1.0000	1.1488	1.2263	1.2232	1.3084	1.6851
2010	6255	1.0000	1.1499	1.2272	1.2236	1.3079	1.6553
2011	6087	1.0000	1.1495	1.2269	1.2215	1.3089	1.6497
2012	6452	1.0000	1.1474	1.2245	1.2213	1.3116	1.6582
2013	6281	1.0000	1.1505	1.2294	1.2225	1.3116	1.6361
2014	6014	1.0000	1.1539	1.2320	1.2231	1.3131	1.6637
2015	5700	1.0000	1.1498	1.2301	1.2214	1.3110	1.6538
2016	5468	1.0000	1.1456	1.2286	1.2188	1.3117	1.6451
2017	5215	1.0000	1.1411	1.2238	1.2140	1.3071	1.6169
2018	4952	1.0000	1.1250	1.2121	1.2036	1.2976	1.5906

Table 2.14: Difference in Medians for Shadow Prices of Equity by Size Quartile

Year	1st Quartile	2nd Quartile	3nd Quartile	4th Quartile
2001	0.0761	0.0975	0.1064	0.0498
2002	0.0673	0.0858	0.0729	0.0317
2003	0.0460	0.0640	0.0596	0.0142
2004	0.0358	0.0630	0.0425	0.0123
2005	0.0291	0.0536	0.0434	-0.0010
2006	0.0344	0.0556	0.0534	-0.0093
2007	0.0285	0.0662	0.0380	-0.0040
2008	0.0291	0.0460	0.0295	-0.0164
2009	0.0310	0.0470	0.0313	0.0006
2010	0.0319	0.0468	0.0410	0.0093
2011	0.0313	0.0444	0.0334	0.0046
2012	0.0292	0.0395	0.0326	0.0079
2013	0.0255	0.0320	0.0282	0.0044
2014	0.0231	0.0321	0.0307	0.0050
2015	0.0313	0.0382	0.0275	-0.0040
2016	0.0347	0.0478	0.0276	-0.0088
2017	0.0386	0.0323	0.0307	0.0021
2018	0.0316	0.0359	0.0376	0.0019

Table 2.15: Correlation Between the Estimated Shadow Prices of Equity and Sizes of Banks

Year	Correlation
2001	-0.0682
2002	-0.0856
2003	-0.0862
2004	-0.0884
2005	-0.0635
2006	-0.0738
2007	-0.0741
2008	-0.0946
2009	-0.0806
2010	-0.0630
2011	-0.0783
2012	-0.0801
2013	-0.0526
2014	-0.0643
2015	-0.1010
2016	-0.1153
2017	-0.0888
2018	-0.0817

Table 2.16: Counts of Banks Having Negative, 0, and Positive Shadow Prices of Equity by Size Quartile (.05 significance)

	— **1st quartile** —			— **2nd quartile** —			— **3rd quartile** —			— **4th quartile** —		
Year	Neg.	0	Pos.	Neg.	0	Pos.	Neg.	0	Pos.	Neg.	0	Pos.
2001	173	1432	372	121	1485	371	114	1469	393	229	1359	389
2002	171	1385	393	133	1406	409	123	1464	361	239	1354	355
2003	179	1391	354	164	1354	405	169	1414	340	262	1323	339
2004	203	1310	352	196	1277	391	211	1324	329	274	1279	312
2005	225	1262	322	208	1250	350	183	1322	303	280	1259	270
2006	227	1212	298	163	1247	326	161	1286	289	236	1223	277
2007	203	1241	267	153	1234	324	158	1246	306	281	1155	275
2008	166	1229	296	164	1204	322	193	1228	269	261	1188	241
2009	178	1165	289	135	1174	322	165	1193	273	228	1139	265
2010	160	1091	313	157	1104	303	165	1098	300	227	1032	305
2011	155	1085	282	143	1059	320	156	1058	307	226	1007	289
2012	179	1112	323	169	1092	351	161	1110	342	228	1100	285
2013	183	1091	297	203	1063	304	207	1058	305	247	1013	310
2014	161	1057	286	155	1066	282	175	1018	310	243	988	273
2015	166	962	297	139	972	314	159	977	289	295	866	264
2016	160	930	277	124	967	276	169	925	273	268	839	260
2017	132	864	308	146	907	251	156	871	276	236	805	263
2018	125	850	263	141	828	269	130	848	260	230	751	257

Table 2.17: Correlation Between the Estimated Shadow Prices of Equity and Equity-Asset Ratios

Year	Correlation
2001	-0.0206
2002	-0.0285
2003	0.0084
2004	0.0138
2005	0.0286
2006	-0.0021
2007	0.0039
2008	0.0626
2009	0.0477
2010	0.0296
2011	0.0104
2012	0.0050
2013	0.0632
2014	0.0453
2015	0.0455
2016	0.0399
2017	0.0472
2018	0.0438

Table 2.18: Correlation Between the Ratio of Shadow Prices of Equity over Price of Deposits and the Ratio of Equity over Deposits

Year	Correlation
2001	-0.0159
2002	-0.0191
2003	-0.0040
2004	-0.0053
2005	0.0036
2006	-0.0167
2007	-0.0015
2008	0.0467
2009	-0.0041
2010	0.0013
2011	-0.0080
2012	-0.0022
2013	0.0017
2014	0.0005
2015	-0.0003
2016	0.0274
2017	0.0026
2018	0.0027

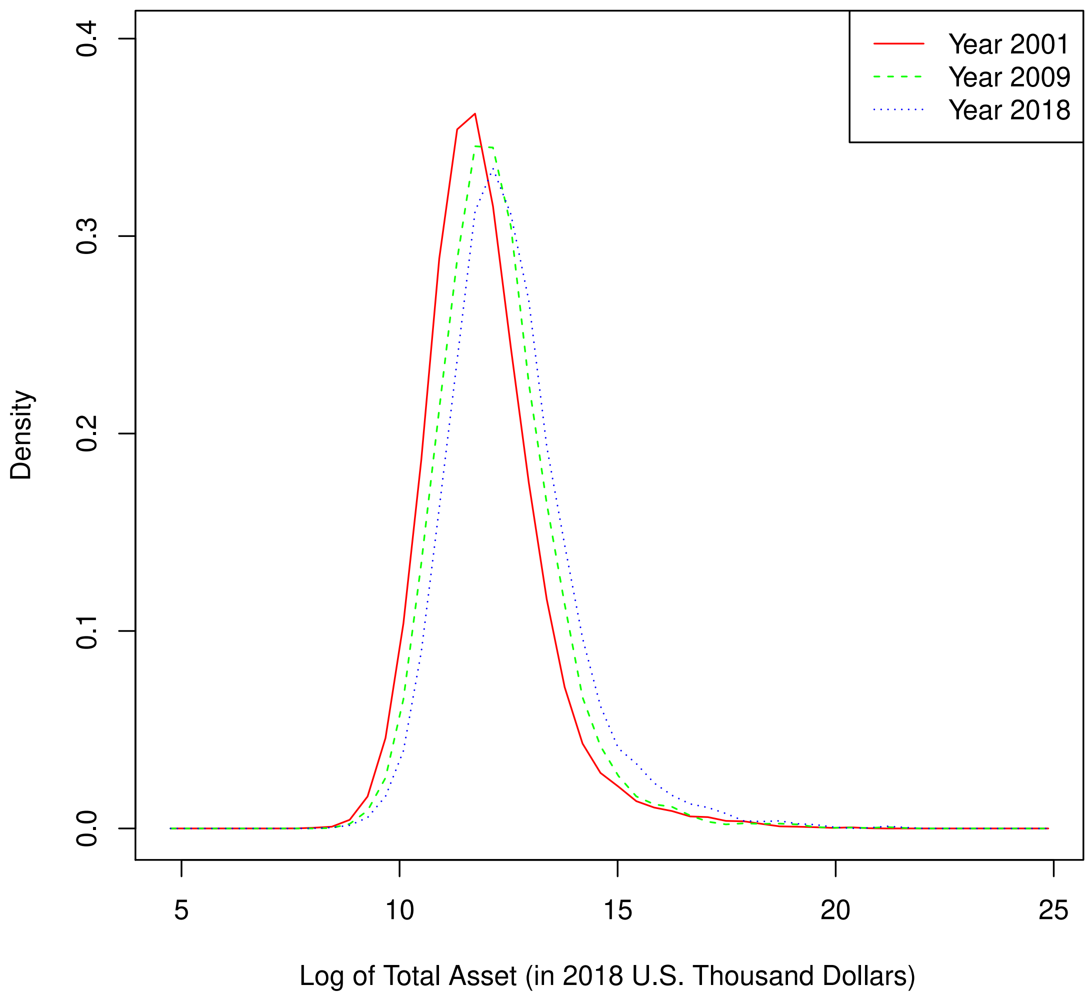
Table 2.19: Differences in Means for Shadow Prices of Equity, Prices of Deposits, and Prices of Loans and Leases for the Top 100 Largest Banks and the Other Banks

	— **Top 100 Largest Banks** —			— **The Other Banks** —		
Year	Equity	Deposits	Loans	Equity	Deposits	Loans
2001	-0.0126	0.0452	0.0834	0.0888	0.0490	0.0863
2002	-0.0206	0.0276	0.0693	0.0772	0.0320	0.0761
2003	-0.0117	0.0205	0.0607	0.0535	0.0234	0.0699
2004	-0.0211	0.0182	0.0555	0.0410	0.0195	0.0654
2005	-0.0127	0.0273	0.0622	0.0305	0.0239	0.0689
2006	-0.0139	0.0366	0.0692	0.0376	0.0328	0.0757
2007	0.0148	0.0386	0.0707	0.0348	0.0374	0.0778
2008	-0.0076	0.0269	0.0639	0.0237	0.0287	0.0691
2009	-0.0305	0.0162	0.0586	0.0308	0.0209	0.0649
2010	-0.0045	0.0122	0.0589	0.0334	0.0150	0.0643
2011	-0.0218	0.0087	0.0539	0.0342	0.0109	0.0621
2012	-0.0166	0.0071	0.0515	0.0341	0.0081	0.0591
2013	0.0023	0.0055	0.0484	0.0201	0.0062	0.0552
2014	-0.0222	0.0048	0.0454	0.0243	0.0052	0.0526
2015	-0.0193	0.0046	0.0461	0.0251	0.0048	0.0513
2016	-0.0077	0.0047	0.0448	0.0297	0.0047	0.0508
2017	-0.0019	0.0054	0.0456	0.0283	0.0050	0.0509
2018	0.0031	0.0082	0.0516	0.0261	0.0065	0.0526

Table 2.20: Differences in Means for Shadow Prices of Equity, Prices of Deposits, and Prices of Loans and Leases for the Top 10 Largest Banks and the Other Banks

	— **Top 10 Largest Banks** —			— **The Other Banks** —		
Year	Equity	Deposits	Loans	Equity	Deposits	Loans
2001	-0.0619	0.0434	0.0773	0.0877	0.0490	0.0863
2002	-0.0628	0.0236	0.0614	0.0761	0.0319	0.0760
2003	-0.0898	0.0164	0.0543	0.0528	0.0234	0.0698
2004	-0.0721	0.0148	0.0474	0.0404	0.0195	0.0653
2005	-0.1212	0.0253	0.0567	0.0301	0.0239	0.0688
2006	0.0583	0.0370	0.0738	0.0368	0.0329	0.0757
2007	0.1325	0.0393	0.0740	0.0343	0.0374	0.0777
2008	0.0327	0.0236	0.0643	0.0232	0.0287	0.0691
2009	-0.0632	0.0109	0.0533	0.0301	0.0209	0.0649
2010	-0.0242	0.0074	0.0573	0.0329	0.0150	0.0642
2011	0.0147	0.0053	0.0414	0.0333	0.0108	0.0620
2012	0.0462	0.0041	0.0419	0.0333	0.0081	0.0590
2013	0.0593	0.0033	0.0400	0.0197	0.0062	0.0551
2014	0.0383	0.0027	0.0379	0.0235	0.0052	0.0525
2015	0.0633	0.0027	0.0359	0.0242	0.0048	0.0512
2016	0.0650	0.0031	0.0374	0.0289	0.0047	0.0507
2017	0.0968	0.0043	0.0400	0.0276	0.0050	0.0508
2018	0.1546	0.0074	0.0447	0.0253	0.0066	0.0526

Figure 2.1: Density of (log) Total Assets of U.S. Commercial Banks in 2001, 2009, and 2018



NOTE: Solid line shows density for 2001; dashed line shows density for 2009; dotted line shows density for 2018.

Chapter 3

Performance of Countries in the Post-Crisis Era

3.1 Introduction

The global financial crisis of 2007–2012 was the worst economic disaster since the 1929 great depression. It began with the subprime crisis of the housing mortgage markets in the U.S. Since securities linked to subprime loans were accumulated in all the banks and all the global financial market, the subprime crisis quickly caused inter-banking crisis. As early as on 14 September 2007, Britain's Northern Rock Bank received a liquidity support facility from the Bank of England, which led to the UK's first bank run in 150 years. In 2008, several other depressed financial institutions were purchased by others (Bear Stearns by JP Morgan Chase, Merrill Lynch by Bank of America), nationalized (Fannie Mae, Freddie Mac, and American International Group), or bankrupted (Washington Mutual, Lehman Brothers). Even though the governments of many countries provided liquidity and enacted large fiscal programs, bank failures led to a credit shortage and blocked the investment, plunging the global economy into a deep recession. Given this disruptive period 2004–2014, it is reasonable to ask what happened to the global economy following 2004, some years before the global financial crisis. Especially, we are interested in the evolution of productivity, efficiency, and technology of the countries during this period.

The traditional approach to the analysis of productivity using non-frontier analysis assumes that all the countries lie in the frontier and are perfectly efficient so that the growth of productivity is purely interpreted as the movement of the frontier or technology. However, the non-frontier analysis does not incorporate the inefficiency part resulting from the constraints or low efficient management, hence the estimation of productivity and technical progress would be biased. On the other hand, the frontier analysis could directly model the inefficiency behavior of the countries when estimating productivity, efficiency, and technical progress.

There exist some literature using parametric and nonparametric frontier methods to estimate technical efficiency and productivity of the countries in the world. Färe et al. (1994) may be the first to use nonparametric frontier method to analyze productivity growth in 17 OECD countries over the period 1979–1988. They decompose Malmquist productivity growth into changes in technical efficiency which measures catching-up effect, and shifts in technology which measures the innovation. They find that U.S. productivity growth is slightly higher than average. Ray and Desli (1997) propose an alternative decomposition of Malmquist productivity growth. Maudos et al. (1999) add human capital as another input to estimate Malmquist productivity and they find that the inclusion of human capital has a significant effect on the accurate measurement of total factor productivity. Chang and Luh (1999) use the same method as in Färe et al. (1994) to check for Asian countries and they find that the United States is not the sole innovator among the 19 APEC member economies. Rather, Hong Kong and Singapore have shown their capability to shift the grand frontier of the APEC economies. Arcelus and Arocena (2000) use nonparametric frontier method to estimate technical efficiency and scale efficiency for 14 OECD countries over 1970–1990 period. They find evidence of convergence, even if at quite different speeds, for total industry, manufacturing, and services. Emrouznejad (2003) proposes a dynamic nonparametric efficiency model to estimate efficiency, and compare the results with the static efficiency model proposed by Färe et al. (1994). They find that dynamic models capture efficiency better than static models. Han et al. (2004) use a varying coefficients frontier production (parametric method) to estimate total factor productivity, and then decompose it into technical efficiency change and technical change for a sample of 45 developed and developing countries over the period 1970–1990. They find that East Asian economies are not outliers in terms of total factor productivity growth. Salinas-Jiménez and Salinas-Jiménez (2007) use nonparametric frontier approach to estimate productivity, technical efficiency and then check the effects of corruption on these two measures for OECD countries. They find that corruption negatively influences the possibilities of growth. Wang et al. (2012)

estimate Malmquist productivity index with and without defense expenditure, and then compare the difference of Malmquist productivity over time to check the effects of defense expenditure on economic productivity in OECD countries. They find that average Malmquist productivity with defense expenditure is higher than that without defense expenditure.

Among the nonparametric methods, Data Envelopment Analysis (DEA) estimators which impose convexity assumption of the production set, have been widely applied to estimate productivity and efficiency in these literature. However, they did not test the convexity of the production set, nor do they test constant returns to scale (CRS) versus variable returns to scale (VRS). Moreover, some of these studies simply report efficiency estimates without any inference and compare the mean efficiency of two groups without correcting the bias of the estimated efficiency. Of course, point estimates without inference are largely meaningless. Hence, the results of these studies are dubious. In addition, these studies only have a few observations and hence they naturally encounter the "curse of dimensionality", which is a serious problem in nonparametric efficiency estimation.[1] Hence, dimension reduction is needed in the context of nonparametric efficiency estimation. Recently, Kneip et al. (2016), using the central limit theorem results from Kneip et al. (2015), develop hypotheses testing the model structure. They provide tests of the convexity of the production set, returns to scale and differences in mean efficiency across groups of producers. Also, Wilson (2018) proposes a new dimension reduction technique that is advantageous in terms of reducing estimation error. Results suggest that Free Disposal Hull (FDH) estimator which does not impose convexity assumption is a viable, attractive alternative to the VRS-DEA in many cases when dimension reduction is used. We are the first to use recently-developed statistical methods to assess what can be learned about efficiency change, productivity growth and technical progress of the countries in the world from the data.

This paper provides evidence on the performance of the countries in the world before, during and after the 2007–2012 global financial crisis. The approach is fully non-parametric and exploits recently developed theoretical results. Estimates of technical efficiency and productivity at two-year intervals from 2004 to 2014 are examined in a statistical paradigm permitting inference and hypothesis testing. Therefore, this paper both (i) contributes to the literature by shedding light on the evolution of efficiency and productivity of the countries in the world before, during and after

[1] Curse of dimensionality means the convergence rates of nonparametric estimators decrease with increasing dimensions (number of inputs and outputs).

the 2007–2012 global financial crisis, and (ii) fills the gap between point estimates and inference in the empirical research on countries' technical efficiency and productivity.

The rest of paper is organized as follows. Estimators of technical efficiency and their properties are discussed in Section 3.2. Section 3.3 discusses various statistical results needed for testing hypotheses about model features. Section 3.4 describes the data we used. Section 3.5 discusses the empirical results of the tests. Major conclusions and directions for future works are presented in Section 3.6.

3.2 The Statistical Model

To establish notation, let $X \in \mathbb{R}_+^p$ and $Y \in \mathbb{R}_+^q$ denote (random) vectors of input and output quantities, respectively. Similarly, let $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ denote fixed, nonstochastic vectors of input and output quantities. The production set

$$\Psi := \{(x, y) \mid x \text{ can produce } y\} \tag{2.1}$$

gives the set of feasible combinations of inputs and outputs. Several assumptions on Ψ are common in the literature. The assumptions of Shephard (1970) and Färe (1988) are typical in microeconomic theory of the firm and are used here.

Assumption 3.2.1. Ψ *is closed.*

Assumption 3.2.2. $(x, y) \notin \Psi$ *if* $x = 0,\ y \geq 0,\ y \neq 0$*; i.e., all production requires use of some inputs.*

Assumption 3.2.3. *Both inputs and outputs are strongly disposable, i.e.,* $\forall\,(x, y) \in \Psi$*, (i)* $\widetilde{x} \geq x \Rightarrow (\widetilde{x}, y) \in \Psi$ *and (ii)* $\widetilde{y} \leq y \Rightarrow (x, \widetilde{y}) \in \Psi$*.*

Here and throughout, inequalities involving vectors are defined on an element-by-element basis, as is standard. Assumption 3.2.1 ensures that the *efficient frontier* (or *technology*) Ψ^∂

$$\Psi^\partial := \left\{(x, y) \mid (x, y) \in \Psi,\ (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for any } \alpha \in (1, \infty)\right\} \tag{2.2}$$

is the set of extreme points of Ψ and is contained in Ψ. Assumption 3.2.2 means that production of any output quantities greater than 0 requires use of some inputs so that there can be no free lunches. Assumption 3.2.3 imposes weak monotonicity on the frontier.

The Farrell (1957) input efficiency measure

$$\theta(x, y \mid \Psi) := \inf \{\theta \mid (\theta x, y) \in \Psi\} \tag{2.3}$$

gives the amount by which input levels can feasibly be scaled downward, proportionately by the same factor, without reducing output levels. The Farrell (1957) output efficiency measure gives the feasible, proportionate expansion of output quantities and is defined by

$$\lambda(x, y \mid \Psi) := \sup \{\lambda \mid (x, \lambda y) \in \Psi\}. \tag{2.4}$$

Both (2.3) and (2.4) provide *radial* measures of efficiency since all input or output quantities are scaled by the same factor θ or λ, holding output or input quantities fixed (respectively). Clearly, $\theta(x, y \mid \Psi) \leq 1$ and $\lambda(x, y \mid \Psi) \geq 1$ for all $(x, y) \in \Psi$.

Alternatively, Färe et al. (1985) provide a hyperbolic, graph measure of efficiency defined by

$$\gamma(x, y \mid \Psi) := \inf \left\{\gamma > 0 \mid (\gamma x, \gamma^{-1} y) \in \Psi\right\}. \tag{2.5}$$

By construction, $\gamma(x, y \mid \Psi) \leq 1$ for $(x, y) \in \Psi$. Just as the measures $\theta(x, y \mid \Psi)$ and $\lambda(x, y \mid \Psi)$ provide measures of the *technical efficiency* of a firm operating at a point $(x, y) \in \Psi$, so does $\gamma(x, y \mid \Psi)$, but along the hyperbolic path from (x, y) to the frontier of Ψ. The measure $\gamma(x, y \mid \Psi)$ gives the amount by which input levels can be feasibly, proportionately scaled downward while simultaneously scaling output levels upward by the same proportion.

All of the quantities and model features defined so far are unobservable, and therefore must be estimated. The set Ψ can be estimated using the free-disposal hull (FDH) estimator introduced by Deprins et al. (1984) or either the variable returns to scale (VRS) or constant returns to scale (CRS) versions of the data envelopment analysis (DEA) estimator proposed by Farrell (1957). But, inference is needed in order to know what might be learned from data, and inference requires a well-defined statistical model.

3.3 Estimation and Inference

Let $\mathcal{S}_n = \{(X_i, Y_i)\}_{i=1}^n$ be a random input-output pairs sample drawn from the density f introduced in Assumption A.1. Given a random sample $\mathcal{S}_n = \{(X_i, Y_i)\}$, the production set Ψ can be estimated by the free disposal hull of the sample observations in $\mathcal{S}_n$,

$$\widehat{\Psi}_{\text{FDH},n} := \bigcup_{(X_i,Y_i)\in\mathcal{S}_n} \left\{(x,y) \in \mathbb{R}_+^{p+q} \mid x \geq X_i,\ y \leq Y_i\right\}, \tag{3.1}$$

proposed by Deprins et al. (1984). Alternatively, Ψ can be estimated by the convex hull of $\widehat{\Psi}_{\text{FDH},n}$, i.e., by

$$\widehat{\Psi}_{\text{VRS},n} := \left\{(x,y) \in \mathbb{R}^{p+q} \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{3.2}$$

where $\boldsymbol{X} = \begin{pmatrix} X_1, \ \ldots, \ X_n \end{pmatrix}$ and $\boldsymbol{Y} = \begin{pmatrix} Y_1, \ \ldots, \ Y_n \end{pmatrix}$ are $(p \times n)$ and $(q \times n)$ matrices of input and output vectors, respectively; $\boldsymbol{i}_n$ is an $(n \times 1)$ vector of ones, and $\boldsymbol{\omega}$ is a $(n \times 1)$ vector of weights. The estimator $\widehat{\Psi}_{\text{VRS},n}$ imposes convexity, but allows for VRS. This is the VRS (DEA) estimator of Ψ proposed by Farrell (1957) and popularized by Banker et al. (1984). The CRS (DEA) estimator $\widehat{\Psi}_{\text{CRS},n}$ of Ψ is obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (3.2). FDH, VRS or CRS estimators of $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ defined in Section 3.2 are obtained by substituting $\widehat{\Psi}_{\text{FDH},n}$, $\widehat{\Psi}_{\text{VRS},n}$ or $\widehat{\Psi}_{\text{CRS},n}$ for Ψ in (2.3)–(2.5) (respectively). In the case of VRS estimators, this results in

$$\widehat{\theta}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\theta,\boldsymbol{\omega}} \left\{\theta \mid y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \theta x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}, \tag{3.3}$$

$$\widehat{\lambda}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \max_{\lambda,\boldsymbol{\omega}} \left\{\lambda \mid \lambda y \leq \boldsymbol{Y}\boldsymbol{\omega},\ x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\} \tag{3.4}$$

and

$$\widehat{\gamma}_{\text{VRS}}(x, y \mid \mathcal{S}_n) = \min_{\gamma,\boldsymbol{\omega}} \left\{\gamma \mid \gamma^{-1} y \leq \boldsymbol{Y}\boldsymbol{\omega},\ \gamma x \geq \boldsymbol{X}\boldsymbol{\omega},\ \boldsymbol{i}_n'\boldsymbol{\omega} = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^n\right\}. \tag{3.5}$$

The corresponding CRS estimators $\widehat{\theta}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{CRS}}(x, y \mid \mathcal{S}_n)$ are obtained by dropping the constraint $\boldsymbol{i}_n'\boldsymbol{\omega} = 1$ in (3.3)–(3.5). The estimators in (3.3)–(3.4) can be computed using linear programming methods, but the hyperbolic estimator in (3.5) is a non-linear program. Nonetheless, estimates can be computed easily using the numerical algorithm developed by Wilson (2011). Substituting $\widehat{\Psi}_{\text{FDH},n}$ into (2.3)–(2.5) (respectively) will yield FDH estimators

$\widehat{\theta}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$, $\widehat{\lambda}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$ and $\widehat{\gamma}_{\text{FDH}}(x, y \mid \mathcal{S}_n)$. However, this leads to integer programming problems, but the estimators can be computed using simple numerical methods.[2]

The statistical properties of these efficiency estimators are well-developed. Kneip et al. (1998) derive the rate of convergence of the input-oriented VRS estimator, while Kneip et al. (2008) derive its limiting distribution. Park et al. (2010) derive the rate of convergence of the input-oriented CRS estimator and establish its limiting distribution. Park et al. (2000) and Daouia et al. (2017) derive both the rate of convergence and limiting distribution of the input-oriented FDH estimator. These results extend trivially to the output orientation after straightforward (but perhaps tedious) changes in notation. Wheelock and Wilson (2008) extend these results to the hyperbolic FDH estimator, and Wilson (2011) extends the results to the hyperbolic DEA estimator.

Kneip et al. (2015) derive moment properties of both the input-oriented FDH, VRS and CRS estimators and establish central limit theorem (CLT) results for mean input-oriented efficiency after showing that the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold unless $(p+q) < 4$ in the CRS case, $(p+q) < 3$ in the VRS case, or unless $p+q < 2$ in the FDH case.[3] Kneip et al. (2016) use these CLT results to establish asymptotically normal test statistics for testing differences in mean efficiency across two groups, convexity versus non-convexity of Ψ, and CRS versus VRS (provided Ψ is weakly convex).[4] All of these results extend trivially (but again, tediously) to the output-oriented FDH, VRS and CRS estimators. These results could also be extended to the hyperbolic VRS and CRS estimators following Wilson (2011). The hyperbolic FDH estimator can be viewed as an input-oriented FDH estimator in a transformed space, hence moment results for the hyperbolic FDH estimator could also be extended trivially (but again, tediously) from the input-oriented FDH estimator. The new CLT results of Kneip et al. (2015) as well as the results from Kneip et al. (2016) on tests of differences in means, convexity versus non-convexity of Ψ, and CRS versus VRS carry over to the hyperbolic FDH estimator.

To summarize, in all cases, the FDH, VRS and CRS estimators are consistent, converge at rate n^{κ} (where $\kappa = 1/(p+q)$ for the FDH estimators, $\kappa = 2/(p+q+1)$ for the VRS estimators and $\kappa = 2/(p+q)$ for the CRS estimators) and possess non-degenerate limiting distributions under the appropriate set of assumptions. In addition, the bias of each of the three estimators is of order

[2]For details, see Kneip et al. (2015) and Wilson (2011).

[3]In other words, standard CLT results hold in the FDH case if and only if $p = 1$ and output is fixed and constant, or $q = 1$ and input is fixed and constant.

[4]If Ψ^{∂} is globally CRS, then the VRS estimator attains the faster convergence rate of the CRS estimator due to the Theorem 3.1 of Kneip et al. (2016).

$O\left(n^{-\kappa}\right)$. Bootstrap methods proposed by Kneip et al. (2008, 2011) and Simar and Wilson (2011) provide consistent inference about $\theta(x, y \mid \Psi)$, $\lambda(x, y \mid \Psi)$ and $\gamma(x, y \mid \Psi)$ for a fixed point $(x, y) \in \Psi$, and Kneip et al. (2015) provide new CLT results enabling inference about the expected values of these measures over the random variables (X, Y).

Additional technical assumptions required for moment properties and central limit theorem results of means of FDH, VRS and CRS estimates, established by Kneip et al. (2015) and used below are given in Appendix A.

3.4 Data and Variable Specification

We calculate the efficiency and productivity for 144 countries over 2004–2014 using data from version 9.0 of the Penn World Table, PWT9.0 (Feenstra et al. (2015)).[5] Following Glass et al. (2016), the output Y is output-side real GDP at chained PPPs (in million 2011 US$, *rgdpo*). As recommended in the documentation, we use *rgdpo* to measure productive capacity across countries rather than expenditure-side real GDP (*rgdpe*) or real GDP at constant 2011 national prices (*rgdpna*). The first input X_1 is the labor input, measured by the number of persons engaged (in millions, *emp*). The second input X_2 is the capital stock at current PPPs (in million 2011 US$, *ck*). Maudos et al. (1999) mentioned that the inclusion of human capital has a significant effect on the accurate measurement of productivity. Therefore, the third input X_3 is human capital stock (*hc*), one index based on years of schooling and returns to education (see Human capital in PWT9.0). The input-output specification is standard (Färe et al. (1994)), reflecting the basic production process of countries. Table 3.1 shows the summary statistics for year 2014.

We assume that all countries operate in the same production set Ψ defined by (2.1), and therefore they face the same frontier in the four-dimensional input-output space. The model described in Section 3.2 is fully non-parametric, and hence quite flexible. The assumptions listed in Section 3.2 impose only minimal restrictions involving free-disposability, continuity, and some smoothness of the frontier, etc. Note that there is no assumption of convexity of Ψ, which is tested below in Section 3.5.

The flexibility of the non-parametric model specified in Section 3.2 comes with a price, however, in terms of the well-known "curse of dimensionality". The convergence rate of non-parametric

[5] For more information about this data, see https://www.rug.nl/ggdc/productivity/pwt/.

efficiency estimators decreases with increasing inputs and outputs. The number of observations in each period that we consider is 144. With $p = 3$ and $q = 1$, the effective parametric sample size defined by Wilson (2018) is $144^{\frac{2}{4}} \approx 12$ for FDH estimators, $144^{\frac{4}{5}} \approx 53$ for VRS estimators and $144^{\frac{4}{4}} \approx 144$ for CRS estimators. With the sample size of 144 and the highest converge rate of $n^{\frac{2}{4}}$, non-parametric estimators should be expected to result in estimation error of order no better than that one would achieve with 144 observations in a typical parametric estimator. Given the relatively small sample size and the four dimensions, it is not surprising that the estimated efficiency for many countries is equal to 1 and hence is not reliable.

To address this, the dimension reduction technique proposed by Wilson (2018) is applied. Considering the $(n \times p)$ and $(n \times q)$ matrices $\mathbf{X}$ and $\mathbf{Y}$ of observed non-negative inputs and outputs, we compute the $(n \times 1)$ vectors of principle components $X^* = \boldsymbol{X}\Lambda_x$ and $Y^* = \boldsymbol{Y}\Lambda_y$, where Λ_x and Λ_y are the $(p \times 1)$ and $(q \times 1)$ eigenvectors corresponding to the largest eigenvalues of $\mathbf{X'X}$ and $\mathbf{Y'Y}$, respectively. The dimensions of both $\mathbf{X}$ and $\mathbf{Y}$ are then reduced to only one dimension. However, we need to examine R_x and R_y, which are the ratios of the largest eigenvalue of the moment matrices $\mathbf{X'X}$ and $\mathbf{Y'Y}$ to the corresponding sums of the eigenvalues for these moment matrices. Wilson (2018) mentions that R_x and R_y provide measures of how close the corresponding moment matrices are to rank-one, regardless of the joint distributions of inputs and outputs.

The eigensystem analysis of input moment matrix is shown in Table 3.2. Since we only have one output, there is no need of dimension reduction for output $\mathbf{Y}$. The columns give the values of R_x from 2004 to 2014. The results are quite similar across years. An eigensystem analysis on the full data yields $R_x \geq 87.43$ percent for all years. It is clear that X^* contains most of the independent information of $\boldsymbol{X}$. Wilson (2018) shows that in many cases, but not in general, this dimension reduction method is advantageous in terms of reducing efficiency estimation error. In addition, dimension reduction could significantly increase the convergence rate of non-parametric efficiency estimators and lead to a more accurate estimation of efficiency. After dimension reduction applied, the convergence rates for FDH, VRS, and CRS are $n^{\frac{1}{2}}$, $n^{\frac{2}{3}}$ and n respectively. The tradeoff is that a small amount of information may be lost, but the mean squared error is reduced. All estimation in the following is done using X^* and Y.

3.5 Empirical Results

3.5.1 Efficiency and Productivity Evolution

As a robustness check to the need for dimension reduction, we estimate the hyperbolic efficiency for each year first using full data with four dimensions, and then using reduced data with only two dimensions. The FDH, VRS and CRS estimators are applied for both cases. Table 3.3 shows the number of observations with estimates equal to one. For each year, the FDH estimator produces more estimates equal to 1 than the VRS estimator, which produces more estimates equal to 1 than the CRS estimator. This is expected since there are more restrictions for the CRS estimator than the VRS estimator, which has more restrictions than the FDH estimator. More importantly, when using the full data, the FDH estimator results in 70.14 percent to 77.08 percent of the observations in a given year with estimates equal to one. The proportions for the VRS are between 15.28 percent and 17.36 percent and for the CRS are between 6.94 percent and 10.42 percent. This is clear evidence of too many dimensions for the given sample size. With dimension reduction, Table 3.3 shows that when using either estimator for any given year, the number of observations with estimates equal to 1 is much smaller than that without dimension reduction. In addition, the numbers using the FDH estimator are at least 3 times those using the VRS estimator, suggesting that the production set Ψ may be non-convex. In addition to large values of R_x discussed in Section 3.4, Table 3.3 provides another piece of evidence that dimension reduction likely reduces estimation error relative to what would be obtained when using the full data without dimension reduction. Therefore, the principal component for the three inputs X^* and the single output Y described in Section 3.4 are used for obtaining all the following results.

The next question is to determine which estimator we should use. As discussed in Section 3.2, in decreasing order of restrictions and rates of convergence lies the CRS, VRS, and FDH estimators. Using the test developed by Kneip et al. (2016), we use the principal component X^* and Y to test the null hypothesis of convexity of the production set Ψ versus the alternative hypothesis that Ψ is non-convex. Two randomly splitted subsamples for a given year are needed for this test. The first subsample of size $n_1 = \lfloor n/2 \rfloor$ is used for computing VRS estimates, and the second subsample of size $n_2 = n - n_1$ is used for computing FDH estimates. As Daraio et al. (2018) suggest, we first randomly shuffle the data using their randomization algorithm, and then take the first n_1 observations as the first subsample for computing VRS estimates, and the remaining n_2 observations as the second subsample for computing FDH estimates. The test statistic given in equation (50) of Kneip et al. (2016) involves the difference of the means of these two sets of estimates, with generalized jackknife estimates of biases and corresponding sample variances, and is asymptotically normally distributed with mean zero and unit variance. The test is a one-sided test since under the null the two means should be roughly similar, but should diverge with increasing departures from the null resulting in the mean of the FDH estimates exceeding the mean of the VRS estimates. The statistic given in equation (50) of Kneip et al. (2016) is defined in terms of input-oriented estimators but extends trivially to output-oriented and hyperbolic estimators. The tests are one-sided and we define the statistics so that "large" positive values indicate rejection of the null hypothesis.

The results of the convexity tests for each year are shown in Table 3.4. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.01. The results reveal that convexity is overwhelmingly rejected except the one in the input direction of 2004. Hence, the results in Table 3.4 provide strong evidence of the non-convexity of the production set Ψ. Also, even if the production set is convex, FDH estimator is still consistent. However, if the production set is non-convex, the VRS estimator is not consistent anymore. Consequently, the FDH estimators are applied for the remainder of the analysis. Our results cast doubts on the results of previous literature, which use DEA estimators to estimate the Malmquist productivity and technical efficiency.

Table 3.5 presents summary statistics of the FDH technical efficiency estimates in the input, output, and hyperbolic orientations. To compare with the input-oriented and hyperbolic-oriented estimates, we report the statistics of the reciprocals of the output-oriented estimates. For each orientation, the closer the estimates are to 1, the more technically efficient the countries. As might be expected, the hyperbolic estimates are more conservative on average, with mean efficiencies ranging from 0.7799 to 0.8252. By contrast, the means of the input-oriented estimates range from 0.6113 to 0.7139, while the means of the output-oriented estimates range from 0.6499 to 0.7146. These differences are due to the geometry of the efficiency measures as discussed by Wilson (2011). The mean efficiency in hyperbolic orientation increased from 2004 to 2010 and then decreased from 2010 to 2014. The pattern of mean efficiency in the input orientation appears to be the same, while the pattern in the output orientation is a little different. The mean efficiency in the output orientation increased from 2004 to 2008 and then decreased from 2008 to 2014.

We use the test described by Kneip et al. (2016, Section 3.1.1) to test for significant differences between the means reported in Table 3.5 from one year to the next, as well as from the first year to the last year. As discussed in Kneip et al. (2015, 2016), even with the reduced dimensionality so that $p + q = 2$, the usual CLT results (e.g., the Lindeberg-Feller CLT) do not hold for means of FDH efficiency estimates. As with the convexity test discussed above, the test statistic given by equation (18) of Kneip et al. (2016) involves not only the difference in sample means of efficiency estimates in a pair of years, but also the corresponding difference in generalized jackknife estimates of bias. The test extends trivially to the output-orientation and the hyperbolic orientation. In each case, the statistic used here is defined so that a positive value indicates that efficiency increases from year 1 to year 2, while a negative value indicates that efficiency decreases from year 1 to year 2.[6] As shown by Kneip et al. (2016), the test statistics are asymptotically normal with zero mean and unit variance. Since our data is balanced panel, there may exist time correlation, which violates the independent assumption of the test for differences of mean efficiency. The technical details dealing with time correlations is given in Appendix B Section B.1.

Table 3.6 gives the results of the tests of significant differences in mean efficiency over time. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. The result from 2004 to 2006 is mixed, where one case (input-oriented) shows that mean efficiency increased, while another case (output-oriented) shows that mean efficiency decreased. From 2006 to 2008, while there was no change of mean efficiency in the input and output orientation, mean efficiency declined significantly in the hyperbolic orientation. This possibly reflects the negative effect of the global financial crisis. The tests provide clear evidence that mean efficiency increased from 2008 to 2010 while there was no change of mean efficiency from 2010 to 2012. From 2012 to 2014, we see that mean efficiency started increasing again, showing that the global economy finally recovered from the global financial crisis. Overall, from 2004 to 2014, even though the statistic in the hyperbolic orientation is not significant, mean efficiency increased significantly in the other two orientations. Therefore, we find that the technical efficiency declined at the start of the global financial crisis (2006–2008), but recovered in the years later (2008–2014), ending higher in 2014 than in 2004.

In order to measure productivity, note that with the dimension reduction to $(p + q) = 2$ dimensions using the principal components X_i^*, Y_i as described in Section 3.4, productivity can be measured by Y_i/X_i^* for country i. Summary statistics for this measure is displayed in Table 3.7. Mean productivity continuously decreased from 2004 to 2010, then continuously increased from

[6] Consequently, the statistic we use for the output orientation is the negative of the statistic appearing in equation (18) of Kneip et al. (2016).

2010 to 2014. The results show that the pattern of mean productivity is similar to the pattern of mean efficiency. Since productivity is measured by a simple ratio that does not involve estimators of efficiency, standard CLT results can be used to test for significant changes in means over time.[7] The results of these tests are shown in Table 3.8. Cells in column 7 are shaded whenever p-value is less than 0.01. It shows that mean productivity continued decreasing significantly from 2004 to 2010 and there was no change from 2010 to 2012 and from 2012 to 2014. Overall, from 2004 to 2014, the data reveals that there was a significant decrease in mean productivity.

To learn more about the difference of distributions of productivity in the two years interval, we use the stochastic dominance tests developed by Linton et al. (2005). Their method is based on subsampling and allows for the observations to be serially dependent. The outcome for the first-order stochastic dominance test is shown in Table 3.9. Cells in columns 3 and 5 are shaded whenever we could not reject the null hypothesis. Denote stochastic dominance in the first and second-order as SD1 and SD2, respectively. The p-value for the null hypothesis that productivity in 2004 SD1 that in 2006, is 0.999, which suggests that 2004 SD1 2006. The p-value for the null hypothesis that productivity in 2006 SD1 that in 2004, is close to 0, which suggests that 2006 does not SD1 2004 and hence the possibility that 2004 and 2006 have the same distribution is ruled out. Combining these two tests, we find that 2004 SD1 2006. Similarly, we also find that 2006 SD1 2008, 2008 SD1 2010 and 2004 SD1 2014. However, we do not find any SD1 over the periods 2010–2012 and 2012–2014. We also test whether there exists any SD2 in the two years interval. The outcome for the second-order stochastic dominance test is shown in Table 3.10. Cells in columns 3 and 5 are shaded whenever we could not reject the null hypothesis. Since SD1 implies SD2, it is not surprising that 2004 SD2 2006, 2006 SD2 2008, 2008 SD2 2010, and 2004 SD2 2014. Moreover, there did not exist SD2 over the periods 2010–2012 and 2012–2014.

The results presented so far provide clear evidence of changes in mean technical efficiency and productivity over the years represented in the sample. To gain further insight, we test whether the frontiers change over time. This involves the test of "separability" developed by Daraio et al. (2018), in which time is treated as a binary "environmental" variable. We examine it using pairs of years 2004–2006, 2006–2008, . . . , 2012–2014 as well as 2004–2014.

Implementation of the separability test of Daraio et al. (2018) involves pooling the data for two periods and then randomly shuffling the observations using the randomization algorithm

[7] However, we need to deal with time correlation. Please see Appendix B Section B.2 for technical details.

presented by Daraio et al.. Then the pooled, randomly shuffled observations are split into two subsamples of equal size (or, if the combined number of observations is odd, one subsample will have one more observation than the other). Using the first subsample, efficiency is estimated as usual for each observation, ignoring which period a particular observation comes from, and the sample mean of the efficiency estimate is computed. The second subsample is split into the set of observations from period 1 and the set of observations from period 2. Efficiency is estimated for the period 1 observations using only the observations from period 1, while efficiency for the period 2 observations is estimated using only those observations from period 2. Then the sample mean of these two sets of efficiency estimates from the two sub-subsamples (of the second subsample) is computed. The resulting test statistic involves differences in the two subsample means as well as differences in the corresponding generalized jackknife estimates of bias. See Daraio et al. (2018) for discussion and details.

Results of the separability tests are shown in Table 3.11. Cells in columns 3, 5 and 7 are shaded whenever p-value is less than 0.10. In every case of periods 2004–2006, 2008–2010, 2010–2012 and 2012–2014, separability is rejected with p-value less than 0.01, and in most cases well less than 0.01. From 2006 to 2008, two statistics are significant at the 1 percent level, while the remaining statistic is not significant at the 10 percent level. Therefore the data provides moderate evidence that the technology changed from 2006 to 2008. Overall, from 2004 to 2014, two statistics are significant at the 10 percent level, while the remaining statistic is not significant at all. Therefore the evidence shows that the technology changed from 2004 to 2014.

In order to learn something about the *direction* in which technology may have shifted, we use new results from Simar and Wilson (2018) who provide CLT results for components of productivity changed measured by Malmquist indices. Simar and Wilson define the Malmquist index in terms of hyperbolic distances, and then consider various decompositions that can be used to identify components of productivity change. In particular, let Ψ^t represent the production set at time $t \in \{1,\ 2\}$ and let $Z_i^t = (X_i^t, Y_i^t)$ denote the i-th firm's observed input-output pair at time t. Then technical change relative to firm i's position at times 1 and 2 is measured by

$$\mathcal{T}_i = \left[\frac{\gamma(Z_i^2 \mid \Psi^1)}{\gamma(Z_i^2 \mid \Psi^2)} \times \frac{\gamma(Z_i^1 \mid \Psi^1)}{\gamma(Z_i^1 \mid \Psi^2)} \right]^{1/2}. \tag{5.1}$$

This is the hyperbolic analog of the output-oriented technical-change index that appears in the decompositions of Ray and Desli (1997), Gilbert and Wilson (1998), Simar and Wilson (1998) and Wheelock and Wilson (1999). The first ratio inside the brackets in (5.1) measures technical change relative to firm i's position at time 2, while the second ratio measures technical change relative to the firm's position at time 1. The measure $\mathcal{T}_i$ is the geometric mean of these two ratios. Values greater than 1 indicate an upward shift in the technology, while values less than 1 indicate a downward shift (a value of 1 indicates no change from time 1 to time 2).

Estimates $\widehat{\mathcal{T}_i}$ are obtained by substituting the hyperbolic FDH estimator for each term in (5.1). Simar and Wilson (2018) develop CLT results for geometric means $\widehat{T}^{1,2}$ of $\mathcal{T}_i$ over firms $i = 1, \ldots, n$, for periods 1 and 2, and these results can be used to test significant differences of the geometric means from 1. Table 3.12 shows the results of tests of technology change for each two-year interval as well as for the entire period 2004–2014. Cells in column 7 are shaded whenever p-value is less than 0.01. All the statistics are significant at the 1 percent level. The geometric mean $\widehat{T}^{1,2}$ is smaller than 1 for each two-year interval from 2004 to 2010, and greater than 1 for each two-year interval from 2010 to 2014. This suggests continuing downward shifts of the technology from 2004 to 2010 and continuing upward shifts of the technology from 2010 to 2014. Over the full period 2004–2014, $\widehat{T}^{1,2}$ is less than 1 significantly at the 1 percent level, suggesting that the technology shifted downward over this period.

3.5.2 Developing Versus Developed Countries

The convergence theory, also known as the catch-up effect, implies that developing countries will tend to grow at faster rates than developed countries. Therefore, developing countries should have higher productivity and efficiency. Our tests developed in Section 3.2 and 3.3 could be used to examine this hypothesis.

According to *the International Monetary Fund's World Economic Outlook Database, October 2018*, the following are considered as developed economies (or advanced economies): United States, Japan, Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Portugal, Slovak Republic, Slovenia, Spain, Canada, United Kindom, Australia, Korea, Singapore, Czech Republic, Macao SAR, Sweden, Denmark, New Zealand, Switzerland, Hongkong SAR, Norway,Taiwan Province of China, Iceland, Puerto Rico, Israel, San Marino. However, our data do not cover Puerto Rico and San Marino. All the other remaining countries are considered as developing economies (or emerging markets and developing economies). Our sample covers 37 developed economies and 107 developing economies. Table 3.13 shows the summary statistics of developing and developed economies for the year 2014. As it is expected, developing economies averagely have lower GDP, capital and human capital but higher labor than developed economies. The annual growth rates of labor, capital, human capital, and GDP are shown in Table 3.14. During the global financial crisis (2008–2009), the total output of the world economy decreased by about 1.12 percent, which mainly caused by developed countries (decreased 4 percent). Over each period from 2004 to 2014, the developing countries almost always had higher annual growth rates of labor, capital, human capital, and GDP. Overall, from 2004 to 2014, developing economies doubled their GDP, while the GDP of developed economies only increased by about 41 percent.

Table 3.15 shows the results of tests of the difference in mean technical efficiency between developing and developed economies. Cells in columns 5, 7 and 9 are shaded whenever p-value is less than 0.01. All the statistics are negative and significant at the 1 percent significance level. This suggests that developing economies had lower technical efficiency than developed economies over each period covered by our data. Table 3.16 provides the results of tests of the difference in productivity between developing and developed economies. Cells in column 7 are shaded whenever p-value is less than 0.01. All the statistics are positive and significant at the 1 percent significance level. This result shows that developing economies had higher productivity than developed economies over each period. Taken together, we find that over 2004–2014, even though developing economies had lower technical efficiency, they had higher productivity than developed economies. This suggests that some developing economies have not fully adopted the current technologies. Our results confirms the convergence theory.

3.6 Summary and Conclusions

Among studies that use either FDH or DEA estimators to estimate efficiency and benchmark the performances of countries, the vast majority use VRS (DEA) estimators which impose convexity on the production set. The test of convexity versus non-convexity of Ψ developed by Kneip et al. (2016) allows researchers to let the data tell them whether DEA estimators are appropriate in a given setting. Here, in the context of countries, convexity is strongly rejected.

Because we reject convexity of the production set, we use FDH estimators which remain consistent when Ψ is not convex, whereas DEA estimators do not. We exploit collinearity in the data to reduce inputs to their first principle components, resulting in a two-dimensional problem. Results from Wilson (2018) indicate that this substantially reduces mean square error of efficiency estimates. Moreover, the simulation evidence provided by Wilson (2018) suggests that when production sets are convex, FDH estimates often have less mean square error than DEA estimators after dimension reduction.

By rigorously comparing estimates and testing differences across the years represented in our data, we find that technical efficiency of 144 countries in the world declined at the start of the global financial crisis (2006–2008) but recovered in the years later (2008–2014). Overall, there was an increase in mean efficiency from 2004 to 2014. The data revealed that productivity continued decreasing from 2004 to 2010. Overall, there was a significant decrease in mean productivity from 2004 to 2014. We also find that the frontier continued shifting downward from 2004 to 2010, and then continued shifting upward from 2010 to 2014. However, the technology had shifted downward from 2004 to 2014. Finally, the data revealed that developing economies had lower technical efficiency but higher productivity than developed economies over this period.

The 2008 global financial crisis indeed had an influential negative effect on efficiency, productivity, and technology of the global economy. Even though the global economy recovered in the years later, however, until 2014, 6 years after the crisis, our results show that the productivity and technology of the global economy had not fully recovered yet. Over this period, developing economies performed better than developed economies in terms of productivity, however, they need to improve their technical efficiency.

Table 3.1: Summary Statistics for Year 2014

Variable	Min	Q1	Median	Mean	Q3	Max
GDP (Y)	$2.7830 \times 10^{+03}$	$3.2340 \times 10^{+04}$	$1.1820 \times 10^{+05}$	$7.1440 \times 10^{+05}$	$4.6660 \times 10^{+05}$	$1.7140 \times 10^{+07}$
Labor (L)	1.2370×10^{-01}	$1.8400 \times 10^{+00}$	$4.9120 \times 10^{+00}$	$2.1850 \times 10^{+01}$	$1.4460 \times 10^{+01}$	$7.9840 \times 10^{+02}$
Capital (K)	$5.1750 \times 10^{+03}$	$1.1140 \times 10^{+05}$	$3.9030 \times 10^{+05}$	$2.6430 \times 10^{+06}$	$1.8000 \times 10^{+06}$	$6.9380 \times 10^{+07}$
Human Capital (H)	$1.1930 \times 10^{+00}$	$2.0180 \times 10^{+00}$	$2.6620 \times 10^{+00}$	$2.5950 \times 10^{+00}$	$3.1560 \times 10^{+00}$	$3.7340 \times 10^{+00}$

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education.

Table 3.2: Eigensystem Analysis by Year

Year	$R_x(\%)$
2004	87.43
2006	87.70
2008	88.02
2010	88.24
2012	88.53
2014	88.76

Table 3.3: Numbers of Observations With Estimated Hyperbolic Technical Efficiency Equal to 1 in Each Year

		Without — Dimension Reduction —			With — Dimension Reduction —		
Year	n	FDH	VRS	CRS	FDH	VRS	CRS
2004	144	105	24	15	27	7	1
2006	144	105	23	11	29	8	1
2008	144	105	22	13	32	8	1
2010	144	111	25	13	36	6	1
2012	144	101	22	10	29	6	1
2014	144	109	25	11	27	6	1

Table 3.4: Results of Convexity Tests (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004	1.4720	7.05×10^{-02}	5.0490	2.22×10^{-07}	3.7056	1.05×10^{-04}
2006	6.2603	1.92×10^{-10}	9.0609	6.47×10^{-20}	7.0398	9.62×10^{-13}
2008	3.5480	1.94×10^{-04}	4.8906	5.03×10^{-07}	3.5020	2.31×10^{-04}
2010	8.4456	1.51×10^{-17}	9.6505	2.45×10^{-22}	9.1008	4.48×10^{-20}
2012	7.5268	2.60×10^{-14}	8.2257	9.71×10^{-17}	8.8921	3.00×10^{-19}
2014	3.0405	1.18×10^{-03}	3.1071	9.45×10^{-04}	3.8103	6.94×10^{-05}

NOTE: The numerator of statistics is the difference of estimated mean VRS estimates minus mean FDH estimates.

Table 3.5: Summary Statistics for FDH Technical Efficiency Estimates (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
			— **Input Orientation** —			
2004	0.1139	0.3785	0.6303	0.6113	0.8899	1.0000
2006	0.1972	0.4616	0.6438	0.6609	0.9033	1.0000
2008	0.2097	0.5135	0.6727	0.6975	0.9557	1.0000
2010	0.2084	0.5239	0.6984	0.7139	0.9953	1.0000
2012	0.2150	0.5038	0.6624	0.6742	0.9218	1.0000
2014	0.1858	0.4150	0.5896	0.6338	0.8684	1.0000
			— **Output Orientation** —			
2004	0.1284	0.4391	0.6477	0.6499	0.9217	1.0000
2006	0.1245	0.5108	0.6970	0.6917	0.9323	1.0000
2008	0.1705	0.5104	0.7415	0.7146	0.9623	1.0000
2010	0.2009	0.5088	0.6706	0.7014	0.9815	1.0000
2012	0.2255	0.4649	0.6501	0.6802	0.9297	1.0000
2014	0.1859	0.4158	0.6589	0.6500	0.8961	1.0000
			— **Hyperbolic Orientation** —			
2004	0.3516	0.6371	0.7840	0.7799	0.9437	1.0000
2006	0.4053	0.6863	0.8300	0.8079	0.9470	1.0000
2008	0.4781	0.6930	0.8312	0.8211	0.9819	1.0000
2010	0.4919	0.6914	0.8278	0.8252	0.9953	1.0000
2012	0.4539	0.6633	0.8129	0.8051	0.9670	1.0000
2014	0.4082	0.6783	0.7840	0.7923	0.9403	1.0000

NOTE: Statistics for the reciprocals of the output efficiency estimates are given to facilitate comparison with the input-oriented and hyperbolic estimates.

Table 3.6: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004–2006	2.0874	3.69×10^{-02}	−2.6490	8.07×10^{-03}	−0.6728	5.01×10^{-01}
2006–2008	0.1634	8.70×10^{-01}	0.6564	5.12×10^{-01}	−2.5097	1.21×10^{-02}
2008–2010	2.2908	2.20×10^{-02}	1.5509	1.21×10^{-01}	2.3430	1.91×10^{-02}
2010–2012	−0.4524	6.51×10^{-01}	−0.7692	4.42×10^{-01}	−0.1788	8.58×10^{-01}
2012–2014	2.8697	4.11×10^{-03}	3.5162	4.38×10^{-04}	3.9343	8.34×10^{-05}
2004–2014	1.8283	6.75×10^{-02}	2.4703	1.35×10^{-02}	1.3329	1.83×10^{-01}

NOTE: The numerator of statistics for each period is the difference of estimated mean efficiency of the second year minus the first year.

Table 3.7: Summary Statistics for Productivity (with Dimension Reduction, $p = q = 1$)

Year	Min	Q1	Median	Mean	Q3	Max
2004	1.5300	2.7583	3.3940	4.0248	4.5488	18.3875
2006	1.4066	2.5806	3.0827	3.6276	4.2816	13.5363
2008	1.3057	2.3642	2.9638	3.3147	3.7893	9.4944
2010	1.2835	2.2572	2.7112	3.0851	3.7210	8.7854
2012	1.2213	2.1914	2.8185	3.1122	3.7039	7.9003
2014	1.2955	2.2069	2.7949	3.1541	3.7377	11.2910

NOTE: Productivity for observation i is defined as Y_i/X_i^*.

Table 3.8: Tests of Differences in Means for Productivity Estimates with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	Mean1	Mean2	Statistic	p-value
2004–2006	144	144	4.0248	3.6276	−5.2576	1.46×10^{-07}
2006–2008	144	144	3.6276	3.3147	−5.0137	5.34×10^{-07}
2008–2010	144	144	3.3147	3.0851	−4.6205	3.83×10^{-06}
2010–2012	144	144	3.0851	3.1122	0.5699	5.69×10^{-01}
2012–2014	144	144	3.1122	3.1541	1.0478	2.95×10^{-01}
2004–2014	144	144	4.0248	3.1541	−4.6766	2.92×10^{-06}

NOTE: The numerator of statistics for each period is the difference of estimated mean productivity of the second year minus the first year.

Table 3.9: First Order Stochastic Dominance Test for Productivity with Respect to Time

	— **Year 1 SD1 Year 2** —		— **Year 2 SD1 Year 1** —	
Period	Statistic	p-value	Statistic	p-value
2004–2006	0.0000	9.99×10^{-01}	0.1250	9.99×10^{-04}
2006–2008	−0.0069	9.99×10^{-01}	0.0972	6.99×10^{-03}
2008–2010	0.0000	9.99×10^{-01}	0.1250	$0.00 \times 10^{+00}$
2010–2012	0.0625	9.19×10^{-02}	0.0417	5.54×10^{-01}
2012–2014	0.0417	2.76×10^{-01}	0.0278	7.60×10^{-01}
2004–2014	−0.0208	9.99×10^{-01}	0.2361	$0.00 \times 10^{+00}$

NOTE: The null hypothesis is that there exists first order stochastic dominance.

Table 3.10: Second Order Stochastic Dominance Test for Productivity with Respect to Time

	— **Year 1 SD2 Year 2** —		— **Year 2 SD2 Year 1** —	
Period	Statistic	p-value	Statistic	p-value
2004–2006	-0.0013	9.84×10^{-01}	0.3105	$0.00 \times 10^{+00}$
2006–2008	-0.0010	9.92×10^{-01}	0.2490	$0.00 \times 10^{+00}$
2008–2010	-0.0002	9.80×10^{-01}	0.1937	$0.00 \times 10^{+00}$
2010–2012	0.0229	4.10×10^{-01}	0.0195	4.21×10^{-01}
2012–2014	0.0130	4.35×10^{-01}	0.0017	7.77×10^{-01}
2004–2014	-0.0018	9.86×10^{-01}	0.7166	$0.00 \times 10^{+00}$

NOTE: The null hypothesis is that there exists second order stochastic dominance.

Table 3.11: Tests for Separability with Respect to Time (with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Period	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004–2006	5.0232	2.54×10^{-07}	2.4078	8.03×10^{-03}	5.4435	2.61×10^{-08}
2006–2008	1.2211	1.11×10^{-01}	5.5903	1.13×10^{-08}	2.9588	1.54×10^{-03}
2008–2010	3.3544	3.98×10^{-04}	4.8669	5.67×10^{-07}	3.7419	9.13×10^{-05}
2010–2012	4.3189	7.84×10^{-06}	5.9273	1.54×10^{-09}	3.1343	8.61×10^{-04}
2012–2014	2.8978	1.88×10^{-03}	4.2091	1.28×10^{-05}	2.7058	3.41×10^{-03}
2004–2014	1.4604	7.21×10^{-02}	−2.0830	9.81×10^{-01}	1.3939	8.17×10^{-02}

NOTE: The numerator of the statistics is the difference of the conditional mean estimates minus the unconditional mean estimates.

Table 3.12: Tests for Technology Change with Respect to Time (with Dimension Reduction, $p = q = 1$)

Period	n_1	n_2	n	$\widehat{T}^{1,2}$	Var	p-value
2004–2006	144	144	144	0.9348	0.0247	1.69×10^{-07}
2006–2008	144	144	144	0.9508	0.0148	1.49×10^{-09}
2008–2010	144	144	144	0.9670	0.0158	8.99×10^{-03}
2010–2012	144	144	144	1.0349	0.0174	9.18×10^{-11}
2012–2014	144	144	144	1.0283	0.0188	1.07×10^{-06}
2004–2014	144	144	144	0.9162	0.0216	4.78×10^{-10}

NOTE: For each period, the number of observations in the first year is n_1, while the number of observations in the second year is n_2. The number of observations existing in both years is n. Mean of the technology ratio $\widehat{T}^{1,2}$ is greater than 1 if and only if the technology shifts upward.

Table 3.13: Summary Statistics of Developing and Developed Economies for Year 2014

Variable	Min	Q1	Median	Mean	Q3	Max
			— Developing Economies —			
GDP (Y)	$2.7830 \times 10^{+03}$	$2.7320 \times 10^{+04}$	$7.7750 \times 10^{+04}$	$5.4090 \times 10^{+05}$	$4.0300 \times 10^{+05}$	$1.7140 \times 10^{+07}$
Labor (L)	1.2370×10^{-01}	$1.8390 \times 10^{+00}$	$5.5490 \times 10^{+00}$	$2.4780 \times 10^{+01}$	$1.4550 \times 10^{+01}$	$7.9840 \times 10^{+02}$
Capital (K)	$5.1750 \times 10^{+03}$	$8.5640 \times 10^{+04}$	$2.2450 \times 10^{+05}$	$1.8760 \times 10^{+06}$	$1.2200 \times 10^{+06}$	$6.9380 \times 10^{+07}$
Human Capital (H)	$1.1930 \times 10^{+00}$	$1.8250 \times 10^{+00}$	$2.4430 \times 10^{+00}$	$2.3420 \times 10^{+00}$	$2.7800 \times 10^{+00}$	$3.4110 \times 10^{+00}$
			— Developed Economies —			
GDP (Y)	$1.0030 \times 10^{+04}$	$1.3560 \times 10^{+05}$	$3.2620 \times 10^{+05}$	$1.2160 \times 10^{+06}$	$1.0450 \times 10^{+06}$	$1.6600 \times 10^{+07}$
Labor (L)	1.8330×10^{-01}	$1.8520 \times 10^{+00}$	$3.9880 \times 10^{+00}$	$1.3380 \times 10^{+01}$	$1.1970 \times 10^{+01}$	$1.4850 \times 10^{+02}$
Capital (K)	$4.0060 \times 10^{+04}$	$4.5960 \times 10^{+05}$	$1.7330 \times 10^{+06}$	$4.8630 \times 10^{+06}$	$3.9450 \times 10^{+06}$	$5.2850 \times 10^{+07}$
Human Capital (H)	$2.4270 \times 10^{+00}$	$3.1180 \times 10^{+00}$	$3.3640 \times 10^{+00}$	$3.3270 \times 10^{+00}$	$3.5940 \times 10^{+00}$	$3.7340 \times 10^{+00}$

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education.

Table 3.14: Annual Growth Rate of Labor, Capital, Human Capital and GDP over 2004-2014

Period	L	K	H	Y
— **All Economies** —				
2004–2005	0.0289	0.1354	0.0082	0.1053
2005–2006	0.0296	0.1300	0.0082	0.0744
2006–2007	0.0315	0.1063	0.0082	0.0737
2007–2008	0.0250	0.0987	0.0081	0.0753
2008–2009	0.0112	0.0661	0.0080	-0.0112
2009–2010	0.0181	0.0877	0.0078	0.0944
2010–2011	0.0213	0.1214	0.0091	0.0939
2011–2012	0.0179	0.0535	0.0090	0.0462
2012–2013	0.0179	0.0594	0.0089	0.0360
2013–2014	0.0182	0.0608	0.0089	0.0342
2004–2014	0.2639	1.5658	0.0897	0.8618
— **Developed Economies** —				
2004–2005	0.0178	0.0975	0.0061	0.0752
2005–2006	0.0252	0.1338	0.0063	0.0303
2006–2007	0.0273	0.1069	0.0063	0.0638
2007–2008	0.0175	0.0816	0.0063	0.0352
2008–2009	-0.0180	0.0328	0.0063	-0.0400
2009–2010	-0.0014	0.0458	0.0062	0.0498
2010–2011	0.0111	0.0600	0.0062	0.0416
2011–2012	0.0055	0.0502	0.0063	0.0283
2012–2013	0.0049	0.0429	0.0063	0.0232
2013–2014	0.0112	0.0319	0.0064	0.0278
2004–2014	0.1154	0.9737	0.0656	0.4145
— **Developing Economies** —				
2004–2005	0.0327	0.1485	0.0090	0.1157
2005–2006	0.0312	0.1287	0.0089	0.0896
2006–2007	0.0330	0.1061	0.0088	0.0771
2007–2008	0.0276	0.1046	0.0088	0.0892
2008–2009	0.0213	0.0776	0.0086	-0.0013
2009–2010	0.0248	0.1022	0.0084	0.1098
2010–2011	0.0248	0.1427	0.0101	0.1120
2011–2012	0.0222	0.0547	0.0099	0.0524
2012–2013	0.0225	0.0652	0.0099	0.0404
2013–2014	0.0206	0.0707	0.0098	0.0364
2004–2014	0.3153	1.7706	0.0980	1.0165

Note: GDP is output-side real GDP at chained PPPs (in million 2011 US$); Labor is number of persons engaged (in millions); Capital is capital stock at current PPPs (in million 2011 US$); Human Capital is one index based on years of schooling and returns to education. The growth rate over 2004–2014 is the accumulated annual growth rate.

Table 3.15: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2004	37	107	-8.1521	3.58×10^{-16}	-11.4682	1.91×10^{-30}	-11.7834	4.75×10^{-32}
2006	37	107	-7.3594	1.85×10^{-13}	-8.2827	1.20×10^{-16}	-8.8034	1.33×10^{-18}
2008	37	107	-4.3511	1.35×10^{-05}	-8.0633	7.43×10^{-16}	-4.9163	8.82×10^{-07}
2010	37	107	-4.5157	6.31×10^{-06}	-7.1762	7.17×10^{-13}	-4.3040	1.68×10^{-05}
2012	37	107	-3.7759	1.59×10^{-04}	-5.0038	5.62×10^{-07}	-4.0607	4.89×10^{-05}
2014	37	107	-8.1837	2.75×10^{-16}	-10.3579	3.85×10^{-25}	-7.5720	3.67×10^{-14}

NOTE: The number of developed economies is n_1, while the number of developing economies is n_2. The numerator of statistics is the difference of estimated mean efficiency of developing economies minus developed economies.

Table 3.16: Tests of Differences in Means for Productivity Estimates with Respect to Type (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2004	37	107	2.9420	4.3993	5.1180	3.09×10^{-07}
2006	37	107	2.5810	3.9895	6.3903	1.66×10^{-10}
2008	37	107	2.3264	3.6565	7.8422	4.43×10^{-15}
2010	37	107	2.1908	3.3944	7.5544	4.21×10^{-14}
2012	37	107	2.2020	3.4269	6.8688	6.48×10^{-12}
2014	37	107	2.2701	3.4598	6.3503	2.15×10^{-10}

NOTE: The number of developed economies is n_1, while the number of developing economies is n_2. The numerator of statistics is the difference of estimated mean productivity of developing economies minus developed economies.

Appendices

Appendix A Additional Assumptions

The assumptions that follow are similar to Assumptions 3.1–3.4 of Kneip et al. (2015) and complete the statistical model. The first two assumptions that follow are needed for both FDH and VRS estimators.

Assumption A.1. *(i) The random variables* (X, Y) *possess a joint density* f *with support* $\mathcal{D} \subset \Psi$*; and (ii)* f *is continuously differentiable on* $\mathcal{D}$*.*

Assumption A.2. *(i)* $\mathcal{D}^* := \{(\theta(x, y \mid \Psi)x,\ y) \mid (x, y) \in \mathcal{D}\} = \{(x,\ \lambda(x, y \mid \Psi)y) \mid (x, y) \in \mathcal{D}\} = \big\{(\gamma(x, y \mid \Psi)x,\ \gamma(x, y \mid \Psi)^{-1}y) \mid (x, y) \in \mathcal{D}\big\} \subset \mathcal{D}$*; (ii)* $\mathcal{D}^*$ *is compact; and (iii)* $f(\theta(x, y \mid \Psi)x, y) > 0$ *for all* $(x, y) \in \mathcal{D}$*.*

The next two assumptions are needed when VRS estimators are used. Assumption A.3 imposes some smoothness on the frontier. Kneip et al. (2008) required only two-times differentiability to establish the existence of a limiting distribution for VRS estimators, by the stronger assumption that follows is needed to establish results on moments of the VRS estimators.

Assumption A.3. $\theta(x, y \mid \Psi)$*,* $\lambda(x, y \mid \Psi)$ *and* $\gamma(x, y \mid \Psi)$ *are three times continuously differentiable on* $\mathcal{D}$*.*

Recalling that the strong (i.e., free) disposability assumed in Assumption 1.3.3 implies that the frontier is weakly monotone, the next assumption strengthens this by requiring the frontier to be strictly monotone with no constant segments. This is also needed to establish properties of moments of the VRS estimators.

Assumption A.4. $\mathcal{D}$ *is* almost strictly convex*; i.e., for any* $(x, y), (\widetilde{x}, \widetilde{y}) \in \mathcal{D}$ *with* $(\frac{x}{\|x\|}, y) \neq (\frac{\widetilde{x}}{\|\widetilde{x}\|}, \widetilde{y})$*, the set* $\{(x^*, y^*) \mid (x^*, y^*) = (x, y) + \alpha((\widetilde{x}, \widetilde{y}) - (x, y))$ *for some* $0 < \alpha < 1\}$ *is a subset of the* interior *of* $\mathcal{D}$*.*

Alternatively, when FDH estimators are used, Assumptions A.3 and A.4 can be replaced by the following assumption.

Assumption A.5. *(i)* $\theta(x, y \mid \Psi)$*,* $\lambda(x, y \mid \Psi)$ *and* $\gamma(x, y \mid \Psi)$ *are twice continuously differentiable on* $\mathcal{D}$*; and (ii) all the first-order partial derivatives of* $\theta(x, y \mid \Psi)$*,* $\lambda(x, y \mid \Psi)$ *and* $\gamma(x, y \mid \Psi)$ *with respect to* x *and* y *are nonzero at any point* $(x, y) \in \mathcal{D}$*.*

Assumption A.5 strengthens strong disposability in the assumption 1.3.3 by requiring that the frontier is strictly monotone and does not possess constant segments (which might be the case, for example, if outputs are discrete as opposed to continuous, as in the case of ships produced by shipyards). Finally, part (i) of Assumption A.5 is weaker than Assumption A.3; here the frontier is required to be smooth, but not as smooth as required by Assumption A.3.[8] Assumptions 1.3.1–A.2 and Assumption A.5 comprise a statistical model appropriate for use of FDH estimators of technical efficiency, while Assumptions 1.3.1–A.4 comprise a statistical model appropriate for use of VRS estimators of technical efficiency.[9] These assumptions are sufficient for establishing consistency of the corresponding estimators. The stronger assumptions here are needed for results on moments and central limit theorems of the corresponding estimators.

[8] Assumption A.5 is slightly stronger, but much simpler than assumptions AII–AIII in Park et al. (2000).

[9] Additional assumptions are needed for CRS efficiency estimators. See Kneip et al. (2015) for additional discussion.

Appendix B Time Correlation in Testing Means

B.1 Correlation in Efficiency

Notice that there may exist time correlation when we testing the differences in mean efficiency over time, which violates the independent assumption of test for differences of mean efficiency in Kneip et al. (2016). Hence we take the following method to deal with time correlation.

let n_0 be the number of observations existing in both periods, n_1 be the number of observations existing in period 1 but not in period 2 and n_2 be the number of observations existing in period 2 but not in period 1. We then use the randomization algorithm in Daraio et al. (2018) to randomly shuffle these n_0 observations. For period 1, we combine the first half $n_{01} = \lfloor n_0/2 \rfloor$ of n_0 observations with these n_1 observations to form the sample, denoted as S_1. Similarly, for period 2, we combine the second half $n_{02} = n_0 - \lfloor n_0/2 \rfloor$ of n_0 observations with these n_2 observations to form the sample, denoted as S_2. By construction, S_1 and S_2 are independent, and now we can use the tests for differences in mean efficiency in Kneip et al. (2016).

B.2 Correlation in Productivity

let n_0 be the number of observations existing in both periods, n_1 be the number of observations existing in period 1 but not in period 2 and n_2 be the number of observations existing in period 2 but not in period 1. Productivity is calculated by the ratio of output over input.

Then to test H_0: $\mu_1 = \mu_2$, versus H_1: $\mu_1 \neq \mu_2$, we can use statistics

$$\hat{T} = \frac{\hat{\mu_2} - \hat{\mu_1}}{\frac{\hat{\sigma}_1^2}{(n_1+n_0)} + \frac{\hat{\sigma}_2^2}{(n_2+n_0)} - 2n_0 \frac{\hat{\sigma}_{12}}{(n_1+n_0)(n_2+n_0)}} \sim N(0,1) \tag{2.2}$$

Where $\hat{\mu_i}$, $i \in \{1,2\}$, is the sample mean for all the observations in period i, $\hat{\sigma}_i^2$ is the sample variance for all the observations in period i and $\hat{\sigma}_{12}$ is the sample covariance for all the n_0 observations exsting in both periods.

Appendix C Additional Results

Table C1–C2 list the numbers of China's commercial banks over the sample period in population and in sample, respectively. Table C3 shows the eigensystem analysis of input and output moment matrices.

Table C4 shows the total assets of the five largest commercial banks over the period 2007–2014 in the data. Total assets of the largest banks increased over the whole period. They amount to about 33.04 quadrillion RMB in 2007 and about 65.57 quadrillion RMB in 2014, for an increase of about 1 time. Moreover, the growth rate from 2008 to 2009 is at least 20 percent and is the highest for each of the five big banks among all the one-year intervals, reflecting the expansion effect of the stimulus.

Table C5 provides the numbers of observations with estimated hyperbolic technical efficiency equal to 1 in each year.

Table C6 shows the robustness check of results of convexity tests when the sample is split unevenly.

Table C7–C10 show the robustness check for the tests of differences in means for FDH technical efficiency and productivity estimates when we use two different definitions for the big and small banks.

Table C1: Number of China's Commercial Banks in Population

Bank's Type	2007	2008	2009	2010	2011	2012	2013	2014
Four Big State Owned Banks (SOCBs)	4	4	4	4	4	4	4	4
Joint Stock Commercial Banks (JSCBs)	13	13	13	13	13	13	13	12
City Commercial Banks (CCBs)	124	136	143	147	144	144	145	133
Rural Commercial Banks (RCBs)	17	22	43	85	212	337	468	665
Foreign Banks	29	32	37	40	40	42	42	41
Total Number	187	207	240	289	413	540	672	855
Total Assets(in Trillion 2010 CNY)	60.57	66.73	84.16	95.30	104.76	120.64	133.74	150.95

NOTE: The main data are from China Banking Regulation Committee (CBRC).

Table C2: Number of China's Commercial Banks in Sample

Bank's Type	2007	2008	2009	2010	2011	2012	2013	2014
Four Big State Owned Banks (SOCBs)	4	4	4	4	4	4	4	4
Joint Stock Commercial Banks (JSCBs)	6	8	9	11	11	13	12	12
City Commercial Banks (CCBs)	7	13	13	23	36	49	59	58
Rural Commercial Banks (RCBs)	1	2	3	8	11	15	18	18
Foreign Banks	6	14	16	19	20	27	30	32
Total Number	24	41	45	65	82	108	123	124
Total Assets(in Trillion 2010 CNY)	38.19	45.27	57.59	66.93	74.52	89.31	98.13	108.90

NOTE: The main data are from Bankscope.

Table C3: Eigensystem Analysis by Year

Year	$R_x(\%)$	$R_y(\%)$
2007	98.25	95.87
2008	98.61	97.10
2009	98.81	97.83
2010	98.75	97.46
2011	98.59	97.15
2012	98.65	97.26
2013	98.95	96.92
2014	98.96	96.47

Table C4: Total Assets of Five Largest Commercial Banks by Year in Billions of 2010 CNY

Name	Total Assets	Name	Total Assets
— **2007** —		— **2011** —	
Industrial & Commercial Bank of China	10,000,250	Industrial & Commercial Bank of China	14,310,270
China Construction Bank Corporation	7,598,527	China Construction Bank Corporation	11,356,070
Bank of China Limited	6,899,545	Bank of China Limited	10,938,100
Agricultural Bank of China Limited	6,109,874	Agricultural Bank of China Limited	10,797,360
Bank of Communications Co Ltd	2,430,408	Bank of Communications Co Ltd	4,263,601
— **2008** —		— **2012** —	
Industrial & Commercial Bank of China	10,420,760	Industrial & Commercial Bank of China	15,840,950
China Construction Bank Corporation	8,069,318	China Construction Bank Corporation	12,617,730
Agricultural Bank of China Limited	7,491,415	Agricultural Bank of China Limited	11,959,890
Bank of China Limited	7,424,482	Bank of China Limited	11,450,840
Bank of Communications Co Ltd	2,860,404	Bank of Communications Co Ltd	4,761,960
— **2009** —		— **2013** —	
Industrial & Commercial Bank of China	12,603,460	Industrial & Commercial Bank of China	16,711,220
China Construction Bank Corporation	10,291,640	China Construction Bank Corporation	13,571,270
Agricultural Bank of China Limited	9,499,433	Agricultural Bank of China Limited	12,863,600
Bank of China Limited	9,359,716	Bank of China Limited	12,256,030
Bank of Communications Co Ltd	3,538,938	Bank of Communications Co Ltd	5,265,664
— **2010** —		— **2014** —	
Industrial & Commercial Bank of China	13,458,620	Industrial & Commercial Bank of China	18,056,470
China Construction Bank Corporation	10,810,320	China Construction Bank Corporation	14,669,570
Bank of China Limited	10,459,870	Agricultural Bank of China Limited	13,995,020
Agricultural Bank of China Limited	10,337,410	Bank of China Limited	13,361,800
Bank of Communications Co Ltd	3,951,593	Bank of Communications Co Ltd	5,491,685

Table C5: Numbers of Observations With Estimated Hyperbolic Technical Efficiency Equal to 1 in Each Year

		Without — Dimension Reduction —			With — Dimension Reduction —		
Year	n	FDH	VRS	CRS	FDH	VRS	CRS
2007	24	24	22	13	23	4	1
2008	41	41	34	16	29	5	1
2009	45	45	31	14	32	6	1
2010	65	65	46	34	47	7	1
2011	82	82	50	42	47	8	1
2012	108	108	73	59	58	5	1
2013	123	123	76	56	67	6	1
2014	124	124	77	58	64	9	1

Table C6: Results of Convexity Tests, Average over 1000 splits, Bootstrap 1000 times (Uneven Split, with Dimension Reduction, $p = q = 1$)

	— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	-3.6389	0.1550	3.1170	0.4820	-3.5848	0.4270
2008	-2.3565	0.4100	1.2186	0.5870	-2.8590	0.2580
2009	-3.4031	0.0920	4.3980	0.0010	-4.4576	0.0050
2010	-5.1630	0.0080	3.6780	0.2440	-4.4724	0.0460
2011	-3.2804	0.0490	0.3984	0.6110	-2.1660	0.2050
2012	-3.0409	0.0810	1.7431	0.2470	-2.9384	0.0650
2013	-5.0122	0.0000	3.7484	0.0030	-6.1833	0.0020
2014	-2.7546	0.0500	3.1398	0.0020	-3.7721	0.0020

Table C7: Robustness Check: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	6	18	−2.3383	1.94×10^{-02}	−2.1641	3.05×10^{-02}	−2.2798	2.26×10^{-02}
2008	11	30	−3.7841	1.54×10^{-04}	−4.0663	4.78×10^{-05}	−4.0089	6.10×10^{-05}
2009	12	33	−5.4692	4.52×10^{-08}	−5.4176	6.04×10^{-08}	−6.5292	6.61×10^{-11}
2010	17	48	−4.9000	9.58×10^{-07}	−3.7321	1.90×10^{-04}	−3.5207	4.30×10^{-04}
2011	21	61	−6.0628	1.34×10^{-09}	−3.9386	8.20×10^{-05}	−6.5101	7.51×10^{-11}
2012	27	81	−8.7340	2.46×10^{-18}	−6.5111	7.46×10^{-11}	−7.2261	4.97×10^{-13}
2013	31	92	−4.3180	1.57×10^{-05}	−1.6459	9.98×10^{-02}	−3.2212	1.28×10^{-03}
2014	31	93	−6.2219	4.91×10^{-10}	−5.2614	1.43×10^{-07}	−4.7034	2.56×10^{-06}

NOTE: We split the total observations of each year into two uneven subsamples by the 75% quantile of total assets in that year. The number of big banks is n_1, while the number of small banks is n_2. The numerator of statistics is the difference of estimated mean efficiency of small banks minus big banks.

Table C8: Robustness Check: Tests of Differences in Means for Productivity Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2007	6	18	0.6444	0.6618	0.3500	7.26×10^{-01}
2008	11	30	0.6245	0.6785	2.0410	4.13×10^{-02}
2009	12	33	0.6241	0.7034	1.7481	8.05×10^{-02}
2010	17	48	0.6065	0.6328	1.3590	1.74×10^{-01}
2011	21	61	0.5870	0.5594	−1.3852	1.66×10^{-01}
2012	27	81	0.5860	0.5543	−1.7786	7.53×10^{-02}
2013	31	92	0.5876	0.5448	−2.4880	1.28×10^{-02}
2014	31	93	0.5987	0.5683	−1.8466	6.48×10^{-02}

NOTE: We split the total observations of each year into two uneven subsamples by the 75% quantile of total assets in that year. The number of big banks is n_1, while the number of small banks is n_2. The numerator of statistics is the difference of estimated mean productivity of small banks minus big banks.

Table C9: Robustness Check: Tests of Differences in Means for FDH Technical Efficiency Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

			— **Input** —		— **Output** —		— **Hyperbolic** —	
Year	n_1	n_2	Statistic	p-value	Statistic	p-value	Statistic	p-value
2007	8	8	—	—	—	—	—	—
2008	14	14	−3.3509	8.05×10^{-04}	−2.9847	2.84×10^{-03}	−3.6668	2.46×10^{-04}
2009	15	15	−4.9624	6.96×10^{-07}	−1.3183	1.87×10^{-01}	−3.8646	1.11×10^{-04}
2010	22	22	−4.2734	1.93×10^{-05}	−2.2921	2.19×10^{-02}	−2.7055	6.82×10^{-03}
2011	28	28	−5.0868	3.64×10^{-07}	−2.9354	3.33×10^{-03}	−6.1081	1.01×10^{-09}
2012	36	36	−5.0359	4.75×10^{-07}	−4.9898	6.05×10^{-07}	−4.2668	1.98×10^{-05}
2013	41	41	−2.4495	1.43×10^{-02}	−1.2766	2.02×10^{-01}	−3.0377	2.38×10^{-03}
2014	42	42	−4.3523	1.35×10^{-05}	−4.2584	2.06×10^{-05}	−4.4456	8.77×10^{-06}

NOTE: We split the total observations of each year into the top one third quantile group and the bottom one third quantile group in that year. The number of big banks is n_1, while the number of small banks is n_2. The numerator of statistics is the difference of estimated mean efficiency of small banks minus big banks. For 2007, there are too few observations for each group, making all of the efficiency estimates in group 1 or 2 equal to 1.

Table C10: Robustness Check: Tests of Differences in Means for Productivity Estimates with Respect to Size (with Dimension Reduction, $p = q = 1$)

Year	n_1	n_2	Mean1	Mean2	Statistic	p-value
2007	8	8	0.6431	0.7385	0.9623	3.36×10^{-01}
2008	14	14	0.6243	0.7022	1.5142	1.30×10^{-01}
2009	15	15	0.6324	0.7487	1.2144	2.25×10^{-01}
2010	22	22	0.6133	0.6434	1.4837	1.38×10^{-01}
2011	28	28	0.5863	0.5470	−1.4343	1.51×10^{-01}
2012	36	36	0.5791	0.5287	−2.1039	3.54×10^{-02}
2013	41	41	0.5893	0.5144	−3.4066	6.58×10^{-04}
2014	42	42	0.6026	0.5436	−2.3099	2.09×10^{-02}

NOTE: We split the total observations of each year into the top one third quantile group and the bottom one third quantile group in that year. The number of big banks is n_1, where the number of small banks is n_2. The numerator of statistics is the difference of estimated mean productivity of small banks minus big banks.

Bibliography

An, Q., H. Chen, J. Wu, and L. Liang (2015), Measuring slacks-based efficiency for commercial banks in China by using a two-stage DEA model with undesirable output, *Annals of Operations Research* 235, 13–35.

Arcelus, F. J. and P. Arocena (2000), Convergence and productive efficiency in fourteen OECD countries: A non-parametric frontier approach, *International Journal of Production Economics* , 13.

Ariff, M. and L. Can (2008), Cost and profit efficiency of Chinese banks: A non-parametric analysis, *China Economic Review* 19, 260–273.

Avkiran, N. K. (2011), Association of DEA super-efficiency estimates with financial ratios: Investigating the case for Chinese banks, *Omega* 39, 323–334.

Baker, D. and T. Mcarthur (2009), The Value of the "Too Big to Fail" Big Bank Subsidy, *Center for Economic and Policy Research (CEPR)* , 6.

Banker, R. D., A. Charnes, and W. W. Cooper (1984), Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30, 1078–1092.

Barros, C. P., Z. Chen, Q. B. Liang, and N. Peypoch (2011), Technical efficiency in the Chinese banking sector, *Economic Modelling* 28, 2083–2089.

Berger, A. N., I. Hasan, and M. Zhou (2009), Bank ownership and efficiency in China: What will happen in the world's largest nation?, *Journal of Banking & Finance* 33, 113–130.

Boucinha, M., N. Ribeiro, and T. Weyman-Jones (2013), An assessment of Portuguese banks' efficiency and productivity towards euro area participation, *Journal of Productivity Analysis* 39, 177–190.

Braeutigam, R. R. and A. F. Daughety (1983), On the estimation of returns to scale using variable cost functions, *Economics Letters* 11, 25–31.

Brewer, E. and J. Jagtiani (2013), How Much Did Banks Pay to Become Too-Big-To-Fail and to Become Systemically Important?, *Journal of Financial Services Research* 43, 1–35.

Chang, C.-C. and Y.-H. Luh (1999), Efficiency change and growth in productivity: the Asian growth experience, *Journal of Asian Economics* 10, 551–570.

Chen, X., M. Skully, and K. Brown (2005), Banking efficiency in China: Application of DEA to pre- and post-deregulation eras: 1993–2000, *China Economic Review* 16, 229–245.

Daouia, A., L. Simar, and P. W. Wilson (2017), Measuring firm performance using nonparametric quantile-type distances, *Econometric Reviews* 36, 156–181.

Daraio, C., L. Simar, and P. W. Wilson (2018), Central limit theorems for conditional efficiency measures and tests of the 'separability condition' in non-parametric, two-stage models of production, *Econometrics Journal* 21, 170–191.

Deprins, D., L. Simar, and H. Tulkens (1984), Measuring labor inefficiency in post offices, in M. M. P. Pestieau and H. Tulkens, eds., *The Performance of Public Enterprises: Concepts and Measurements*, Amsterdam: North-Holland, pp. 243–267.

Dong, Y., M. Firth, W. Hou, and W. Yang (2016), Evaluating the performance of Chinese commercial banks: A comparative analysis of different types of banks, *European Journal of Operational Research* 252, 280–295.

Dong, Y., R. Hamilton, and M. Tippett (2014), Cost efficiency of the Chinese banking sector: A comparison of stochastic frontier analysis and data envelopment analysis, *Economic Modelling* 36, 298–308.

Du, K., A. C. Worthington, and V. Zelenyuk (2018), Data envelopment analysis, truncated regression and double-bootstrap for panel data with application to Chinese banking, *European Journal of Operational Research* 265, 748–764.

Duygun, M., M. Shaban, R. C. Sickles, and T. Weyman-Jones (2015), How a regulatory capital requirement affects banks' productivity: an application to emerging economies, *Journal of Productivity Analysis* 44, 237–248.

Emrouznejad, A. (2003), An alternative DEA measure: a case of OECD countries, *Applied Economics Letters* 10, 779–782.

Färe, R. (1988), *Fundamentals of Production Theory*, Berlin: Springer-Verlag.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*, Boston: Kluwer-Nijhoff Publishing.

Färe, R., S. Grosskopf, M. Norris, and Z. Zhang (1994), Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review* 84, 66–83.

Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015), The Next Generation of the Penn World Table, *American Economic Review* 105, 3150–3182.

Fethi, M. D., M. Shaban, and T. Weyman-Jones (2012), Turkish Banking Recapitalization and the Financial Crisis: An Efficiency and Productivity Analysis, *Emerging Markets Finance and Trade* 48, 76–90.

Fiordelisi, F., N. Radić, and T. Weyman-Jones (2018), Bank performance and the shadow price of capital, *Working Paper* .

Gilbert, A. and P. W. Wilson (1998), Effects of deregulation on the productivity of korean banks, *Journal of Economics and Business* 50, 133–155.

Glass, A. J., K. Kenjegalieva, and R. C. Sickles (2016), A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers, *Journal of Econometrics* 190, 289–300.

Gu, H. and J. Yue (2011), The Relationship between Bank Efficiency and Stock Returns:, *World Journal of Social Sciences* 1, 95–106.

Han, G., K. Kalirajan, and N. Singh (2004), Productivity, Efficiency and Economic Growth: East Asia and the Rest of the World, *The Journal of Developing Areas* 37, 99–118.

Härdle, W. (1990), *Applied Nonparametric Regression*, Cambridge: Cambridge University Press.

Härdle, W. and E. Mammen (1993), Comparing nonparametric versus parametric regression fits, *Annals of Statistics* 21, 1926–1947.

Hasannasab, M., D. Margaritis, and C. Staikouras (2019), The financial crisis and the shadow price of bank capital, *Annals of Operations Research* 282, 131–154.

Hughes, J. P. (1999), Incorporating risk into the analysis of production, *Atlantic Economic Journal* 27, 1–23.

Hughes, J. P., L. J. Mester, and C.-G. Moon (2001), Are scale economies in banking elusive or illusive? Evidence obtained by incorporating capital structure and risk-taking into models of bank production q, *Journal of Banking & Finance* , 40.

Ji, K., W. Song, and R. Wang (2012), Research on China's Commercial Banks Rating and Ranking Based on DEA, *American Journal of Operations Research* 02, 122–125.

Kneip, A., B. Park, and L. Simar (1998), A note on the convergence of nonparametric DEA efficiency measures, *Econometric Theory* 14, 783–793.

Kneip, A., L. Simar, and P. W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory* 24, 1663–1697.

— (2011), A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators, *Computational Economics* 38, 483–515.

— (2015), When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory* 31, 394–422.

— (2016), Testing hypotheses in nonparametric models of production, *Journal of Business and Economic Statistics* 34, 435–456.

Laurenceson, J. and Z. Yong (2008), Efficiency amongst China's banks: a DEA analysis five years after WTO entry, *China Economic Journal* 1, 275–285.

Lee, T.-H. and S.-H. Chih (2013), Does financial regulation affect the profit efficiency and risk of banks? Evidence from China's commercial banks, *The North American Journal of Economics and Finance* 26, 705–724.

Linton, O., E. Maasoumi, and Y.-J. Whang (2005), Consistent Testing for Stochastic Dominance under General Sampling Schemes, *Review of Economic Studies* 72, 735–765.

Liu, W., Z. Zhou, C. Ma, D. Liu, and W. Shen (2015), Two-stage DEA models with undesirable input-intermediate-outputs, *Omega* 56, 74–87.

Luo, D. and S. Yao (2010), World financial crisis and the rise of Chinese commercial banks: an efficiency analysis using DEA, *Applied Financial Economics* 20, 1515–1530.

Maudos, J., J. M. Pastor, and L. Serrano (1999), Total factor productivity measurement and human capital in OECD countries, *Economics Letters* 63, 39–44.

Park, B. U., S.-O. Jeong, and L. Simar (2010), Asymptotic distribution of conical-hull estimators of directional edges, *Annals of Statistics* 38, 1320–1340.

Park, B. U., L. Simar, and C. Weiner (2000), FDH efficiency scores from a stochastic point of view, *Econometric Theory* 16, 855–877.

Radić, N. (2015), Shareholder value creation in Japanese banking, *Journal of Banking & Finance* 52, 199–207.

Ray, S. C. and E. Desli (1997), Productivity growth, technical progress, and efficiency change in industrialized countries: Comment, *American Economic Review* 87, 1033–1039.

Restrepo, D. A., S. C. Kumbhakar, and K. Sun (2013), Are U.S. Commercial Banks Too Big? , 46Working Paper.

Salinas-Jiménez, M. d. M. and J. Salinas-Jiménez (2007), Corruption, efficiency and productivity in OECD countries, *Journal of Policy Modeling* 29, 903–915.

Santos, J. and J. A. C. Santos (2014), *Evidence from the Bond Market on Banks' "Too-Big-to-Fail " Subsidy*.

Sealey, C. W. and J. T. Lindley (1977), Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions, *Journal of Finance* .

Shephard, R. W. (1970), *Theory of Cost and Production Functions*, Princeton: Princeton University Press.

Simar, L., I. Van Keilegom, and V. Zelenyuk (2017), Nonparametric least squares methods for stochastic frontier models, *Journal of Productivity Analysis* 47, 189–204.

Simar, L. and P. W. Wilson (1998), Productivity growth in industrialized countries. Discussion Paper #9810, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

— (2007), Estimation and inference in two-stage, semi-parametric models of productive efficiency, *Journal of Econometrics* 136, 31–64.

— (2011), Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis* 36, 33–53.

— (2018), Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices. Discussion Paper 2018/21, Institut de Statistique Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

— (2020), Hypothesis Testing in Nonparametric Models of Production using Multiple Sample Splits, *Journal of Productivity Analysis* Forthcoming.

Sufian, F. and M. S. Habibullah (2009), Asian financial crisis and the evolution of korean banks efficiency: A DEA approach, *Global Economic Review* 38, 335–369.

— (2011), Opening the Black Box on Bank Efficiency in China: Does Economic Freedom Matter?, *Global Economic Review* 40, 269–298.

Tan, Y. and C. Floros (2013), Risk, capital and efficiency in Chinese banking, *Journal of International Financial Markets, Institutions and Money* 26, 378–393.

Ueda, K. and B. Weder di Mauro (2013), Quantifying structural subsidy values for systemically important financial institutions, *Journal of Banking & Finance* 37, 3830–3842.

Wang, K., W. Huang, J. Wu, and Y.-N. Liu (2014), Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA, *Omega* 44, 5–20.

Wang, T.-P., S. H.-P. Shyu, and H.-C. Chou (2012), The impact of defense expenditure on economic productivity in OECD countries, *Economic Modelling* 29, 2104–2114.

Weyman-Jones, T. (2016), *Efficiency, Competition and the Shadow Price of Capital*, London: Palgrave Macmillan UK, pp. 147–167.

Wheelock, D. C. and P. W. Wilson (1999), Technical progress, inefficiency, and productivity change in U. S. banking, 1984–1993, *Journal of Money, Credit, and Banking* 31, 212–234.

— (2008), Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations, *Journal of Econometrics* 145, 209–225.

— (2011), Are credit unions too small?, *Review of Economics and Statistics* 93, 1343–1359.

— (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.

— (2018), The evolution of scale-economies in U.S. banking, *Journal of Applied Econometrics* 33, 16–28.

— (2019), Nonparametric Estimation of Lerner Indices for U.S. Banks Allowing for Inefficiency and Off-Balance Sheet Activities, Technical report.

Wilson, P. W. (2011), Asymptotic properties of some non-parametric hyperbolic efficiency estimators, in I. Van Keilegom and P. W. Wilson, eds., *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, Berlin: Springer-Verlag, pp. 115–150.

— (2018), Dimension reduction in nonparametric models of production, *European Journal of Operational Research* 267, 349–367.

Yao, S., Z. Han, and G. Feng (2008), Ownership reform, foreign competition and efficiency of chinese commercial banks: A nonparametric approach, *The World Economy* 31, 1310–1326.

Zha, Y., N. Liang, M. Wu, and Y. Bian (2016), Efficiency evaluation of banks in China: A dynamic two-stage slacks-based measure approach, *Omega* 60, 60–72.