

HONG KONG INSTITUTE FOR MONETARY RESEARCH

**RATIONAL COST INEFFICIENCY IN CHINESE
BANKS**

Kent Matthews, Zhiguo Xiao and Xu Zhang

HKIMR Working Paper No.29/2009

September 2009



Hong Kong Institute for Monetary Research
(a company incorporated with limited liability)

All rights reserved.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

Rational Cost Inefficiency in Chinese Banks

Kent Matthews

Cardiff University, Wales

Hong Kong Institute for Monetary Research

and

Zhiguo Xiao

Fudan University, Shanghai

and

Xu Zhang

Citigroup (China)

Cardiff University, Wales

September 2009

Abstract

According to a frequently cited finding by Berger et al (1993), X-inefficiency contributes 20% to cost-inefficiency in western banks. Empirical studies of Chinese banks tend to place cost-inefficiency in the region of 50%. Such estimates would suggest that Chinese banks suffer from gross cost inefficiency. Using a non-parametric bootstrapping method, this study decomposes cost-inefficiency in Chinese banks into X-inefficiency and allocative-inefficiency. It argues that allocative inefficiency is the optimal outcome of input resource allocation subject to enforced employment constraints. The resulting analysis suggests that allowing for rational allocative inefficiency; Chinese banks are no better or worse than their western counterparts.

Keywords: Bank Efficiency, China, X-inefficiency, DEA, Bootstrapping

JEL Classification: D23, G21, G28

Corresponding author: K Matthews
Cardiff Business School
Colum Drive,
Cardiff, CF10 3EU
Wales, UK
E-mail address: MatthewsK@cf.ac.uk
Tel.: +44(0)292-875855

The research for this work was carried out while the first author was a Visiting Fellow at the Hong Kong Institute for Monetary Research of the HKMA, April 2009. We are grateful to an anonymous referee, Patrick Minford, Kul Luintel and participants of the HKIMR research seminar for helpful comments.

The views expressed in this paper are those of the authors, and do not necessarily reflect those of the Hong Kong Institute for Monetary Research, its Council of Advisors, or the Board of Directors.

1. Introduction

Bank efficiency in China has become a popular subject of research in recent years. A number of studies of Chinese banking efficiency have been published in Chinese scholarly journals¹ but to date there have been only a handful of studies that are available to non-Chinese readers.² The consensus of finding among Chinese scholars is that the state-owned commercial banks tend on average to exhibit the lowest levels of efficiency and the joint stock commercial banks show a faster growth in performance and efficiency.

Cost inefficiency relative to 'best practice' is usually blamed on bad management and poor motivation. Following Leibenstein (1966) this efficiency gap is termed 'X-inefficiency'. In an oft cited study of bank efficiency Berger *et al.* (1993) argue that 20% of bank costs is due to X-inefficiency. Recent studies of bank efficiency in China have estimated cost inefficiency in the region of 50%.³ Such figures are in stark contrast to the expectations of conventional inefficiency derived from the Berger *et al.* (1993) study. It implies that either Chinese bank management is grossly inefficient or that the estimates of cost efficiency have failed to take into account policy objectives and/or policy constraints that enter the decision making process.

This research has three objectives. First it aims to decompose the measure of cost inefficiency in Chinese banks into technical inefficiency (sometimes viewed as X-inefficiency), and allocative inefficiency. This paper argues that while the underutilization of factors is consistent with the notion of X-inefficiency, but the wrong factor-mix is indicative of long-standing employment constraints imposed on the banking system in the pre-reform period. Insofar as allocative inefficiency can be explained as the result of official employment constraints, the implied cost inefficiency cannot be viewed as a management deficiency but a rational outcome of optimizing behaviour. The decomposition of cost inefficiency into X-inefficiency (technical inefficiency) and allocative inefficiency allows us to examine their evolution over the sample period.

Second, the measures of cost inefficiency and its decomposition are obtained using the familiar non-parametric method of Data-Envelopment-Analysis (DEA). The problem with the standard DEA approach

¹ For example Qing and Ou, (2001); Xu, Junmin, and Zhensheng, (2001); Wei and Wang, (2000); Xue and Yang, (1998) and Zhao (2000) have used non-parametric methods while Liu and Song (2004), Zhang, Gu and Di (2005), Sun (2005) and Qian (2003) have used parametric methods.

² Recent exceptions are studies using non-parametric methods by Chen *et al.* (2005), and Yao *et al.* (2008) and parametric methods by Fu and Heffernan (2009). Other recent studies published in English are, Lin and Zhang (2009), Berger *et al.* (2009), Fu and Heffernan (2008), Matthews *et al.* (2007)

³ Fu and Heffernan (2007), Shen *et al.* (2008), Matthews *et al.* (2007)

is that it does not lend itself to statistical inference.⁴ This paper aims to provide an inferential capability to the point-estimates of inefficiency through the use of bootstrapping methods.

Third, the bootstrap estimates of inefficiency are used to test various hypotheses regarding the levels, trends and convergence in X-inefficiency and allocative inefficiency. Over time, as the profit motive replaces other (social and economic) imperatives the levels of X-inefficiency and allocative inefficiency should decline. The opening up of the Chinese banking market and threat of entry of foreign banks into the Chinese market may have led to improved management, which should result in improved technical efficiency and lower cost-inefficiency as incumbent banks attempt to cut costs and consolidate their balance sheets.

This paper is organized on the following lines. The next section provides a brief motivation and discusses the model of rational allocative inefficiency. Section 3 outlines the method, reviews the literature and discusses the data. Section 4 discusses the results and section 5 concludes.

2. A Model of Rational Allocative Efficiency

At a first glance the Chinese banking system appears fragmented and diverse. In 2007 it consisted of 8,877 institutions, including 3 policy banks, 5 large state-owned commercial banks (SOCB), 12 joint-stock commercial banks (JSCB), 124 city commercial banks (CCB), 29 locally incorporated foreign bank subsidiaries and the rest made up of urban and rural credit cooperatives and other financial institutions.⁵ In contrast, its neighbour India has only 482 institutions but this includes 59 nationwide state-owned and private banks and 29 foreign banks. While the 28 public sector banks dominate the market in India with 72% of the share of assets, in China the 5 state-owned or state-controlled banks command 53 per cent of the market.⁶

Chinese banks can be characterized as historically having low return on assets and low net interest margins (despite having wide interest spreads), a high non-performing loan ratio, a high cost-income ratio and overstuffed. A number of good descriptions of the Chinese banking system exist⁷ and what follows is a brief statement of the elements relevant to the issue of efficiency.

⁴ See Simar and Wilson (2008)

⁵ CBRC Annual Report 2007

⁶ <http://www.cbrc.gov.cn> (figure relates to end 2007)

⁷ For example Garcia-Herreo *et al.* (2006)

Up until 1995, control of the banking system remained firmly under the government and its agencies.⁸ Under state control, the banks in China served the socialist plan of directing credits to specific projects dictated by political preference rather than commercial imperative. An important but relatively unknown feature of the pre-reform banking system was that the banks were compelled to employ all banking graduates of the universities set up by the People's Bank of China, party officials and retirees of the People's Liberation Army who had completed their tour of duty. The overhang of Party officials and former PLA officers employed in the banks during the pre-reform period contributes to the overall picture of overstaffing.

We develop a model of allocative inefficiency based on staffing targets provided by the central authorities. Assume that the bank produces a single earnings asset (A). In reality this will consist of a combination of commercial loans, mortgages, government bonds, short-term bills, etc. We assume that this earning asset is produced by the inputs deposits (D), labour (L) and fixed capital assets (\bar{K}):⁹

$$A = D^\alpha \left(L_1^\gamma L_2^{(1-\gamma)} \right)^\beta \bar{K}^{(1-\alpha-\beta)} \quad (1)$$

The price of inputs are, the cost of deposits (r), the cost of labour (w) and the cost of fixed assets (ρ). The bank can hire two types of labour $\{L_1$ and $L_2\}$. The first type (L_1) are bank workers who have a higher marginal product than the second type (L_2) who are bureaucrats. However the bank is constrained to pay the same wage to both types of workers. The objective of the bank manager is to minimise costs subject to an output target:

$$\text{Min } \rho\bar{K} + wL_1 + wL_2 + rD - \lambda \left(D^\alpha \left(L_1^\gamma L_2^{(1-\gamma)} \right)^\beta \bar{K}^{(1-\alpha-\beta)} - A \right) \quad (2)$$

The bank is constrained to employ some type 2 labour but clearly in an unconstrained world the bank would only employ type 1 workers.¹⁰

The solution for output in the unconstrained case is given by:

⁸ According to La Porta, *et al.* (2002), 99% of the 10 largest commercial banks were owned and under the control of the government in 1995.

⁹ This uses the assumption of the intermediation approach that recognises that the outputs are the interest earning assets while deposits and borrowed funds are included with capital labour as inputs. See Sealey and Lindley (1977).

¹⁰ From the FOC of (1) the ratio of the marginal products of the two types of labour is given by $\frac{\gamma(A/L_1)}{(1-\gamma)(A/L_2)} = 1$. Since this contradicts the assumption of type 1 labour having a higher marginal product than type 2 labour it follows that $\gamma = 1$ and $L_2 = 0$.

$$A = \left(\frac{\alpha}{\beta}\right)^{\alpha} w^{\alpha} r^{-\alpha} \left(\frac{1-\gamma}{\gamma}\right)^{(1-\gamma)\beta} \gamma^{-\alpha} L_1^{\alpha+\beta} \bar{K}^{(1-\alpha-\beta)} \quad (3)$$

In the constrained case, the bank has to employ a certain number of type 2 labour given by the central government so that $L_2 = \bar{L}_2$.

The objective function is now:

$$\text{Min } \rho \bar{K} + wL_1 + wL_2 + rD - \lambda_1 \left(D^{\alpha} \left(L_1^{\gamma} L_2^{(1-\gamma)} \right)^{\beta} \bar{K}^{(1-\alpha-\beta)} - A \right) - \lambda_2 (L_2 - \bar{L}_2) \quad (4)$$

The first order conditions are:

$$\begin{aligned} r - \lambda_1 \alpha \left(\frac{A}{D} \right) &= 0 \\ w - \lambda_1 \beta \gamma \left(\frac{A}{L_1} \right) &= 0 \\ w - \lambda_1 \beta (1-\gamma) \left(\frac{A}{L_2} \right) - \lambda_2 &= 0 \end{aligned}$$

The marginal wage premium for type 2 labour is given by λ_2 .

The output function is now:

$$A = \left(\frac{\alpha}{\beta}\right)^{\alpha} w^{\alpha} r^{-\alpha} \gamma^{-\alpha} L_1^{\alpha+\beta\gamma} \bar{L}_2^{(1-\gamma)\beta} \bar{K}^{(1-\alpha-\beta)} \quad (5)$$

Denoting the input of type 1 labour in the unrestricted case as L_{1U} and the same for the restricted case as L_{1R} , from (5) and (3) we have the relationship described by:

$$\left(\frac{1-\gamma}{\gamma}\right)^{(1-\gamma)\beta} L_{1U}^{\alpha+\beta} = L_{1R}^{\alpha+\beta\gamma} \bar{L}_2^{(1-\gamma)\beta} \quad (6)$$

The allocative inefficiency generated by the additional constraint in the restricted case is described by:

$$\frac{L_{1R} + \bar{L}_2}{L_{1U}} \quad (7)$$

Expression (7) must be strictly greater than unity for an allocative inefficiency to exist. Because the marginal productivity of the type 2 labour is less than type 1 labour, the type 1 labour displaced by the constraint of having to employ a fixed amount of type 2 labour is less than one-for-one if the target level of output (earnings assets) is to be maintained.

Figure 1 describes the situation for the case of the 2 variable inputs deposits and labour (physical capital is fixed by assumption). The isoquant is given by qq and the bank cost constraint by pp . The point 'e' describes the cost minimum factor composition as in the unrestricted case which uses type 1 labour only. Point R describes the constrained case which uses both types of labour. The cost inefficiency generated by the allocative inefficiency is described by the ratio OP/OR .

Substituting (6) into (7) and rearranging we have

$$CE = \frac{L_{1R} + \bar{L}_2}{L_{1U}} = \left(\frac{1-\gamma}{\gamma} \right)^{\frac{(1-\gamma)\beta}{\alpha+\beta}} \left[\left(\frac{L_{1R}}{\bar{L}_2} \right)^{\frac{(1-\gamma)\beta}{\alpha+\beta}} + \left(\frac{\bar{L}_2}{L_{1R}} \right)^{\frac{\alpha+\gamma\beta}{\alpha+\beta}} \right] \quad (8)$$

We can show that for any values of $\alpha, \beta, \gamma, L_{1R}, \bar{L}_2$, $CE > 1$.¹¹

Qualitative evidence of overstaffing is scant and was gleaned from confidential interviews with individual bank managers. However an examination of the ratio of fixed assets to employees and deposits per employee at the beginning and end of the sample period is indicative. Table 1 shows the mean values of the ratios for the four large state-owned banks and the others in the sample for 1997 and 2007.

¹¹ Define $\theta = \frac{\beta}{\alpha+\beta}$, $a = \frac{L_{1R}}{\bar{L}_2}$, then

$$CE = \left(\frac{1-\gamma}{\gamma} \right)^{(1-\gamma)\theta} \left[a^{(1-\gamma)\theta} + a^{(1-\gamma)\theta-1} \right] = \left(a \frac{1-\gamma}{\gamma} \right)^{(1-\gamma)\theta} \left[1 + \frac{1}{a} \right]$$

After analyzing the properties of the function $CE(\gamma, \theta, a)$ we have: (1) $CE > 1$; (2) If α, β, γ are fixed, then when $a = \frac{\gamma}{1-\gamma}$, CE achieves its minimum value $\frac{1}{\gamma}$; (3) If we allow all $\alpha, \beta, \gamma, L_{1R}, \bar{L}_2$ to vary, then under the following situations,

CE will approach its minimum value 1: (i) $\theta \rightarrow 0$, i.e., $\beta \rightarrow 0$; (ii) $a \rightarrow \infty$; (iii) $a^\theta \rightarrow 1$.

The average of fixed assets per employee has doubled for the big four and increased by one-third for the remainder banks over this period indicating some branch expansion but also reduction in staffing. However, because of the accounting difficulties of comparing fixed asset values over time it may be more appropriate to concentrate on the deposits per employee. Quite clearly, deposits per employee have tripled over the period reflecting the massive expansion in deposits in the decade to 2007 but also the relative reduction in staffing. These figures suggest that if there was an allocative inefficiency in the banks in 1997, there has been a strong attempt to reduce it by 2007.

A more indicative picture is gained from Figure 2 which examines the average \$ deposit per employee of a sample of large commercial banks in the UK and that of a number of far-Eastern economies. We can interpret the figures for the foreign banks as an external benchmark.

The figure suggests that by 2007, except for the big-4 banks in China, the other banks in general have converged on external benchmarks as indicated by foreign banks.

Having demonstrated that allocative inefficiency can be generated from rational decision making we now turn to the methodology of measuring inefficiency.

3. Methodology

In reality, banks are multi-output enterprises and one of the conventional ways of modelling the efficiency of banks is the non-parametric method of Data Envelopment Analysis (DEA), which is the extension by Charnes *et al.* (1978) (CCR)¹² of the single input-output model of Farrell (1957) to a multiple input-output generalisation. Technical efficiency (TE) is measured as the ratio of projected output (on the efficient frontier) to actual input used. There are a number of papers that describe the methodology of DEA as applied to banking,¹³ and therefore will not be elaborated here. A diagrammatic explanation illustrates the main concepts.

In Figure 1, Technical efficiency is measured by the ratio OR/OQ (Technical inefficiency is given by RQ/OQ). Cost inefficiency (CI) is measured by PQ/OQ which in turn can be decomposed into X-inefficiency (or Technical inefficiency (RQ/OQ)) and allocative inefficiency (PR/OQ).

DEA constructs a non-parametric frontier of the best practices amongst the decision-making units (DMUs). An efficiency score for each DMU is measured in relation to this frontier. DEA is relatively insensitive to

¹² Charnes *et al.* (1978) popularised the DEA method. According to Tavares (2002) who produces a bibliography of DEA (1978-2001), there are 3203 DEA authors whose studies cover a wide range of fields. Banxia.com also compiles DEA papers from 1978 to the present.

¹³ The most recent being Drake (2004)

model specification (input or output orientation) and functional form;¹⁴ however the results are sensitive to the choice of inputs and outputs. The weakness of the DEA approach is that it assumes data are free from measurement errors. Furthermore, since efficiency is measured in a relative way, its analysis is confined to the sample used. This means that an efficient DMU found in the analysis cannot be compared in a straightforward way with other DMUs outside of the sample.

One of the criticisms levelled at the standard DEA approach is that it produces estimates of efficiency but nothing can be said about the sensitivity (finite sample bias, confidence interval) of the estimator to sampling variation.¹⁵ In a practical sense what this means is that if a DMU has a score of 0.95, it is 5% less efficient than the benchmark but nothing can be said about statistical significance – meaning is the 5% inefficiency statistically significant in any meaningful way. Without the capability for statistical inference, non-parametric methods would be weak alternatives to parametric methods of estimating efficiency. However, uncertainties also exist in the estimation of efficiency using DEA. The most obvious uncertainty is what comes from measurement error. Measurement error in the context of data on Chinese banks is particularly marked. There are three potential sources of error: firstly, differences between local bank's accounting procedures and those of international bodies; secondly, differences between local bank's accounting conventions; and thirdly, researcher assumptions relating to the generation of missing observations. Other uncertainties arise from the estimation of the efficiency frontier; changes to the inputs and/or outputs can cause large differences in the resulting scores. Furthermore there may be errors in the sampling variation caused by the difficulty in obtaining a sufficiently large and consistent sampling frame.

The bootstrap procedure for non-parametric frontier models is set out in Simar and Wilson (1998, 2000a, 2000b). The efficiency scores calculated with the original data are used to construct pseudo data. The bootstrap procedure is based on the idea that there exists a Data Generating Process (DGP), which can be determined by Monte Carlo simulation. By using the estimated distribution of the DGP to generate a large number of random samples, a set of pseudo estimates of the efficiency scores $\hat{\theta}_i$ are obtained. However this 'naive' bootstrap yields biased and inconsistent estimates (Simar and Wilson, 2000a). We operate the heterogeneous bootstrap procedure that produces consistent values of $\hat{\theta}_i$ from a kernel density estimate as given in Simar and Wilson (2000b). Briefly stated, we have observations (x_i, y_i) $i = 1, \dots, n$ which is assumed to be *i.i.d.* random sample from a probability density function $f(x, y)$ on the production set Ψ . The idea of the bootstrap is to first estimate $f(x, y)$ using data (x_i, y_i) . Denoting the estimated density by $\hat{f}(x, y)$. The **next** step is to randomly draw B samples using $\hat{f}(x, y)$. The bootstrap **can be** conducted in one of two ways. The direct approach uses Cartesian coordinates to

¹⁴ Hababou (2002) and Avkiran (1999) provide a relatively thorough discussion of the merits and limits of the DEA.

¹⁵ Simar and Wilson (1998)

estimate $\hat{f}(x, y)$ directly. This approach is difficult to implement and in general has not been followed by researchers. The indirect approach uses polar coordinates and takes the following four steps.

- Step 1 Translate the Cartesian coordinate data $\{(x_i, y_i), i = 1, \dots, n\}$ into polar coordinate data $\{(\theta_i, \eta_i, y_i), i = 1, \dots, n\}$, where $\hat{\theta}_i$ denotes efficiency and η_i denotes angles.
- Step 2 Use $\{(\theta_i, \eta_i, y_i), i = 1, \dots, n\}$ to estimate their density $f(\theta, \eta, y)$ and denote the estimated density by $\hat{f}(\theta, \eta, y)$.
- Step 3 Randomly draw B sets of random samples from $\hat{f}(\theta, \eta, y)$. Denote the b^{th} bootstrap sample by $\{(\theta_{ib}^*, \eta_{ib}^*, y_{ib}^*), i = 1, \dots, n\}$.
- Step 4 Translate the b^{th} bootstrap sample $\{(\theta_{ib}^*, \eta_{ib}^*, y_{ib}^*), i = 1, \dots, n\}$ into Cartesian coordinate data $\{(x_{ib}^*, y_{ib}^*), i = 1, \dots, n\}$.

The difference between the homogeneous bootstrap and the heterogeneous bootstrap appears in step 1. The homogeneous bootstrap method assumes that the distribution of efficiency $\{\theta\}$ is unrelated to the distribution of (η, y) . That is $f(\theta|\eta, y) = f(\theta)$. The heterogeneous bootstrap does not maintain this assumption. Following the Simar-Wilson method, 1000 bootstrap values of the individual DMU for the efficiency scores are generated in each year.¹⁶ The appendix provides a description of the algorithm.

Most studies of banking efficiency have focussed on the developed economies.¹⁷ While there have been some studies of other Far Eastern economies,¹⁸ the number is small in comparison. Indeed, from Berger and Humphrey's (1997) survey of 130 studies of frontier analysis in 21 countries, only 8 were about developing and Asian countries (including 2 in Japan). Studies on US financial institutions were the most common, accounting for 66 out of 116 single country studies.

A number of efficiency studies of Chinese banks have emerged in recent years, using both DEA and stochastic frontier analysis.¹⁹ The consensus of finding from the DEA studies is threefold. First, because of the continued banking reform programme technical inefficiency has been declining over time. Second,

¹⁶ Recent bootstrapping applications to DEA have been conducted by Löthgren and Tambour (1999); in the case of banking efficiency by Casu and Molyneux (2003); and in the case of Chinese rural credit cooperatives, Dong and Featherstone (2006).

¹⁷ See for example Drake and Hall (2003), Cavallo and Rossi (2002), Elyasiani and Rezvanian (2002), Maudos *et al.* (2002), Drake (2001) Altunbas and Molyneux (1996) and Molyneux and Forbes (1993)

¹⁸ See Rezvanian and Mehdiian (2002), Hardy and di Patti (2001), Karim (2001), Laevan (1999), Katib and Matthews (1999), Chu and Lim (1998), Bhattacharyya *et al.* (1997) and Fukuyama (1995)

average bank efficiency is lower in the state owned banks (SOBs) than in the joint stock banks. Third, the gap between the two has been narrowing in recent years.

Studies of bank efficiency have used the terms technical efficiency and X-efficiency interchangeably as if they were the same thing. While similar in concept they are not necessarily the same. The concept of technical efficiency derives its basis in the neo-classical theory of the firm and assumes profit maximising behaviour. A firm or a bank may be technically inefficient for technical reasons such as low training or low human capital levels of managers and workers, or the use of inferior or out-of-date technology. The diffusion of new technology is not instantaneous and some firms or banks may lag behind others in the acquisition and utilisation of new technology. With further training and updating of capital, the firm or bank can expect to move towards the efficient frontier described by the isoquant in Figure 1. X-inefficiency is not caused by the variability of skills or the time variability of technology diffusion but by the use and organisation of such skills and technology.

Berger, Hunter and Timme (1993) argue that X-inefficiency constitutes 20% or more of bank costs. Poor motivation and weak pressure resulting in under utilization of factors of production are parts of what Leibenstein (1975) describes as 'organisational entropy'. X-inefficiency arises as a result of low pressure for performance. Some institutions would be protected by government regulation that would reduce the external pressure of competition. But even with a higher degree of pressure from the environment, firms may have organisational deficiencies so that management signals and incentives are lost in the hierarchy of the organisation.

This study employs annual data (1997-2007) for 14 banks: the five state-owned banks (SOB), and nine joint-stock commercial banks (JSB). The total sample consisted of 154 bank year observations. The main source of the data was Fitch/Bankscope, and individual annual reports of banks. The choice of banks was based on the fact that they face a common market and compete nationwide.

Two approaches are normally taken in determining what constitutes bank input and output. Under the intermediation approach (Sealey and Lindley, 1977), bank assets measure outputs and liabilities measure inputs. In contrast, inputs in the production approach are physical entities such as labour and physical capital and revenue flows represent outputs. In this study, we adopt a hybrid of the two approaches. We use three inputs and three outputs for the estimation of technical efficiency. Inputs are the number of employees (*LAB*), fixed assets (*FA*) and total deposits (*DEP*). Outputs are total loans (*LOANS*), other earning assets (*OEA*), and non-interest income (*NII*). Although the latter variable remains undeveloped in China, it is selected to reflect the growing contribution of non-interest income to banks' total income.

¹⁹ In addition to the papers cited in footnote 1, other studies by Chinese scholars that have used non-parametric techniques include Fang *et al.* (2004) and studies using parametric methods include Liu and Liu (2004), Sun (2005), Qian (2003), Chi, Sun and Lu (2005), Yao, Feng and Jiang (2004)

The inputs for the construction of cost-efficiency additionally require the factor prices of the relevant inputs above. We distinguish between the price of labour (PL), price of fixed capital (PK) and the price of funds (PF). The price of labour is obtained as the ratio of personnel expenses to the number of employees. The price of fixed capital is obtained as operating expenses less personnel expenses divided by fixed assets (less depreciation). The price of funds is obtained from the ratio of interest paid to total funds.

Table 2 presents the summary statistics of the input and output data for 1997 and 2007 as a snapshot indicator of the scale of the variables used. The high standard deviation is an indication of the dominance of the 5 state owned banks. The table shows how fast earnings assets have grown over this period. The total stock of loans has grown at an average of 12 per cent a year. Other earning assets have grown at an average rate of 20% a year, in part reflecting the activities of the asset management companies that swapped tranches of the NPLs of the big 4 SOBs for bonds in 1999 and 2001. The most remarkable growth is in non-interest earnings which have grown at an average rate of 28% a year, reflecting an increasing source of profit for banks that have traditionally depended on the banking book for the generation of income.

Other points to note are that net employment has grown by an average of 0.6% a year but average labour cost has grown by a remarkable 14% a year, reflecting the increasing skill premium paid to workers in this sector. A further point to note is reduced relative dispersion of the variables (coefficient of variation) which also indicates an increased convergence of the nationwide banks on each other.

Having outlined the methodology and the data we now examine the empirical results from the bootstrap method.

4. Empirical Results

Tables 3a-b illustrate the results of the bootstrap method for two representative years; 1997 and 2007 in the case of technical efficiency (X-efficiency, XE) and cost efficiency (CE). The tables show the biased estimates, the bootstrapped bias-adjusted estimates and the 95% confidence intervals for 2000 bootstraps.

Simar and Wilson (2000a, 2000b) show that the bias correction will introduce extra noise that may result in a mean-square error (MSE) greater than the MSE of the bias-unadjusted bootstrap values. In the limit the bias corrected MSE will be four times that of the uncorrected estimate and Simar and Wilson caution

against the bias correction unless the ratio $\left(\frac{\frac{1}{3}bia\hat{s}^2}{\hat{\sigma}^2}\right)$ is greater than unity, where $bia\hat{s}$ is the bias

correction and $\hat{\sigma}^2$ is the sample variance of the uncorrected bootstrap values. The statistic ρ_i is defined

as $\rho_i = \left(\frac{\frac{1}{3} \text{bias}^2}{\hat{\sigma}^2} \right)$ and is shown in the final column. A statistic of $\rho_i > 1$ implies that the bias correction is valid.

The Tables report the median of the 2000 bootstrap values of cost efficiency (CE) and X-efficiency (Technical efficiency) (XE) as the standard for previous studies.²⁰ The confidence intervals of the bootstrap values in Table 3a-b support the significance of the bias. The value of ρ indicates the validity of the bias correction which in all cases was always one-sided.²¹

We now turn to the third objective of this paper and that is to evaluate the levels, trends and convergence in the two types of inefficiency. Table 4 shows the full period sample means and weighted means of the cost inefficiency (CI) and X-inefficiency (XI) estimates obtained as $CI = (1 - CE)$ and $XI = (1 - XE)$ respectively. The allocative inefficiency estimate was obtained as the residual of the cost inefficiency CI and X-inefficiency XI ($AI = CI - XI$).²²

Weighting the inefficiency scores provide a more accurate picture of the average levels of inefficiency but in reality the difference between the pure average and the weighted average is of second-order magnitude. It can be seen from Table 4 that the weighted average cost inefficiency over the period for the SOCBs is slightly smaller than the average for the JSCBs. The table also indicates that the weighted average allocative inefficiency is comparable for both groups of banks.

Using the mean of the un-weighted scores, we test if there is a significant difference between the average inefficiency of banks that have a foreign stakeholding from those that do not. Since the distribution of the inefficiency scores may not be standard normal we apply a non-parametric test (Mann-Whitney) with the results shown in Table 5.

The results of the non-parametric tests indicate a significant difference. Banks that have a foreign stakeholding have lower average X-inefficiency and allocative inefficiency. But is inefficiency in Chinese banks as a whole declining and if so can anything be said about the speed at which the inefficient banks are converging on the benchmark efficient banks? Figure 3 and 4 show the yearly weighted average technical inefficiency and allocative inefficiency for the SOCBs and JSCBs as a group. The short dotted line is the weighted average measure of inefficiency for the SOCBs and the long dotted line is the same for the JSCBs. Figure 3 shows a rise in technical inefficiency in the years 1999-2001 which roughly

²⁰ See for example Dong and Featherstone (2006). The argument for reporting the median rather than the mean is that the distribution of the efficiency scores may not be standard normal. In reality there was little difference between the two.

²¹ Out of 154 bootstrap results for each bank-year in only seven cases was the bias correction invalid. In such cases the bootstrap value was used for consistency.

²² Strictly $AE = CE/XE$ but as we are dealing with the median values of an unknown distribution it was convenient to define $AI = CI - XI = XE - CE$. The alternative measure is $AI' = (1 - AE) = (XE - CE)/XE$.

coincides with the activities of the Asset Management Companies that transferred tranches of NPLS of the big 4 SOCBs and returned them as face value bonds in 1999-2000.²³

It is possible that the portfolio switch of the NPLs from the loan book to other earning assets of the big-4 balance sheets distorted the average efficiency measure for the SOCBs. However, this explanation is less convincing for the JSCBs which show a similar pattern for the allocative inefficiency in Figure 4. Also the asset management companies conducted a further transfer of NPLs from the big-4 banks of around \$120 billion in 2003 which does not appear to have affected the average inefficiency figures as seen in the figures. But what is clear is that there is a discernible negative trend in both types of inefficiency for both groups of banks.

Using the concept of *beta-convergence* from the growth convergence literature (Barro, 1991), we can obtain a measure of the speed of convergence to a common level of inefficiency by regressing the change in the level of inefficiency on the lag of inefficiency and environmental and bank specific variables to allow for convergence to different levels of inefficiency.²⁴ However it is shown by Simar and Wilson (2007) that the estimated inefficiencies may be serially correlated. They propose a double bootstrap procedure to adjust for the bias caused by the inherent correlation among the estimated inefficiencies. The problem of potential bias is further compounded by the existence of the lagged inefficiency score. Developing a valid bootstrap procedure for estimating beta-convergence is computationally intractable. However in an attempt to deal with the potential serial correlation we present estimates of the rate of decline of inefficiency controlled for individual bank factors using a panel GLS *Heteroskedastic-Autocorrelation Consistent* (HAC) estimator (Table 6). The dependant variable was the yearly change in the specific type of inefficiency. Bank specific variables were lagged one period to avoid potential endogeneity.

The key covariate is the lag in the level of inefficiency. A negative coefficient defines the common speed of convergence of inefficiency. The bank specific variable that proved significant was the logarithm of total assets as a measure of size. Environmental variables were a zero-one dummy variable identifying SOCBs (SOB) and the proportion of the bank owned by foreign financial enterprises (FOR). An interaction term between the lag in inefficiency and the SOB dummy defined differing speed of adjustment between SOCBs and JSCBs. The third but last row is the log likelihood. The second but last row is an F test for autocorrelation in panel data (Wooldridge, 2002) and the final row is a likelihood ratio test for heteroskedasticity in panel data.

The last two rows of Table 6 indicate that the use of the HAC estimator was appropriate. Autocorrelation in the panel could not be rejected at the 1% level for the regressions of both types of inefficiency, and

²³ During 1999-2000 4 AMCs bought up an aggregate of roughly \$205 billion in NPLs in return for 10-year bonds paying a fixed rate of 2.25%

heteroskedasticity could not be rejected at the 1% for the X-inefficiency and 10% for allocative inefficiency regressions.

The results of Table 6 indicate that controlling for ownership, larger banks are associated with lower levels of both types of inefficiency. Banks that have a foreign stake are associated with lower levels of both types of inefficiency, confirming the finding reported in Table 4. State-owned banks are associated with higher levels of both types of inefficiency and in particular higher levels of allocative inefficiency. Importantly, the negative coefficient on the lagged measure of inefficiency shows significant decline in both measures of inefficiency. The interaction term of lagged inefficiency with SOB suggests that the state-owned banks reduce inefficiency at a lower speed than the JSCBs,²⁵ but that they are reducing both types at roughly the same speed. The results suggest that the joint-stock commercial banks are reducing allocative inefficiency at a faster rate than X-inefficiency.

5. Conclusion

This paper has used non-parametric methods to conduct an analysis of inefficiency in a sample of Chinese banks. The estimates of bank inefficiency were obtained using a bootstrapping method to enable statistical inference. We have partitioned cost inefficiency into X-inefficiency and allocative inefficiency. Our findings suggest that Chinese banks have been improving performance by reducing both types of inefficiency. However, the state-owned banking sector has higher levels of both types of inefficiency and is also reducing both types of inefficiency at a slower rate than joint-stock commercial banks. This suggests that state-owned banks are more constrained by social and political objectives in their downsizing strategy than JSCBs.

We confirm the findings from nonparametric and stochastic frontier based studies of Chinese banks that average cost inefficiency are in the region of 50%. Inefficiency in Chinese banking is made up of both X-inefficiency and allocative inefficiency. We have argued in this paper that given the social and political constraints that Chinese banks had to operate in, allocative inefficiency was symptomatic of rational decision making dictated by social employment objectives.

However, we must still interpret the results with caution. Not all of allocative inefficiency can be attributed to over-staffing and not all of over-staffing can be explained by past employment objectives. Overstaffing caused by political and social constraints is observationally equivalent to rent-seeking behaviour by bank managers (Matthews *et al.*, 2007). It is also possible that poor management decisions that may have contributed to X-inefficiency could also have contributed to allocative inefficiency.

²⁴ See also Fung (2006)

Yet, the argument of this paper is that there have been significant improvements in bank efficiency. The 2007 weighted average of all banks X-inefficiency and allocative inefficiency is 16% and 20% respectively. If the Berger *et al.* (1993) finding that 20% of all bank costs are due to X-inefficiency represents a common benchmark for banking markets in general then the message of this paper is that Chinese banks are not out of line.

²⁵ Because of the problem of multi-collinearity between SOB and the interaction term in the X-inefficiency regression, a grid search using constrained estimates that maximised the log likelihood is reported in Model (3) of Table 5.

References

- Altunbas, Y. and P. Molyneux (1996), "Cost Economies in EU Banking System," *Journal of Economics and Business*, 48(3): 217-30.
- Avkiran, N. K. (1999), "An Application Reference for Data Envelopment Analysis in Branch Banking: Helping the Novice Researcher," *International Journal of Bank Marketing*, 17(5): 206-20.
- Barro, R. J. (1991), "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106: 407-43.
- 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 and Finance*, in print.
- Berger, A. N. and D. B. Humphrey (1997), "Efficiency of Financial Institutions: International Survey and Directions for Future Research," *European Journal of Operational Research*, 98: 175-212.
- Berger, A. N., W. C. Hunter and S. G. Timme (1993), "The Efficiency of Financial Institutions: A Review and Preview of Research Past, Present and Future," *Journal of Banking and Finance*, 17: 221-49.
- Bhattacharyya, A., C. A. K. Lovell and P. Sahay (1997), "The Impact of Liberalization on the Productive Efficiency of Indian Commercial Banks," *European Journal of Operational Research*, 98: 332-45.
- Casu, B. and Molyneux (2003), "A Comparative Study of Efficiency in European Banking," *Applied Economics*, 35(17): 1865-76.
- Cavallo, L. and S. P. S. Rossi (2002), "Do Environmental Variables Affect the Performance and Technical Efficiency of the European Banking Systems? A Parametric Analysis Using the Stochastic Frontier Approach," *The European Journal of Finance*, 8: 123-6.
- Charnes, A., W. W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, 2: 429-44.
- Chen, X., M. Skully and K. Brown (2005), "Banking Efficiency in China: Applications of DEA to Pre-and Post-Deregulation Eras: 1993-2000," *China Economic Review*, 16: 2229-45.
- Chi, G., X. Sun and D. Lu (2005), "Empirical Study of the Cost Efficiency of China Commercial Banks," *Economic Research Journal*, 6: 104-14 (Chinese language).

- Chu, S. F. and G. H. Lim (1998), "Share Performance and Profit Efficiency of Banks in an Oligopolistic Market: Evidence from Singapore," *Journal of Multinational Financial Management*, 8: 155-68.
- Drake, L. (2001), "Efficiency and Productivity Change in UK Banking," *Applied Financial Economics*, 11: 557-71.
- Drake, L. (2004), "Costs and Efficiency in Banking: A Survey of the Evidence from the US, the UK and Japan," in A. W. Mullineux and V. Murinde, eds., *Handbook of International Banking*, Cheltenham: Edward Elgar.
- Drake, L. and M. J. B. Hall (2003), "Efficiency in Japanese Banking: An Empirical Analysis," *Journal of Banking and Finance*, 27: 891-917.
- Dong, F. and A. Featherstone (2006), "Technical and Scale Efficiencies for Chinese Rural Credit Cooperatives: A Bootstrapping Approach in Data Envelope Analysis," *Journal of Chinese Economic and Business Studies*, 4(1): 57-75.
- Elyasiani, E. and R. Rezvanian (2002), "A Comparative Multiproduct Cost Study of Foreign-Owned Bank and Domestic-Owned US Banks," *Applied Financial Economics*, 12: 271-84.
- Fang, C., W. Sun, Z. Wang and H. Wang (2004), "State-Owned Bank's Efficiency Measurement and their Behaviour Character Study," *Journal of Qualitative and Technical Economics*, 4: 42-5 (Chinese language).
- Fu, X. and S. Heffernan (2007), "Cost X-efficiency in China's Banking Sector," *China Economic Review*, 18: 35-53.
- Fu, X. and S. Heffernan (2008), "The Effects of Reform on China's Bank Structure and Performance," *Journal of Banking and Finance*, in print.
- Fukuyama, H. (1995), "Measuring Efficiency and Productivity Growth in Japanese Banking: A Non-Parametric Frontier Approach," *Applied Financial Economics*, 5: 95-117.
- Fung, M. K. (2006), "Scale Economies, X-Efficiency, and Convergence of Productivity among Bank Holding Companies," *Journal of Banking and Finance*, 30: 2857-74.
- Garcia-Herrero, A., G. Segio and Santabarbara (2006), "China's Banking Reform: An Assessment of its Evolution and Possible Impact," *CESifo Economic Studies*, 52(2): 304-63.

- Hababou, M. (2002), "Tutorial in DEA," <http://members.tripod.com/moezh/DEAtutorial/DEAtutorial.html>.
- Hardy, D. C. and B. E. di Patti (2001), "Bank Reform and Bank Efficiency in Pakistan," IMF Working Paper No.01/138, Washington D.C.: International Monetary Fund.
- Jahanshahloo, G. R., M. Soleimani-damaneh and A. Mostafaei (2008), "A Simplified Version of the DEA Cost Efficiency Model," *European Journal of Operational Research*, 184: 814-5.
- Karim, M. Z. A. (2001), "Comparative Bank Efficiency across Select ASEAN Countries," *ASEAN Economic Bulletin*, 18(3): 289-304.
- Katib, M. N. and K. Matthews (1999), "A Non-Parametric Approach to Efficiency Measurement in the Malaysian Banking System," *The Singapore Economic Review*, 44(2): 89-114.
- Laeven, L. (1999), "Risk and Efficiency in East Asian Banks," World Bank, Policy Research Working Papers No.2255.
- La Porta, R., F. Lopez-de-Silanes and A. Shleifer (2002), "Government Ownership of Banks," *Journal of Finance*, 57: 265-302.
- Leibenstein, H. (1966), "Allocative Efficiency vs. X-Efficiency," *American Economic Review*, 56(3): 392-415.
- Leibenstein, H. (1975), "Aspects of the X-Efficiency Theory of the Firm," *The Bell Journal of Economics*, 6(2): 580-606.
- Lin, X. and Y. Zhang (2008), "Bank Ownership Reform and Bank Performance in China," *Journal of Banking and Finance*, in print.
- Liu, C. and W. Song (2004), "Efficiency Analysis in China Commercial Banks Based on SFA," *Journal of Financial Research*, 6: 138-42 (Chinese language).
- Liu, Z. and C. Liu (2004), "Efficiency Analysis in China Commercial Banks Based on DFA Approach," *Journal of Qualitative and Technical Economics*, 4: 42-5 (Chinese language).
- Löthgren, M. and M. Tambour (1999), "Testing Scale Efficiency in DEA Models: A Bootstrapping Approach," *Applied Economics*, 31: 1231-7.

- Matthews, K., J. Guo and X. Zhang (2007), "Rational Inefficiency and Non-Performing Loans in Chinese Banking," *China Finance Review*, 3(1): 55-75.
- Matthews, K., X. Zhang and J. Guo (2009), "Non-Performing Loans and Productivity in Chinese Banks 1997 – 2006," *The Chinese Economy*, 42(2): 30-47.
- Maudos, J., J. M. Pastor, F. Perez and J. Quesada (2002), "Cost and Profit Efficiency in European Banks," *Journal of International Financial Markets, Institutions and Money*, 12: 33-58.
- Molyneux, P. and W. Forbes (1993), "Market Structure and Performance in European Banking," Institute of European Finance, RP 93/5.
- Qian, Q. (2003), "On the Efficiency Analysis of SFA in China Commercial Banks," *Social Science of Nanjing*, 1: 41-6 (Chinese language).
- Qing, W. and Y. Ou (2001), "Chinese Commercial Banks: Market Structure, Efficiency and its Performance," *Economic Science*, 100871: 34-45 (Chinese language).
- Rezvanian, R. and S. Mehdian (2002), "An Examination of Cost Structure and Production Performance of Conventional Banks in Singapore," *Journal of Banking and Finance*, 26: 79-98.
- Sealey, C. W. and J. T. Lindley (1977), "Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions," *Journal of Finance*, 32(4): 1251-66.
- Shen, Z., H. Liao and T. Weyman-Jones (2008), "Cost Efficiency Analysis in Banking Industries of Ten Asian Countries and Regions," Department of Economics, Loughborough University Working Paper.
- Simar, L. and P. W. Wilson (1998), "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Non-Parametric Frontier Models," *Management Science*, 44: 49-61.
- Simar, L. and P. W. Wilson (2000a), "Statistical Inference in Non-Parametric Frontier Models: The State of the Art," *Journal of Productivity Analysis*, 13: 49-78.
- Simar, L. and P. W. Wilson (2000b), "A General Methodology for Bootstrapping in Non-Parametric Frontier Models," *Journal of Applied Statistics*, 27: 779-802.
- Simar, L. and P. W. Wilson (2007), "Estimation and Inference in Two-Stage Semi-Parametric Models of Production Processes," *Journal of Econometrics*, 136: 31-64.

- Sun, Z. (2005), "Frontier Efficiency Analysis for the State-Owned Banks in China," *Industrial Economics Research*, 3: 41-7 (Chinese language).
- Tavares, G. (2002), *A Bibliography of Data Envelopment Analysis (1978-2001)*, RUTCOR Research Report, Rutgers University, USA.
- Wooldridge, J. (2002), *Econometric Analysis of Cross-Section and Panel Data*, Cambridge MA: MIT Press.
- Wei, Y. and L. Wang (2000), "The Non-Parametric Approach to the Measurement of Efficiency: The Case of China Commercial Banks," *Journal of Financial Research*, 3: 88-96 (Chinese language).
- Xu, Z., Z. Junmin and J. Zhenseng (2001), "An Analysis of the Efficiency of State-Owned Banks with Examples," *South China Financial Research*, 16(1): 25-7 (Chinese language).
- Xue, F. and D. Yang (1998), "Evaluating Bank Management and Efficiency Using the DEA Model," *Econometrics and Technological Economy Research*, 6: 63-6 (Chinese language).
- Yao, S., H. Zhongwei and W. Genfu (2008), "Ownership Reform, Foreign Competition and Efficiency of Chinese Commercial Banks: A Non-Parametric Approach," *The World Economy*: 1310-26.
- Yao, S., G. Feng and C. Jiang (2004), "The Empirical Analysis of Efficiency of Chinese Banks," *Economic Research Journal*, 8: 4-15 (Chinese language).
- Zhao, X. (2000), "State-Owned Commercial Banks Efficiency Analysis," *Journal of Economic Science*, 6: 45-50 (Chinese language).
- Zhang, C., F. Gu and Q. Di (2005), "Cost Efficiency Measurement in China Commercial Banks Based on Stochastic Frontier Analysis," *Explorations in Economic Issues*, 6: 116-9 (Chinese language).

Table 1. Average Fixed Assets per Employee and Deposits per Employee

Year	Bank Group	Fixed assets per employee mill rmb	Deposits per employee mill rmb
1997	Big-4	0.1	6.1
	Other	0.3	14.2
2007	Big-4	0.2	18.4
	Other	0.4	45.3

Table 2. Output-Input Variables 1997 - 2007 (Million RMB)

Variable	Description	Mean 1997	SD 1997	Mean 2007	SD 2007
LOANS RMB mill	Total stock of loans	430033	657201	1296424	1351614
OEA RMB mill	Investments	205103	301626	1227591	1419119
NII RMB mill	Net Fees and Commissions	862	1922	9994	12851
LAB	Total Employed	105138	175233	111960	157645
DEP RMB mill	Total stock of Deposits	604013	891353	2309760	2568177
FA RMB mill	Fixed assets	12831	19398	27374	33704
PL	Unit price of labour	0.0631	0.0380	0.2353	0.0915
PF	Unit price of funds	0.0502	0.0202	0.0214	0.0057
PK	Unit price of fixed assets	0.6528	0.5282	0.7409	0.2333

Sources: Fitch/Bankscope and author calculations from web sources.

Table 3a. Median Bootstrap Estimates of Cost Efficiency 1997 & 2007 CRS

Bank	Efficiency Score	Bias corrected	Lower bound	Upper bound	ρ
1997					
ABOC	0.7941*	0.4856	0.4212	0.5683	22.9
BOC	0.9502*	0.6942 [#]	0.1441	0.8789	0.64
BCOMM	0.9154*	0.3099	-0.1241	0.6710	2.79
CCB	0.8938*	0.7606	0.7069	0.7863	11.3
ICBC	1.000*	0.4158	0.0362	0.7490	2.86
CITIC	1.000*	0.4425	0.0381	0.8305	2.17
CMB	1.000*	0.5093	0.0559	0.7882	2.10
CMBCL	1.000*	0.6721	0.2203	0.8547	1.37
EVERBRT	1.000*	0.5537	0.2113	0.6604	1.75
GDB	0.8325*	0.5242	0.0507	0.8457	1.73
HUAXIA	1.000*	0.3325	0.0191	0.7127	3.99
IBCL	1.000*	0.8518	0.7406	0.9442	2.64
SDB	0.8409*	0.6142	0.5430	0.6782	9.08
SPB	1.000*	0.2533	0.0162	0.6641	5.88
2007					
ABOC	0.6848*	0.1851	0.1095	0.3288	25.8
BOC	0.8578*	0.6725	0.5419	0.7138	5.25
BCOMM	0.7737*	0.4566	0.3745	0.5412	14.6
CCB	0.8250*	0.5929	0.4190	0.6472	4.81
ICBC	0.8200*	0.7056	0.3610	0.8566	1.14
CITIC	0.9451*	0.7630	0.4132	0.9078	1.06
CMB	0.9962*	0.6674	0.3918	0.8560	2.55
CMBCL	1.000*	0.6691	0.3750	0.8294	1.91
EVERBRT	0.9536*	0.6771	0.4342	0.7768	1.81
GDB	0.8922*	0.6452	0.3744	0.8284	2.00
HUAXIA	0.9293*	0.6268	0.3374	0.8359	2.93
IBCL	1.000*	0.5846	0.5024	0.6370	12.5
SDB	1.000*	0.6426	0.2745	0.8274	2.07
SPB	0.9712*	0.7057 [#]	0.2856	0.8835	0.98

* significant bias at the 95% level of confidence. # Bias correction invalid

Table 3b. Median Bootstrap Estimates of Technical efficiency 1997 & 2007 CRS

Bank	Efficiency Score	Bias corrected	Lower bound	Upper bound	ρ
1997					
ABOC	0.8914*	0.9809	0.7671	0.8803	1.07
BOC	1.0000*	0.8304	0.6119	0.7796	9.12
BCOMM	1.0000*	0.6889	0.5804	0.7624	9.95
CCB	0.9699*	0.6430	0.8521	0.9572	1.13
ICBC	1.0000*	0.9109	0.5887	0.7397	13.2
CITIC	1.0000*	0.6500	0.6122	0.8156	6.50
CMB	1.0000*	0.6882	0.6056	0.7821	8.21
CMBCL	1.0000*	0.6801	0.5844	0.7292	14.8
EVERBRT	0.8932*	0.6423	0.6028	0.8058	2.69
GDB	1.0000*	0.7007	0.6056	0.8060	6.79
HUAXIA	1.0000*	0.6822	0.5516	0.7370	10.0
IBCL	1.0000*	0.6152	0.7097	0.9079	2.64
SDB	0.9809*	0.8165 [#]	0.7755	0.9530	0.89
SPB	1.0000*	0.8925	0.5692	0.6922	19.9
ABOC	0.9088*	0.8274	0.7543	0.8715	1.92
BOC	1.0000*	0.8744	0.7879	0.9283	2.76
BCOMM	0.9033*	0.8080	0.7200	0.8539	1.80
CCB	1.0000*	0.8749	0.7719	0.9515	1.58
ICBC	1.0000*	0.7431	0.6421	0.8580	3.81
CITIC	1.0000*	0.7836	0.6703	0.8881	2.89
CMB	1.0000*	0.7521	0.6695	0.8547	5.39
CMBCL	1.0000*	0.7555	0.6671	0.8677	4.33
EVERBRT	0.9415*	0.7863	0.6912	0.8659	2.63
GDB	0.9796*	0.7443	0.6425	0.8839	2.95
HUAXIA	1.0000*	0.7017	0.6277	0.8093	7.66
IBCL	0.9726*	0.9000 [#]	0.7989	0.9423	0.97
SDB	1.0000*	0.7518	0.6669	0.8363	5.77
SPB	1.0000*	0.7469	0.6410	0.8673	3.33

* significant bias at the 95% level of confidence. # Bias correction invalid

Table 4. Mean Inefficiency 1997 – 2007

Statistic	Cost-inefficiency	X-inefficiency	Allocative - Inefficiency
Mean all	55.0	30.4	24.6
Weighted mean	50.5	25.9	24.6
Mean SOB	52.9	25.9	27.0
W - mean SOB	50.1	25.1	25.1
Mean JSCB	55.0	30.4	24.6
W – mean JSCB	55.2	30.8	24.4

Table 5. Non-Parametric Test for Difference in Means

Statistic	X-inefficiency	Allocative inefficiency
Mean (Foreign stake)	24.5	13.9
Mean (all other)	30.1	29.1
Z – statistic	4.15***	5.55***

Table 6. Panel GLS (Panel Heteroskedastic-Autocorrelation Consistent Estimates). Dependant Variable is the Year-on-Year Change in Inefficiency. P Values in Parenthesis

Variable	Change in X-inefficiency = ΔXI_t (observations = 140)			Change in Allocative inefficiency = ΔAI_t (observations = 140)		
	1	2	3	4	5	6
Model						
Intercept	0.353*** (0.000)	0.354*** (0.000)	0.354*** (0.000)	0.923*** (0.000)	0.567*** (0.000)	1.01*** (0.000)
XI_{t-1}	-0.444*** (0.000)	-0.485*** (0.000)	-0.447*** (0.000)		-	-
AI_{t-1}	-	-	-	-0.617*** (0.000)	-0.735*** (0.000)	-0.704*** (0.000)
SOB	0.021** (0.035)	-	0.020** (0.036)	0.173*** (0.000)	-	0.137*** (0.000)
SOB* XI_{t-1}	-	0.055** (0.049)	0.004** (0.036)	-	-	-
SOB* AI_{t-1}	-	-	-	-	0.375*** (0.000)	0.202** (0.002)
$\ln(TA)_{t-1}$	-0.018*** (0.000)	-0.017*** (0.000)	-0.018*** (0.000)	-0.062*** (0.000)	-0.031*** (0.000)	-0.068*** (0.000)
FOR	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.003*** (0.000)	-0.004*** (0.000)	-0.003*** (0.000)
L^*	1021.7	1037.2	1041.5	891.6	898.0	898.1
$F_{1,13}$	42.6***	48.4***	48.4***	27.7***	39.3***	39.3***
χ^2_{13}	100.2***	100.2***	100.2***	22.0*	21.7*	21.6*

*** significant at the 1%, ** significant at the 5%, * significant at the 10%.

Figure 1. Rational Allocative Inefficiency

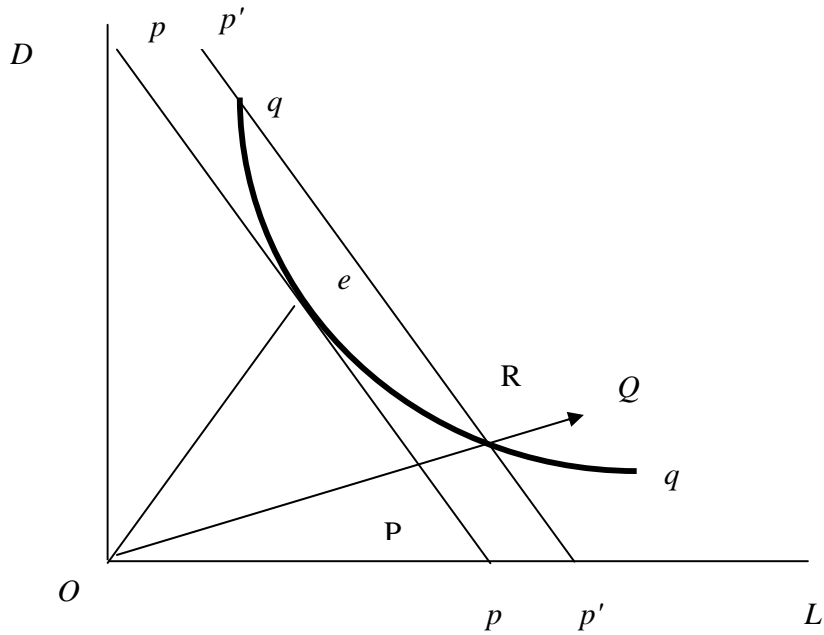


Figure 2. China and Other Economies Average Deposits per Employee

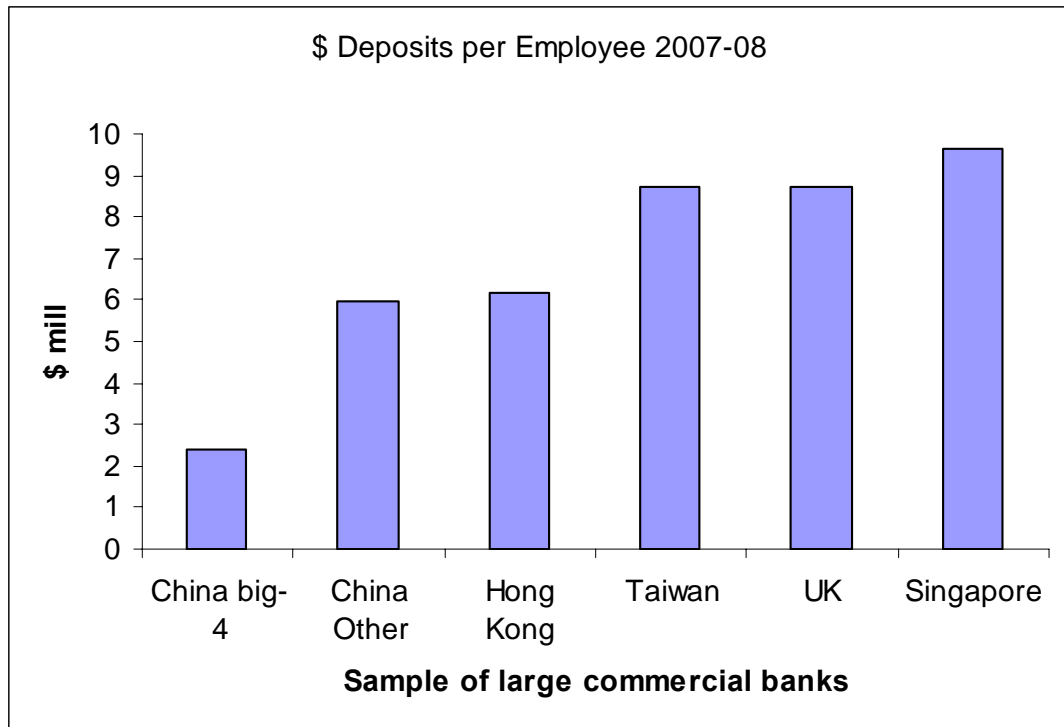


Figure 3. Weighted Average Tech Inefficiency

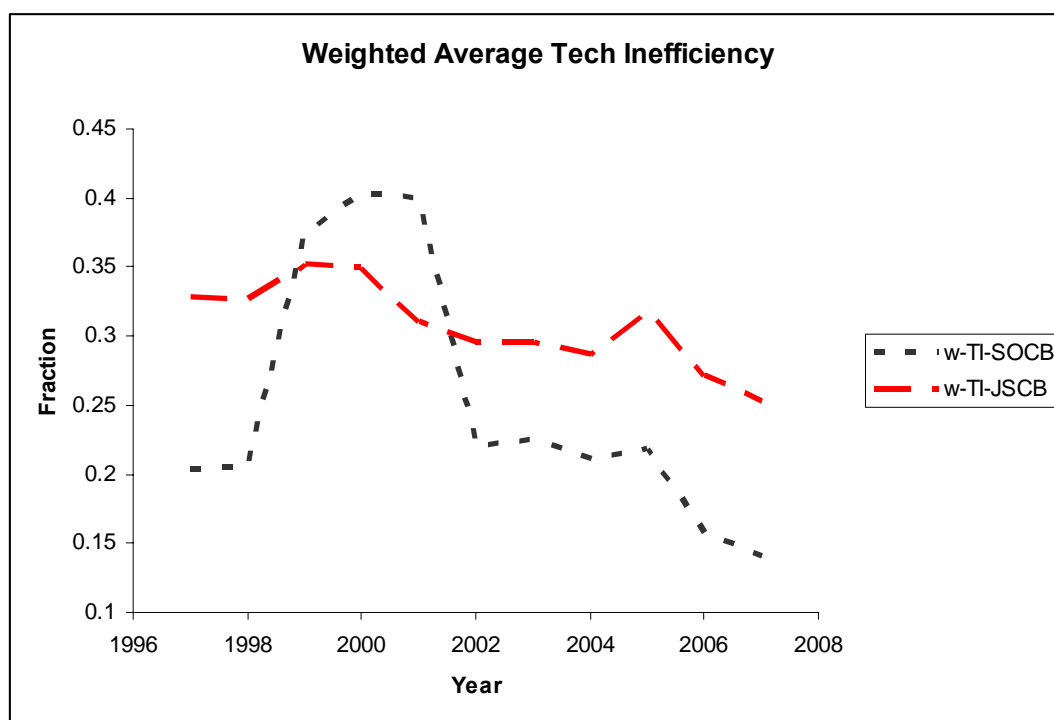
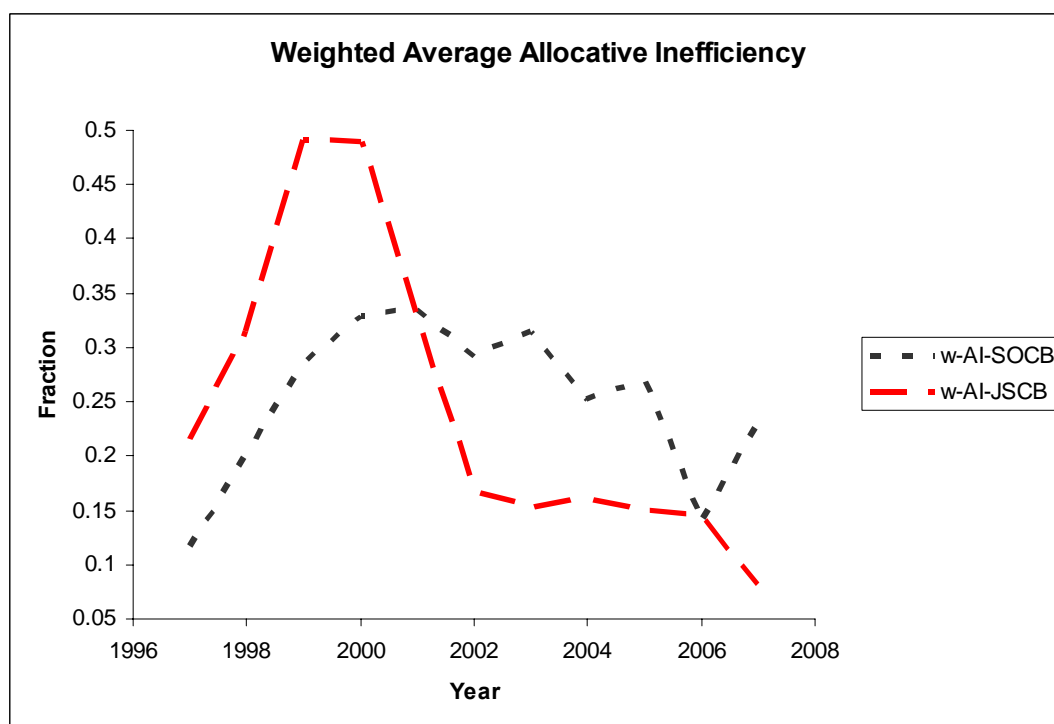


Figure 4. Weighted Average Allocative Inefficiency



Appendix. The Heterogeneous Bootstrap Algorithm

In this paper, we implement the heterogeneous bootstrap algorithm of Simar and Wilson (2000b) to compute the bias-corrected technical efficiency and cost efficiency scores. Specifically, we follow the following steps:

Note: In our application $n = 14, p = 3, q = 3$.

Step 1. Compute the technical efficiency scores using the original data. To be consistent with Simar and Wilson (2000b), we denote these technical efficiency estimates by $\hat{\delta}_i = \hat{\delta}_i(x_i, y_i), i = 1, \dots, n$. The $\hat{\delta}_i$ s are computed using the linear programming described in equation (17) of Simar and Wilson (2000). Note that by definition, $\hat{\delta}_i \geq 1, \forall i = 1, \dots, n$.

Step 2. For each $i = 1, \dots, n$, translate the data (x_i, y_i) into its polar coordinate representation $(y_i, \eta_i, \hat{\delta}_i)$ by defining $\eta_i = (\eta_{i1}, \eta_{i2})$ as (1) $\eta_{ik} = \frac{\pi}{2}, \forall k = 1, 2$, if $x_{i1} = 0$; and (2) $\eta_{ik} = \arctan\left(\frac{x_{k+1}}{x_k}\right), \forall k = 1, 2$, if $x_{i1} > 0$. Define matrices Z and Z_R by letting their i^{th} row be $(y_{i1}, y_{i2}, y_{i3}, \eta_i, \hat{\delta}_i)$ and $(y_{i1}, y_{i2}, y_{i3}, \eta_i, 2 - \hat{\delta}_i)$ respectively. Define matrices Z_1 and Z_{R1} by letting their i^{th} row be $(-y_{i1}, y_{i2}, y_{i3}, \eta_i, \hat{\delta}_i)$ and $(-y_{i1}, y_{i2}, y_{i3}, \eta_i, 2 - \hat{\delta}_i)$ respectively. Define matrices Z_2, Z_{R2}, Z_3 and Z_{R3} similarly. Define matrices Z_{12} and Z_{R12} by letting their i^{th} row be $(-y_{i1}, -y_{i2}, y_{i3}, \eta_i, \hat{\delta}_i)$ and $(-y_{i1}, -y_{i2}, y_{i3}, \eta_i, 2 - \hat{\delta}_i)$ respectively. Define matrices Z_{13}, Z_{R13}, Z_{23} and Z_{R23} similarly.

Define matrices Z_{123} and Z_{R123} by letting their i^{th} row be $(-y_{i1}, -y_{i2}, -y_{i3}, \eta_i, \hat{\delta}_i)$ and $(-y_{i1}, -y_{i2}, -y_{i3}, \eta_i, 2 - \hat{\delta}_i)$ respectively. For simplicity we rename the above 16 matrices as

Z_1, Z_2, \dots, Z_{16} respectively. Form the augmented $16n \times 6$ matrix $\tilde{Z} = \begin{bmatrix} Z_1 \\ \dots \\ Z_{16} \end{bmatrix}$.

Step 3. Compute the estimated covariance matrices of $Z_k, k=1, \dots, 16$. Denote them by $\hat{\Sigma}_k, k=1, \dots, 16$ respectively. Obtain the upper triangular matrices $L_k, k=1, \dots, 16$ through Cholesky decomposition: $\hat{\Sigma}_k = L_k L_k'$.

Step 4. Choose an appropriate bandwidth as suggested by Simar and Wilson (2000). We use the normal reference rule for its simplicity. This choice was also supported by Simar and Wilson's (2000b) simulation study.

Step 5. Draw n rows with replacement from \tilde{Z} to form a new $n \times (p+q)$ matrix \tilde{Z}^* . Denote the row means of \tilde{Z}^* by \bar{z}^* . The dimension of \bar{z}^* is $n \times 1$.

Step 6. Generate an $n \times (p+q)$ random matrix ε such that its entries are i.i.d. standard normal random variates. Denote its i^{th} row by ε_i . Construct a new $n \times (p+q)$ matrix ε^* by defining its i^{th} row $\varepsilon_i^* = \varepsilon_i L_k$, if the i^{th} row of \tilde{Z}^* was drawn from Z_k .

Step 7. Define the $n \times (p+q)$ matrix Γ by

$$\Gamma = \frac{1}{\sqrt{1+h^2}} (M\tilde{Z}^* + h\varepsilon^*) + l_n \otimes \bar{z}^*$$

where l_n is the $n \times 1$ vector of ones, and $M = I_n - \frac{1}{n} l_n l_n'$, with I_n the identity matrix of order n , and \otimes the Kronecker product.

Step 8. Denote the i^{th} row of Γ by γ_i , with $\gamma_i = (\gamma_{i1}, \gamma_{i2}, \gamma_{i3})$, where γ_{i1} is q -dimensional, γ_{i2} is $(p-1)$ -dimensional, and γ_{i3} is 1-dimensional. Construct an $n \times (p+q)$ matrix Z^* by defining its i^{th} row $z_i^* = (z_{i1}^*, z_{i2}^*, z_{i3}^*)$ as

$$z_i^* = \begin{cases} (\gamma_{i1}, \gamma_{i2}, \gamma_{i3}) & \text{if } \gamma_{i3} \geq 1 \\ (\gamma_{i1}, \gamma_{i2}, 2 - \gamma_{i3}) & \text{otherwise} \end{cases}$$

The so-constructed Z^* is the bootstrap data in polar coordinate.

Step 9. Translate the polar coordinate data Z^* back to Cartesian coordinate data $X^* = \{(x_i^*, y_i^*), i = 1, \dots, n\}$. Specifically, let $y_i^* = z_{i1}^*$, $\eta_i^* = z_{i2}^*$, $\delta_i^* = z_{i3}^*$, and define \tilde{x}_i by $\tilde{x}_{i1} = 1, \tilde{x}_{ik} = \tan(\eta_{ik}^*)$, $k = 1, \dots, p-1$. Define $x_i^* = \frac{\delta_i^*}{\hat{\delta}_i(\tilde{x}_i, y_i^*)} \tilde{x}_i$. X^* is our bootstrap sample data.

Repeat Step 5-8 if the linear programming of obtaining $\hat{\delta}_i(\tilde{x}_i, y_i^*)$ has no solution.

Step 10. For a given point (x_i, y_i) , using the bootstrap data X^* as the reference data to compute the technical efficiency score $\hat{\delta}^*(x_i, y_i)$ and cost efficiency score $\hat{\vartheta}^*(x_i, y_i)$ (along with the data for inputs prices). We use model (7) in Jahanshahloo *et al.* (2008) to compute the cost efficiency score $\hat{\vartheta}^*(x_i, y_i)$.

Step 11. Repeat Steps 5-10 B times. We obtain B bootstrap technical efficiency estimates $\{\hat{\delta}_b^*(x_1, y_1), \dots, \hat{\delta}_b^*(x_n, y_n)\}, b = 1, \dots, B$ and B bootstrap cost efficiency estimates $\{\hat{\vartheta}_b^*(x_1, y_1), \dots, \hat{\vartheta}_b^*(x_n, y_n)\}, b = 1, \dots, B$.

Step 12. Compute the bias-corrected estimates of technical efficiency and cost efficiency, and bootstrap confidence intervals.