Performance of Chinese Banks over 2007–2014

Paul W. Wilson   Shirong Zhao *

Feb 2020

Preliminary Draft - Not for Citation

Abstract

This paper examines 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.

Keywords: technical efficiency, productivity, technology, Chinese commercial banks, non-parametric efficiency estimators

---

*Wilson: Department of Economics and School of Computing, Division of Computer Science, Clemson University, Clemson, South Carolina 29634–1309, USA; email pww@clemson.edu.

Zhao: Department of Economics, Clemson University, Clemson, South Carolina 29634–1309, USA; email shironz@clemson.edu.

We thank the Cyber Infrastructure Technology Integration group at Clemson University for operating the Palmetto cluster used for our computations. We received helpful comments and suggestions from Howard Bodenhorn and Matthew Lewis. In addition, We thank Sarah Wilson for her careful edits. All remaining errors are our own. JEL classification nos.: G21 L11 C12 C13 C14.

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

[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.

(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. 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

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

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

[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.

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, 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 2 provides background on the Chinese banking industry. Estimators of technical efficiency and their properties are discussed in Section 3. Section 4 discusses various statistical results needed for testing hypotheses about model features. Section 5 describes the Chinese commercial banks data, specially the input and output variables. Section 6 discusses the empirical results of the tests. Major conclusions and directions for future works are presented in Section 7. Additional details on model assumptions, data, and results are provided in separate Appendices A–C, which are available online.

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.

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

[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.

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

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 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.

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 3.1. Ψ *is closed.*

Assumption 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 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 3.1 ensures that the *efficient frontier* (or *technology*) Ψ^{∂}

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

is the set of extreme points of Ψ and is contained in Ψ. Assumption 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 3.3 imposes weak monotonicity on the frontier.

The Farrell (1957) input efficiency measure

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

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

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

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.

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 $\boldsymbol{X} = \big(X_1,\ \ldots,\ X_n\big)$ and $\boldsymbol{Y} = \big(Y_1,\ \ldots,\ Y_n\big)$ 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 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

[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).

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 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.

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 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 3 is fully non-parametric, and hence quite flexible. The assumptions listed in Section 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 6.

The flexibility of the non-parametric model specified in Section 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

[13]The numbers of China's commercial banks across years in population and sample are provided in Tables C.1–C.2 of the separate Appendix C.

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 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 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]

[14]The details about eigensystem analysis of input and output moment matrices are shown in Table C.3 of the separate Appendix C.

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

[16]See Wheelock and Wilson (2018).

Table 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.

6 Empirical Results

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 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 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 5 are used for obtaining all the following results.

The next question is to determine which estimator we should use. As discussed in Section

[17] Additional details about the number of observations with estimates equal to one are given in the Table C.5 of the separate Appendix C.

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, some have noticed that p-values resulting from the tests vary across different random splits of the original sample. Simar and Wilson (2019) 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 (2019), 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 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 2 provide strong evidence of

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

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 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 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 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

[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 C.6 of the separate Appendix C. The results in Table C.6 are consistent with that in Table 2.

[20]We also use the KS-statistics developed in Simar and Wilson (2019). 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.

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 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 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 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

[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).

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 5, productivity can be measured by Y_i^*/X_i^* for bank i. Summary statistics for this measure is displayed in Table 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 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 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 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.

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 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 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 C.7–C.10 of the separate Appendix C. The results are quite consistent with our baseline estimates.

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 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 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.

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.

References

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.

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.

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.

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., 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.

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.

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

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

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

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.

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.

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.

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

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

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.

— (2019), Hypothesis Testing in Nonparametric Models of Production using Multiple Sample Splits Working Paper.

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.

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.

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.

— (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.

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.

Table 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 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 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 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 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 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 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 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 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 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 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}

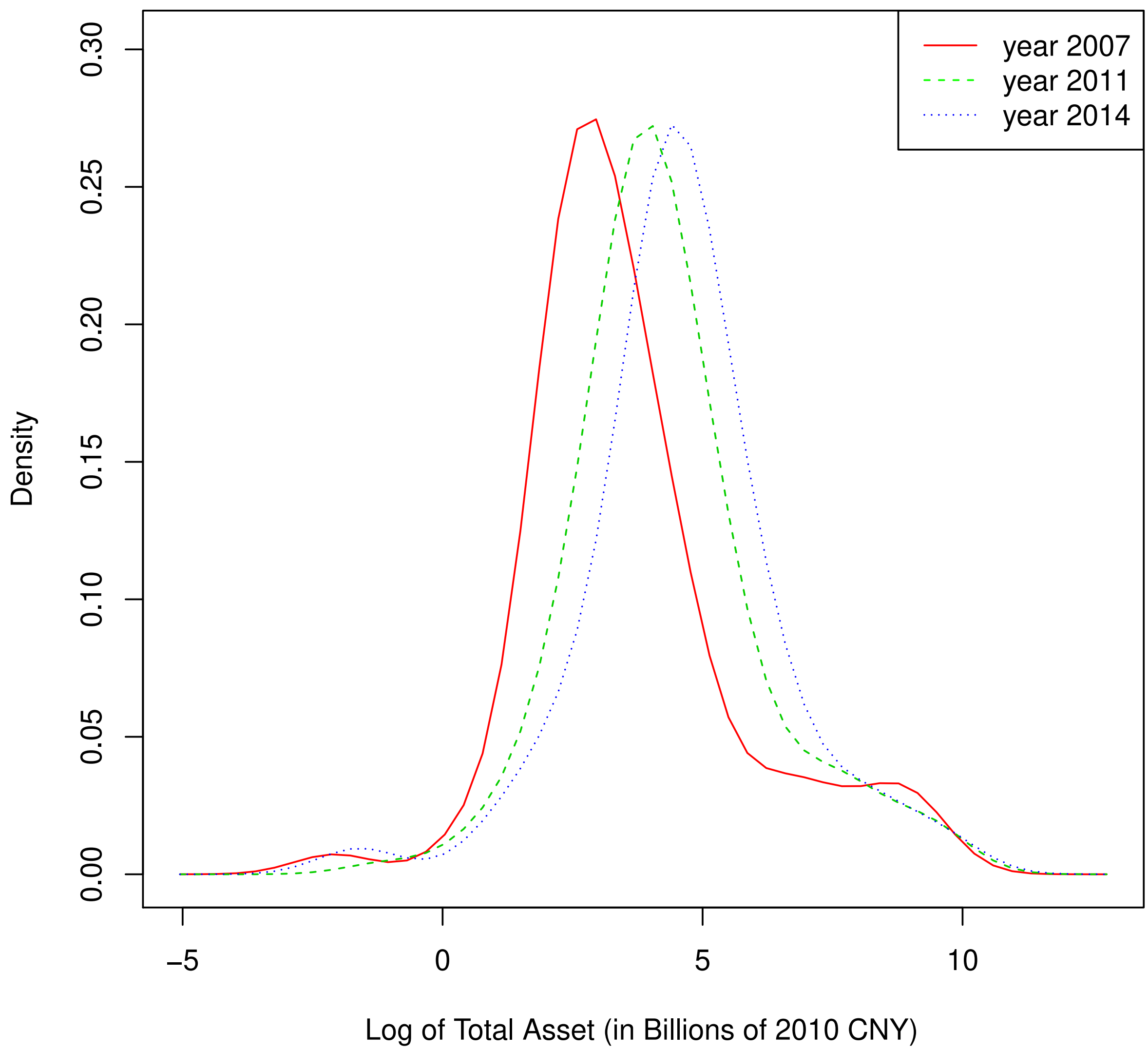
NOTE: The number of domestic banks is n_1, while the number of foreign banks is n_2. The numerator of statistics is the difference of estimated mean efficiency of foreign banks minus domestic banks.

Table 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: 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.