Research Article

A Network DEA Model with Super Efficiency and Undesirable Outputs: An Application to Bank Efficiency in China

Jianhuan Huang,[1,2] Juanjuan Chen,[1] and Zhujia Yin[3]

[1] *School of Economics and Trade, Hunan University, Changsha 410079, China*
[2] *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100090, China*
[3] *School of Economics and Management, Changsha University of Technology & Science, Changsha 410004, China*

Correspondence should be addressed to Zhujia Yin; yinzhujia@126.com

There are two typical subprocesses in bank production—deposit generation and loan generation. Aiming to open the black box of input-output production of banks and provide comprehensive and accurate assessment on the efficiency of each stage, this paper proposes a two-stage network model with bad outputs and supper efficiency (US-NSBM). Empirical comparisons show that the US-NSBM may be promising and practical for taking the nonperforming loans into account and being able to rank all samples. Applying it to measure the efficiency of Chinese commercial banks from 2008 to 2012, this paper explores the characteristics of overall and divisional efficiency, as well as the determinants of them. Some interesting results are discovered. The polarization of efficiency occurs in the bank level and deposit generation, yet does not in the loan generation. Five hypotheses work as expected in the bank level, but not all of them are supported in the stage level. Our results extend and complement some earlier empirical publications in the bank level.

1. Introduction

Since it was first developed by Charnes et al. [1], DEA (data envelopment analysis) has been widely used to measure the performance of DMUs (decision making unites) that convert multi-inputs into multi-outputs, such as bank performance [2, 3], company performance [4, 5], hospital web security [6], production planning [7], energy consumption productivity [8], bankruptcy assessment [9], electricity distribution [10], R&D performance [11], agricultural economics [12], airport performance [13], and other applications [14]. In traditional DEA models, DMU is treated as a "black box," in which the inputs enter and outputs exit, neglecting the intervening steps [15]. What goes on inside the DMU is ignored or unknown. Yet some production systems have a network structure, such as when production by a subprocess results in an intermediate output that is an input to another subprocess [16] and the bank production which takes on a typical structure of two stages, that is, the deposit generation and loan generation [17]. Then bank managers are likely to know more information from subprocess or divisional efficiency than from bank level

efficiency. To open the "black box" and get greater insight into the production process, the network DEA model is constructed to analyze the network structure of production by researchers, such as Färe and Grosskopf [18], Lewis and Sexton [15], Sexton and Lewis [19], and Tone and Tsutsui [20]. However, to obtain a comprehensive and accurate measurement of bank efficiency, the extant models need to be updated. The first motivation of this paper is to fill up this approach gap, which is explained in detail as follows.

Färe et al. [21] first introduced network DEA model, which was improved and extended by other researchers. Lewis and Sexton [15] propose a network DEA model for multistage system which is an extension of the two-stage DEA model proposed by Sexton and Lewis [19]. Their studies solve a DEA model for each node independently. Yet these radial models are inconsistent with most practical production processes and ignore input slacks and output slacks, for they stand on the assumption that inputs and outputs undergo proportional changes. Tone and Tsutsui [20] propose a network slacks-based measure model (NSBM) to evaluate efficiency when inputs and outputs might change

nonproportionally. However, to the best of our knowledge, none of these network DEA models can distinguish all DMUs because the efficiency scores of DMUs on the frontier are one. The super efficiency model [22] provides a solution to rank the efficiency of DMUs on the frontier, which yet has not been combined to network model so far. Another issue is that undesirable/bad outputs such as nonperforming loans (NPLs) of banks are very common in production process but are usually ignored in the above literatures. Färe et al. [21] incorporate polluting outputs in the black box DEA model and model firm inefficiency using the directional output distance function. Huang et al. [23] propose a model named US-SBM which combines super efficiency, undesirable outputs, and slacks-based measure (SBM) together. Yet they are still unable to open the black box. Fukuyama and Weber [16] propose a slacks-based inefficiency measure for a two-stage system with bad outputs and analyze the source of inefficiency, which also does not consider the super efficiency. Aiming to solve the above mentioned gap, this paper extends the NSBM to a new model called US-NSBM by combining NSBM with super efficiency and undesirable outputs. The latter takes the undesirable outputs such as NPLs into account. Comparisons among models show that the US-NSBM may be more promising and practical.

Another motivation of this paper is the gap of empirical study on the efficiency of the Chinese banks in the mainland. First, while the efficiency of the Chinese commercial bank has been studied by a wealth of literatures, they employ "black box" DEA models with no consideration of intervening steps [24–26]. Consequently, less information of division-specific guidance on improving the efficiency of each stage is provided. Second, there is rising interest in using network DEA models to study the efficiency of financial institutions, such as Japanese banks [16], banks holding companies in the USA [17], Bangladesh banks [27], and nonlife insurance companies in Taiwan [28], a province of China. However, to the best of our knowledge, network DEA models have not been used to analyze banks in the mainland of China in published study so far. Although the characteristics of the proposed models and the bank efficiency are often observed in those literatures, the determinants of the divisional efficiency are seldom explored. This paper seeks to fill the above empirical gaps by applying the proposed US-NSBM model to measure the overall efficiency and divisional efficiency of banks with consideration of NPLs. Based on previous studies [16, 17], this paper divides the bank production into two stages of network structure and treats deposits and other raised funds as intermediates, that is, the outputs of deposit generation and the inputs of loan generation. Using the overall efficiency or divisional efficiency of each stage as explained variable, we run regression models to explore the determinants of efficiency both at bank level and division level. The comparisons of them stress the difference on the effect mechanism of different stages, which offers deeper insight into the sources of efficiency in the deposit generation and loan generation.

The remainder of the paper is organized as follows. In the next section, a new network SBM model is proposed to measure the bank efficiency by extending previous studies, and

the properties of the new model are also discussed. The third section compares the efficiency of Chinese banks measured by black box models and that by network models. It shows that the new model is promising and practical for it combines super efficiency and undesirable outputs with the NSBM model. Then by using the new model and data of Chinese commercial banks from 2008 to 2012, the characteristics of overall and divisional efficiency and the determinants of them are explored based on statistical comparison and regression analysis in this section. The conclusions and the limitations are presented in Section 4.

2. Methodology and Mathematical Models

2.1. Basic Models

2.1.1. Black Box Models

SBM Model with Undesirable Outputs (U-SBM). Assume that N DMUs have three types of variables, that is, inputs, desirable (good) outputs, and undesirable (bad) outputs, which are denoted as three vectors, namely, $x \in R^m$, $y^g \in R^{v_1}$, and $y^b \in R^{v_2}$, respectively, among which m, v_1, and v_2 represent the numbers of the variables. Define the matrices as follows: $X = [x_1, \ldots, x_N] \in R^{m \times N}$, $Y^g = [y_1^g, \ldots, y_N^g] \in R^{v_1 \times N}$, and $Y^b = [y_1^b, \ldots, y_N^b] \in R^{v_2 \times N}$. Set λ as a weighting vector, and assume that $X > 0$, $Y^g > 0$, and $Y^b > 0$. The production possibility set is as follows:

$$P = \left\{ (x, y^g, y^b) \mid x \geq X\lambda, y^g \leq Y^g\lambda, y^b \geq Y^b\lambda, \lambda \geq 0 \right\}. \quad (1)$$

Tone [29] extends the SBM model [30] to a new one that deals with undesirable outputs. Given that DMU_o is efficient in the presence of undesirable outputs, Tone [29] defines the nonoriented SBM efficiency as the optimal objective function value of the following program under the variable returns to scale (VRS) assumption:

$$
\begin{aligned}
[\text{U-SBM}] \quad \rho_o^* = \min &\left(\left(1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_i^-}{x_{io}}\right) \right.\\
&\times \left(1 + \frac{1}{s_1 + s_2}\right.\\
&\left.\left. \times \left(\sum_{r=1}^{v_1}\frac{s_r^g}{y_{ro}^g} + \sum_{r=1}^{v_2}\frac{s_r^b}{y_{ro}^b}\right)\right)^{-1}\right)\\
\text{s.t.} \quad & x_o - \sum_{j=1}^{n}\lambda_j x_j - s^- = 0\\
& \sum_{j=1}^{n}\lambda_j y_j^g - y_o^g - s^g = 0\\
& y_o^b - \sum_{j=1}^{n}\lambda_j y_j^b - s^b = 0\\
& \sum_{j=1}^{N}\lambda_j = 1\\
& \lambda, s^-, s^g, s^b \geq 0,
\end{aligned}
$$
$$\quad (2)$$

where s^- represents the inputs slack vectors and s^g (s^b) are the desirable (undesirable) outputs slack vectors. To solve the efficiency, the fractional program [U-SBM] is transformed into a linear programming problem by the Charnes-Cooper transformation [29, 31]. When studying the bank efficiency in the US, Holod and Lewis [17] point out that bank managers seek to simultaneously decrease input levels and increase output levels, so it would be better to evaluate the nonoriented efficiency. Moreover, they assumed variable returns to scale since they believed it to be unfair to compare "large" banks to "small" banks and vice versa. Following them, we also measure the nonoriented efficiency under the VRS assumption.

SBM Model with Undesirable Outputs and Super Efficiency (US-SBM). Charnes et al. [1] put forward the famous DEA model called CCR. However, in their model, the efficiency scores of efficient DMUs are 100% and thus cannot be distinguished. Same cases can be seen in Kao and Hwang [28] and Fukuyama and Weber [16]. Andersen and Petersen [22] propose the super efficiency model to solve this problem. The difference between the super efficiency model and the standard efficiency model lies in the fact that the DMU_o in the reference set of the super efficiency model is excluded (which is denoted by $j \neq o$). In the super efficiency model, the efficiency scores of inefficient DMUs are in accordance with those of the standard efficiency model, whereas for efficient DMUs in an input-oriented model, for example, if the super efficiency value is 130%, the DMU remains relatively efficient in the entire DMUs set, even if its inputs are increased proportionally by 30%. The super efficiency model makes it possible to rank efficient DMUs, thus providing further analysis with tangible and more accurate evidence.

Huang et al. [23] extend the U-SBM models to a model named the US-SBM model. The US-SBM model combines U-SBM with super efficiency. If DMU_o is US-SBM-efficient, the nonoriented efficiency is evaluated by solving the following program under the VRS assumption:

$$[\text{US-SBM}] \quad \rho_o^* = \min \left(\left(1 + \frac{1}{m} \sum_{i=1}^{m} \frac{s_i^-}{x_{io}} \right) \right.$$

$$\times \left(1 - \frac{1}{s_1 + s_2} \right.$$

$$\left. \left. \times \left(\sum_{r=1}^{v_1} \frac{s_r^g}{y_{ro}^g} + \sum_{r=1}^{v_2} \frac{s_r^b}{y_{ro}^b} \right) \right)^{-1} \right)$$

$$\text{s.t.} \quad x_o - \sum_{j=1, \neq o}^{n} \lambda_j x_j + s^- \geq 0$$

$$\sum_{j=1, \neq o}^{n} \lambda_j y_j^g - y_o^g - s^g \geq 0$$

$$y_o^b - \sum_{j=1, \neq o}^{n} \lambda_j y_j^b - s^b \geq 0$$

$$1 - \frac{1}{s_1 + s_2} \left(\sum_{r=1}^{v_1} \frac{s_r^g}{y_{ro}^g} + \sum_{r=1}^{v_2} \frac{s_r^b}{y_{ro}^b} \right) \geq \varepsilon$$

$$\sum_{j=1, \neq o}^{N} \lambda_j = 1$$

$$\lambda, s^-, s^g, s^b \geq 0. \tag{3}$$

The term ε is non-Archimedean infinitely small. When the scores of SBM-efficient DMUs are measured, the growth of undesirable outputs may exceed 100%. This may make the denominator of the object function negative, potentially leading to objective function boundlessness; that is, the optimal value approaches negative infinity. To avoid this result, the fourth constraint is appended, thus limiting the denominator of the objective function to a positive number.

2.1.2. Network SBM Model (NSBM). When the black box is opened, the intermediate products and divisions (stages) in the production process should be considered. Assume that N DMUs ($j = 1, \ldots, N$) consist of K divisions. Let m_k and v_k be the number of inputs and outputs of Division k ($k = 1, \ldots, K$), respectively, and let ζ_k be the number of intermediate products. Denote the link leading from Division k to Division h by (k, h) and the set of links by L. The observed data are $\{X_j^k \in R_+^{m_k}\}$ (input resources to DMU_j at Division k), $\{Y_j^k \in R_+^{v_k}\}$ (output products from DMU_j at Division k), and $\{Z_j^{(k,h)} \in R_+^{t(k,h)}\}$ (linking intermediate products from Division k to Division h), where $t(k, h)$ is the number of items in Link (k, h). Note that $Z_j^{(k,h)}$ denotes outputs from k and inputs to h.

Tone and Tsutsui [20] evaluated the nonoriented overall efficiency of DMU_o as follows:

$$[\text{NSBM}] \quad \rho_o^* = \min \frac{\sum_{k=1}^{K} w^k \left[1 - (1/m_k) \left(\sum_{i=1}^{m_k} \left(s_i^{k-}/x_{io}^k \right) \right) \right]}{\sum_{k=1}^{K} w^k \left[1 + (1/v_k) \left(\sum_{r=1}^{v_k} \left(s_r^{k+}/y_{ro}^k \right) \right) \right]}$$

$$\text{s.t.} \quad x_o^k - \sum_{j=1}^{n} \lambda_j^k x_j^k - s^{k-} = 0$$

$$\sum_{j=1}^{n} \lambda_j^k y_j^k - y_o^k - s^{k+} = 0$$

$$z^{(k,h)} \lambda^h = z^{(k,h)} \lambda^k$$

$$\sum_{j=1}^{N} \lambda_j^k = 1$$

$$\sum_{k=1}^{K} w^k = 1$$

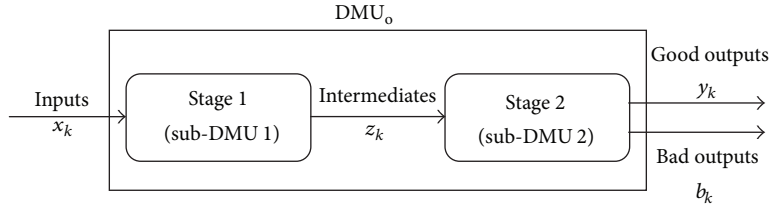$$\lambda^k, s^{k-}, s^{k+}, w^k \geq 0, \tag{4}$$

FIGURE 1: Two-stage network of bank production.

where $\lambda^k \in R_+^n$ is the intensity vector corresponding to Division k and s^{k-} (s^{k+}) are the input (output) slack vectors. w^k is the relative weight of Division k which is determined corresponding to its importance. Note that bad outputs are not considered in this model.

Tone and Tsutsui [20] then defined divisional efficiency score by

$$\rho_k = \frac{1 - (1/m_k)\left(\sum_{i=1}^{m_k}\left(s_i^{k-*}/x_{io}^k\right)\right)}{1 + (1/v_k)\left(\sum_{r=1}^{v_k}\left(s_r^{k+*}/y_{ro}^k\right)\right)}, \tag{5}$$

where s^{k-*} and s^{k+*} are the optimal input slacks and output slacks for (4).

2.2. A New Network SBM Model to Measure Bank Efficiency

2.2.1. Bank Production. Allowing the bank production to have a network structure has intuitive appeal [16]. There are two typical subprocesses in bank production—deposit generation and loan generation. In the first stage, banks use inputs such as labor, fixed assets, and equity capital to absorb deposits and other funds. In the second stage, those deposits are used to produce loans and other earning assets such as securities investments. However, traditional DEA models treat bank production as a "black box" with no consideration of the underlying process [17]. Therefore, it is difficult for bank managers to indentify which stage is more efficient. Network DEA allows managers to open the "black box" and analyzes the efficiency of each stage by defining the two stages as sub-DMUs. Deposits and other funds are defined as intermediates; that is, they are outputs produced by the sub-DMU 1 (stage 1) and also are inputs for the sub-DMU 2 (stage 2). Figure 1 illustrates the two-stage network model.

2.2.2. A Two-Stage Network Model with Bad Outputs and Supper Efficiency (US-NSBM). Neither the undesirable outputs nor the super efficiency is considered in the NSBM model proposed by Tone and Tsutsui [20]. Thus, it could be difficult to evaluate the bank efficiency accurately. And the efficient DMUs on the frontier are indistinguishable. This paper extends the NSBM model to US-NSBM model by incorporating the undesirable outputs and supper efficiency into NSBM model. Let $Y^g = [y_1^g, \ldots, y_N^g] \in R^{v_1 \times N}$ be desirable (good) outputs matrix, and let $Y^b = [y_1^b, \ldots, y_N^b] \in R^{v_2 \times N}$ be undesirable (bad) outputs matrix. Following Holod and Lewis [17], nonoriented model is applied to emphasize

the importance of simultaneously decreasing inputs and increasing outputs in a bank, which is often adopted by managers. We evaluate the nonoriented overall efficiency of DMU_o as follows:

$$[\text{US-NSBM}] \quad \rho_o^* = \min\left(\left(\sum_{k=1}^{K} w^k \left[1 + \frac{1}{m_k}\left(\sum_{i=1}^{m_k}\frac{s_i^{k-}}{x_{io}^k}\right)\right]\right)\right.$$

$$\times \left(\sum_{k=1}^{K} w^k\right.$$

$$\times \left[1 - \frac{1}{v_{1k} + v_{2k}}\right.$$

$$\left.\left.\times \left(\sum_{r=1}^{v_{1k}}\frac{s_r^{gk}}{y_{ro}^{gk}} + \sum_{r=1}^{v_{2k}}\frac{s_r^{bk}}{y_{ro}^{bk}}\right)\right]\right)^{-1}\right)$$

$$\text{s.t.} \quad x_o^k - \sum_{j=1,\neq o}^{n} \lambda_j^k x_j^k + s^{k-} \geq 0$$

$$\sum_{j=1,\neq o}^{n} \lambda_j^k y_j^{gk} - y_o^{gk} + s^{gk} \geq 0$$

$$y_o^{bk} - \sum_{j=1,\neq o}^{n} \lambda_j^b y_j^{bk} + s^{bk} \geq 0$$

$$1 - \frac{1}{v_{1k} + v_{2k}}\left(\sum_{r=1}^{v_{1k}}\frac{s_r^{gk}}{y_{ro}^{gk}} + \sum_{r=1}^{v_{2k}}\frac{s_r^{bk}}{y_{ro}^{bk}}\right) \geq \varepsilon$$

$$z^{(k,h)}\lambda^h = z^{(k,h)}\lambda^k$$

$$\sum_{j=1,\neq o}^{N} \lambda_j^k = 1$$

$$\sum_{k=1}^{K} w^k = 1$$

$$\lambda^k, s^{k-}, s^{gk}, s^{bk}, w^k \geq 0. \tag{6}$$

This program can be solved by transforming it to a linear program with the Charnes and Cooper [31] transformation. For a two-stage (divisions) bank production procession, that is, $K = 2$, the situation is different from what is defined in the NSBM model. The bad outputs such as NPLs are parts of

stage 2 instead of stage 1. So the divisional efficiency score of each stage is computed as follows:

$$\rho_o^{*1} = \frac{1 + (1/m_1)\left(\sum_{i=1}^{m_1}\left(s_i^{1-*}/x_{io}^k\right)\right)}{1 - (1/\zeta)\left(\sum_{r=1}^{\zeta}\left(s_r^{1+*}/z_{ro}\right)\right)},$$

$$\rho_o^{*2}$$

$$= \frac{1 + (1/\zeta)\left(\sum_{r=1}^{\zeta}\left(s_r^{1+*}/z_{ro}\right)\right)}{1 - \left(1/\left(v_{12}+v_{22}\right)\right)\left(\sum_{r=1}^{v_{12}}\left(s_r^{g*}/y_{ro}^g\right) + \sum_{r=1}^{v_{22}}\left(s_r^{b*}/y_{ro}^b\right)\right)},$$

(7)

where ζ is the number of intermediates. s_i^{1-*} (s_i^{1+*}) is the optimal slacks in the inputs (outputs) of stage 1 (2). s_r^{g*} (s_r^{b*}) is the optimal slacks in good (bad) outputs of stage 2 for (6). And v_{12} (v_{22}) is the number of good (bad) outputs of the stage 2.

2.3. Several Properties of the US-NSBM Model.
In this section, several properties of the US-NSBM models are discussed.

Theorem 1. *A DMU is overall efficient if and only if it is efficient for all divisions.*

Proof. Similar theorem has been proved in the NSBM model by Tone and Tsutsui [20]. Adding a new output such as *NPLs* will not change the theorem. So for the US-NSBM model, its main difference from NSBM lies in the consideration of super efficiency which has no influence on model except ranking the efficient DMUs. Given that the inputs and outputs are fixed, the number of efficient and inefficient DMUs measured by NSBM is the same as that measured by US-NSBM. Therefore, Theorem 1 also goes on the US-NSBM. Note that the division efficiency is defined as the ratio of inputs/outputs only if overall efficiency is the optimal for (6). In other words, if one of divisions is inefficient, the DMU is overall inefficient. Table 4 in Section 3.2 provides empirical evidence that 16 banks are overall efficient if and only if all divisions are efficient. □

Theorem 2. *Every division has at least one divisionally efficient DMU under the variable returns-to-scale assumption.*

Proof. Similar to Tone and Tsutsui [20], we sort the n DMUs in the Division k in ascending order in input values using Input i as the ith key and then sort the resultant in descending order in output values, using Output r as the $m_k + r$th key. The slacks of DMU listed at the lexicographical minimum (top) are zero for every feasible λ^k under the VRS assumption. So the division has at least one efficient DMU regardless of the orientation. □

3. An Application to Efficiency Analysis of Chinese Bank

3.1. Data and Input-Outputs.
The samples in this paper are commercial banks in the China mainland. The data of banks are drawn from BVD (Bankscope) from 2008 to 2012. Due to lack of data, only 37 samples are available for each year (the procedure of choosing samples is as follows. First, drop all banks with missing value. Second, drop all banks which were set up after 2007. Third, keep data since 2008 which leads to the biggest number of samples. Forth, only keep samples with panel data. We get 37 banks and 185 observations from the original 139 banks). In addition, the macroeconomics data such as growth of GDP (gross domestic product), GDP deflator, and growth of $M2$ (money and quasi money) are drawn from the World Bank.

Following previous studies [16, 17], the inputs and outputs are set as follows. The inputs are fixed assets (x_1), equity (x_2), and personnel expenses (x_3) (note that personnel expenses are used in this paper, which is different from the number of employees used by Holod and Lewis [17]. For one thing, only 17 samples are available for each year when using the number of employees as input, as many banks did not report the indicator in the BVD. For another, the number of employees does not reflect the real cost or investment because banks have different policies on the wages, benefits, and trainings). The deposits and short-term funding (z_1) and other raised funds (z_2 = *total funding minus deposits and short-term funding*) are treated as intermediates. The desirable outputs are gross loans (y_1) and other earning assets (y_2). The latter is defined as the sum of securities, federal funds sold, and trading assets. The importance of accounting for bad loans as part of the lending process cannot be understated [16]. So nonperforming loans (*NPLs*) are used as undesirable output, which includes more than 90 days past due loans, nonaccrual loans, and restructured loans. All financial data are deflated using GDP deflator with a base = 100 in 2008. Descriptive statistics of the inputs and outputs of 185 observations in pooled sample are provided in Table 1.

3.2. Model Comparisons.
The comparisons among models will help to identify which is the best. As for black box model, two cases should be considered. In the first case, deposits are treated as inputs and in the other case as outputs. Both of them are based on US-SBM model [23]. With regard to network models, NSBM proposed by Tone and Tsutsui [20] and US-NSBM introduced in this paper are included, using initial inputs, intermediates, and final outputs mentioned in Section 3.1. To have a better comparison with Tone and Tsutsui [20] models, this paper sets weight of one-half for each subprocess: $w1 = w2 = 0.5$, following Fukuyama and Weber [16]. It implies that two divisions are equally important. Similar to Fukuyama and Weber [16], the observed input and output vectors we use are $g^x = x_o$, $g^y = y_o$, and $g^b = b_o$, for it is unfair to compare "large" banks to "small" banks and vice versa. So we get four models, that is, black box (input), black box (output), NSBM, and US-NSBM, accordingly. The main difference between black box models and US-NSBM is that the latter uses two-stage network model and regards z_1 and z_2 as intermediates. The main difference between NSBM and US-NSBM lies in that the latter considers supper efficiency and bad output (NPLs).

Holod and Lewis [17] suggest that the assumption of variable returns to scale (VRS) was a better alternative to

TABLE 1: Pooled descriptive statistics of inputs and outputs (185 observations).

Variable	Mean	Std. Dev.	Min.	Max.
Subprocess 1 inputs				
x_1 = fixed assets	2489.101	5914.042	2.561	28267.010
x_2 = equity	17342.500	38230.240	80.385	212196.100
x_3 = personnel expenses	1612.693	3565.357	6.613	18097.020
Intermediates				
z_1 = deposits and short-term funding	257641.300	550848.900	1286.135	2897070.000
z_2 = other raised funds	4941.384	12923.420	0.010	106278.600
Final outputs				
y_1 = gross loans	146832.400	314206.600	892.766	1655451.000
y_2 = other earning assets	130770.400	282761.900	287.213	1537414.000
b = NPLs	1860.672	4234.151	0.038	19615.920

Note: all are in million USD.

TABLE 2: Pooled descriptive statistics of bank efficiency measured by different models.

Model	Obs	Mean	Std. Dev.	Min.	Max.	Coefficient of variation	Number of efficient DMUs
Black box (input)	185	0.663	0.400	0.000	1.446	0.603	76
Black box (output)	185	7.177	85.060	0.389	1158.000	11.852	126
NSBM	185	0.425	0.270	0.000	1.000	0.635	12
US-NSBM overall	185	0.611	0.228	0.264	1.507	0.373	16
US-NSBM 1	185	0.752	0.186	0.361	1.199	0.247	26
US-NSBM 2	185	0.802	0.155	0.323	1.507	0.193	18

Note: US-NSBM 1 and US-NSBM 2 denote US-NSBM stage 1 and US-NSBM stage 2 respectively (the same below).

constant returns to scale. Thus a software named MAXDEA is used to estimate the bank efficiency (inefficiency) under the assumption of VRS. The model comparisons are reported in Tables 2–4.

Table 2 presents the pooled descriptive statistics of bank efficiency measured by different models. Surprisingly, there are 126 (76) efficient DMUs among 185 observations when z_1 and z_2 are treated as outputs (inputs) in the black box models, which means that almost 2/3 (1/2) of DMUs are efficient. At least one outlier (1158) occurs when z_1 and z_2 are considered as outputs, which leads to the abnormal mean (7.177) and standard. deviation (85.06). And there are only 16 (12) efficient DMUs in the US-NSBM (NSBM). It shows that the black box models may not provide accurate results. The maximum efficiency value measured by NSBM is 1, which means that it cannot rank all efficient banks, while the maximum by US-NSBM is bigger than 1, which indicates that the efficient DMUs are recognizable. Note that the minimum of NSBM is zero, which means that at least one DMU is totally inefficient or without output. This is not true for the minimum of outputs being 287.213 million \$ (see Table 1). This outlier does not occur in the results of US-NSBM and the minimums of overall/divisional efficiency are bigger than 0.2 (0.264/0.361/0.323). The above findings indicate that more reasonable and accurate measurements could be provided by US-NSBM.

Using US-NSBM as a base model, the correlations and difference comparisons are reported in Table 3. There are significant high correlations between US-NSBM and other models (Spearman's rank coefficients are bigger than 0.5).

However, both paired t-test and sign rank test show that there are also significant differences among the four models. There are similar significant differences between the divisional efficiency of stage 1 (2) measured by US-NSBM and that measured by NSBM. Therefore, the US-NSBM cannot be replaced by other models.

Table 4 reports efficient observations measured by US-NSBM, including overall or divisional ones. 28 observations of 15 banks are overall or divisionally efficient in some years. Among them, 16 observations are overall efficient, and 26(18) observations are divisionally efficient in the stage 1 (2). The 16 observations are overall efficient if and only if both of the divisions are efficient. These findings not only provide evidence of Theorems 1 and 2, but also show that, when a DMU is overall or divisionally efficient in the US-NSBM, it is efficient in the black box (output) model.

3.3. Characteristics of Bank Efficiency in China.
The US-NSBM model is used to study the bank efficiency in the following parts as a consequence of providing a better bank-efficiency measurement indicated in the above comparisons. Columns 6, 7, and 8 of Table 4 show that Industrial & Commercial Bank of China (ICBC) is overall and divisionally efficient from 2010 to 2012 and Hana Bank (China) Company Ltd. is overall efficient and divisionally efficient at stage 2 from 2008 to 2009, as well as divisionally efficient at stage 1 from 2008 to 2009 and 2011 to 2012. The results indicate that ICBC is efficient in both deposit generation and loan generation, while Hana Bank is only efficient in the deposit generation for the observed year. Although some banks such as HSBC

Table 3: Correlations and difference comparisons of bank efficiency measured by different models.

Groups	Correlations		Difference comparisons		
	Pearson	Spearman's rank	t-test (t value)	Mean difference	Sign rank test (Z value)
Black box (input) versus US-NSBM	0.722***	0.752***	2.502**	0.052	2.389**
Black box (output) versus US-NSBM	0.292***	0.666***	1.051	6.566	11.779***
NSBM versus US-NSBM	0.574***	0.535***	−10.858***	−0.185	−9.763***
NSBM 1 versus US-NSBM 1	0.843***	0.855***	12.036	0.091	9.787***
NSBM 2 versus US-NSBM 2	0.283***	0.290***	−15.863***	−.313	−10.780***

Note: *** or ** denotes significance at the level of 1% or 5%, respectively.

Table 4: Efficient observations measured by US-NSBM model.

Year	Bank name	Black box (input)	Black box (output)	NSBM	US-NSBM		
					Overall	Stage 1	Stage 2
2009		1.014	1.018	0.893	0.965	1.000	0.965
2010	Industrial and Commercial Bank of China	1.016	1.112	0.912	1.004	1.003	1.001
2011		1.007	1.006	0.941	1.003	1.000	1.003
2012		1.188	1.118	1.000	1.118	1.000	1.118
2009	Agricultural Bank of China Limited	1.136	1.067	0.968	0.812	1.000	0.812
2012		1.013	1.016	1.000	1.012	1.000	1.012
2012	Bank of Communications Co. Ltd.	1.028	1.035	1.000	1.021	1.009	1.012
2011	China Merchants Bank Co. Ltd.	1.007	1.030	0.377	0.723	0.723	1.000
2012	China Minsheng Banking Corporation	1.046	1.050	1.000	1.046	1.000	1.046
2011	Shanghai Pudong Development Bank	1.073	1.137	0.380	1.112	1.000	1.112
2012		1.117	1.041	1.000	1.046	1.000	1.046
2011	Bank of Beijing Co. Ltd.	1.201	1.141	0.447	1.126	1.022	1.102
2012		1.061	1.193	1.000	1.048	1.000	1.048
2011	HSBC Bank (China) Co. Ltd.	1.259	1.336	0.031	1.085	1.000	1.085
2012		1.203	1.168	1.000	1.173	1.000	1.173
2011	Hankou Bank	1.446	1157.860	0.003	1.507	1.000	1.507
2012	Citibank (China) Co. Ltd.	1.171	1.291	1.000	1.114	1.000	1.114
2008	Nanyang Commercial Bank (China) Limited	1.056	1.133	0.286	0.519	1.000	0.519
2009		0.003	1.163	0.260	0.464	1.000	0.464
2008		1.232	1.193	1.000	0.663	1.000	0.663
2009	Bank of Fuxin Co. Ltd.	1.033	1.040	0.493	0.619	1.000	0.619
2012		1.000	1.000	0.000	0.727	0.727	1.000
2008	First Sino Bank	1.062	1.051	0.203	0.687	1.000	0.687
2008	United Overseas Bank (China) Limited	1.242	1.209	0.957	0.917	1.000	0.917
2008		1.358	1.359	0.624	1.199	1.199	1.000
2009	Hana Bank (China) Company Ltd.	1.019	1.058	1.000	1.017	1.017	1.000
2011		1.050	1.049	0.188	0.878	1.000	0.878
2012		1.105	1.203	0.131	0.899	1.000	0.899
	Mean	1.077	42.431	0.646	0.947	0.989	0.957

Bank are also efficient in both deposit generation and loan generation, most banks such as Nanyang Commercial Bank are only efficient in one of them.

The histograms of overall efficiency and divisional efficiency at each stage are shown in Figures 2(a), 2(b), and 2(c), respectively, which picture the frequency distribution of the efficiency scores of all samples (care is needed in comparing divisional scores mutually [20], because the number of inputs and outputs may differ division by division and DEA scores

are affected by the inputs/outs and benchmark. It is easy to say that stage 1 (1) of a DMU is highly efficient, while it is hard to judge which stage is more efficient. For example, the score of certain DMU is 0.9 in stage 1 and that in stage 2 is 0.8, yet we should simply not make the conclusion that the efficiency in stage 1 is higher than that in stage 2. More evidence is needed). There are significant differences among the distribution of them. First, few banks are efficient in the whole production or in the loan generation, that is, stage 2,
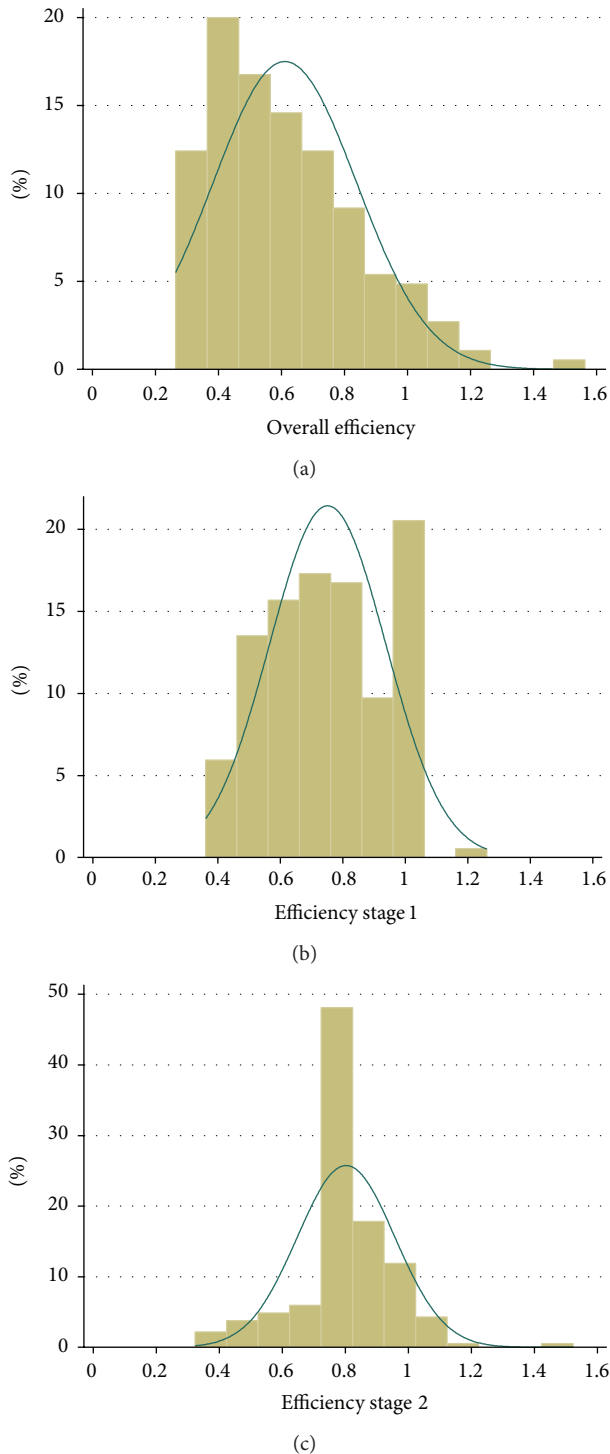
(a)

(b)

(c)

FIGURE 2: (a) Histogram of overall efficiency. (b) Histogram of efficiency stage 1. (c) Histogram of efficiency stage 2.

yet more banks are efficient in the deposit generation, that is, stage 1. As described in Figure 2(a), only 8.65% of the samples are overall efficiency, and 20% of the samples are in the interval [0.35, 0.45], which shows that most banks are overall inefficient in the observation period. As shown in Figure 2(c), only 9.72% is bigger than or equal to 1, which

demonstrates the inefficiency in the loan generation of many banks. However, more than 20% of the samples are in the interval [0.95, 1.05] in the deposit generation, which indicates that a number of banks are efficient in the deposit generation relatively (Figure 2(b)). Second, polarization occurs in the deposit generation. More than 50% of samples are smaller than 0.75 in the deposit generation (Figure 2(b)), while about 50% are in the interval [0.75, 0.85] in the loan generation (Figure 2(c)). The coefficient of variation in the deposit generation is 0.247, which is much bigger than that in the loan generation (0.193; see Table 2). Similarly, the overall efficiency also shows polarization. However, the divisional efficiency of stage 2 clusters around 0.8, which indicates that the banks do not show significant differences in the loan generation.

3.4. The Determinants of Chinese Banks' Efficiency in Two Stages

3.4.1. Hypothesis and Variables. Lots of factors may exert influence on the bank efficiency. This paper focuses on the five types of determinants and goes further to the subprocess or divisional level. The hypotheses are formulated as follows.

(1) "Risk taking" hypothesis: high bank efficiency may be a result of managers bearing more risk. From the bank level, under this hypothesis, a loss of earnings from any source reduces the bank efficiency. Once the efficiency of a bank is reduced, the bank responds to moral hazard incentives by increasing the riskiness of its loan portfolio, which results in higher loans, higher nonperforming loans, and higher efficiency score on average in the future. Thus we identify "risk taking" hypothesis by the relationship between net loan to total assets and bank efficiency. Fortunately, previous empirical study [32] has shown that net loan to total assets is positively and significantly related to cost efficiency of commercial banks, saving banks, and cooperative banks. It indicates that bank risk has a significant effect on bank efficiency. However, if we go further to the division or subprocess level, the coin may not be this side. As to the deposit generation, risk taking may bring higher raised funds such as deposits, which may improve the efficiency of this stage. Yet to the loan generation, risk taking may bring higher loans and NPLs. The former may help to improve the efficiency, while the latter goes on the contrary. So the results may be mixed if the bad output is considered in this stage.

(2) "Assets liquidity" hypothesis: banks with higher assets liquidity may be regarded as more powerful, solvent, and healthy, which helps to attract investment and customers and tends to increase bank efficiency. Banks with more liquidity may also imply that bank managers are proficient in assets management and therefore can make more good outputs. Previous study [32, 33] found a positive relationship between bank liquidity and cost efficiency for a large sample of European banks. The situation at the subprocess level

such as deposit/loan generation may be similar to the bank level.

(3) "Interest margin" hypothesis: higher net interest margin may exert a negative impact on bank efficiency because the banks tend to be in a not-so-fierce competition. As an indicator of competitive behavior in the banking industry, low net interest margin means a competitive market, where managers are stimulated to make every effort to obtain more good outputs and therefore bank efficiency is improved, and vise visa. Previous study [34] indicates that managerial efficiency of EU banks is negatively and significantly related to net interest margins. As to the subprocess level, a similar result is expected.

(4) "Shareholders behind" hypothesis: banks supported by powerful, rich, and professional shareholders may have the advantage to absorb deposits, to grasp business opportunities, and to access to large customers, which helps to improve efficiency. In this paper, two kinds of obvious supports are identified. The first factor is that whether a bank is a foreign bank. Extant literature [35, 36] shows that foreign banks are more efficient than domestic banks in China, which indicates that foreign banks perform better in both profit efficiency and cost efficiency due to shareholders' superior management skills. The other factor is that whether it is a listed bank. Being a listed company is conducive to improving bank's ability to gain more investment such as initial public offer and debt issuing and increasing the reputation. The supervision from public market forces bank managers to work efficiently as well. This kind of support is helpful to the bank efficiency. We suppose that this support mechanism also works for the deposit and loan generation.

(5) "Scale effect" hypothesis: banks with larger size are probably more efficient because of two reasons. One is economies of scale; that is, the more deposits or loans are produced, the lower per-unit fixed cost is because these costs are shared over a larger number of products. Altunbas et al. [32] and Sufian [37] found that banks with larger size were more efficient for economies of scale. The other is that large banks are likely to be easier to absorb deposits and get access to customers in consideration of its safety, powerfulness, and convenience. As to the subprocess level, we expect that "scale effect" shows itself in the deposit and loan generation in Chinese banks.

Other country level factors, such as macroeconomic conditions, monetary policy, and financial system structure, may also have an influence on bank efficiency (note that there are other factors, such as macroeconomic factors at the country level, including macroeconomic conditions, monetary policy, and financial system structure). Following previous studies [38–40], the annual growth rate of GDP (*g_gdp*) is applied to gauge the economy development. And similar to Wang and Zhu [26], we employ annual growth rate of money and quasi money (*g_m2*) as an indicator for the monetary policy formulated by central bank which may have a significant effect on the bank efficiency. The influence of the market structure is considered, which is measured by Herfindahl-Hirschman index (HHI) [41]. The HHI index is defined as the sum of squares of individual bank loan in the total banking sector loan for China in a certain year [42, 43] (*hhi_gross_loan*). The data of the three factors comes from the World Bank Group.

3.4.2. Regression Model. Based on the test proposed by Breusch and Pagan [42] (see [44]) and Hausman test, random panel model is proper to this study. To deal with heteroskedasticity, the FGLS (feasible generalized least squares; see [44]) approach is employed to estimate the regression model (in this case, some literatures employ Tobit model for regression since they hold that there are two limits in OLS model as the efficiency interval is [0, 1]. Hoff [45] makes a specialized comparison on the selection between OLS and Tobit model and points out that the judgment of two limits is incorrect because it is possible for the efficiency score to value at 1 but impossible to value at 0. Hoff then holds that there is no significant disparity between the regression results of OLS and Tobit model whether the efficiency interval is [0, 1] or not, so OLS can be adopted in most cases. This paper uses super efficiency model, and the DMU efficiency score on the frontier can be greater than 1, which excludes the limit that efficiency score can value at no greater than 1, so Tobit model is not adopted). The estimated equation is as follows:

$$y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{kit} + \sum_{r=1}^{R} C_r x_{(K+r)it} + u_i + w_{it}, \qquad (8)$$

where y_{it}, that is, the dependent variable, is the overall/divisional efficiency of bank i ($i = 1, 2 \ldots, N$) at period t ($t = 1, 2 \ldots, T$). β_0 is intercept term. x_{kit} is the kth ($k = 1, 2 \ldots, K$) observed variables of bank i at period t, and β_k is the coefficient of the kth observed variable. In this paper, six observed variables are set (see Table 5). $x_{(K+r)it}$ is the rth ($r = 1, 2 \ldots, R$) control variables, and C_r is the coefficient of control variables, respectively. Three control variables are set, that is, *g_gdp* (growth of GDP), *g_m2* (growth of *M2*), and *hhi_gross_loan* (HHI) computed by gross loans. u_i is the unobserved individual effects and w_{it} is the error term.

3.4.3. Empirical Results. Table 6 presents the descriptive statistics of the variables for the pooled sample of 185 observations.

The Pearson correlations of independent variables are provided in Table 7, which shows that only the absolute value of correlation coefficient between the *list* and *ln_ta* is bigger than 0.5 (0.774). The implication is that the multicollinearity of them should be considered. The variables *list* and *ln_ta* appear in the regression models separately.

Table 8 reports the regression results estimated by FGLS approach. For each model, the log_likelihood and Wald tests are also reported, which shows that the regression model is significant as a whole. Both model 1 and model 2

TABLE 5: Definition of variables used to test the hypotheses.

Variable description	Definition	Hypothesis test
Net loan to total assets	$netloan\_ta = \dfrac{net\ loan}{total\ assets}$	"Risk taking" (+)
Liquid assets to customer and ST funding	$liqassets\_stfund = \dfrac{liquid\ assets}{customer\ deposit + other\ short\ term\ funding}$	"Assets liquidity" (+)
Net interest margin	$net\_int\_margin = \dfrac{interest\ income}{earning\ assets}$	"Interest margin" (−)
Shareholder	$Sharehold = 1$ for foreign banks, 0 otherwise	"Shareholders behind" (+)
List	$List = 1$ for publicly listed banks, 0 otherwise	
Total assets	$ln\_ta = \ln(total\ assets)$	"Scale effect" (+)

Note: data comes from Bankscope database. All ratios are expressed in percentage points except defined specially. The expected coefficient signs at bank level are shown in parenthesis.

TABLE 6: Pooled descriptive statistics of variables.

Variable	Mean	Std. Dev.	Min.	Max.	Unit
Overall efficiency	61.080	22.790	26.400	150.700	%
Efficiency of stage 1	75.170	18.610	36.100	119.900	%
Efficiency of stage 2	80.250	15.500	32.300	150.700	%
netloan_ta	50.900	9.355	31.420	83.250	%
liqassets_stfund	32.840	9.551	10.930	56.970	%
net_int_margin	2.671	0.766	1.110	6.112	%
sharehold	0.270	0.445	0.000	1.000	
list	0.341	0.475	0.000	1.000	
ta	261151.000	550026.000	1592.000	2789000.000	Million $
g_gdp	9.260	0.845	7.800	10.400	%
g_m2	18.530	5.157	13.610	27.680	%
hhi_gross_loan	0.098	0.007	0.091	0.111	

regard overall efficiency as dependent variable. The difference between the two models is that the variable *list* enters model 1, while the variable *ln_ta* enters model 2. Similar arrangement is also used for models 3 (5) and 4 (6) with the efficiency of stage 1 (2) as dependent variables.

The coefficient of *netloan_ta* is positive and significant in model 1, which indicates that "risk taking" hypothesis is true in the bank level; that is, bank managers' efforts to bear more risks may help to improve bank efficiency. Similarly, the coefficient of it is positive and significant in models 3 and 4, which provide strong evidence that "risk taking" hypothesis is true in the deposits generation. However, the coefficient is insignificantly negative in models 5 and 6, which means that the hypothesis is not supported in this subprocess. The above results present that the banks' efforts to attract funds are helpful to the efficiency in the first stage, yet radical behavior such as relaxing the policy on loans may not work. The negative coefficient implies that the *NPLs* may increase more quickly when banks issue loans in a radical way.

The coefficients of *liqassets_stfund* are positive and significant in the six models, which indicates that "assets liquidity" hypothesis is true not only in the bank level, but also in the subprocess level. High liquidity of assets does help banks to improve their overall and divisional efficiency.

As expected, the coefficient of *net_int_margin* is negative and significant in model 2, which indicates that "interest margin" hypothesis is supported in the bank level. The results of models 5 and 6 are similar, which means that the hypothesis also works in the second stage. However, the situation is different for stage 1. The coefficient of *net_int_margin* is positive (negative) in model 3 (4) indistinctively, which shows that the "interest margin" hypothesis does not work in this stage. Although this finding is different from previous studies, it can be explained. One more thing that should be noted is that interest margin is decided mainly by the central bank of China instead of the bank market. Low competition related to higher net interest margin may weaken the banks' efforts to issue loans. However, more deposit and other funds are still needed for banks to enlarge the source of loans and investments, which may not be loosen even in a market with low competition.

As to the "shareholders behind" hypothesis, the coin goes on two sides. First, the coefficients of *sharehold* are significantly positive in models 1 to 4 (except model 2),

TABLE 7: Correlation of dependent variables.

	netloan_ta	liqassets_stfund	net_int_margin	sharehold	list	ta	ggdp	g_m2
netloan_ta	1							
liqassets_stfund	−0.454	1						
net_int_margin	0.143	−0.257	1					
sharehold	0.259	0.240	−0.471	1				
list	−0.098	−0.092	−0.023	−0.437	1			
ln_ta	−0.217	−0.053	−0.104	−0.478	0.774	1		
g_gdp	0.160	−0.297	0.015	0.000	−0.004	−0.114	1	
g_m2	0.273	−0.378	−0.114	0.000	−0.019	−0.142	0.336	1
hhi_gross_loan	0.302	−0.411	0.124	0.000	−0.025	−0.200	0.309	0.422

TABLE 8: Regression results.

Variable	Overall efficiency		Efficiency of stage 1		Efficiency of stage 2	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
netloan_ta	0.362**	0.314	0.591***	0.555***	−0.063	−0.085
	(−0.184)	(−0.195)	(−0.156)	(−0.163)	(−0.115)	(−0.120)
liqassets_stfund	0.660***	0.492**	0.543***	0.418**	0.323***	0.245**
	(−0.186)	(−0.195)	(−0.158)	(−0.164)	(−0.117)	(−0.120)
net_int_margin	−2.326	−6.227***	2.557	−0.338	−6.442***	−8.253***
	(−2.394)	(−2.380)	(−2.029)	(−1.993)	(−1.504)	(−1.460)
sharehold	12.483***	5.532	18.376***	13.217***	−4.006	−7.234**
	(−4.764)	(−4.760)	(−4.039)	(−3.988)	(−2.992)	(−2.921)
list		22.257***		16.507***		10.334***
		(−3.435)		(−2.878)		(−2.108)
ln_ta	7.175***		5.323***		3.331***	
	(−0.873)		(−0.740)		(−0.549)	
g_gdp	0.043	−0.967	0.979	0.230	−0.634	−1.103
	(−1.709)	(−1.799)	(−1.449)	(−1.507)	(−1.073)	(−1.104)
g_m2	0.004	−0.306	0.515*	0.285	−0.466**	−0.610***
	(−0.316)	(−0.328)	(−0.268)	(−0.275)	(−0.198)	(−0.201)
hhi_gross_loan	138.769	−106.561	194.097	12.107	28.760	−85.150
	(−213.555)	(−223.597)	(−181.055)	(−187.321)	(−134.150)	(−137.203)
Constant	−66.509**	61.524**	−78.720***	16.264	67.413***	126.861***
	(−32.368)	(−28.782)	(−27.442)	(−24.113)	(−20.333)	(−17.661)
Log_likelihood	−795.508	−805.357	−764.967	−772.609	−709.495	−715.006
Wald_chi2	115.548	85.191	93.810	71.701	167.052	146.690

Note: we use the FGLS estimator with standard error (reported in brackets). ***, **, or *denotes significance at the level of 1%, 5%, or 10%, respectively. There are 185 observations for each model. Sample period is 2008–2012.

which indicates that the hypothesis is supported in the bank level and deposit generation; that is, foreign banks show better efficiency than others. But the situation reverses in the loan generation as the coefficient of *sharehold* is significantly negative in model 6 and also negative in model 5, which presents lower efficiency of foreign banks in this subprocess. It is explained that the foreign banks are probably more cautious to issue loans in China in recent years, especially after the world financial crisis. Second, as expected, the coefficients of *list* are positive and significant in all models, which indicates that being a listed company does help banks to improve their efficiency in the bank level and subprocess level.

Finally, the "scale effect" hypothesis also works as expected both in the bank level and subprocess level. The coefficients of *ln_ta* are positive in all models significantly, which implies that larger banks have been more efficient in China so far.

By the way, most of the coefficients of the control variables such as *g_gdp* and *hhi_gross_loan* are not significant, which shows that macrofactors such as the growth of GDP and market structure do not exert important influence on the bank efficiency. However, the growth of M2 shows positive influence on the deposit generation and negative influence on the loan generation significantly. This is reasonable on account that more money to the economy brought by higher

TABLE 9: Hypotheses tested based on regression results.

Number	Hypothesis	Overall production	Deposit generation	Loan generation
1	Risk taking	Y*	Y*	—
2	Assets liquidity	Y*	Y*	Y*
3	Interest margin	Y*	—	Y*
4	Shareholders behind			
	Foreign	Y*	Y*	N*
	List	Y*	Y*	Y*
5	Scale effect	Y*	Y*	Y*

Note: Y* indicates that the results significantly support the hypotheses proposed in Section 3.4.1. N* indicates that the results refuse the hypothesis significantly. And "—" indicates the hypothesis can neither be supported nor refused.

growth of $M2$ helps banks to absorb deposits and other funds, though the loans are unlikely to keep up with the speed of deposits and the $NPLs$ may increase with the rise of loans.

Table 9 sums our regression analysis to show whether the hypotheses are supported or not in different levels of bank production. It presents that all five hypotheses work as expected in the bank level. However, the situation in the divisional level is mixed. It is interesting that "risk taking" hypothesis does not work for the loan generation, and foreign banks show lower efficiency in this subprocess. Moreover, it is far beyond our expectations that "interest margin" hypothesis fails to work in deposit generation.

3.4.4. Robust Test. To test the robustness of the above conclusions, further work has been done from the following aspects. First, bank efficiency measured under the assumption of CRS is regarded as dependent variables, then runs regression, and analyzes the results again. Second, Tobit model is applied to make regression analysis. Third, some other independent variables are used, such as the ratio of debt to total assets and liquid assets to total assets. Finally, the lagged dependent variable is employed to avoid possible endogeneity between efficiency and determinants. It is found that these changes do not lead to much difference, denoting good robustness of the conclusion.

4. Conclusions

When the black box of traditional DEA model is opened, the super efficiency and undesirable outputs should be considered in the network model to provide more accurate and comprehensive measurement of bank efficiency. This paper extends the NSBM model proposed by Tone and Tsutsui [20] to a new two-stage network model named as US-NSBM by means of combing it with super efficiency and undesirable outputs. Based on the data of Chinese commercial banks from 2008 to 2012, we make empirical comparisons between black box models and network models to show that the proposed US-NSBM model may be promising and practical. This paper divides the bank production into two stages of network structure and treats deposits and other raised funds as intermediates. When the overall efficiency or divisional

efficiency at each stage is served as explained variable, regression models are constructed to explore the determinants of efficiency both at bank and division level.

Some interesting results are discovered. First, statistical analysis presents that the polarization of efficiency occurs in the bank level and deposit generation, yet it does not occur in the loan generation. Second, all hypotheses proposed in this paper work as expected in the bank level, whereas not all of them are supported in the subprocess level. For example, "risk taking" hypothesis is not true for the loan generation, and foreign banks show lower efficiency in this subprocess. It also shows that the "interest margin" hypothesis does not work in deposit generation. The empirical results shed insight into the sources of efficiency in the deposit generation and loan generation. It implicates that more efforts should be paid to loan generation especially for foreign banks and the reform of interest margin policy should be put forward by the government.

The proposed US-NSBM model can be applied to other financial holding companies and DMUs with network structure production. Nevertheless, there are some limitations and further research aspects. First, the dynamic structure should be taken into account in the network model. Second, the comparison among Chinese bank and banks of other countries could be carried out to study the difference on the determinants of divisional efficiency. Third, it is also worthwhile to explore the characteristics and determinants of insurance companies' efficiency in the divisional level.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.

[2] L. M. Seiford and J. Zhu, "Profitability and marketability of the top 55 U.S. commercial banks," *Management Science*, vol. 45, no. 9, pp. 1270–1288, 1999.

[3] N. K. Avkiran, "Association of DEA super-efficiency estimates with financial ratios: investigating the case for Chinese banks," *Omega*, vol. 39, no. 3, pp. 323–334, 2011.

[4] J. H. Huang and J. J. Chen, "A new FA-DEA method based on classification and weighting," in *Proceedings of the 5th International Conference on Business Intelligence and Financial Engineering*, pp. 154–158, Lanzhou, China, August 2012.

[5] G. J. N. Seyed and R. N. Hamid, "A two-stage DEA to analyze the effect of entrance deregulation on Iranian insurers: a robust approach," *Mathematical Problems in Engineering*, vol. 2012, Article ID 423524, 24 pages, 2012.

[6] H. Y. Kao, T. K. Chang, and Y. C. Chang, "Classification of hospital web security efficiency using data envelopment analysis and support vector machine," *Mathematical Problems in Engineering*, vol. 2013, Article ID 542314, 8 pages, 2013.

[7] J. Du, L. Liang, Y. Chen, and G.-B. Bi, "DEA-based production planning," *Omega*, vol. 38, no. 1-2, pp. 105–112, 2010.

[8] X. Han, X. Xue, J. Ge, H. Wu, and C. Su, "Measuring the productivity of energy consumption of major industries in China: a DEA-based method," *Mathematical Problems in Engineering*, vol. 2014, Article ID 121804, 12 pages, 2014.

[9] I. M. Premachandra, Y. Chen, and J. Watson, "DEA as a tool for predicting corporate failure and success: a case of bankruptcy assessment," *Omega*, vol. 39, no. 6, pp. 620–626, 2011.

[10] M. V. Pereira De Souza, M. Diallo, R. Castro Souza, and T. K. N. Baidya, "The cost efficiency of the brazilian electricity distribution utilities: a comparison of bayesian SFA and DEA models," *Mathematical Problems in Engineering*, vol. 2010, Article ID 593059, 20 pages, 2010.

[11] W. Zhong, W. Yuan, S. X. Li, and Z. Huang, "The performance evaluation of regional R&D investments in China: an application of DEA based on the first official China economic census data," *Omega*, vol. 39, no. 4, pp. 447–455, 2011.

[12] F. J. André, I. Herrero, and L. Riesgo, "A modified DEA model to estimate the importance of objectives with an application to agricultural economics," *Omega*, vol. 38, no. 5, pp. 371–382, 2010.

[13] M.-M. Yu, "Assessment of airport performance using the SBM-NDEA model," *Omega*, vol. 38, no. 6, pp. 440–452, 2010.

[14] W. W. Cooper, L. M. Seiford, and J. Zhu, *Handbook on Data Envelopment Analysis*, Springer, Boston, Mass, USA, 2nd edition, 2011.

[15] L. Lewis and T. R. Sexton, "Network DEA: efficiency analysis of organizations with complex internal structure," *Computers and Operations Research*, vol. 31, no. 9, pp. 1365–1410, 2004.

[16] H. Fukuyama and W. L. Weber, "A slacks-based inefficiency measure for a two-stage system with bad outputs," *Omega*, vol. 38, no. 5, pp. 398–409, 2010.

[17] D. Holod and H. F. Lewis, "Resolving the deposit dilemma: a new DEA bank efficiency model," *Journal of Banking and Finance*, vol. 35, no. 11, pp. 2801–2810, 2011.

[18] R. Färe and S. Grosskopf, "Network DEA," *Socio-Economic Planning Sciences*, vol. 34, no. 1, pp. 35–49, 2000.

[19] T. R. Sexton and H. F. Lewis, "Two-stage DEA: an application to major league baseball," *Journal of Productivity Analysis*, vol. 19, no. 2-3, pp. 227–249, 2003.

[20] K. Tone and M. Tsutsui, "Network DEA: a slacks-based measure approach," *European Journal of Operational Research*, vol. 197, no. 1, pp. 243–252, 2009.

[21] R. Färe, S. Grosskopf, C. K. Lovell, and S. Yaisawarng, "Derivation of shadow prices for undesirable outputs: a distance function approach," *The Review of Economics and Statistics*, vol. 75, no. 2, pp. 374–380, 1993.

[22] P. Andersen and N. C. Petersen, "A procedure for ranking efficient units in data envelopment analysis," *Management Science*, vol. 39, no. 10, pp. 1261–1264, 1993.

[23] J. H. Huang, X. G. Yang, G. Cheng, and S. Y. Wang, "A comprehensive eco-efficiency model and dynamics of regional eco-efficiency in China," *Journal of Cleaner Production*, vol. 67, pp. 228–238, 2014.

[24] J. H. Zhang, "A DEA to study Chinese commercial bank efficiency and its empirical analysis from 1997 to 2001," *Journal of Finance Research*, no. 273, pp. 11–25, 2003 (Chinese).

[25] B. Wang and N. Zhu, "Efficiency and total factor productivity of Chinese listed commercial banks under the constraint of NPLs—an empirical study based on SBM directional distance function," *Journal of Finance Research*, no. 1, pp. 110–130, 2011 (Chinese).

[26] B. Wang and N. Zhu, "Total factor productivity growth in China banking industry under the constraint of non-performing loans," *Economic Research Journal*, no. 5, pp. 32–45, 2011 (Chinese).

[27] S. Akther, H. Fukuyama, and W. L. Weber, "Estimating two-stage network Slacks-based inefficiency: an application to Bangladesh banking," *Omega*, vol. 41, no. 1, pp. 88–96, 2013.

[28] C. Kao and S.-N. Hwang, "Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan," *European Journal of Operational Research*, vol. 185, no. 1, pp. 418–429, 2008.

[29] K. Tone, "Dealing with undesirable outputs in DEA: a Slacks-Based Measure (SBM) approach," *Nippon Opereshonzu, Risachi Gakkai Shunki Kenkyu Happyokai Abusutorakutoshu*, vol. 2004, pp. 44–45, 2004.

[30] K. Tone, "A slacks-based measure of efficiency in data envelopment analysis," *European Journal of Operational Research*, vol. 130, no. 3-4, pp. 498–509, 2001.

[31] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Research Logistics Quarterly*, vol. 9, no. 3-4, pp. 181–186, 1962.

[32] Y. Altunbas, S. Carbo, E. P. M. Gardener, and P. Molyneux, "Examining the relationships between capital, risk and efficiency in European banking," *European Financial Management*, vol. 13, no. 1, pp. 49–70, 2007.

[33] F. Wen and Z. Liu, "A copula-based correlation measure and its application in chinese stock market," *International Journal of Information Technology and Decision Making*, vol. 8, no. 4, pp. 787–801, 2009.

[34] A. Kasman, G. Tunc, G. Vardar, and B. Okan, "Consolidation and commercial bank net interest margins: evidence from the old and new European Union members and candidate countries," *Economic Modelling*, vol. 27, no. 3, pp. 648–655, 2010.

[35] A. N. Berger, I. Hasan, and M. Zhou, "Bank ownership and efficiency in China: what will happen in the world's largest nation?" *Journal of Banking and Finance*, vol. 33, no. 1, pp. 113–130, 2009.

[36] C. X. Huang, C. L. Peng, X. H. Chen, and F. H. Wen, "Dynamics analysis of a class of delayed economic model," *Abstract and Applied Analysis*, vol. 2013, Article ID 962738, 12 pages, 2013.

[37] F. Sufian, "Determinants of bank efficiency during unstable macroeconomic environment: empirical evidence from Malaysia," *Research in International Business and Finance*, vol. 23, no. 1, pp. 54–77, 2009.

[38] D. Hauner, "Explaining efficiency differences among large German and Austrian banks," *Applied Economics*, vol. 37, no. 9, pp. 969–980, 2005.

[39] A. Kasman and C. Yildirim, "Cost and profit efficiencies in transition banking: the case of new EU members," *Applied Economics*, vol. 38, no. 9, pp. 1079–1090, 2006.

[40] F. Pasiouras, "International evidence on the impact of regulations and supervision on banks' technical efficiency: an application of two-stage data envelopment analysis," *Review of Quantitative Finance and Accounting*, vol. 30, no. 2, pp. 187–223, 2008.

[41] M. Dietsch and A. Lozano-Vivas, "How the environment determines banking efficiency: a comparison between French and Spanish industries," *Journal of Banking and Finance*, vol. 24, no. 6, pp. 985–1004, 2000.

[42] T. S. Breusch and A. R. Pagan, "The Lagrange multiplier test and its applications to model specification in econometrics," *The Review of Economic Studies*, vol. 47, no. 1, pp. 239–253, 1980.

[43] C. X. Huang, G. Xu, X. H. Chen, and F. H. Wen, "Measuring and forecasting volatility in Chinese stock market using HAR-CJ-M model," *Abstract and Applied Analysis*, vol. 2013, Article ID 143194, 13 pages, 2013.

[44] W. H. Greene, *Econometric Analysis*, Prentice Hall, London, UK, 7th edition, 2012.

[45] A. Hoff, "Second stage DEA: comparison of approaches for modelling the DEA score," *European Journal of Operational Research*, vol. 181, no. 1, pp. 425–435, 2007.