# Efficiency Analysis of Chinese Banks

## Master Thesis for Business Analytics And Quantitative Marketing


-

*Authors:*
Ang Qie (423595)

-

*Supervisor:*
Prof. Dennis Fok

-

*Coreader:*
Prof. Richard Paap

Econometrics & Operations Research
Erasmus School of Economics
Erasmus University Rotterdam


July 14, 2016

# Abstract

This thesis investigates the efficiency scores of Chinese banks by employing stochastic frontier analysis (SFA) and data envelopment analysis (DEA). Meanwhile the study tries to examine which situation the results produced by one of both frontier models are more reliable. I estimate the cost efficiency based on a panel data set of 22 selected Chinese commercial banks over the period from 2009 until 2014. The result suggests that SFA and DEA yield a consistent trend on efficiency scores over the period, indicating the cost efficiency of Chinese banks did not show a significant improvement during the time period. However, rank correlations indicate both approaches produce contrary results at individual since both approaches are completely different. This thesis examines the two main differences between DEA and SFA which lie on measurement errors and heterogeneity for efficiency term which DEA can not account for. Furthermore, based on the fact that both frontier approaches cannot provide a coherent overview of the performance of banks, I conclude that other instruments such as traditional performance measures should be used in order to evaluate the accuracy of frontier approaches.



**Keywords:** Cost efficiency, Stochastic frontier analysis, Data envelopment analysis, Chinese banks.

# Contents

# 1 | Introduction

The stability of the global banking market was damaged by the world financial crisis in 2008. Share prices of some famous banks experienced dramatic decreases from that period. For instance, Citigroup and the Royal Bank of Scotland (RBS) encountered drastic price cuts by more than 95% from 2007 to January 2009 (Luo et al.[19]). However, the Chinese banking sector was not influenced seriously by such global crisis. Instead of making huge financial investments like American and European banks, Chinese banks are much more conservative. After the financial crisis in 2008, the Chinese banks have been leading the ranking of the world banks. By early 2009, the three biggest state-owned commercial banks (SOCBs) had won the world's largest banks positions in market capitalisation over American and European banking giants. In 2015, China had three banks in the top five places from *The Banker* ranking list.[1]

One may be curious about how Chinese banks managed to become the top performers after the financial crisis, why Chinese banks are so different from the rest of the world. Let us first look at the Chinese banking system. Prior to 1978, the Chinese banking system was characterized by a mono-banking system, which was dominated by People's Bank of China (PBC). PBC was the only financial institution back that time, which performed the duties of both central bank and commercial bank. After 1978, there have been two stages of reform, one is from 1979 to 1992, and another is from 1993 to present. Stage one started with the establishment of a "two tier" banking system, four SOCBs were introduced in order to take over the commercial operations of PBC. Between 1985 and 1992, joint stock commercial banks (JSCBs) were created, which aimed to promote more competition and profit maximization. Since 1993, city commercial banks (CCBs) were launched through the second stage of reform. Their service are mainly served to small and medium sized enterprises and residents where they

[1] http://www.thebanker.com/Top-1000/The-Banker-Top-1000-World-Banks-2015-ranking-WORLD-Press-IMMEDIATE-RELEASE

locate. Usually they have strong ties with local governments. Nowadays, the Chinese banking market is mainly dominated by SOCBs, JSCBs and CCBs. And the competition among these banks is still very intensive.

Therefore, a good bank performance is crucial for bankers to survive in this competitive environment. One way to measure the bank performance is by estimating efficiency. This paper applies both a non-parametric approach (mathematical programming), and a parametric approach (econometrics) to investigate the efficiency of Chinese commercial banks. The parametric approach is stochastic frontier analysis (SFA). The efficiency is calculated through a cost or profit function, where variables are determined by input prices, number of variable outputs, random error and inefficiency. To estimate the stochastic frontier model, a specific functional form in terms of translog cost function is often proposed. The non-parametric approach is data envelopment analysis (DEA), which is to utilize mathematical linear programming to establish an efficient frontier (production frontier) which can capture all the observed banks, such that all the banks lie on or below the efficient frontier. The efficient frontier is formed as piecewise linear combinations that link all the best practise observations. The linear programming contains the maximization of the weighted output/weighted input ratio. Both techniques are basically frontier research which aim to construct an efficient frontier based on linear programming or econometric tools in order to examine efficient and inefficient units. One advantage of DEA over SFA is that it does not require any concrete functional forms, but DEA can not capture the random errors which SFA is able to. As both DEA and SFA have their own advantages and limits, it is worthwhile to conduct both approaches for the Chinese banking system.

Despite the large literature which applied DEA and SFA to the American and European banking system, quite few studies have been conducted in the Chinese banking system. As mentioned before, Chinese banking system is quite different from American and European banking system. Therefore, the main purpose of this paper is to estimate the Chinese banks' efficiency by both DEA and SFA. According to the results obtained from both methods, we can analyze the distribution trend of the Chinese banking system. The data I use are 22 Chinese commercial banks over the period 2009-2014. These 22 Chinese banks contain 5 SOCBs, 9 JSCBs and 8 CCBs. In addition, previous studies focus on comparing the consistent or inconsistent results obtained from these two techniques. However, there is no consensus that both approaches should lead to the consistent results reported by Berger and Humphrey [7]. Therefore, this paper tries to examine whether both approaches yield consistent results from their distribution trends. First, I use both techniques to estimate the efficiency in order to conduct the efficiency analysis. Accordingly, the performances of these 22 banks over 2009-2014 based on two different approaches can be examined. Next, given the merits and limitations of both techniques, my research analyzes which methods can present more applicable and reliable results for bankers to follow in the future. To the best of my knowledge, this is first research using both SFA and DEA to measure the Chinese banks' efficiencies over the time period after financial crisis.

The remainder of this paper is structured as follows. In Ch. 2 related literature on bank efficiency based on two different approaches is discussed. In Ch. 3 the proposed methods to evaluate the bank performance are introduced. The data is described and analyzed in Ch. 4. Then, in Ch. 5 the results of the proposed methods are present and compared. Last, in Ch. 6 these results are discussed and some suggestions for further research are provided.

# 2 | Related work

Bankers and scholars have been analyzing how efficiently banks can transform their multiple inputs into various financial outputs over the last two decades. Therefore, there is growing literature which apply DEA or SFA to efficiency analysis. However, only a quite limited literature use both techniques. First, I discuss the literature on the application of DEA to bank efficiency. Next, I discuss the past research on measuring the bank efficiency based on SFA. Last, I discuss the literature that combine both techniques for bank efficiency analysis.

The first paper that applied DEA to banking industry was provided by Sherman and Gold [25]. They found that DEA results provide useful insights to improve efficiency. Their paper inspired a lot of scholars to continue work on this topic. In Chinese banking studies, Zhang [17] used the basic and an improved model based on DEA to estimate Chinese commercial banks from 1997 to 2001. The results found that the JSCBs are the most efficient banks while CCBs are the most inefficient ones. Ariff and Can [2] applied DEA approach to estimate the cost and profit efficiency of 28 commercial banks in China over the period 1995-2004. Then, a Tobit regression was used to examine the influence of different factors on bank efficiency The results of their research found that the JSCBs are more cost and profit efficient than SOCBs. Chen et al. [28] measured the cost efficiency of 43 Chinese banks over the period 1993 to 2000. Their objective was to identify the change in Chinese banks' efficiency after the government program of deregulation in 1995. Their obtained results indicated that the SOCBs and CCBs are more efficient than JSCBs.

The stochastic frontier analysis was first proposed by Aigner et al.[1] and Meeusen and Van den Broeck[20]. Efficiency studies of banks based on SFA (parametric approach) have also been applied to Chinese banking system. Liu and Song [27] measured cost efficiency of four SOCBs and ten JSCBs over the period 1996-2001 by using SFA. The results show that SOCBs performed worse than JSCBs because of their lower cost efficiency. At the same year, another interesting Chinese banks' efficiency analysis based on SFA was provided by Yao et al.[26]. They used a panel data of 22 banks over the period 1995—2001 to investigate the effects of ownership structure on cost efficiency based SFA. Their results suggested that JSCBs are 11%—18% more efficient than SOCBs. Fu and Herfernan [29] investigated cost efficiency in Chinese banking system over the period 1985-2002 by employing SFA. Their results also found that JSCBs are more efficient than the SOCBs. Unlike previous studies which only focused on comparing the efficiency, they also pointed out increased privatisation, greater foreign bank participation, and liberalised interest rates should be able to help Chinese banks to increase their cost efficiencies.

The most relevant studies to this paper are the combination of DEA and SFA for banking efficiency analysis. However, the literature in this direction is quite limited. One of the early studies that focused on using both methods was Ferrier and Lovell [13]. They estimated the cost efficiency of 575 American banks for the year 1984 by applying both the SFA and DEA techniques. They found that the average efficiency scores estimated by DEA is a little higher than SFA, 0.8 and 0.74,respectively. They also suggested to take other factors into account when the efficiency scores are not correlated between these two methods. A European study was provided by Sheldon [24]. His research investigated the cost efficiency of Swiss banks by employing SFA and DEA from the period 1987 to 1991. His results presented that the average cost efficiency scores estimated by DEA was about 0.56 while SFA yielded only 0.039 average efficiency. He concluded that both techniques may not yield consistent results. In [22], Resti provided different results. He estimated cost efficiency for a panel data of 270 Italian banks. Based on his results, he found that two techniques do not differ significantly. Moreover, his analysis indicated that efficiency scores estimated by SFA are higher than DEA, which was contrary to previous research in Ferrier and Lovell[13] and Sheldon[24]. Therefore, based on different data set and different variables' settings, the results obtained from two methods are different.

Beccalli et al.[6] measured cost efficiency of stock-market listed European banks in 1999 and 2000. Their objective was to investigate the relationship between bank performance and their share prices by using both SFA and DEA. They found that stock prices are positively related to cost efficiency. Fiorentino et al. [14] measured 34,192 observations for all German universal banks over the period 1993-2004 to examine whether both techniques yield the similar results. Their results showed that SFA efficiency scores are higher than DEA efficiency scores. There are currently only two papers that combine both techniques for the Chinese banking industry. Luo et al.[19] proposed the first study that evaluates the Chinese banks' efficiency by using two different techniques. Their results showed that the Chinese banking reform had made significant progress during the period 1999-2008. A comparison research of whether both techniques can lead to similar results was provided by Dong et al.[11]. They investigated the cost efficiency and scale efficiency based on a panel data set of Chinese commercial banks from 1994 to 2007. Their results indicated moderate consistency between these two approaches. In addition, they suggested that cross checking of both approaches lead more convincing measurements of bank performance. In other words, they suggest to use both methods together so that you compare which method yield more convincing results.

**Table 2.1:** Summary of Chinese banking studies

| Author | Data period | Method | Results. |
|---|---|---|---|
| Zhang (2003) | 1997-2001 | DEA | JSCBs are the most efficient banks |
| Chen et al.(2005) | 1995-2004 | DEA | SOCBs and CCBs are more efficient than JSCBs |
| Arrif (2008) | 1993-2000 | DEA | JSCBs are more efficient than SOCBs |
| Liu (2004) | 1996-2001 | SFA | JSCBs are more efficient than SOCBs |
| Yao et al.(2004) | 1995-2001 | SFA | JSCBs are more efficient than SOCBs |
| Fu (2007) | 1985-2002 | SFA | JSCBs are more efficient than SOCBs |
| Luo et al.(2011) | 1999-2008 | DEA,SFA | The two sets of efficiency results are not comparable |
| Dong et al.(2014) | 1994-2007 | DEA,SFA | The results show moderate consistence between two methods |

To summarize, according to **Table 2.1**, we can see that most studies show JSCBs are more efficient than SOCBs and CCBs, but only one paper show the contrary. This might be the reason that Chen et al. [28] selected different banks' sample than other papers as their data sample involves the trust and investment banks. Moreover, they used DEA for 43 banks over the period 1993-2000. Since DEA has already been criticized for its sensitivity to the outliers, other factors affecting efficiency may not be accounted for. We can notice the only two studies combining both methods for Chinese banks lead to different results. This might be the reason that Luo et al.[19] used production function while Dong et al.[11] used cost function to measure the Chinese banks' efficiencies. In addition, most of studies which combine both parametric and non-parametric approaches focus on analyzing whether both techniques lead to consistent results or not. And there is no consensus whether these two approaches lead to the same conclusion. Additionally, only two studies evaluate the Chinese banking system by using both techniques. And both of their research focus on the time period before global financial crisis. However, to the best of my knowledge, no studies have investigated benchmarking analysis of Chinese banks using both approaches after the period of financial crisis. Therefore, this study involves a benchmarking analysis of Chinese banks over the period 2009-2014 by using DEA and SFA.

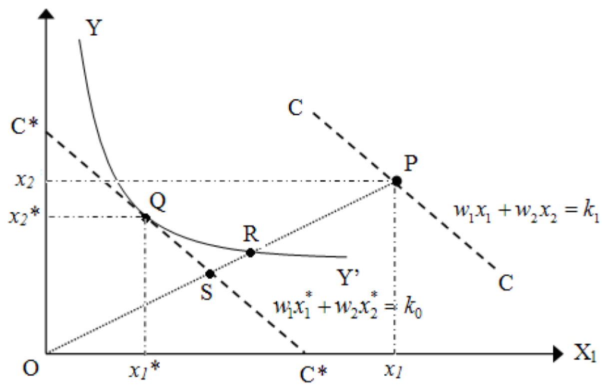
# 3 | Method

In this chapter, I describe two approaches which can be applied to estimate the Chinese banks' efficiencies: one is DEA using mathematical linear programming and another is SFA based on econometric tools. In the literature, a firm is called fully efficient if it uses various inputs to produce the output levels that maximize profits or minimize possible costs. A firm is said to be cost efficient if it is able to transform inputs at minimum cost into maximum outputs. As can be seen from previous chapter that most of papers measure cost efficiency for banking industry. Hence, the objective of this research is to focus on analyzing the cost efficiency.

## 3.1 | Efficiency Concepts

The objective of this section is to provide a number of efficiency concepts, which were first proposed by Farrel [12]. His research plays a fundamental role for the both parametric and non-parametric approaches. The efficiency measurements can be best illustrated by a isoquant diagram.

**Figure 3.1:** Technical, Allocative and Cost Efficiency




Cooper et al. (2007, p.258)

In **Figure 3.1**, the two inputs and single output isoquant diagram is presented. We can see horizontal axis is $X_1$ while vertical axis is $X_2$, $YY'$ is represented as an efficient frontier which indicates minimum combinations of inputs that are used to produce a unit of output. Every firm that lies on $YY'$ is considered as efficient unit, such as firm $Q$. Firm $P$ is identified as an inefficient firm since it lies on the right side of the efficient frontier. In other words, it uses more inputs to produce a unit of output. Hence, the technical inefficiency of firm $P$ can be defined as $RP$, which represents the amount by which all inputs can be decreased without reducing the amount of output. Generally, the technical efficiency (TE) is given by a ratio $OR/OP$, which represents the percentage by which all inputs can be decreased. The value of TE is always between zero and one.

If the input price ratio is also known, which is represented by the isocost line $CC$, allocative efficiency (AE) can be computed. We can see the cost of firm $P$ can be further reduced if we move $CC$ to line $C^*C^*$, where the minimal cost is achieved at the given output level. The relative distance of $S$ and $R$, which is the ratio of $OS/OR$, can be defined as AE of the firm $P$. The ratio indicates a cost decrease can be achieved by a firm which move from technically efficient but not allocatively efficient point $R$ to both efficient point $Q$.

The cost efficiency can be defined by the ratio of optimal cost ($wx^*$) to the actual cost ($wx_1$), which is the ratio $OS/OP$. Therefore, a firm which can choose inputs based on their prices in order to minimize total cost can be identified as a cost efficient firm. In addition, cost efficiency can be seen as the product of TE and AE. Cost efficiency value 1 means that the firm is fully cost efficient while 0.8 indicates that the firm is 20% less cost efficient than the best practice firms (benchmarks) under the same condition.

It is obvious that measuring cost efficiency and identifying the benchmarks by graphical techniques is impossible for banks as they have multiple inputs and outputs. Therefore, parametric (SFA) and non-parametric (DEA) approaches are proposed to solve this problem.

## 3.2 | Data Envelopment Analysis

DEA developed by Charnes et al. [8] is a linear programming based technique, which aims to improve the efficiency measurement proposed by Farrell whose ideas were limited in number of inputs and outputs. The essence of DEA is to utilize mathematical linear programming to measure the performance of decision making units (DMUs) which try to convert multiple inputs into multiple outputs. In other words, this technique forms an envelopment surface and measures the efficiency relative to the efficient frontier. The envelopment form of linear programming is derived from multiplier form, which is proposed by Charnes et al. [8]. They first define the efficiency as a maximum ratio of weighted outputs to weighted inputs, which is the same definition as productivity. As in this efficiency ratio, we can see it is possible to change the efficiency by either changing input or output level. In my research, I assumes bank managers are only able to have control in inputs. Hence, the input-oriented model is used for DEA. The input-oriented DEA model objects to maximize the proportional reduction in inputs as much as possible so as to achieve relative efficiency, given the same output level. An efficiency ratio that equals to 1 means that there is no wastage of inputs to produce the given amount of outputs, which is the perfect situation for any DMUs. Therefore, the linear programming aims to maximize efficiency ratio for each DMU, subject to the conditions that the efficiency ratio of the other DMUs should be less or equal to 1. The linear programming (multiplier form) is defined as :

$$\begin{aligned} \text{maximize } h_o &= \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \\ \text{s.t } \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} &\leq 1 \quad , \quad j = 1, 2, .., n, \\ u_r, v_i &\geq 0 \quad , \quad r = 1, 2, .., s; i = 1, 2, .., m \end{aligned}$$

In this formulation, "o" denotes a focal DMU (i.e., each bank, in turn, becomes a focal bank when its efficiency score is being computed). $x_{ij}$ is the observed amount of the $ith$ input of the $jth$ DMU, $y_{rj}$ is the observed amount of the $rth$ output of the $jth$ DMU. The $u_r$ and $v_i$ are non-negative weights which are determined by the above linear programming. However, one may find out infinite number of solutions by solving such programming approach, if $(u^*, v^*)$ is a solution, then $(\alpha u^*, \alpha v^*)$ is another solution for any non-negative $\alpha$. To avoid such problem, Charnes et al. [8] imposed the constraint $\sum_{i=1}^{m} v_i x_{io} = 1$, which provides:

$$
\begin{aligned}
\text{maximize } z_o = \sum_{r=1}^{s} u_r y_{ro} & \\
\text{s.t } \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0, & \quad j = 1, 2, .., n, \\
\sum_{i=1}^{m} v_i x_{io} = 1 & \\
u_r, v_i \geq 0 \qquad\qquad , & \quad r = 1, 2, .., s; i = 1, 2, .., m
\end{aligned}
$$

In order to derive an equivalent envelopment form, the duality in linear programming is used:

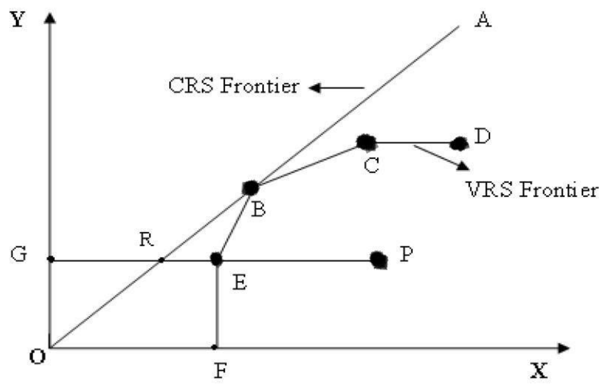
$$
\begin{aligned}
\text{minimize } z_o = \theta_o & \\
\text{s.t } \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_o x_{io}, & \quad i = 1, 2, .., m, \\
\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \quad , & \quad r = 1, 2, .., s, \\
\lambda_j \geq 0 &
\end{aligned}
$$

The objective function tries to minimize the efficiency $\theta$ subject to the constraints such that the weighted sum of the inputs of the other DMUs is less than or equal to the inputs of the DMU being evaluated and the weighted sum of the outputs of the other DMUs is larger than or equal to the DMU being evaluated. The weights $\lambda$ are non-negative values. Optimal solutions $(\theta, \lambda)$ are obtained by solving above linear programming N times, once for each DMU. The value of $\theta$ is called technical efficiency, according to the definition of Farrell we mentioned before. The value of $\theta$ is always less than or equal to 1 based on the constraints and the efficiency score $\theta$ computed for each DMU is relative to other DMUs. Accordingly, DMU for which $\theta = 1$ is considered as technically efficient firm. The optimal $\lambda$ identify benchmarking points (best performers) which locate on the efficient frontier when the problem seeks the reduction of inputs. Moreover, $\lambda$ can identify the shape of DEA frontier. Different constraints on $\lambda$ could lead to different DEA models. The assumption of this model is constant return to scale, which means that all DMUs are operating at an optimal scale. Therefore, this model is also called CRS model. However, Chinese banking industry are not fully developed and perfect competition is unlikely. Therefore, constant return to scale are not suitable for such case but variable return to scale (VRS) can be considered. The VRS model proposed by Banker et al. [3] can be written formally as:

$$
\begin{aligned}
&\text{minimize } \theta_o \\
&\quad \text{s.t } \sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_o x_{io}, \quad i = 1, 2, .., m, \\
&\qquad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \quad , \quad r = 1, 2, .., s, \\
&\qquad \sum_{j=1}^{n} \lambda_j = 1 \geq 0 \\
&\qquad \lambda_j \geq 0 \qquad\qquad , \quad j = 1, 2, .., n
\end{aligned}
$$

The difference compared to CRS model is that they added convexity condition for the weights $\lambda$, which ensures that an inefficient DMU is only compared to the similar sized efficient DMUs. By adding such constraint, the VRS frontier have piecewise linear and convex characteristics which are different from the conical hull of CRS model. These two models can be better explained by a single input and output diagram.

**Figure 3.2:** CRS and VRS model




Cooper et al. (2007)

The above two envelopment linear programming form can be linked to **Figure 3.2**, we can see two different frontiers due to two different linear programming approaches. OA represents a CRS frontier and DCBEF represents a VRS frontier. Firms $P, E, B, C, D$ are located below the CRS frontier. Both models aim to compute the efficiency of firms. If we look at the firm P, it locates below the VRS frontier, which mean it is not cost efficient. Under the VRS assumption, firm P can identify firm E as benchmark by reducing the input to F so to produce the same amount of output. Therefore, the technical efficiency $\theta_p$ under VRS assumption is $GE/GP$. As I mentioned before, $\lambda$ aims to identify the best performers, for firm P being evaluated, $\lambda_E = 1$ and the rest of $\lambda$s are all zero.

The above models only use quantity data to measure the TE. However, if the input prices are also known, the CE can be estimated by the following cost minimization DEA model:

$$\begin{aligned}
&\text{minimize } w_{io}x_{io}^{*} \\
&\quad \text{s.t} \sum_{j=1}^{n} \lambda_j x_{ij} - x_{io}^{*} \leq 0, \quad i = 1, 2, .., m, \\
&\quad \sum_{j=1}^{n} \lambda_j y_{rj} - y_{ro} \geq 0, \quad r = 1, 2, .., s, \\
&\quad \sum_{j=1}^{n} \lambda_j = 1 \\
&\quad \lambda_j \geq 0 \qquad , \quad j = 1, 2, .., n
\end{aligned}$$

where $x_{io}^{*}$ is the cost-minimizing vector of inputs for the *oth* DMU, given the input prices $w_{io}$ and outputs $y_{ro}$. The main differences between cost minimization DEA model and technical efficiency DEA model are as follows. First, the objective function is changed from minimizing the technical efficiency into minimizing the cost of inputs. Since the input prices are given, the purpose of this objective function is to find the minimum combination of inputs. Accordingly, the first constraint is changed to adapt the objective function.

According to **Figure 3.1**, the above linear programming aims to examine the minimum input cost point S for firm P. Then, the CE of the firm P is defined as the ratio of the minimum cost to observed cost, which is $OS/OP$. Another important task for DEA is to identify the benchmarks. In this case, the benchmark for firm P is firm Q since firm P tries to optimize their combinations of inputs based on the optimal information of firm Q. If multiple inputs and outputs are used, then multiple benchmarks will be possible.

A general CE is defined as follows:

$$CE = \frac{w_{ij}x_{ij}^{*}}{w_{ij}x_{ij}} \tag{3.1}$$

As mentioned before, input-oriented DEA model allows to reduce inputs without changing outputs to achieve efficiency. These inputs reduction or savings are defined as input slacks. The input slacks can be seen as an important indicator to help bank managers to improve their banks' performances. The input slacks under cost minimization DEA model can be calculated by the first constraint of the above cost minimization DEA model.

## 3.3 | Stochastic Frontier Analysis

The stochastic frontier analysis was first proposed by Aigner et al.[1] and Meeusen and Van den Broeck[20], since then, it has also been used to evaluate bank efficiency. The essential idea of SFA can be better illustrated by a stochastic cost function model:

$$\ln TC_{it} = \ln C(y_{it}, w_{it}) + \epsilon_{it} = \ln C(y_{it}, w_{it}) + v_{it} + u_{it} \tag{3.2}$$

Where $TC_{it}$ represents the total cost, $C(y_{it}, w_{it})$ stands for cost function, $y_{it}$ is output variables and $w_{it}$ is input prices. $v_{it}$ is the two sided iid random error which represents measurement errors or other random shocks that could affect the value of the output variable. The random error $v_{it}$ together with cost function $C(y_{it}, w_{it})$ are constructed as an stochastic cost frontier, which is determined by $C(y_{it}, w_{it}) \exp(v_{it})$. Moreover, $v_{it}$ can be positive or negative and so the stochastic frontier outputs vary relative to the deterministic cost frontier $C(y_{it}, w_{it})$ (Fiorentino et al.[14]). $u_{it}$ is a non-negative inefficiency which indicates the distance between each firm and the efficient frontier and is usually assumed as a half-normal distribution, $N^+(0, \sigma_u^2)$. In addition, $v_{it}$ and $u_{it}$ are independently distributed from each other and are also time and bank specific. CE is estimated relative to the efficient cost frontier, which is defined by the ratio of optimal cost to the actual cost. This model works fine to estimate CE. But it fails to incorporate the possibility that the inefficiency term can be explained by a set of environmental variables. Therefore, early papers such as

Pitt and Lee [21] try to add the explanation of inefficiency effects by doing the second regression. In other words, they first estimate the stochastic frontier function and predict the inefficiency effects under the assumption that these inefficiency effects are identically distributed. Then the second stage is to regress inefficiency effects on some factors such as environmental variables. However, this contradicts the assumption of identically distributed $u_{it}$ in the stochastic frontier. In order to avoid such situation, Battese and Coelli [5] proposed a one-stage regression to estimate the stochastic frontier and an inefficiency model simultaneously. The inefficiency model is defined as follows:

$$u_{it} = z_{it}\delta + W_{it} \tag{3.3}$$

where random error $W_{it}$ is a truncated normal distribution with mean zero and variance $\sigma^2$, such that the point of truncation is $-z_{it}\delta$ and $W_{it} \geq -z_{it}\delta$. These assumptions are consistent with $u_{it}$ being a non-negative truncation of the $N(z_{it}\delta, \sigma^2)$ distribution (Battese and Coelli [5]). $z_{it}$ is explanatory variables influencing inefficiency, such as ownership, market share, etc. $\delta$ is a vector of parameter to be estimated. Hence, $u_{it}$ is a non-negative truncated normal distribution with mean $z_{it}\delta$ and variance $\sigma^2$. To incorporate inefficiency model (3.2) into model (3.1), the CE can be defined as:

$$CE = \frac{C(y_{it}, w_{it})\exp(v_{it})}{C(y_{it}, w_{it})\exp(v_{it} + z_{it}\delta + W_{it})} = \exp(-z_{it}\delta - W_{it}) \tag{3.4}$$

The estimation of the cost efficiency is based on it conditional expectation. The conditional expectation of $\exp(-u)$ given $e = v - u$ is

$$E(\exp(-u)|e) = \{\exp(-\mu_* + 1/2\sigma_*^2)\} * \{\Phi[(\mu_*/\sigma_*) - \sigma_*]/\Phi(\mu_*/\sigma_*)\} \tag{3.5}$$

where

$$\mu_* = \frac{\sigma_v^2 z\delta - \sigma^2 e}{\sigma_v^2 + \sigma^2} \tag{3.6}$$

and

$$\sigma_*^2 = \sigma^2\sigma_v^2/(\sigma^2 +_v^2) \tag{3.7}$$

The complete proof or derivation can be found in the Appendix of Battese and Coelli[4]

In the literature, the Cobb-Douglas form and the translog cost functional form have been widely used for input to output studies. Kumbhakar and Lovell [18] indicated that translog cost functional form fits better than Cobb-Douglas form. Moreover, most of literature measuring banks use translog cost functional form. Therefore, this paper adopts translog cost functional form as well. The translog functional form proposed by Christensen et al.[9] is a second order approximation of any unknown function. Thus, it provides flexibility relative to a parametric functional form. With the help of translog cost functional form, the model can be specified as follows, the time index is dropped for simple notation

$$\begin{aligned} \ln \frac{TC_o}{w_{or}} =& \beta_0 + \sum_{i=1}^{s} \beta_{oi} \ln(y_{oi}) + \sum_{m=1}^{r-1} \alpha_{om} \ln(\frac{w_{om}}{w_{o3}}) + \frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{s} \theta_{oij} \ln(y_{oi}) \ln(y_{oj}) \\ &+ \frac{1}{2}\sum_{m=1}^{r-1}\sum_{n=1}^{r-1} \eta_{omn} \ln(\frac{w_{om}}{w_{o3}}) \ln(\frac{w_{on}}{w_{o3}}) + \sum_{i=1}^{s}\sum_{m=1}^{r-1} \gamma_{oim} \ln(y_{oi}) \ln(\frac{w_{om}}{w_{o3}}) \\ &+ u_o + v_o \end{aligned} \tag{3.8}$$

Where

$$u_o = z_0 + \sum_{l=1}^{d} z_{ol}\delta_l + W_o \tag{3.9}$$

In this model, $o$ denotes as a focal firm, and we have s outputs $y_{oi}$, r input prices $w_{om}$ and d environmental variables $z_{ol}$ which may influence the distance of each bank to the efficient frontier. s denotes as number of outputs, r is number of inputs and d represents for the number of environmental variables. $\beta, \alpha, \theta, \eta, \gamma$ are parameters to be estimated. The total cost and input price terms are normalised by the last input price $w_{or}$, which show that a linear homogeneity restriction is imposed on the cost function to ensure price homogeneity. In other words, if all input prices are multiplied by the same positive scalar, the cost minimizing term will not be changed. Last, translog cost function requires that cross price effects must be symmetric. This is because the function is continues and twice differential, the second cross derivatives are symmetric. Therefore, the following symmetry restrictions need to be added:

$$\theta_{oij} = \theta_{oji}, \eta_{omn} = \eta_{onm}$$

# 4 | Data

For this research, I use a panel data set which contains information on 22 Chinese banks over the period 2009-2014, totally 132 observations. The data are mainly drawn from *Bankscope: world banking information source*. For missing data like number of employees for each bank in *Bankscope*, I collect from annual reports of individual banks. As mentioned before, the sample is classified into three groups: SOCBs, JSCBs, CCBs. The SOCBs group include five banks, which are the most powerful banks in China. I select nine out of total 12 as JSCBs group. These nine JSCBs accounted for 80% of JSCBs total market share in 2010. Moreover, the eight most representative CCBs are selected as the last group. As it can be seen from literature, some studies select 30 or even more banks for their research. There are two main reasons why I choose 22 banks for my research. First, these 22 banks can, to some extent, reflect overall performance of Chinese banking industry. Second, it takes too much time to manually collect data from all other banks.

In the literature, there are two best known approaches for the measurement of bank efficiency provided by Sealey and Lindley[23], which are the intermediation approach and the production approach. The intermediation approach considers banks as intermediates that transform the capital, labour and deposits as inputs into loans and earning assets as outputs. The production approach supposes that banks use inputs such as capital and labour to produce a number of deposits and loans. The main difference is that the intermediation approach treats deposits as input variables while the production approach treats deposits as output. The intermediation approach focuses on the overall costs of banks, which is relevant for addressing questions that are related to the affected banks (Ferrier and Lovell[13]). Additionally, Berger and Humphrey [7] mention that the intermediation approach is more appropriate for measuring the efficiency of banks while the production approach is more relevant for evaluating branch level efficiency. Last, most literature use the intermediation approach as it is better suited to capture the decisions that are taken to minimize the cost of the financing mix (Fortin and Leclerc[15]). Therefore, I also adopt the intermediation approach for my research. Under this approach, the outputs are specified as total loans ($y_1$), which include gross loans and reserves for impaired loans. The other earning assets ($y_2$) are comprised of advances to banks, derivatives and other securities. The first input is specified as the total deposits with short term funding ($x_1$) which include total customers deposits, deposits from banks and short-term borrowings. The rest of inputs are defined as number of employees ($x_2$) and fixed assets ($x_3$). The assumption that needs to make is all input variables are under control of the banks.

Accordingly, the input prices are defined as follows. The price of deposits ($w_1$) is defined as the ratio of total interest expenses on the total deposits with short term funding. The price of labour ($w_2$) is calculated by the personal expenses to the number of employees. Last is the price of fixed assets ($w_3$), also called physical capital, which is measured by the ratio of other operating expenses to the fixed assets. In line with literature, the environmental variables for inefficiency effects are defined as follows. The sample has three groups which have different ownership. I consider SOCBs as base group, and Dummy1 ($z_1$) for JSCBs while Dummy2 ($z_2$) for CCBs. Market share ($z_3$) is defined as the ratio of an individual bank's total assets to the total assets of all listed 22 banks in a given year. **Table 4.1** presents a summary of all the variables.

**Table 4.1:** Descriptive Statistics of Input and Output Variables between 2009 and 2014.

| Variables | Obvs | Mean | Std.Dev. |
|---|---|---|---|
| Total Cost $TC$ (in CNY million) | 132 | 226,098 | 1,347,228 |
| Total Loans $y_1$ (in CNY million) | 132 | 1,842,332 | 2,566,471 |
| Other Earning Assets $y_2$ (in CNY million) | 132 | 1,184,251 | 1,541,992 |
| Number of employees $x_1$ | 132 | 87276 | 142747 |
| Deposits and Borrowings $x_2$ (in CNY million) | 132 | 3,249,267 | 4,442,114 |
| Fixed Assets $x_3$ (in CNY million) | 132 | 30,777 | 49,988 |
| Price of Labour $w_1$ (in CNY million) | 132 | 0.294 | 0.079 |
| Price of Deposits $w_2$ (%) | 132 | 23% | 18% |
| Price of Fixed Assets $w_3$ (%) | 132 | 77% | 40% |
| Dummy1 for JSCBs $z_1$ | 132 | 0.409 | 0.494 |
| Dummy2 for CCBs $z_2$ | 132 | 0.364 | 0.483 |
| Market Share $z_3$ | 132 | 0.17 | 0.231 |

It is interesting to notice that the standard derivation of total cost, output and input variables are very high, showing there are large variation in these variables. This is because Chinese banks have developed very fast during the period from 2009 to 2014. Moreover, the difference between large sized banks and small-sized banks might also be the reason.

# 5 | Evaluation

## 5.1 | Empirical results of SFA

The cost efficiency of Chinese banks over 2009-2014 based on SFA are computed by the software Frontier 4.1. Frontier 4.1 is used to obtain maximum likelihood estimates of the parameters of a variety of stochastic cost frontiers. However, the programme does not deal with transformations of variables for translog cost functions. Therefore, I use R software to first construct the logarithms of inputs, outputs and their interaction terms for translog cost functions. After that, Frontier 4.1 is used to obtain maximum likelihood estimates and cost efficiency of each bank in our sample. The parameter estimates of stochastic cost function using maximum likelihood are provided in TABLE 5.1.

We can see the estimated translog cost function has many interaction terms between input prices and output variables from Table 5.1. However, these terms are not directly interpretable. Table 5.1 shows that the parameter estimates of output and input variables are positive, which mean cost function is non-decreasing in outputs and input prices. These are the theoretical requirements for cost function which we satisfy. Two output variables, one input price and several interaction terms are statistically significant at 5% level. In order to examine whether increases in input price or outputs would increase total cost, the average marginal elasticities need to be computed. The average marginal elasticities for TOL, OEA, POL and POD are 1.61, 0.334, 1.22 and 0.09, respectively at 5% significance level. This is in line with my expectations, an increase in input prices and outputs increases total costs. Moreover, the results imply that the total costs are more sensitive to price of labour than price of deposits, which is in line with the real situation of banks. Moreover, we can notice the environmental variables influencing the efficiency term are all statistically significant, which imply they all have influences on cost efficiency. The nega-

**Table 5.1:** Maximum Likelihood Parameter Estimates for SFA Functions

| Variables | Parameter | Coefficient | Standard errors |
|---|---|---|---|
| Constant | $\beta_0$ | 3.877*** | 0.916 |
| ln(TOL) | $\beta_1$ | 0.929*** | 0.281 |
| ln(OEA) | $\beta_2$ | 0.837*** | 0.191 |
| ln(POL/POFA) | $\alpha_1$ | 1.302 | 0.106 |
| ln(POD/POFA) | $\alpha_2$ | 0.536*** | 0.155 |
| 0.5*ln(TOL)$^2$ | $\theta_{11}$ | -0.725 | 0.946 |
| ln(TOL)*ln(OEA) | $\theta_{12}$ | -1.021 | 0.913 |
| 0.5*ln(OEA)$^2$ | $\theta_{22}$ | -0.027 | 0.609 |
| 0.5*ln(POL/POFA)$^2$ | $\eta_{11}$ | -1.377** | 0.637 |
| ln(POL/POFA)ln(POD/POFA) | $\eta_{12}$ | 1.63* | 0.86 |
| 0.5*ln(POD/POFA)$^2$ | $\eta_{22}$ | 0.27 | 0.591 |
| ln(TOL*ln(POL/POFA) | $\gamma_{11}$ | -0.107* | 0.08 |
| ln(TOL)*ln(POD/POFA) | $\gamma_{12}$ | -0.196*** | 0.061 |
| ln(OEA)*ln(POL/POFA) | $\gamma_{21}$ | 0.111 | 0.088 |
| ln(OEA)*ln(POD/POFA) | $\gamma_{22}$ | 0.145** | 0.072 |
| intercept | $\delta_0$ | 0.354*** | 0.071 |
| JSCBs $z_1$ | $\delta_1$ | -0.207*** | 0.064 |
| CCBs $z_2$ | $\delta_2$ | -0.301*** | 0.074 |
| Market Share $z_3$ | $\delta_3$ | 0.425*** | 0.109 |
| Sigma$^2$ | $\sigma^2$ | 0.61*** | 0.12 |
| Gamma | $\gamma$ | 0.99*** | 0.002 |

*Notes*: Outputs are Total Loans (TOL), Other Earning Assests (OEA); Input prices are Price of Labour (POL), Price of Deposits (POD), Price of Fixed Assets (POFA). Sigma$^2$ denotes the total amount of variance in the model. Gamma gives the ratio of variance of the inefficiency term over the total amount of variance. $***$ significance at 1%, $**$ significance at 5%, $*$ significant at 10%

tive signs of two dummy variables show that JSCBs are 20.7% more efficient than SOCBs while CCBs are 30.1% more efficient than SOCBs. In addition, the positive sign of market share indicate the banks with larger market share will have lower cost efficiency. Last, the variance parameters of the stochastic cost function are represented by Sigma squared and Gamma. According to the Table 5.1, the Sigma squared is 0.61 which indicates a good fit and correctness of the distribution form assumed for the composite error term. The estimate for the Gamma is close to one, indicating that the inefficiency effects are highly significant in the analysis of the total costs of the banks. This means that 99% of the variation in banks' total costs are due to cost efficiency.

Table 5.2 provides a statistic of estimated cost efficiency scores for all banks over the period 2009-2014. The most cost efficient banks are CHONGQING and HANGZHOU bank from CCBs group with mean efficiency 0.953. Moreover, HANGZHOU bank became fully cost efficient one with efficiency score 1 in 2014, which is the only fully cost efficient bank in my sample. ABC from SOCBs group is found to be the least efficient bank with mean cost efficiency 0.482, implying that ABC could potentially reduce input costs by approximately 52% by using its inputs more efficiently at the given level of output. Most of banks in my sample exhibit growing patterns during 2009-2014, but several banks are quite different. CITIC bank experienced a small efficiency decrease from 2009 (0.88) to 2011 (0.82), then a rapid increase was achieved in 2012 with cost efficiency score 0.97. The reason can be explained by their large decrease in total loans from 2011 to 2012, which reduced the total costs and yielded higher efficiency score. However, they showed a large decline from 0.97 to 0.86 in 2013. The result might due to the fact that their price of labour, total loans and other earning assets were largely increased, which led a rapid growth in total costs and low efficiency. PINGAN bank was the most cost efficient bank in JSCBs group with mean efficiency score 0.942. Although their good performances were quite stable compared to other banks, they still experienced a large decrease in 2013, around 10% cost efficiency was decreased from last year. After 2013, their cost efficiency appeared to be rapidly recovered to previous level. The cost efficiency of HARBIN bank first decreased from 0.74 to 0.70. However, after 2010, they showed a significant increase from cost efficiency 0.7 to 0.97, suggesting their managerial progress in saving cost was remarkable during these five years.

Figure 5.1 provides the efficiency distributions per bank type and their average efficiency distribution based on SFA. According to Figure 5.1, the average cost efficiency dropped from 2009 to 2010, then it increased from 0.74 in 2010 to 0.8 in 2013, however, it reduced down to 0.78 in 2014. It is worth noting that average efficiency declined after world financial crisis after 2009. This may suggest that the world financial crisis in 2008 may have led a negative impact on Chinese banking total costs. Next, we can observe that SOCBs have the similar distribution pattern as Average CE. Moreover, they appear to be less cost efficient than other two groups with mean cost efficiency around 0.55, suggesting SOCBs could use inputs more efficiently so as to reduce costs by around 45%, for a given amount of outputs. It is obviously seen that there is a huge gap between SOCBs and other two groups. This might be the reason that SOCBs are five largest banks in China and they face more restrictions than medium or small sized banks. Moreover, due to the large number of branches of SOCBs, it could be imagined that SOCBs might face more difficulties than JSCBs and CCBs to have efficient management in their total costs. JSCBs and CCBs are close to each other in cost efficiency. According to Figure 5.1, both groups starting from similar positions experienced a significant decrease after 2009. Since 2010, the cost efficiency of CCBs showed a rapid rise in 2011 and remained stable increase over the next few years while the cost efficiency of JSCBs did not increase until 2013, with a share rise from 0.837 to 0.925. However, a significant decline in cost efficiency from 2013 to 2014 can be observed for JSCBs. The trend of CCBs may indicate they have gained efficiency benefits by using input variables more efficiently which might result in new management system since 2010. In summary, CCBs are the most cost efficient group in my sample. Moreover, JSCBs are more cost efficient than SOCBs, which is consistent with previous papers provided in literature chapter.

**Figure 5.1:** Average efficiency scores based on SFA over time

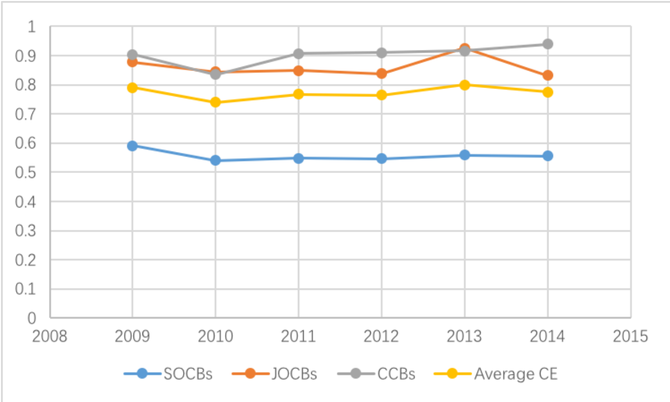



**Table 5.2:** Cost efficiency scores based on SFA for all banks over the period from 2009 to 2014

| | Bank | Year | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Mean |
| SOCBs | ICBC | 0.58 | 0.54 | 0.54 | 0.55 | 0.54 | 0.52 | 0.545 |
| | CCB | 0.6 | 0.55 | 0.56 | 0.52 | 0.58 | 0.59 | 0.567 |
| | ABC | 0.49 | 0.44 | 0.46 | 0.48 | 0.5 | 0.52 | 0.482 |
| | BOC | 0.58 | 0.53 | 0.52 | 0.54 | 0.54 | 0.54 | 0.542 |
| | BOCOM | 0.71 | 0.65 | 0.66 | 0.65 | 0.63 | 0.62 | 0.653 |
| JSCBs | PUDONG BANK | 0.89 | 0.85 | 0.87 | 0.85 | 0.86 | 0.84 | 0.86 |
| | CITIC BANK | 0.88 | 0.90 | 0.82 | 0.97 | 0.86 | 0.85 | 0.88 |
| | EVERBRIGHT BANK | 0.88 | 0.82 | 0.81 | 0.78 | 0.79 | 0.79 | 0.812 |
| | HUAXIA BANK | 0.85 | 0.81 | 0.81 | 0.8 | 0.79 | 0.8 | 0.81 |
| | MINSHENG BANK | 0.87 | 0.83 | 0.83 | 0.75 | 0.75 | 0.73 | 0.793 |
| | MERCHANT BANK | 0.79 | 0.74 | 0.75 | 0.75 | 0.76 | 0.8 | 0.765 |
| | GUANGFA BANK | 0.83 | 0.81 | 0.83 | 0.79 | 0.81 | 0.8 | 0.812 |
| | ZHESHANG BANK | 0.93 | 0.9 | 0.96 | 0.89 | 0.91 | 0.9 | 0.915 |
| | PINGAN BANK | 0.96 | 0.93 | 0.96 | 0.96 | 0.88 | 0.96 | 0.942 |
| CCBs | BEIJING BANK | 0.94 | 0.88 | 0.87 | 0.89 | 0.91 | 0.93 | 0.903 |
| | NINGBO BANK | 0.83 | 0.85 | 0.97 | 0.97 | 0.91 | 0.98 | 0.918 |
| | CHONGQING BANK | 0.97 | 0.89 | 0.96 | 0.97 | 0.99 | 0.94 | 0.953 |
| | SHANGHAI BANK | 0.91 | 0.83 | 0.89 | 0.94 | 0.95 | 0.96 | 0.913 |
| | NANJING BANK | 0.92 | 0.85 | 0.86 | 0.80 | 0.76 | 0.8 | 0.832 |
| | HANGZHOU BANK | 0.99 | 0.85 | 0.97 | 0.96 | 0.95 | 1 | 0.953 |
| | HUISHANG BANK | 0.92 | 0.82 | 0.88 | 0.90 | 0.95 | 0.93 | 0.9 |
| | HARBIN BANK | 0.74 | 0.70 | 0.85 | 0.86 | 0.90 | 0.97 | 0.837 |

## 5.2 | Empirical results of DEA model

The empirical results of DEA are computed by MaxDEA software. MaxDEA is a free software which deals with various DEA models. Unlike other free version of DEA software such as DEAfrontier which can only deal with 20 DMUs, the free version of MaxDEA has no limits in number of DMUs. The basic cross-sectional efficiency scores over the period from 2009 until 2014 under DEA cost efficiency measures are provided in Table 5.3. As we can see there exists nine banks with mean cost efficiency scores 1 during 2009-2014, which can be identified as the most efficient banks or benchmarking banks in my data sample, while the least efficient bank is HARBIN bank with the lowest mean CE 0.827. There are many more banks that are fully efficient compared to SFA.

In order to show how to explain the results of DEA, let us consider BOC from SOCBs group in Table 5.3. This bank reached cost efficiency 1 in 2009, which indicates BOC was the benchmark and used the minimum cost in inputs for producing the observed outputs. However, when we look at year 2010, the cost efficiency of BOC was not 1, which mean they were not a benchmark bank in that year. They might still use the minimum amount of inputs for producing the observed outputs, but the minimum possible costs were not guaranteed by the proportion of inputs. Therefore, there existed the space for potential cost saving in case of BOC. The cost efficiency score 0.988 indicates that BOC could decrease its cost by 1.2%. Moreover, BOC should consider ICBC and CCB as benchmarks according to Table 5.4. Slacks 1 is the input slack which we define in DEA method for number of employees, optimal 1 is the $x^*_{1BOC}$ we try to minimize for cost minimization DEA model. Hence, the potential cost saving of BOC could be obtained by increasing 24809 number of employees, by increasing total deposits to optimal value (projection deposits) 9871443 mil RMB and by reducing fixed assets to 83648 mil RMB, while a given input prices need to be maintained. BOC can be allowed to achieve minimum cost, yielding a shift on efficiency frontier by this optimal combination of inputs. The similar benchmarking analysis can be applied to other banks in my data sample.

**Table 5.3:** Cost efficiency scores based on DEA for all banks over the period from 2009 to 2014

| | Bank | Year | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Mean |
| SOCBs | ICBC | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | CCB | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | ABC | 0.865 | 0.805 | 0.795 | 0.868 | 0.97 | 1 | 0.884 |
| | BOC | 1 | 0.988 | 0.973 | 0.978 | 0.953 | 0.978 | 0.978 |
| | BOCOM | 0.974 | 0.99 | 1 | 0.956 | 0.958 | 0.957 | 0.973 |
| JSCBs | PUDONG BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | CITIC BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | EVERBRIGHT BANK | 0.925 | 0.903 | 0.929 | 0.93 | 0.912 | 0.922 | 0.92 |
| | HUAXIA BANK | 0.993 | 0.871 | 0.885 | 0.828 | 0.882 | 0.888 | 0.891 |
| | MINSHENG BANK | 0.998 | 0.934 | 0.969 | 1 | 0.876 | 0.892 | 0.945 |
| | MERCHANT BANK | 0.953 | 0.931 | 1 | 0.87 | 0.948 | 0.937 | 0.94 |
| | GUANGFA BANK | 0.835 | 0.87 | 0.977 | 0.802 | 0.894 | 0.887 | 0.878 |
| | ZHESHANG BANK | 0.942 | 0.996 | 1 | 0.999 | 1 | 1 | 0.99 |
| | PINGAN BANK | 1 | 0.93 | 0.988 | 0.99 | 1 | 0.922 | 0.972 |
| CCBs | BEIJING BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | NINGBO BANK | 0.797 | 0.91 | 0.97 | 0.944 | 0.985 | 1 | 0.934 |
| | CHONGQING BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | SHANGHAI BANK | 0.866 | 0.861 | 0.935 | 0.935 | 0.958 | 1 | 0.926 |
| | NANJING BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | HANGZHOU BANK | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | HUISHANG BANK | 0.877 | 0.959 | 0.968 | 0.956 | 1 | 0.981 | 0.957 |
| | HARBIN BANK | 0.772 | 0.845 | 0.865 | 0.817 | 0.823 | 0.841 | 0.827 |

Figure 5.2 provides the efficiency distributions per bank type and their average efficiency distribution from DEA. According to Figure 5.2, average CE to the entire samle first showed a increasing rate from 0.948 in 2009 to 0.962 in 2011, then it dropped to 0.949 in 2013, finally it grew up to 0.952 in 2014. The results indicate that Chinese banking sector based on my sample performed best in 2011 during the period from 2009 to 2014. When we look at three groups,

**Table 5.4:** Benchmarks and input slacks for SOCBs in 2010

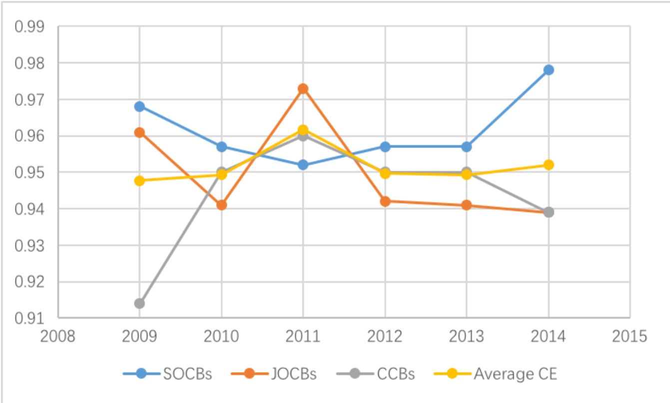
| banks | CE | benchmarks | slacks1 | optimal1 | slacks2 | optimal2 | slacks3 | optimal3 |
|---|---|---|---|---|---|---|---|---|
| ICBC | 1 | ICBC | 0 | 397339 | 0 | 12289615 | 0 | 103412 |
| CCB | 1 | CCB | 0 | 313867 | 0 | 9845295 | 0 | 83434 |
| ABC | 0.805 | ICBC,PUDONG | -155993 | 288454 | -249991 | 9258334 | -46032 | 75359 |
| BOC | 0.988 | ICBC,CCB | 24809 | 314760 | 557151 | 9871443 | -39920 | 83648 |
| BOCOM | 0.99 | CCB,CITIC | 7840 | 92756 | 55764 | 3640759 | -7727 | 26184 |

*Notes*: slacks1, 2, 3 represent input slack for number of employees, total deposits and fixed assets, respectively; optimal1, 2, 3 indicate the optimal value for number of employees, total deposits and fixed assets, respectively.

we notice these three groups show completely different cost efficiency distribution. However, it is worth noting that both JSCBs and SOCBs groups showed a significant decrease while CCBs experienced a sharp increase since 2009. This might suggest CCBs had seized the opportunities to grow rapidly while big and medium-size banks suffered from world financial crisis after 2008. CCBs showed a very similar pattern as Average CE distribution except they started with very low cost efficiency scores 0.913 in 2009 and ended with cost efficiency scores 0.939 in 2014. The cost efficiency of JSCBs showed a decrease from 2009 to 2010, suggesting JSCBs might suffer from their bad managerial strategies in input controls. After that, the cost efficiency of JSCBs increased rapidly to 0.972 in 2011, but then they showed a continuous decrease from 2011 to 2014. The cost efficiency of SOCBs experienced a decline from 2009 to 2011. However, since 2011, they exhibited a continuous increase until the end of my sample.

This result shows that SOCBs are the most cost efficient banks in my sample, indicating big banks utilize their resources better than small banks. This result is in line with previous study in DEA (Dong [10]), but contrary to the previous results provided by SFA. Therefore next, it is interesting to have a comparison between these two results.

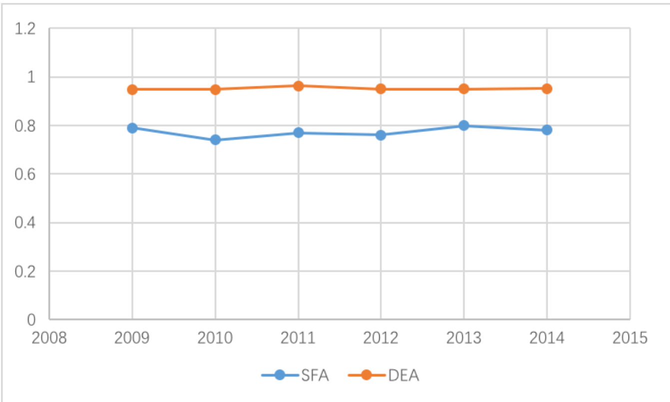
**Figure 5.2:** Average efficiency scores based on DEA over time




## 5.3 | Comparison between SFA and DEA

This section analyzes the comparison of results between SFA and DEA, which focuses on efficiency distributions and efficiency rankings.

**Figure 5.3:** Average efficiency scores between SFA and DEA over time




According to Figure 5.4, both DEA and SFA results show that average cost efficiency scores were roughly stable from 2009 to 2014, suggesting that the performance of Chinese banks sector did not show a significant change over the defined years. Therefore, both models provide the same conclusion for the performance of Chinese banks in my sample. However, the differences between these two models are quite obvious. There exists a big gap of efficiency scores between

DEA and SFA where the DEA CE scores are higher than SFA CE scores. More importantly, the results obtained from DEA indicate that big banks (JSCBs) are more cost efficient than small banks while the results according to SFA show contrary conclusion. It is not surprising that both models yield different results because DEA fails to account for heterogeneity of inefficiency effects for each bank. In other words, the results obtained from DEA are under the condition where every bank faces the same environmental situation to transform inputs into outputs while SFA does account such environmental specifics of different banks. Therefore, I also try to estimate cost efficiency of SFA model without incorporating environmental variables, which is model (3.2) provided before. It turns out that average CE scores are higher than the average CE scores of SFA accounting for time varying inefficiency effects (0.83 vs 0.77), but average efficiency scores are still lower than DEA model (0.83 vs 0.95). Thus, we can see heterogeneity plays an important role to capture the time-varying effects which causes cost efficiency around 6% lower. However, heterogeneity is not the only reason that causes the different results of DEA and SFA. From a theoretical point of view, DEA and SFA are very different which might be another important reason that can be explained the different results. The SFA approach allows banks to depart from efficient frontier because of random shocks or statistical noises while DEA does not account for random shocks. DEA uses linear combinations of inputs and outputs to come up with best performers, then the frontier is constructed by connecting all the best performers under convexity hull restriction. So if measurement errors occur which means the frontier might be shifted a little bit up, then the efficient banks might become inefficient. In other words, the cost efficiency estimated by DEA might be overestimated. Fiorentino et al. [14] argue the fact that DEA does not account for random shocks may be the reason of the different results between both models, which is consistent with my argument. On the other hand, I find out it is very likely that SFA model underestimates the cost efficiency. As discussed before, stochastic frontier is constructed by the cost function and measurement errors. The cost function is translog cost function which has been widely accepted as one of the most effective cost function to estimate banks. Hence, the question left is where the measurement errors come from. The measurement errors come from data itself. As shown in Data chapter, the standard deviation for total cost, output and input variables are very high, due to the fact that Chinese banks have developed very fast during 2009-2014. However, SFA model does not think so. It considers that the data set has a lot of measurement errors or statistical noises. In such case, the frontier could be shifted up due to the large measurement errors, the distance of the observed bank to the frontier will be larger, which mean the cost efficiency will be lower. Therefore, the large measurement errors might directly cause negative impact on efficiency measure.

In order to examine the consistency of efficiency ranks, a Spearman's rank correlation test is computed. The Spearman's rank correlation is used to evaluate how well the relationship between two ranking variables by using a monotonic function. The value 1 or -1 represent a perfect Spearman correlation, indicating each of the variables is a perfect monotone function of the other. The rank order correlation between two methods is quite low, at only -0.16. The negative value suggests there is a negative relationship between both rankings. In other words, SFA and DEA produce contrary ranking results. That is why SFA suggests that JSCBs are more cost efficient than SOCBs while the results from DEA are other way around. Berger and Humphrey [7] point out the differences between DEA and SFA are due to the fact that they tend to have different degrees of dispersion and rank banks differently. Both conflicts are exactly what I report above.

## 5.4 | Efficiency and Accounting-based Performance Measures

Accounting-based performance measures are widely used by bank managers while parametric and non-parametric approaches are more applied in academic world as they require more technical knowledge. Therefore, we can use traditional performance measures as reference to our parametric and non-parametric approaches in order to evaluate which approach is closer to traditional way in my data sample. I choose return on average assets (ROAA) and return of average equity (ROAE) as traditional performance measures, which are generally represented as the profitability of banks. Hence, higher values of both indicators are meant to imply more efficient use of bank assets or equity. Table 5.6 reports the correlations between the cost efficiency scores computed by the two frontier models and the two traditional performance measures. The results show that neither DEA or SFA have high correlations with traditional performance measures. This is in line with other research reported by Bauer et al.[14] and Dong [10] as frontier measures contain much more information than accounting-based performance. Moreover, it is acknowledged known that various state restrictions are imposed on Chinese bank sector. Accordingly, the bank management cannot fully control input and output variables, which also leads to the conflict between frontier estimations and financial indicators. The most positive information we can obtain from Table 5.6 is that the correlation between ROAA and DEA is highly significant with positive sign, suggesting that DEA measure can capture the characteristics of bank profit performance. However, on the other hand, SFA measure is negatively correlated with ROAA at lower significant level. Therefore, DEA approach for my data sample is closer to traditional accounting-based performance measures than SFA approach.

**Table 5.5:** Correlations between Frontier Efficiencies and traditional Performance Measures

| | SFA | DEA |
|---|---|---|
| ROAA | -0.19* | 0.12*** |
| ROEE | -0.05 | -0.04 |

$***$ significance at 1%, $*$ significant at 10%

# 6 | Conclusion

This paper aims to evaluate the cost efficiency of 22 selected Chinese commercial banks over the period 2009 to 2014 by employing the parametric SFA and the non-parametric DEA approaches. Moreover, I analyze the results between SFA and DEA by comparing the efficiency distribution and ranking correlation. Last, I compare two frontier efficiency results to traditional accounting-based performance measures.

In the case of the SFA model, the average cost efficiency scores of Chinese banks from 2009 until 2014 is around 0.77, which is an inefficiency level 0.23. As my sample banks are categorized by bank types (that is SOCBs, JSCBs and CCBs), the results of SFA show JSCBs and CCBs are more cost efficient than SOCBs. In other words, medium and small-sized banks are more cost efficient than large banks. According to the results of DEA model, the average cost efficiency of Chinese banks over the period 2009 to 2014 is around 0.95, which is largely higher than the results obtained from the SFA model. Moreover, the results of DEA show large banks are more cost efficient than small banks. It is not surprising to see both approaches yield inconsistent results since it has been already argued in Berger and Mester [7]. The two major reasons are summarized as follows. First, DEA uses a deterministic frontier to estimate the efficiency scores while SFA uses a stochastic frontier which is constructed by the cost function and measurement errors. As SOCBs in China have large number of branches, the measurement errors or some statistical noises might appear to be larger than JSCBs or CCBs. This might be the one of the reasons that the results of SFA show JSCBs are less cost efficient. Secondly, the SFA model is also able to capture the heterogeneity for inefficient term (such as ownership, market share, etc) while the DEA model does not. As proved before, heterogeneity plays an important role to estimate the cost efficiency. The estimation from inefficiency model also explains why JSCBs are more cost efficient. However, it is very likely that SFA underestimates the cost efficiency since there are large variations in data set. These large variations in data will bring a lot of measurement errors to the SFA model. That in turn might yield lower cost efficiency scores. On the other hand, the cost efficiency scores estimated by the DEA model ignore the possibility of measurement errors, which might cause the overestimated results as discussed before. Moreover, DEA ignores the fact there are different environmental factors among different banks every year, which might cause the result not very accurate. Overall, the results are satisfactory as both approaches find that the average cost efficiency of the selected Chinese banks exhibited a roughly stable change during the period from 2009 to 2014, suggesting the performance of Chinese banks did not show a significant change over the defined period. I believe this behavior is more important to determine the consistency of both models than the levels of efficiency as two approaches have completely different way to estimate the efficiency. Additionally, I also find DEA model is closer to traditional accounting-based performance measures than SFA approach. Therefore, I incline to the DEA results as it is closer to the reality than SFA. One might doubt that if it is true that Chinese banks did not show a significant improvement during 2009-2014 as they appeared to develop very well in reality. The answer is yes as it can be seen from DEA results, the average cost efficiency scores of Chinese banks started from 0.948 in 2009 to 0.952 in 2014, indicating that Chinese banks maintained at very high cost efficient level over the time.

Another purpose of this thesis is to determine the choice of both approaches for bank managers to apply in the future. The choice of technique is an important aspect for estimating efficiency for bank managers or researchers. One might say SFA model is more convincing as it account for measurement errors and unobserved heterogeneity influencing the inefficient term. However, the SFA method is implemented by using a specific functional form to estimate the efficiency frontier. If the specific functional form is mis-specified or not applicable to the research area, the whole results might be biased. Moreover, the result of

SFA in this study can not bring any insights to managerial level. In other words, managers simply do not know what they can do to improve the management with the results of SFA. For example, 0.97 estimated cost efficiency score from SFA only tell managers that reducing 3% input costs can achieve efficiency for a given amount of outputs but not to tell which input variables should be reduced. On the other hand, DEA model can bring rich benchmarking analysis by comparing your own firm to the identified best performers in order to learn the best practices. However, it is under the assumption that every bank faces the same environmental and technical conditions to transform inputs into outputs. Thus, if researchers want to analyze the efficiency distribution which could capture a lot of information over a large time period, then SFA is a good option. DEA could be a great option if researchers want to obtain benchmarking analysis for a relatively small time period and measurement errors are not considered as a serious problem. To conclude, this study suggests that if the bank manager wants to use an efficiency analysis for measuring the banking performance with the frontier approach, then the efficiency assessment should be used in conjunction with other instruments such as comparing to traditional performance measures, because the frontier performance measures cannot provide a coherent overview of the performance of banks.

Some limitations and potential future work are summarized as follows. First, in this paper I only analyze the cost efficiency of Chinese banks. It is also very interesting to analyze the profit efficiency as it can capture the inefficiencies on the output level. Thus, it is valuable to conduct an analysis of cost and profit efficiency. Second, when I specified the inefficiency model, I assume the used three variables for which managers have no control. If we can find relevant variables which managers are able to control for inefficiency term, we can find solutions for inefficient banks to become efficient. Moreover, since the large standard derivations in data will bring large measurement errors, which might cause the SFA results underestimated. The way to improve such problem is to add the time trend variable to the cost function model in order to control for the large variations in data over the years. Last, this thesis models the intercept term of cost function as constant over the time. The possible way to improve such model is to treat the stochastic frontier model in a 'true' fixed or random effects formulation proposed by Greene[16]. Greene's model tries to distinguish all time invariant unobserved heterogeneities from the inefficiency term by integrating such bank specific constant term in the SFA model.

# Bibliography


[1] Aigner, D., Lovell, C.K., Schmidt, P.: Formulation and estimation of stochastic frontier production function models. journal of Econometrics 6(1), 21–37 (1977)

[2] Ariff, M., Luc, C.: Cost and profit efficiency of chinese banks: A non-parametric analysis. China Economic Review 19(2), 260–273 (2008)

[3] Banker, R.D., Charnes, A., Cooper, W.W.: Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management science 30(9), 1078–1092 (1984)

[4] Battese, G.E., Coelli, T.J.: A stochastic frontier production function incorporating a model for technical inefficiency effects, vol. 69. Department of Econometrics, University of New England Armidale (1993)

[5] Battese, G.E., Coelli, T.J.: A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical economics 20(2), 325–332 (1995)

[6] Beccalli, E., Casu, B., Girardone, C.: Efficiency and stock performance in european banking. Journal of Business Finance & Accounting 33(1-2), 245–262 (2006)

[7] Berger, A.N., Mester, L.J.: Inside the black box: What explains differences in the efficiencies of financial institutions? Journal of Banking & Finance 21(7), 895–947 (1997)

[8] Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. European journal of operational research 2(6), 429–444 (1978)

[9] Christensen, L.R., Jorgenson, D.W., Lau, L.J.: Transcendental logarithmic production frontiers. The review of economics and statistics pp. 28–45 (1973)

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

[11] Dong, Y., Meng, C., Firth, M., Hou, W.: Ownership structure and risk-taking: Comparative evidence from private and state-controlled banks in china. International Review of Financial Analysis 36, 120–130 (2014)

[12] Farrell, M.J.: The measurement of productive efficiency. Journal of the Royal Statistical Society. Series A (General) 120(3), 253–290 (1957)

[13] Ferrier, G.D., Lovell, C.K.: Measuring cost efficiency in banking: econometric and linear programming evidence. Journal of econometrics 46(1), 229–245 (1990)

[14] Fiorentino, E., Karmann, A., Koetter, M.: The cost efficiency of german banks: a comparison of sfa and dea. Available at SSRN 947340 (2006)

[15] Fortin, M., Leclerc, A.: Should we abandon the intermediation approach for analyzing banking performance? Cahier de recherche/Working Paper 7, 01 (2007)

[16] Greene, W.: Fixed and random effects in stochastic frontier models. Journal of productivity analysis 23(1), 7–32 (2005)

[17] Jianhua, Z.: Dea method on efficiency study of chinese commercial banks and the positivist analysis from 1997 to 2001 [j]. Journal of Finance 3, 11–25 (2003)

[18] Kumbhakar, S., Lovell, K.: Stochastic production frontier. Cambridge University Press. Kumbhakar, SC, & Sarkar, S.(2003). Deregulation, ownership and productivity growth in the banking industry: Evidence from India. Journal of Money Credit and Banking 35(3), 403424 (2000)

[19] Luo, D., Yao, S., Chen, J., Wang, J.: World financial crisis and efficiency of chinese commercial banks. The World Economy 34(5), 805–825 (2011)

[20] Meeusen, W., Van den Broeck, J.: Efficiency estimation from cobb-douglas production functions with composed error. International economic review pp. 435–444 (1977)

[21] Pitt, M.M., Lee, L.F.: The measurement and sources of technical inefficiency in the indonesian weaving industry. Journal of development economics 9(1), 43–64 (1981)

[22] Resti, A.: Evaluating the cost-efficiency of the italian banking system: What can be learned from the joint application of parametric and nonparametric techniques. Journal of Banking & Finance 21(2), 221–250 (1997)

[23] Sealey, C.W., Lindley, J.T.: Inputs, outputs, and a theory of production and cost at depository financial institutions. The Journal of Finance 32(4), 1251–1266 (1977)

[24] Sheldon, G.: Economies, inefficiences and technical progress in swiss banking. In: The competitiveness of financial institutions and centres in Europe, pp. 115–132. Springer (1994)

[25] Sherman, H.D., Gold, F.: Bank branch operating efficiency: Evaluation with data envelopment analysis. Journal of banking & finance 9(2), 297–315 (1985)

[26] Shujie, Y., Chunxia, F.G.J.: The empirical analysis of efficiency of chinese banks. Economic Research Journal 8, 000 (2004)

[27] Weilan, L.C.S.: A study on the efficiency of china's commercial bank based on sfa [j]. Journal of Finance 6, 016 (2004)

[28] Xiaogang, C., Skully, M., Brown, K.: Banking efficiency in china: Application of dea to pre-and post-deregulation eras: 1993–2000. China Economic Review 16(3), 229–245 (2005)

[29] Xiaoqing Maggie, F., Heffernan, S.: Cost x-efficiency in china's banking sector. China Economic Review 18(1), 35–53 (2007)