



# Bank efficiency estimation in China: DEA-RENNA approach

Jorge Antunes<sup>1</sup> · Abdollah Hadi-Vencheh<sup>2</sup> · Ali Jamshidi<sup>2</sup> · Yong Tan<sup>3</sup> · Peter Wanke<sup>1</sup>

Accepted: 7 May 2021  
© The Author(s) 2021

## Abstract

The current study proposes a new DEA model to evaluate the efficiency of 39 Chinese commercial banks over the period 2010–2018. The paper also, in the second stage, investigates the inter-relationships between efficiency and some bank-specific variables (i.e. bank profitability, bank size, expenses management, traditional business and non-traditional business) under the Robust Endogenous Neural Network Analysis. The findings suggest that the sample of Chinese banks experiences a consistent increase in the level of bank efficiency up to 2015; the efficiency score is 0.915, after which the efficiency level declines and then experiences a slight volatility, while finally ending up with an efficiency score of 0.746 by the end of 2018. We also find that among different bank ownership types, the state-owned banks have the highest efficiency, the rural commercial banks are found to be least efficient and the foreign banks experience the strongest volatility over the examined period. The second-stage analysis shows that bank size exerts a positive influence on the development of non-traditional banking business and a proactive expense management, bank size and non-traditional businesses have a positive impact on efficiency levels, while bank profitability, traditional businesses and expenses management have negative influences on bank efficiency.

**Keywords** DEA · Robust endogenous neural network analysis · Banking · China

## 1 Introduction

According to the statistics from the world bank (data.worldbank.org), the GDP growth rate in China experiences a consistent decline from 2010 to 2019. The figure decreases from 10.636 to 6.109, however, this slow-down in economic growth does not prevent the banking sector from playing an even more important role in allocating credits to different economic sectors. We can see from the data provided by the global economy, the proportion of

---

✉ Yong Tan  
a.y.tan@hud.ac.uk

<sup>1</sup> COPPEAD Graduate Business School, Federal University of Rio de Janeiro, Rua Paschoal Lemme, 355, Rio de Janeiro, Brazil

<sup>2</sup> Department of Mathematics, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

<sup>3</sup> Department of Accounting, Finance and Economics, Huddersfield Business School, University of Huddersfield, Huddersfield, Queensgate HD1 3DH, UK

banking sector in the country's GDP keeps increasing consistently from 2010 up to 2017, and the percentage increases from 125.55% to the peak of 175.54% by the end of 2017.

This statistic is within expectations because, similar to other developing countries, China does not have a well-developed stock market, and the allocation of credits heavily relies on the banking sector. To be able to fulfill the function of the banking industry to serve the economy, solving the question of how to improve resource allocation in the banking production process will be very helpful to achieve this specific bank function. In other words, banks need to find an effective way to minimize the input investment or maximize their output production. The empirical research has been consistently making great efforts in achieving this goal through analyzing the efficiency level of various economic sectors. The widely used methods to estimate efficiency can be classified into two streams: one is the parametric stochastic frontier analysis, which has been applied to the non-banking sector (Charoenrat & Harvie, 2014; Lin, & Long, 2015; Wanke et al., 2020) as well as the banking sector (Fang et al., 2019; Tan & Floros, 2019; Tan et al., 2017). However, the parametric stochastic frontier approach suffers from several limitations: (1) it does not work very well with small samples; (2) it needs to specify a specific function form and different specifications will lead to variations of efficiency scores (Charnes et al., 1995). The second stream of the method is the non-parametric data envelopment analysis (DEA). This method has been applied widely to various sectors of the economy and various attempts have been made to make relevant developments based on the traditional model.<sup>1</sup> Compared to the parametric stochastic frontier analysis, the data envelopment analysis also may be criticized from a variety of perspectives: (1) the selection of inputs and outputs will affect the efficiency scores; (2) no statistical noise is assumed by DEA and the results from DEA are sensitive to measurement errors and extreme observations. However, the empirical studies show that using DEA to estimate efficiency can generate more robust results (Seiford & Thrall, 1990).

Data Envelopment Analysis (DEA) (Charnes et al., 1978) is a prominent technique for evaluating the relative efficiency of a set of entities called decision making units (DMUs) with homogeneous structures. Efficiency in DEA is assessed by measuring how far a DMU is from the best practices (or worst practices if an optimistic manner is followed). DEA uses an axiomatically defined production possibility set with an efficiency measure assigned to it. The efficiency measure estimates the distance between the under-evaluation DMU and the efficient frontier. A variety of efficiency measures have been developed in DEA. An efficiency measure should possess some properties to be suitable for performance assessment and ranking DMUs. *Units invariance*, *monotonicity* and *translation invariance* are some basic requirements of an efficiency measure (Cooper et al., 1999; Fare & Lovell, 1978). Units invariance means that estimating efficiency is independent of the units by which inputs and outputs are measured. Monotonicity implies that the efficiency measure strictly decreases in response to worsening inputs or outputs of an inefficient DMU. Another vital property of an efficiency measure is translation invariance. A DEA model is said to be translation invariant if translating the original input and/or output data values results in a new problem that has the same optimal solution for the envelopment form as the old one.

Efficiency in DEA is evaluated by either radial or non-radial projection. Radial measures proportionally increase outputs (output-oriented) or decrease inputs (input-oriented)

<sup>1</sup> Please see the literature review section for detail.

to project a DMU onto the efficient frontier. In terms of non-radial projection, each of the inputs and/or outputs can change independently. Non-radial measures are of great importance in DEA because non-proportional expansions or contractions are possible in many applications implying that only non-radial benchmarking and comparison are sensible. It is worth noting that the radial models ignore the effect of non-radial slacks in the efficiency (Tone & Tsutsui, 2010). The DMUs cannot be identified as completely efficient when the efficiency is one, but the slacks are not zero. Furthermore, for some cases—such as banking efficiency evaluation—some factors (inputs/outputs) are partly substitutional and may not change proportionally. Therefore, radial models have some shortcomings in evaluating banking efficiency. In this paper, a DEA method is proposed which provides a non-radial efficiency assessment in the banking sector.

The evaluation of performance including efficiency analysis will not only provide a picture regarding how the inputs and output resources are allocated in the production process, but also relevant policies can be made specifically on how to better allocate the factors of production through increasing or decreasing certain types of products. However, this policy implication is mainly limited within the production process; all the other factors outside of the production process are ignored. Therefore, a growing number of empirical studies have been engaging in a second-stage regression analysis to investigate the determinants of performance. In the banking sector specifically, efforts have been made to use various econometric techniques to evaluate the determinants of bank performance, in particular in the area of efficiency determinants examination through Tobit regression (Defung et al., 2016), fractional logic regression (Tan & Anchor, 2017), and Bootstrapped truncated regression (Wanke et al., 2016), while a step further on this has been made in the empirical literature to investigate the inter-relationships between efficiency and other variables concerned through simultaneous equation modelling under either seemingly unrelated regression analysis (Altunbas et al., 2007), three-stage least square estimator (Konara et al., 2019) or Grainger-causality test (Fiordelisi et al., 2011). However, the existing literature concerning the investigation of efficiency determinants, in particular, the inter-relationship between efficiency and other variables concerned has two issues: (1) all the empirical studies only examined the efficiency with a very limited number of variables focusing on bank risk. All the other variables which are very important in banking operating, such as bank profit, traditional and non-traditional business, expense management and bank size are ignored. The inclusion of these would not only provide more robust results, but also provide more accurate and concrete policies; (2) although the existing modelling framework addressed the issue of endogeneity to a larger or smaller extent, this issue has not been addressed thoroughly, which will influence the robustness of the results.

The current study significantly contributes to the empirical literature in efficiency analysis as well as the inter-relationship between efficiency and other variables concerned in the following two ways: (1) we propose a new non-radial DEA model to estimate efficiency using a sample of Chinese banks as the sample; our newly proposed innovative DEA model will provide the most accurate results related to the efficiency level in the Chinese banking industry; (2) unlike the existing studies, in a second-stage analysis, we propose a Robust Endogenous Neural Network Analysis to investigate the inter-relationships between a comprehensive set of bank-specific variables, including bank efficiency, bank size, bank profitability, expense management, traditional banking business and non-traditional bank business. The proposed new modelling framework will not address the endogeneity issue in a much better way than the existing literature, but the inclusion of comprehensive bank-specific variables in the modelling framework will facilitate a more accurate policy making process.

Our results show that the Chinese banking industry experiences a consistent increase in the level of efficiency up to 2015, after which a slight decline in the efficiency level is observed, with slight volatility over the last three years. The Chinese banking efficiency ends up with an efficiency score of 0.746 by the end of 2018. Also, we find that state-owned commercial banks have the highest level of efficiency, while rural commercial banks are the least efficient banking group. We notice that foreign banks experience the highest volatility in the efficiency level over the examined period. The second-stage analysis shows that bank size exerts a positive influence on the development of non-traditional banking business and proactive expense management. Bank size and non-traditional businesses have a positive impact on efficiency levels, while bank profitability, traditional businesses and expenses management have a negative influence on bank efficiency.

The current paper is structured as below: Sect. 2 reviews the empirical studies of DEA applications in various sectors of the economy as well as the DEA applications in the banking sector in general and Chinese banking sector specifically; Section 3 presents the non-radial DEA model as well as the Robust Endogenous Neural Network Analysis (RENNA). Section 4 provides the results and discussion, while the conclusion is given in Sect. 5.

## 2 Literature review

The evaluation of performance in various economic sectors is very important for policy-making purposes and, consequently, it attracts attention from various parties, including the government officials and regulatory authorities as well as academic researchers. The performance assessment in the empirical literature can be generally divided into two groups, with the first one focusing on the use of accounting ratios and the second stream concentrating on the use of parametric or non-parametric operational research methods. In particular, the use of non-parametric data envelopment analysis has been gaining popularity in estimating efficiency in various economic sectors, including the high technology industry (An et al., 2020); the hotel sector (Assaf et al., 2010; Yin et al., 2020; Yu & Chen, 2020); insurance industry (Barros et al., 2010; Cummins et al., 2010; Eling & Jia, 2019; Eling & Luhnen, 2010); telecommunications industry (Bayraktar et al., 2012; Sahoo & Sahoo, 2020); transportation industry (Chang et al., 2013; Merkert & Hensher, 2011; Stefaniec et al., 2020; Wu & Goh, 2010; Yu, 2010); iron and steel industry (He et al., 2013; Lee et al., 2019); the tourism industry (Pestana et al., 2011; Song & Li, 2019); the agricultural sector (Picazo-Tadeo et al., 2011; Wang & Feng, 2020); the hospital sector (Rouyendegh et al., 2019); petroleum industry (Sueyoshi & Goto, 2012; Wang et al., 2019); the education sector (Aparicio et al., 2019; Thanassoulis et al., 2011); and the energy industry (Wang et al., 2013; Wanke et al., 2020; Zhou et al., 2012, 2013). Also, the DEA models have been applied to analyse the efficiency level in various business-related operational issues, including supply chain management (Azadi et al., 2015; Malesios et al., 2020; Mirhedayatian et al., 2014); sustainability and environmental performance (Allevi et al., 2019; Chen & Delmas, 2011; Lee & Saen, 2012); innovation system (Guan & Chen, 2012; Kalapouti et al., 2020; Zhong et al., 2011); and the manufacturing industry (Margaritis & Psillaki, 2010; Walheer & He, 2020). In particular, the study of Psillaki et al. (2010) evaluates the efficiency level of different industries in France, including textile and wood, as well as research and development.

The evaluation of efficiency in the banking sector has been evaluated by empirical research through the traditional DEA model and various developments. The traditional

DEA model has been applied by Chortareas et al. (2012) to evaluate the efficiency in the European banking sector over the period 2000–2008. The study also engages in a second stage regression analysis to further investigate the impact of bank supervision and regulation on bank efficiency under the truncated regressions, generalized linear models as well as a fractional logic estimator. Slightly different from Chortareas et al. (2012), the efficiency level of the European banking industry is also assessed by Chortareas et al. (2013) under the traditional DEA model, while they further examine the impact of financial freedom on bank efficiency under the Simar and Wilson (2007) bootstrapped truncated regression analysis. Studies also attempt to evaluate bank efficiency in emerging markets, not solely in the European banking industry. Konara et al. (2019) using a sample of banks from eight emerging market economies to investigate the relationship between FDI and bank efficiency over the period 2009–2013. A traditional DEA model is used to derive various efficiency estimates, including technical, pure technical, scale, cost and revenue efficiencies, while the relationship test is facilitated by the three-stage least square estimator. Traditional DEA models are also used by Staub et al. (2010) to evaluate the efficiency level (i.e. cost, technical and allocative efficiencies) in the Brazilian banking industry. In the second-stage analysis, three different panel data specifications are used to analyze the determinants of bank efficiency, including a dynamic model, a Tobit model and an autoregressive model. The empirical studies have made efforts to build on the traditional DEA models and relevant advancements in the modelling framework have been proposed, such as the network DEA model (Kumar & Gulati, 2010; Liu et al., 2018, 2020; Paradi et al., 2011).

There are some empirical research studies analyzing the efficiency level in the Chinese banking industry specifically. Using a sample of Chinese commercial banks over the period 2003–2009, Tan and Floros (2013) evaluate the efficiency level under the traditional DEA models. Various performance indicators, including technical, pure technical and scale efficiencies as well as Malmquist productivity index, are examined and the authors further test the inter-relationship between risk, efficiency and capitalization using a three-stage least square estimator.

Taking a sample of Chinese commercial banks as the sample over the period 2002–2009, Chang et al. (2012) proposed an input slack-based productivity index, which possesses the advantages of being able to disaggregate the total productivity index into the productivity change of each input. The findings show that the productivity growth of Chinese banks is dominated by capital productivity, which is mainly derived from the technical progress in capital use.

Nine super-efficiency DEA models have been proposed and used by Avkiran (2011) to explore the association between bank efficiency and performance ratios. The results show that there is a weak relationship between the two. However, profit margin and return on assets are found to have a significant relationship with the efficiency estimates, with the former having an even stronger effect.

A new three-stage network DEA model is proposed by Fukuyama and Tan (2020) to evaluate three different types of efficiencies, including input efficiency, stability efficiency and output efficiency. The study considers the role of market power in loans and deposits in the production process and treats loan loss provisions as the good intermediate product in a network model. Besides, Tobit regression is used in the second stage to examine the impact of non-performing loans from three different geographical areas (Eastern, Central and Western) on different types of efficiencies, and the results are also cross-checked by fractional logit regression and Bootstrapped truncated regression. The findings show that a higher level of financial risk from the Central area has a significant and positive impact on input inefficiency, while the Western area has an opposite effect.

A dynamic two-stage DEA model considering the carryover characteristics of the non-performing loans is proposed by Zha et al. (2016) to analyze the efficiency level for a sample of Chinese banks. The proposed method benefits itself from the ability to identify the source of inefficiencies more effectively. The findings show that both the productivity stage and profitability stage need to improve the allocation of resources to further improve the overall level of efficiency in the Chinese banking industry. City commercial banks are found to have higher levels of pure technical efficiency than joint-stock commercial banks. Although the state-owned commercial banks have the highest level of pure technical efficiency, their overall efficiency level is lower than that of city commercial banks.

Instead of considering the carryover characteristics of the non-performing loans, Zhou et al. (2019) treat the non-performing loans as the undesirable output to develop a multi-period, multi-stage DEA model in which not only is a different variable considered as carryover product, but the share inputs are also introduced in the model. Using a sample of Chinese banks over the period 2014–2016 as the model application object, the findings show that there is a level of disparity in the level of bank efficiency and all the banks in the sample are found to be inefficient. It is interesting to note that for different types of banks, the source of inefficiency is derived from different stages.

The network DEA models and the traditional DEA models mainly use the secondary databases or the annual financial statement of the banks to collect data of the variables in the production processes. Matthews (2013) made the first attempt to use qualitative data in the evaluation of performance in a network DEA model in the Chinese bank efficiency evaluation. More specifically, questionnaires are given to the Chinese bank managers to collect information about risk management and risk practice, through which two metrics are generated. The metric related to risk organization is used as internal intermediate input and the one related to risk practice is treated as an external input. The findings suggest that there is no significant relationship between the risk metrics and the bank performance measure; however, including these measures in the model would allow bank performance to be explained in a better way than would excluding them.

One important and interesting area in the efficiency analysis using the DEA model is to choose the appropriate input and output variables in the production process. This issue has been addressed by Luo et al. (2012) by proposing a method based on the concept of cash value-added. It is defined by the method that if a variable is tested to be positively related to the bank's cash flow, it will be treated as output, otherwise, it will be regarded as input. This method has been applied to a sample of Chinese banks.

### 3 Methodology and data

#### 3.1 The multi-period aggregative efficiency method

Most of the DEA studies have dealt primarily with cross-sectional data and measured relative efficiencies in a single period, usually 1 year. Exceptions are window analysis, dynamic models, and Malmquist-type indices of productivity. Looking beyond the difference between their model details, we recognize that their common goal is to account for the changing patterns of efficiency performances over several periods. However, these approaches, while vital and practically useful, do not consider an aggregated measure of efficiency for multiple period production systems. Consider a simple example where three DMUs produce different amounts of two outputs and consume the same unit amounts of

single input in two periods,  $t = 1, 2$ . Using an ordinary DEA, we can obtain the efficiency ratings of each DMU for individual periods. Suppose that the first DMU is efficient for both periods, the second DMU is efficient for the period  $t = 1$  but inefficient for the period  $t = 2$ , and the last DMU is inefficient for both periods. These efficiency ratings obtained for individual periods are needed for efficiency valuations but reflect partial performances of multi-period production units.

This brings into play an important question as to how we can achieve a multi-period aggregative efficiency, shortly referred to as MAE. To carry out a panel data analysis, we assume underlying production technologies in which all of a period's input is expected to go into producing the output for the same period. We do not consider a special production system where current input amounts might be used to produce future period outputs. Specifically, the MAE is measured in such a manner that a DMU's performance in a particular period is compared with the performance of all DMUs in the same period. The same way is taken in DEA to obtain the efficiency ratings in the above example, but each of these ratings signifies a single period efficiency.

### 3.2 The proposed DEA model

Suppose we have  $n$  DMUs (banks) and there are  $T$  periods,  $t = 1, 2, \dots, T$ , and in each period,  $DMU_j, j = 1, 2, \dots, n$ , consume  $m$  inputs,  $x_{ij}^t, i = 1, 2, \dots, m$ , to produce  $s$  outputs,  $y_{rj}^t, r = 1, 2, \dots, s$ . In order to compute the multi-period aggregative efficiency, shortly referred to as MAE in the context of time serial data, we propose the following DEA model, related to  $DMU_p, p = 1, 2, \dots, n$ :

$$\begin{aligned}
 E_p = \max \quad & \frac{T - \frac{1}{mT} \sum_{t=1}^T \sum_{i=1}^m \frac{s_i^{t-}}{x_{ip}^t}}{T + \frac{1}{sT} \sum_{t=1}^T \sum_{r=1}^s \frac{s_r^{t+}}{y_{rp}^t}} \\
 \text{s.t.} \quad & \sum_{j=1}^n \lambda_j^t x_{ij}^t + s_i^{t-} = x_{ip}^t, \forall i, \forall t \\
 & \sum_{j=1}^n \lambda_j^t y_{rj}^t - s_r^{t+} = y_{rp}^t, \quad \forall r, \forall t \\
 & \lambda_j^t \geq 0, \quad \forall j, \forall t \\
 & s_i^{t-} \geq 0, \quad \forall i, \forall t \\
 & s_r^{t+} \geq 0, \quad \forall r, \forall t
 \end{aligned} \tag{1}$$

**Definition 1** Let  $E_p^*$  be the optimal value of objective function of the above model. Then,  $DMU_p$  is called efficient if  $E_p^* = 1$ .

This condition is equivalent to all slack variables  $s_i^{t-}$  and  $s_r^{t+}$  equal to zero for all periods, and it means no input excesses and no output shortfalls in any optimal solution (Esmailzadeh & Hadi-Vencheh, 2013). The proposed model is a fractional programming model and can be transformed into a linear programming (LP) model by Charnes and Cooper (1962) transformation. By this change, we have the following LP model:



### 3.3 Linear model

$$\begin{aligned}
 E_p &= \max MT - \frac{1}{mT} \sum_{i=1}^T \sum_{i=1}^m \frac{s_i^{t-}}{x_{ip}^t} \\
 \text{s.t. } &\sum_{j=1}^n \lambda_j^t x_{ij}^t + s_i^{t-} = M x_{ip}^t, \quad \forall i, \forall t \\
 &\sum_{j=1}^n \lambda_j^t y_{rj}^t - s_r^{t+} = M y_{rp}^t, \quad \forall r, \forall t \\
 &MT + \frac{1}{sT} \sum_{i=1}^T \sum_{r=1}^s \frac{s_r^{t+}}{y_{rp}^t} = 1 \\
 &\lambda_j^t \geq 0, \quad \forall j, \forall t \\
 &s_i^{t-} \geq 0, \quad \forall i, \forall t \\
 &s_r^{t+} \geq 0, \quad \forall r, \forall t
 \end{aligned} \tag{2}$$

where  $M = \frac{1}{T + \frac{1}{sT} \sum_{i=1}^T \sum_{r=1}^s \frac{s_r^{t+}}{y_{rp}^t}}$ ,  $\lambda_j^t = M \lambda_j^t$ ,  $s_i^{t-} = M s_i^{t-}$  and  $s_r^{t+} = M s_r^{t+}$ . In this case the MAE of  $DMU_p$  is  $E_p^* = MT - \frac{1}{mT} \sum_{i=1}^T \sum_{i=1}^m \frac{s_i^{t-}}{x_{ip}^t}$  and the efficiency of  $DMU_p$  at time period  $t, t = 1, 2, \dots, T$ , equals to  $E_p^* = \frac{M - \frac{1}{m} \sum_{i=1}^m \frac{s_i^{t-}}{x_{ip}^t}}{M + \frac{1}{s} \sum_{r=1}^s \frac{s_r^{t+}}{y_{rp}^t}}$ .

### 3.4 Robust endogenous neural network analysis (RENNA)

Artificial Neural Networks (ANNs) are computational algorithms that are based on the human thinking paradigm. ANNs are formed of processing units (neurons) that are weight connected. These connections motivate the estimation of non-linear models by using a training data set. Athanassopoulos and Curram (1996) give the first literature insight on combining ANNs and DEA for predicting efficiency levels. Other ANN applications in DEA can be found in existing literature (Wu et al., 2006). Examples can be found in Santin and Delgado (2004), Emrouznejad and Shale (2009), Misiunas et al. (2016), Shokrollahpour et al. (2016), Olanrewaju et al. (2012), Bashiri et al. (2013) and Modhej et al. (2017).

In this research, a specific focus is placed on the MLP network that has been extensively studied in forecasting applications (Mubiru & Banda, 2008). Within the ambit of an MLP, neurons are pooled in layers and only forward connections are allowed. These features provide a robust architecture capable of learning upon any kind of continuous nonlinear mapping. A typical MLP is represented in Fig. 1.

MLP constituents encompass neurons, weights, and transfer functions. An input  $x_j$  is transmitted via connections that multiply their respective strength by  $w_{ij}$  weights, yielding the product  $x_j w_{ij}$  used in the transfer function  $f$  to compute a specific output  $y_i$  given as  $y_i = f\left(\sum_{j=1}^n x_j w_{ij}\right)$ .  $i$  is the neuron index in the hidden layer and  $j$  is the input index to the MLP. The modification of the weights of each connection observing some orderly fashion is known as training. During the training, an input is assigned to the network along with the desired output and the weights are adjusted so that the MLP catches-up with the desired output value.



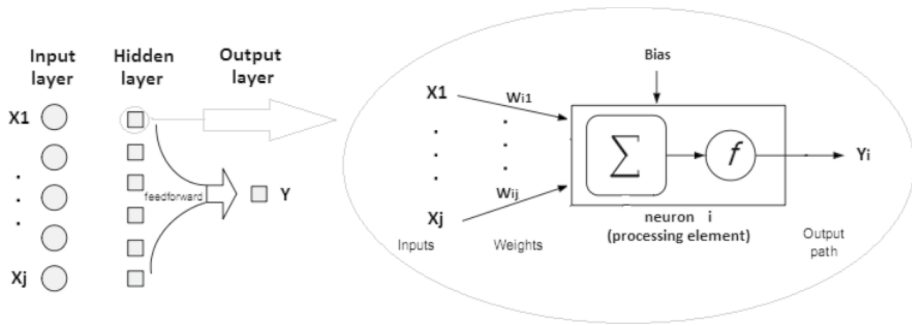


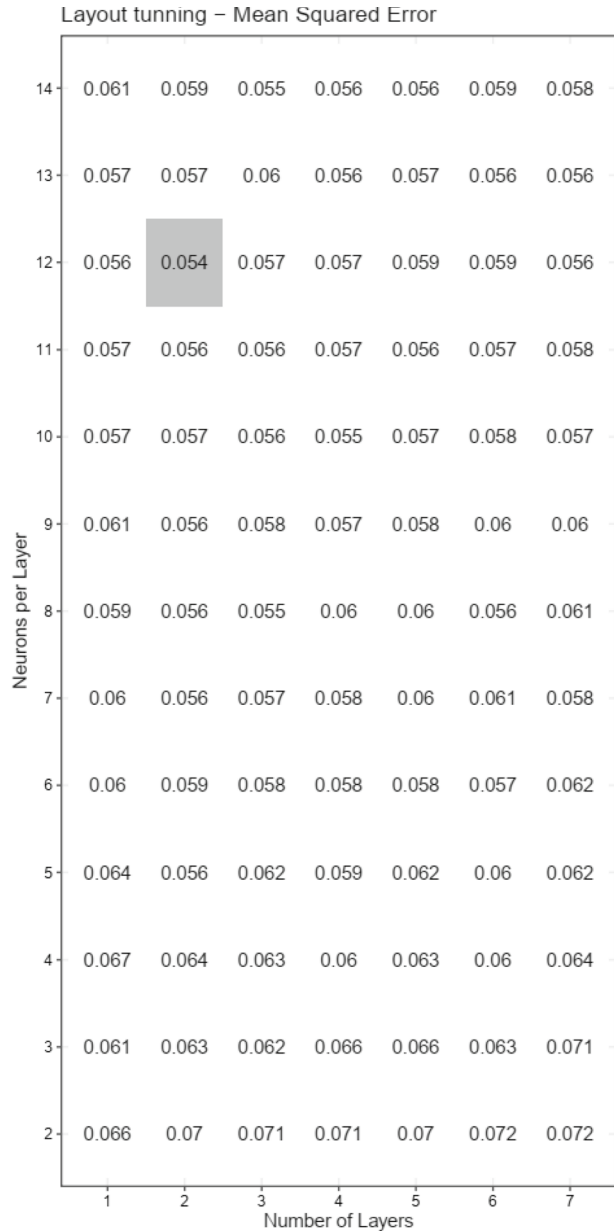
Fig. 1 MLP framework

As regards RENNA, the robustness analysis is structured in two steps. In the first step, MLP is employed on unveiling endogeneity among banking efficiency, size, profitability, expense management, and business type (traditional vs. non-traditional). Hence, this paper departs from previous research in the banking sector by using the MLP network structure to explore endogeneity between these variables in terms of the following linear models:

- Model 1: Efficiency  $\sim f(\text{Size, Profitability, Expense Management, Traditional Business, Non-traditional Business})$ .
- Model 2: Size  $\sim f(\text{Efficiency, Profitability, Expense Management, Traditional Business, Non-traditional Business})$ .
- Model 3: Profitability  $\sim f(\text{Size, Efficiency, Expense Management, Traditional Business, Non-traditional Business})$ .
- Model 4: Expense Management  $\sim f(\text{Size, Profitability, Efficiency, Traditional Business, Non-traditional Business})$ .
- Model 5: Traditional Business  $\sim f(\text{Size, Profitability, Expense Management, Efficiency, Non-traditional Business})$ .
- Model 6: Non-traditional Business  $\sim f(\text{Size, Profitability, Expense Management, Traditional Business, Efficiency})$ .

In fact, neural networks were employed here to the detriment of other approaches, such as linear regressions, due to its ability to unveil non-linear relationships between dependent variables and their predictors. This feature is particularly useful as long as the ordinary covariance matrix embeds in itself a partial measurement of all possible associative functions among variables, solely relying on linear correlation. As long as non-linear associative measures between variables are not captured by ordinary covariance matrices, neural networks were chosen as the cornerstone for analyzing the distributional profile of the residuals of each one of these six regressions, in the search for evidence that could support endogeneity claims based on the hidden non-linear association between efficiency levels and their predictors related to the Chinese banking industry.

Hence, the underlying idea of running these six models is to explore their respective residuals, obtained by considering each variable as the dependent one. In the first step of the analysis, once the joint variances of residuals are minimized by finding optimal weights for each model, it would be possible to affirm that the resulting weighted combination of predictive models yielded minimal feedback effects. Hence, on the second step of the

**Fig. 2** MLP optimal architecture

analysis, it becomes possible to address issues of cause-and-effect relationships among the dependent variables of each model, by identifying the respective conditional distributions of residuals with higher weights in a second optimization round.

The results for the optimal MLP architecture are depicted in Fig. 2, where an exhaustive combination of neurons and layers were tested for the MSE error combination of models (1)–(6). Two layers and 12 neurons in each yielded a minimal combined MSE of 0.054. In other words, a layer can be regarded as a collection of neurons that, while performing the

same computational steps, assures non-linear relationships among inputs and outputs to be properly captured. The fundamental computational steps performed by neurons are to (i) calculate the weighted sum of inputs and weights, (ii) map the forecasting bias and (iii) execute an activation function among inputs and outputs.

The relative importance of Models 1–6 in explaining the feedback process between contextual variables and efficiency levels in the Chinese banking industry, together with the endogenous nature of these variables, were explored, respectively, by the variances of each model and the covariances between models. Variances and covariances of the residuals ( $R_i$ ) of these eight models are simultaneously minimized by a non-linear stochastic optimization problem, as presented in Model (3), where  $w_i$  stands for the weights – which range from 0 to 1—assigned, respectively, to the residual vectors of each one of the six models previously described. The values of  $w$  are optimized so that the variance ( $Var$ ) and covariance ( $Covar$ ) of the pooled residuals are minimal. Model (3) was solved through differential evolution (DE). DE is a research stream of genetic algorithms, also emulating natural selection and evolution. Readers should refer to Ardia et al (2011) and Mullen et al. (2011) for further details. Results are discussed in the next section.

$$\begin{aligned} \min & \left[ Var \left( \sum_{i=1}^6 w_i * R_i \right) + \left( 2 * \sum_{i,j=1}^6 Covar(w_i * w_j * R_i * R_j), i \neq j, j < i \right) \right] \\ \text{s.t.} & \sum_{i=1}^6 w_i = 1 \\ & 0 \leq w_i \leq 1 \forall i \end{aligned} \quad (3)$$

Residuals of the MLP models were bootstrapped 100 times allowing the collection of a distributional profile of  $w$  for the most accurate prediction of efficiency scores and contextual variables.

Subsequently, in the second step, a full combinatorial set of conditional distributions of residuals ( $CR_k$ ) was computed. The previous 100 bootstrapped replications for the individual residuals ( $R_i$ ) of each model served as cornerstones for this computation, where  $C R_k \sim f(R_i/R_j)$  for all  $i$  and  $j$ ,  $i \neq j$ , and  $K = i * j - i = 6 * 6 - 6 = 30$ . Similarly, DE was employed for a minimal *Var-Covar* model to diagnose whether conditional distributions of each residual pairs presented significant differences in terms of directions (in this case, weights assigned to  $f(R_i/R_j)$  variance would differ from those assigned to  $f(R_j/R_i)$  variance). Besides, higher-order joint variations of conditional residuals were explored by means of joint covariance weights so that additional hidden effects could be explored, as depicted in the model (4).

$$\begin{aligned} \min & \left[ Var \left( \sum_{k=1}^{30} w_k * CR_k \right) + \left( 2 * \sum_{k,l=1}^{30} Covar(w_k * w_l * CR_k * CR_l), k \neq l, k < l \right) \right] \\ \text{s.t.} & \sum_{k=1}^{30} w_k = 1 \\ & 0 \leq w_k \leq 1 \forall k \end{aligned} \quad (4)$$

In terms of the data we use in the current paper, we collect a sample of 39 Chinese commercial banks over the period 2010–2018. The data is collected from the Fitchconnect database. The sample of Chinese banks is further divided into five groups according to the

**Table 1** Descriptive statistics of the variables used in the current study

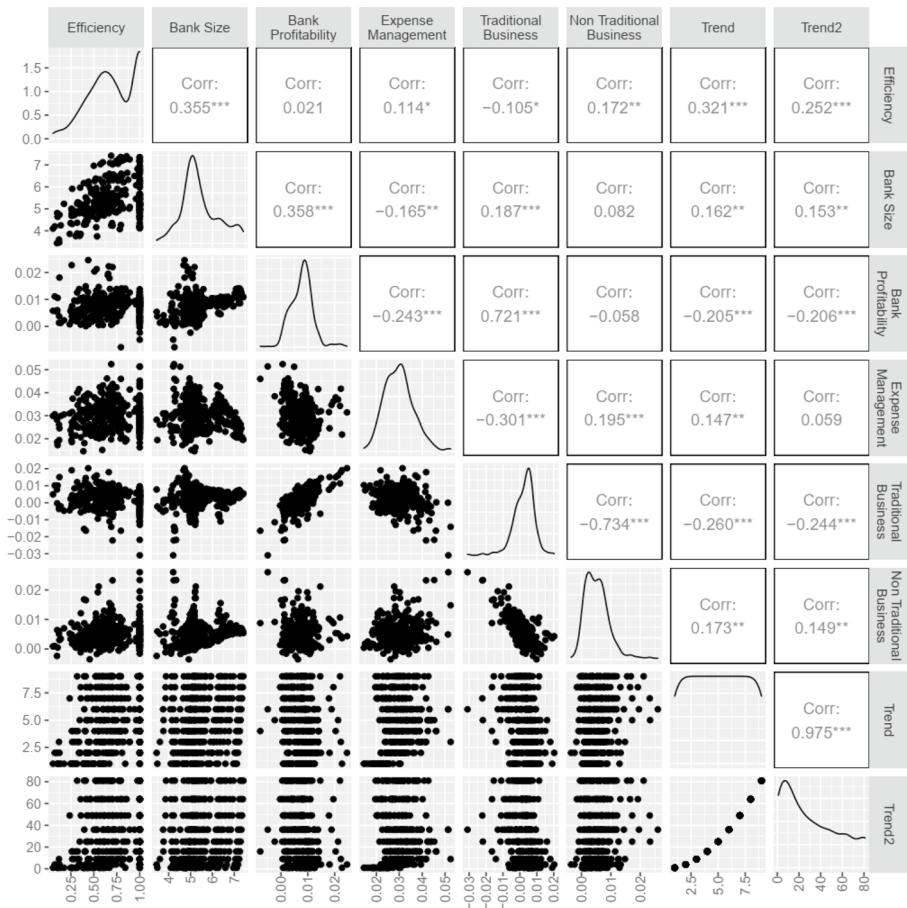
Variables	Observations	Mean	SD	Minimum	Maximum
<i>Inputs</i>					
Total deposits	351	1,647,518	3,972,533	1475.1	2.23e+07
Fixed assets	351	11,557.2	29,720.5	2.168	145,421
Personnel expenses	351	8951.617	21,951.02	24.9	111,354
<i>Outputs</i>					
Total securities	351	464,301.4	1,097,267	57.9	6,348,656
Gross loans	351	1,003,506	2,546,259	660.1	1.46e+07
<i>Contextual variables</i>					
Bank size	351	5.43	0.84	3.45	7.42
Bank profitability	351	0.0077	0.0042	-0.0077	0.025
Expense management	351	0.03	0.006	0.015	0.052
Traditional bank business	351	0.0024	0.0061	-0.031	0.02
Non-traditional bank business	351	0.0053	0.0043	-0.0035	0.026

All the inputs and outputs variables are in the unit of million RMB. The contextual variables are in ratios. Bank size is measured by the natural logarithm of total assets; bank profitability is measured by the ratio of net income to total assets; expense management is measured by the ratio of (interest expenses + non-interest expenses)/total assets; traditional bank business is measured by the ratio of interest income to total assets; non-traditional bank business is measured by the ratio of non-interest operating income to total assets.

ownership types. They are 15 city commercial banks, 2 rural commercial banks, 14 foreign banks, 4 state-owned commercial banks and 4 joint-stock commercial banks. Regarding the use of inputs and outputs in our first-stage efficiency analysis, three inputs are considered, including fixed assets, total deposits, and personnel expenses, while total securities and gross loans are treated as outputs. In the second-stage Robust Endogenous Neural Network Analysis, we include a series of bank-specific variables, including bank size (natural logarithm of total assets), bank profitability (return on assets), expense management (the sum of total interest expenses and non-interest expenses over total assets); traditional bank business (the ratio of interest income to total assets); and non-traditional bank business (the ratio of non-interest operating income to total assets). Table 1 shows the descriptive statistics of the variables in the current study.

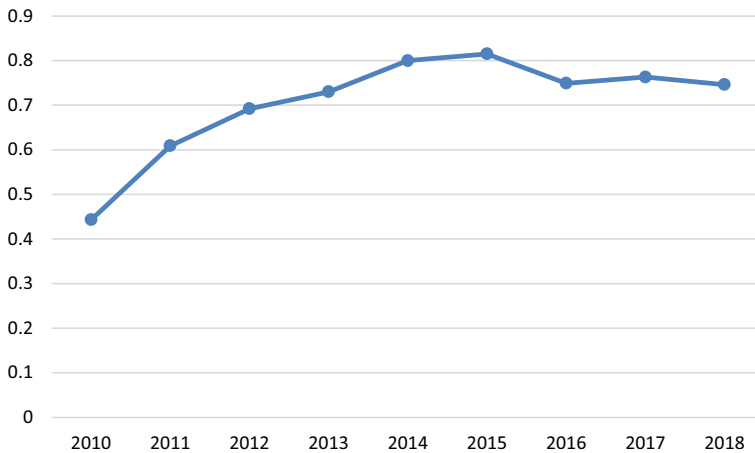
## 4 Analysis and discussion of results

Efficiency score results are depicted in Fig. 3, altogether with their joint-variation profile with contextual variables related to size, profitability, expense management, traditional business, and non-traditional business. The choice of our contextual variables is based on the following considerations. Bank size is supposed to influence the efficiency level because large banks benefit from the cost advantage gained from economies of scale and economies of scope (Fu and Sio, 2011). Girardone et al. (2004) find that profitable banks have a higher ability to control all aspects of bank costs, therefore we include profitability as one of the contextual variables. Expense management is mainly related to the management of interest expenses and non-interest expenses in the banking operation. Good control of these costs is supposed to have a positive influence on the efficiency level. It seems that the empirical literature has not focused on the examination regarding the impact of



**Fig. 3** Correlogram and joint variation profile between efficiency scores and contextual variables in Chinese banks

traditional business on bank efficiency, but various efforts have been made to evaluate the impact of income diversification on bank efficiency (Alhassan, 2015; Nguyen, 2018). We argue that higher proportions of traditional business indicate a lower degree of diversification and further influence the efficiency level. The empirical literature has attempted to investigate the influence of non-traditional activity on bank efficiency by incorporating this item in the first stage of efficiency analysis (Gulati & Kumar, 2011; Lozano-Vivas and Pasiouras, 2010). We include this variable as one of the contextual variables to see whether we will get different results by incorporating it in a second-stage analysis. One can observe in upper diagonal significant correlation relationships, marked with “\*”, for most cases, which may possibly be an indicator of feedback processes among efficiency and different contextual variables, thus denoting possible endogeneity. The main diagonal shows the empirical distribution profile for Efficiency and Contextual Variables individually as well as the lower diagonal shows scatterplots relationship between variables in the given column and row.

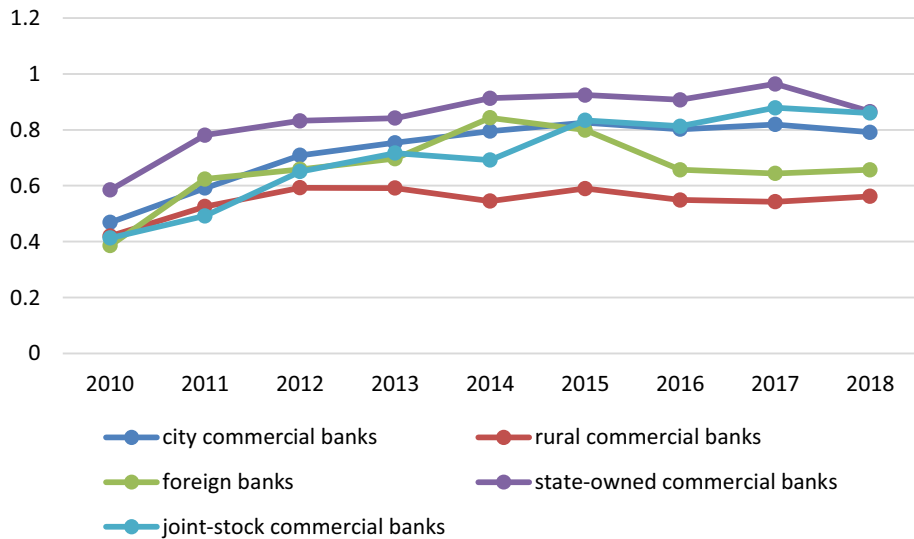


**Fig. 4** Efficiency level in the Chinese banking industry

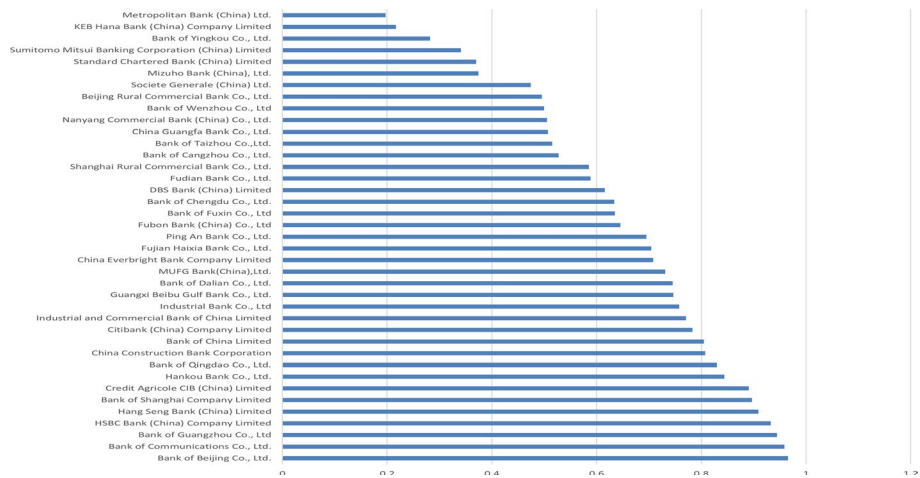
In order to further look at the efficiency level of the banks in the sample in a clearer manner, we generate the figure below (Fig. 4), which shows the efficiency level in the Chinese banking industry over the period 2010 to 2018. The findings show that the Chinese banking industry experiences stronger volatility in the efficiency scores over the period, while in general there is a consistently increasing trend for the efficiency level up to 2015, after which we can see that the efficiency level declines in 2016, and although a slight increase in the efficiency level in 2017 is observed compared to 2016, the efficiency score experiences another drop in 2018, compared to the previous year, reaching 0.746. Our results are in contrast with results from Fukuyama and Tan (2020). The difference is mainly attributed to the fact that not only do we use a different method in the efficiency analysis, but also the inputs and outputs used in the method are different.

Not only do we focus on the general efficiency level of the Chinese banking industry as a whole, but we also look at the efficiency level of different ownership types of Chinese banks. The figure below (Fig. 5) shows the efficiency scores of different ownership types of Chinese banks over the examined period. We can see that the state-owned commercial banks have the highest level of efficiency, and in general, the rural commercial banks are the least efficient banking group. There is no clear comparison about the efficiency level among joint-stock commercial banks, city commercial banks and foreign banks, while it is observed that the foreign banks experience the strongest volatility in the efficiency level compared to all the other banking groups in the sample. Zhou et al. (2019), over the same period (2014–2016), report that city commercial banks have a higher efficiency than state-owned banks, while joint-stock banks have the lowest efficiency, the difference in our results can be mainly explained by the number of banks included in the sample, the method adopted as well as the period of investigation.

It is even more interesting to look at the efficiency level of individual banks in the sample. The figure below (Fig. 6) shows the efficiency level of all the 39 banks in the sample over the examined period. Combining the number of the bank and the specific bank name in the “Appendix”, we can see that the top four best performers in the sample are Bank of Beijing, Bank of Communication, Bank of Guangzhou and HSBC (China), while the worst four bank performers include Metropolitan Bank (China), KEB Hana Bank (China),



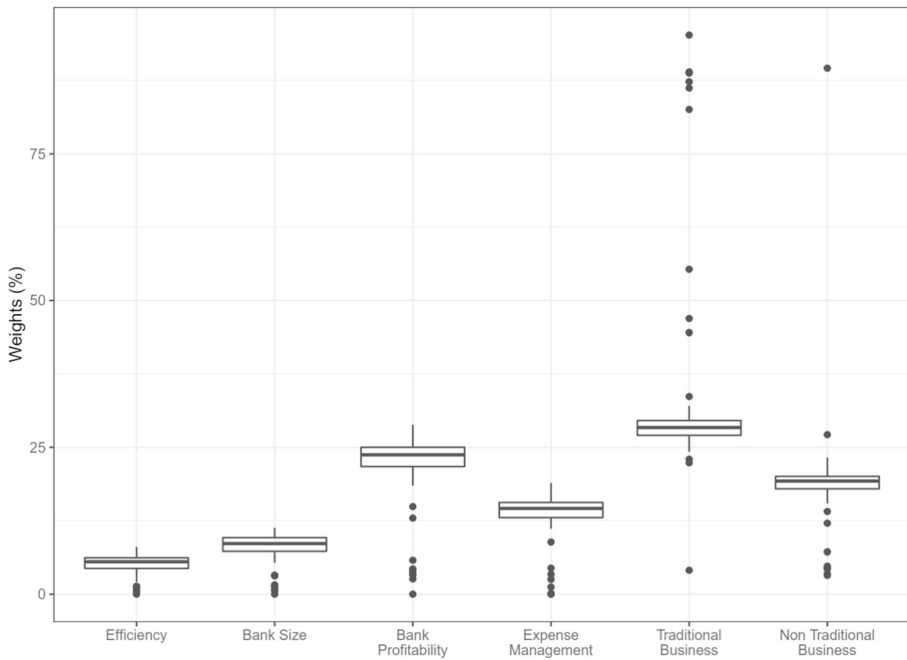
**Fig. 5** Efficiency level in the Chinese banking industry with different ownership types



**Fig. 6** Efficiency level of individual banks in the sample

Bank of Yingkou and Sumitomo Mitsui Banking Corporation (China). We can see that both of these two extreme groups of banks include two ownership types (city commercial banks and foreign banks), Bank of communication is a state-owned commercial bank. This finding generates feasible policies for the banks, especially the ones with very low levels of efficiency to improve their performance by communicating and learning from the ones with the very best performance within the same ownership type. For other state-owned commercial banks, the bank of communication can be a good model to learn in terms of its resource allocation.





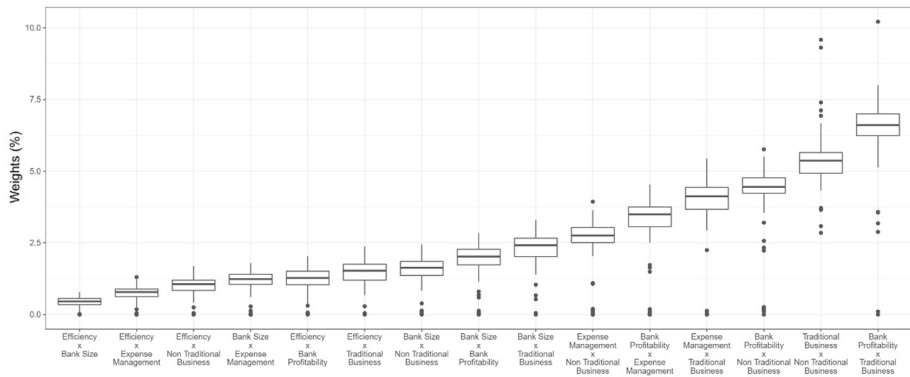
**Fig. 7** Relative importance of models (1)–(6)

#### 4.1 Marginal impact of efficiency scores

Notwithstanding, the relative importance of models (1)–(6) for achieving minimal residual variance is quite heterogeneous (cf. Fig. 7). In fact, business type and bank profitability account for more than 50% of total weight, while efficiency and bank size account for less than 20%. These results suggest that the eventual feedback process (endogeneity) among banking efficiency, size, and expense management is obfuscated by more profitable business types. Results for the relative importance of the interaction pairs in explaining overall residual variance are presented in Fig. 4 and also confirm this issue. Putting into perspective the results obtained for 100 bootstrap replications, while the efficiency scores present marginal relevance in comparison to other variables related to business models and banking profitability, the feedback between size, profitability, and efficiency is also negligible, which should denote market regulation. The economies of scale and scope achieved by large banks and the resulting improvement in the level of bank profitability are offset by the level of intense competition in the banking industry, therefore their impact on bank efficiency is not obvious.

#### 4.2 Low endogeneity levels

The joint feedback effect of all models pairs on overall residual variance is lower than 15% (cf. Fig. 8), pretty low when compared to the maximal possible endogeneity effect, achievable when all six models would account individually for 16.67% of the total residual



**Fig. 8** Endogeneity weights for pairs of models (1)–(6)

**Table 2** Median weights for conditional residual distribution of [ROW] given that [COLUMN]

Models	Efficiency	Bank size	Bank profitability	Expense management	Traditional business	Non traditional business
Efficiency		0.0617	0.0538	0.0576	0.0604	0.0604
Bank Size	0.0153		0.0217	0.0227	0.0207	0.0238
Bank profitability	0.0224	0.0293		0.0177	0.0219	0.0284
Expense management	0.0181	0.0342	0.0276		0.0226	0.0236
Traditional business	0.0203	0.0217	0.0238	0.0305		0.0268
Non traditional business	0.0169	0.0359	0.0264	0.0285	0.0210	

variation. Under the case of this equal weight, the maximal joint effect of  $83.36\% = 2 \times 41.68\% = 2 \times 16.67\% \times 16.67\% \times 15$ , where 15 is the number of combinations taken two by two obtained from models (1)–(6).

### 4.3 Cause–effect relationships

In summary, efficiency in Chinese banks appears to be weakly endogenous with size and profitability due to imperfect, oligopolistic, and strongly regulated industry. Readers should observe that the median expected weight for each conditional distribution would be 0.033 (1/30), considering a balanced bi-directional relationship among variables. According to Table 2, efficiency is a consequence of bank size, bank profitability, and its business model; rather than acting as causation on such characteristics, as long as conditional residual distribution weights are significantly higher than 0.033 at  $p < 0.05$ . Bank size, on the other hand, indirectly impacts efficiency levels, by favoring expense management and non-traditional business models, thus signifying a scale barrier. The final cause-effect framework of Chinese banks is given in Fig. 9. Our results related to the impact of bank size on the efficiency level is in line with Tan and Floros (2013).

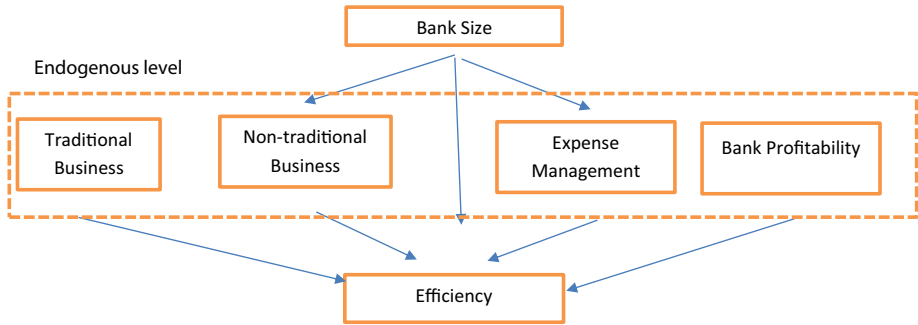


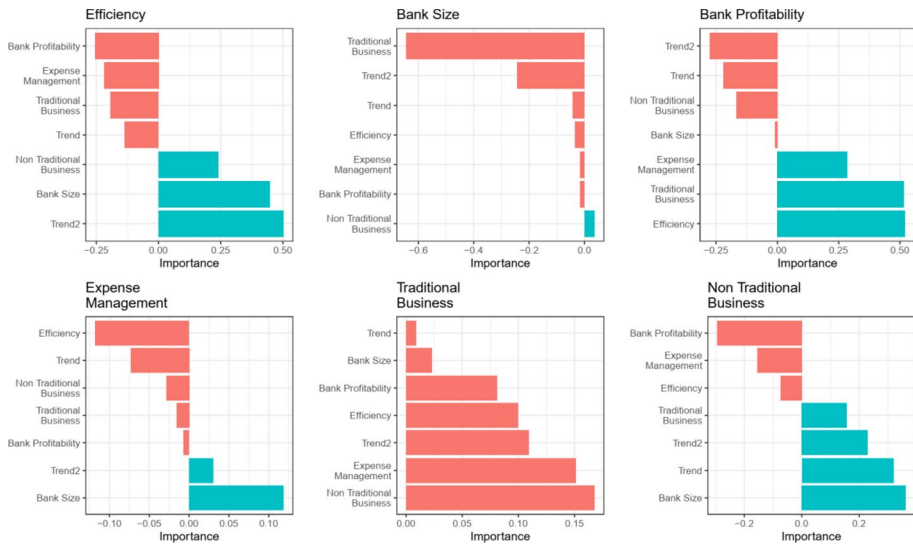
Fig. 9 Cause-effect and endogeneity framework

#### 4.4 Endogenous relationships

RENNA, however, allowed the confirmation of endogenous processes or feedbacks among business type, expense management, and bank profitability, thus suggesting that market segmentation and product differentiation drivers are mutually reinforced over time, once Chinese banks are big enough to start alternative banking venues. It is important to note that, while non-traditional business appears to present negative feedback with banking profitability at first sight, on the other hand, the non-traditional business can be regarded as a positive efficiency driver as long as it gears a more than proportional increase of outputs relative to inputs in Chinese banking operations. This positive impact is in contrast with the findings of Lozano-Vivas and Pasiouras (2010). We explain the difference by the fact that we focused on the Chinese banking industry and that we use a second-stage analysis for the investigation, while Lozano-Vivas and Pasiouras (2010) incorporate the non-traditional activities in the first-stage efficiency analysis. These new business models, however, are still less profitable than traditional ones because, compared to the traditional interest generating activities, the fee-based activities are more likely to generate revenues with higher levels of volatilities due to the fact that there is a weaker bank-customer relationship in the non-traditional banking businesses compared to the traditional ones, and customers are easier to switch to another financial provider. Besides, unlike the traditional loan businesses, the increase in the cost of additional loans will be only the interest expenses. In comparison, fee-based businesses will significantly increase the banks' fixed cost (Li & Zhang, 2013). Due to the risk and cost considerations, the banks would be more conservative in engaging in this business category and that explains the low profitability from the non-traditional businesses.\*\*

#### 4.5 Relationship signs

Olden's sensitivity analysis results are presented in Fig. 10 and reveal the signs of the relationships between variables, and their relative importance in terms of joint variation. Bank size exerts a positive impact on the development of non-traditional business and more proactive expense management. While the positive impacts of bank size and non-traditional business are transferred to efficiency levels, it is worth noting the countervailing negative forces played by expense management, bank profitability, and traditional business



**Fig. 10** Results for Olden's sensitivity analysis in terms of signs and relative importance for models (1)–(6)

on efficiency levels. This implies that physical and monetary resources in Chinese banks have been parsimoniously used in generating more loans. The positive sign of trend square term suggests that this behaviour is increasing over time, possibly due to higher regulation of Chinese banks after the global financial crisis of 2007/2008. The negative influence of bank profitability on bank efficiency is in contrast with the findings of Tan and Floros (2018) in the Chinese banking industry. This difference is mainly attributed to the fact that we not only use a different dataset, but also the methods in the efficiency analysis as well as the second-stage analysis are different. The empirical literature does not focus on the investigation related to the impact of traditional banking business on bank efficiency; instead, few attempts have been made to examine the influence of income diversification on bank efficiency. Our findings from this perspective show that income diversification has a positive impact on bank efficiency, which is different from the findings of Alhassan (2015) in terms of the Ghanaian banking industry. Finally, the negative influence of expense management on the efficiency level is in line with the consideration that interest expenses and non-interest expenses are the main inputs in banking operation and large amounts of expenses (poor expense management) will reduce the efficiency level (Khankhoje & Sathye, 2008).

## 5 Conclusion

Using a sample of 39 Chinese commercial banks over the period 2010–2018 with different ownership types, including state-owned commercial banks, joint-stock commercial banks, city commercial banks, rural commercial banks, and foreign banks, we analyze the efficiency level using a proposed non-radial DEA model. In the second stage, a Robust Endogenous Neural Network Analysis is proposed and used for the first time to examine the inter-relationships between efficiency, bank size, bank profitability, expense management, traditional banking business and non-traditional banking business.

We fill in a gap in the empirical literature by not only proposing an innovative non-radial DEA model, but also using a Robust Endogenous Neural Network Analysis with a comprehensive set of bank-specific variables which can thoroughly address the endogeneity issue and provide more accurate and concrete policy implications for the banking regulatory authority.

Our findings show that, although in general, the Chinese banking industry experiences strong volatility in the level of efficiency over the examined period, consistent growth in the efficiency level has been observed up to 2015, after which a slight decline and volatility has been noticed and the Chinese banking industry ends up with an efficiency score of over 0.7 by the end of 2018. Comparing different ownership types, the results suggest that state-owned commercial banks are the most efficient banking group, while the rural commercial banks have the lowest level of efficiency. It is further found that foreign banks experience the strongest volatility in the efficiency level. The second-stage analysis shows that bank size exerts a positive influence on the development of non-traditional banking business and proactive expense management, bank size and non-traditional businesses have a positive impact on efficiency levels, while bank profitability, traditional businesses and expenses management have a negative influence on bank efficiency.

The current study generates important policy implications to the Chinese banking regulatory authority: (1) Chinese banks should reduce the volume of deposit-loan business; (2) the Chinese banks are encouraged to expand non-traditional banking businesses, which would be facilitated by financial innovation; (3) Chinese banks should engage in an improved practice of expense management.

In terms of future research, our newly proposed innovative multi-period aggregative efficiency method can be applied to other banking sectors or non-banking sectors and, in particular, our method can be used for comparison with other alternative efficiency evaluation methods. In terms of our second-stage Robust Endogenous Neural Network Analysis, this advanced method can be applied to test the inter-relationships between efficiency, competition and stability in the banking sector and the results can be compared with other empirical banking studies addressing the same issue but under other econometric methods.

## Appendix

See Table 3.

**Table 3** Names of banks in the sample

Name	DMU
Bank of Cangzhou Co., Ltd	1
Bank of Chengdu Co., Ltd	2
Bank of Dalian Co., Ltd	3
Bank of Fuxin Co., Ltd	4
Bank of Guangzhou Co., Ltd	5
Bank of Qingdao Co., Ltd	6
Bank of Shanghai Company Limited	7
Bank of Taizhou Co., Ltd	8
Bank of Wenzhou Co., Ltd	9
Bank of Yingkou Co., Ltd	10
Beijing Rural Commercial Bank Co., Ltd	11
Citibank (China) Company Limited	12
Credit Agricole CIB (China) Limited	13
DBS Bank (China) Limited	14
Fubon Bank (China) Co., Ltd	15
Fudian Bank Co., Ltd	16
Fujian Haixia Bank Co., Ltd	17
Guangxi Beibu Gulf Bank Co., Ltd	18
HSBC Bank (China) Company Limited	19
Hang Seng Bank (China) Limited	20
Hankou Bank Co., Ltd	21
KEB Hana Bank (China) Company Limited	22
MUFG Bank(China),Ltd	23
Metropolitan Bank (China) Ltd	24
Mizuho Bank (China), Ltd	25
Nanyang Commercial Bank (China) Co., Ltd	26
Shanghai Rural Commercial Bank Co., Ltd	27
Societe Generale (China) Ltd	28
Sumitomo Mitsui Banking Corporation (China) Limited	29
Bank of Beijing Co., Ltd	30
Bank of China Limited	31
Bank of Communications Co., Ltd	32
China Construction Bank Corporation	33
China Everbright Bank Company Limited	34
China Guangfa Bank Co., Ltd	35
Industrial Bank Co., Ltd	36
Industrial and Commercial Bank of China Limited	37
Ping An Bank Co., Ltd	38
Standard Chartered Bank (China) Limited	39

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alhassan, A. L. (2015). Income diversification and bank efficiency in an emerging market. *Managerial Finance*, 41, 1318–1335.
- Allevi, E., Basso, A., Bonenti, F., Oggioni, G., & Riccardi, R. (2019). Measuring the environmental performance of green SRI funds: A DEA approach. *Energy Economics*, 79, 32–44.
- Altunbas, Y., Carbo, S., Gardener, E. P. M., & Molyneux, P. (2007). Examining the relationship between capital, risk and efficiency in European Banking. *European Financial Management*, 13, 49–70.
- An, Q., Meng, F., Xiong, B., Wang, Z., & Chen, X. (2020). Assessing the relative efficiency of Chinese high-tech industries: A dynamic network data envelopment analysis approach. *Annals of Operations Research*, 290, 707–729.
- Aparicio, J., Cordero, J. M., & Ortiz, L. (2019). Measuring efficiency in education: The influence of imprecision and variability in data on DEA estimates. *Socio-Economic Planning Sciences*, 68, 100698. <https://doi.org/10.1016/j.seps.2019.03.004>
- Ardia, D., Boudt, K., Carl, P., Mullen, K., & Peterson, B. G. (2011). Differential evolution with DEoptim: An application to non-convex portfolio optimization. *The R Journal*, 3, 27–34.
- Assaf, A., Barros, C. P., & Josiassen, A. (2010). Hotel efficiency: A bootstrapped metafrontier approach. *International Journal of Hospitality Management*, 29, 468–475.
- Athanassopoulos, A. D., & Curram, S. (1996). A comparison of data envelopment analysis and artificial neural networks as tools for assessing the efficiency of decision making units. *Journal of the Operational Research Society*, 47, 1000–1017.
- Avkiran, N. K. (2011). Association of DEA super-efficiency estimates with financial ratios: Investigating the case of Chinese banks. *Omega*, 39, 323–334.
- Azadi, M., Jafarian, M., Saen, R. F., & Mirhedayatian, S. M. (2015). A new fuzzy DEA model for evaluation of efficiency and effectiveness of suppliers in sustainable supply chain management context. *Computers and Operations Research*, 54, 274–285.
- Barros, C. P., Nektarios, M., & Assaf, A. (2010). Efficiency in the Greek insurance industry. *European Journal of Operational Research*, 205, 431–436.
- Bashiri, M., Farshbaf-Geranmayeh, A., & Mogouie, H. (2013). A neuro-data envelopment analysis approach for optimization of uncorrelated multiple response problems with smaller the better type controllable factors. *Journal of Industrial Engineering International*, 9, 30. <https://doi.org/10.1186/2251-712X-9-30>
- Bayraktar, E., Tatoglu, E., Turkyilmaz, A., Delen, D., & Zaim, S. (2012). Measuring the efficiency of customer satisfaction and loyalty for mobile phone brands with DEA. *Expert Systems with Applications*, 39, 99–106.
- Chang, T., Hu, J., Chou, R. Y., & Sun, L. (2012). The source of bank productivity growth in China during 2002–2009: A disaggregation view. *Journal of Banking and Finance*, 36, 1997–2006.
- Chang, Y., Zhang, N., Danao, D., & Zhang, N. (2013). Environmental efficiency analysis of transportation system in China: A non-radial DEA approach. *Energy Policy*, 58, 277–283.
- Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9, 181–185.
- Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1995). *Data envelopment analysis: Theory methodology and applications*. Kluwer.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
- Charoenrat, T., & Harvie, C. (2014). The efficiency of SMEs in Thai Manufacturing: A stochastic frontier analysis. *Economic Modelling*, 43, 372–393.



- Chen, C., & Delmas, M. (2011). Measuring corporate social performance: An efficiency perspective. *Productions and Operations Management*, 20, 789–804.
- Chortareas, G. E., Girardone, C., & Ventouri, A. (2012). Bank supervision, regulation, and efficiency: Evidence from the European Union. *Journal of Financial Stability*, 8, 292–302.
- Chortareas, G. E., Girardone, C., & Ventouri, A. (2013). Financial freedom and bank efficiency: Evidence from the European Union. *Journal of Banking and Finance*, 37, 1223–1231.
- Cooper, W. W., Park, K. S., & Pastor, J. T. (1999). RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. *Journal of Productivity Analysis*, 11, 5–42.
- Cummins, J. D., Weiss, M. A., Xie, X., & Zi, H. (2010). Economies of scope in financial services: A DEA efficiency analysis of the US insurance industry. *Journal of Banking and Finance*, 34, 1525–1539.
- Defung, F., Salim, R., & Bloch, H. (2016). Has regulatory reform had any impact on bank efficiency in Indonesia? A two-stage analysis. *Applied Economics*, 48, 5060–5074.
- Eling, M., & Jia, R. (2019). Efficiency and profitability in the global insurance industry. *Pacific-Basin Finance Journal*, 57, 101190. <https://doi.org/10.1016/j.pacfin.2019.101190>
- Eling, M., & Lühnen, M. (2010). Efficiency in the international insurance industry: A cross-country comparison. *Journal of Banking and Finance*, 34, 1497–1509.
- Emrouznejad, A., & Shale, E. A. (2009). A combined neural network and DEA for measuring efficiency of large-scale data sets. *Computers and Industrial Engineering*, 56, 249–254.
- Esmailzadeh, A., & Hadi-Vencheh, A. (2013). A super-efficiency model for measuring aggregative efficiency of multi-period production systems. *Measurement*, 46(10), 3988–3993.
- Fang, J., Lau, C. K. M., Lu, Z., Tan, Y., & Zhang, H. (2019). Bank performance in China: A perspective from bank efficiency, risk-taking and market competition. *Pacific-Basin Finance Journal*, 56, 290–309.
- Fare, R., & Lovell, C. A. K. (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19, 150–162.
- Fiordelisi, F., Marques-Ibanez, D., & Molynux, P. (2011). Efficiency and risk in European banking. *Journal of Banking and Finance*, 35, 1315–1326.
- Fu, M. X., & Sio, E. U. (2011). Economies of scale and scope in Macau's banking sector. *Banks and Bank Systems*, 6, 90–97.
- Fukuyama, H., & Tan, Y. (2020). Deconstructing three-stage overall efficiency into input, output and stability efficiency components with consideration of market power and loan loss provision: An application to Chinese banks. *International Journal of Finance and Economics*. <https://doi.org/10.1002/ijfe.2185>
- Girardone, C., Molyneux, P., & Gardener, E. P. M. (2004). Analyzing the determinants of bank efficiency: The case of Italian Banks. *Applied Economics*, 36, 215–227.
- Guan, J., & Chen, K. (2012). Modelling the relative efficiency of national innovation systems. *Research Policy*, 41, 102–115.
- Gulati, R., & Kumar, S. (2011). Impact of non-traditional activities on the efficiency of Indian banks: An empirical investigation. *Macroeconomics and Finance in Emerging Market Economies*, 4, 125–166.
- He, F., Zhang, Q., Lei, J., Fu, W., & Xu, X. (2013). Energy efficiency and productivity change of China's iron and steel industry: Accounting for undesirable outputs. *Energy Policy*, 54, 204–213.
- Kalapouti, K., Petridis, K., Malesios, C., & Dey, P. K. (2020). Measuring efficiency of innovation using combined data envelopment analysis and structure equation modeling: Empirical study in EU regions. *Annals of Operations Research*, 294, 297–320.
- Khankhoje, D., & Sathye, M. (2008). Efficiency of rural banks: The case of India. *International Business Research*, 1, 140.
- Konara, P., Tan, Y., & Johnes, J. (2019). FDI and heterogeneity in bank efficiency: Evidence from emerging markets. *Research in International Business and Finance*, 49, 100–113.
- Kumar, S., & Gulati, R. (2010). Measuring efficiency, effectiveness and performance of Indian public sector banks. *International Journal of Productivity and Performance Management*, 59, 51–74.
- Lee, C., Wang, K., & Sun, W. (2019). Allocation of emissions permit for China's Iron and Steel industry in an imperfectly competitive market: A Nash equilibrium DEA model. *IEEE Transactions on Engineering Management*, 68, 548–561.
- Lee, K., & Saen, R. F. (2012). Measuring corporate sustainability management: A data envelopment analysis approach. *International Journal of Production Economics*, 140, 219–226.
- Li, L., & Zhang, Y. (2013). Are there diversification benefits of increasing noninterest income in the Chinese banking industry? *Journal of Empirical Finance*, 24, 151–165.
- Lin, B., & Long, H. (2015). A stochastic frontier analysis of energy efficiency of China's chemical industry. *Journal of Cleaner Production*, 87, 235–244.

- Liu, X., Sun, J., Yang, F., & Wu, J. (2018). How ownership structure affect bank deposits and loan efficiencies: An empirical analysis of Chinese commercial banks. *Annals of Operations Research*, 290, 983–1008.
- Liu, X., Yang, F., & Wu, J. (2020). DEA considering technological heterogeneity and intermediate output target setting: The performance analysis of Chinese commercial banks. *Annals of Operations Research*, 291, 605–626.
- Lozano-Vivas, A., & Pasirouras, F. (2010). The impact of non-traditional activities on the estimation of bank efficiency: International evidence. *Journal of Banking and Finance*, 34, 1436–1449.
- Luo, Y., Bi, G., & Liang, L. (2012). Input/output indicator selection for DEA efficiency evaluation: An empirical study of Chinese commercial banks. *Expert Systems with Applications*, 39, 1118–1123.
- Malesios, C., Dey, P. K., & Abdelaziz, F. B. (2020). Supply chain sustainability performance measurement of small and medium sized enterprises using structural equation modelling. *Annals of Operations Research*, 294, 623–653.
- Margaritis, D., & Psillaki, M. (2010). Capital structure, equity ownership and firm performance. *Journal of Banking and Finance*, 34, 621–632.
- Matthews, K. (2013). Risk management and managerial efficiency in Chinese banks: A network DEA framework. *Omega*, 41, 207–215.
- Merkert, R., & Hensher, D. A. (2011). The impact of strategic management and fleet planning on airline efficiency: A random effects Tobit model based on DEA efficiency scores. *Transportation Research Part A: Policy and Practice*, 45, 686–695.
- Mirhedayatian, S. M., Azadi, M., & Saen, R. F. (2014). A novel network data envelopment analysis model for evaluating green supply chain management. *International Journal of Production Economics*, 147, 544–554.
- Misiunas, N., Oztekin, A., Chen, Y., & Chandra, K. (2016). DEANN: A healthcare analytic methodology of data envelopment analysis and artificial neural networks for the prediction of organ recipient functional status. *Omega*, 58, 46–54.
- Modhej, D., Sanei, M., Shoja, N., & Hosseinzadeh Lotfi, F. (2017). Integrating inverse data envelopment analysis and neural network to preserve relative efficiency values. *Journal of Intelligent and Fuzzy Systems*, 32, 4047–4058.
- Mubiru, J., & Banda, E. (2008). Estimation of monthly average daily global solar irradiation using artificial neural networks. *Solar Energy*, 82, 181–187.
- Mullen, K. M., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2011). DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, 40, 6. <https://doi.org/10.18637/jss.v040.i06>
- Nguyen, T. L. A. (2018). Diversification and bank efficiency in six ASEAN countries. *Global Finance Journal*, 37, 57–78.
- Olanrewaju, O., Jimoh, A., & Kholopan, P. (2012). Integrated IDA–ANN–DEA for assessment and optimization of energy consumption in industrial sectors. *Energy*, 46, 629–635.
- Paradi, J. C., Rouatt, S., & Zhu, H. (2011). Two-stage evaluation of bank branch efficiency using data envelopment analysis. *Omega*, 39, 99–109.
- Pestana, B. C., Laurent, B., Nicolas, P., Elisabeth, R., Bernardin, S., & Assaf, A. G. (2011). Performance of French destinations: tourism attraction perspectives. *Tourism Management*, 32, 141–146.
- Picazo-Tadeo, A. J., Gómez-Limón, J. A., & Reig-Martínez, E. (2011). Assessing farming eco-efficiency: A data envelopment analysis approach. *Journal of Environmental Management*, 92, 1154–1164.
- Psillaki, M., Tsolas, I. E., & Margaritis, D. (2010). Evaluation of credit risk based on firm performance. *European Journal of Operational Research*, 201, 873–881.
- Rouyendegh, B. D., Oztekin, A., Ekong, J., & Dag, A. (2019). Measuring the efficiency of hospitals: A fully-ranking DEA–FAHP approach. *Annals of Operations Research*, 278, 361–378.
- Sahoo, D. K., & Sahoo, P. K. (2020). Efficiency, productivity dynamics and determinants of productivity growth in Indian telecommunication industries: An empirical analysis. *Journal of Public Affairs*. <https://doi.org/10.1002/pa.2353>
- Santin, D., & Delgado, F. J. (2004). The measurement of technical efficiency: A neural network approach. *Applied Economics*, 36, 627–635.
- Seiford, L. M., & Thrall, R. M. (1990). Recent developments in DEA: The mathematical programming approach to frontier analysis. *Journal of Econometrics*, 46, 7–38.
- Shokrollahpour, E., Hosseinzadeh Lotfi, F., & Zandieh, M. (2016). An integrated data envelopment analysis–artificial neural network approach for benchmarking of bank branches. *Journal of Industrial Engineering International*, 12, 137–143.
- Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136, 31–64.

- Song, M., & Li, H. (2019). Estimating the efficiency of a sustainable Chinese tourism industry using bootstrap technology rectification. *Technological Forecasting and Social Change*, 143, 45–54.
- Staub, R. B., Souza, G. D. S., & Tabak, B. M. (2010). Evolution of bank efficiency in Brazil: A DEA approach. *European Journal of Operational Research*, 202, 204–213.
- Stefaniec, A., Hosseini, K., Xie, J., & Li, Y. (2020). Sustainability assessment of inland transportation in China: A triple bottom line-based network DEA approach. *Transportation Research Part D: Transport and Environment*, 80, 102258. <https://doi.org/10.1016/j.trd.2020.102258>
- Sueyoshi, T., & Goto, M. (2012). Data envelopment analysis for environmental assessment: Comparison between public and private ownership in petroleum industry. *European Journal of Operational Research*, 216, 668–678.
- Tan, Y., & Anchor, J. (2017). The impacts of risk-taking behavior and competition on technical efficiency: Evidence from the Chinese banking industry. *Research in International Business and Finance*, 41, 90–104.
- Tan, Y., & Floros, C. (2013). Risk, capital and efficiency in Chinese banking. *Journal of International Financial Markets, Institutions and Money*, 26, 378–393.
- Tan, Y., & Floros, C. (2018). Risk, competition and efficiency in banking: Evidence from China. *Global Finance Journal*, 35, 223–236.
- Tan, Y., & Floros, C. (2019). Risk, competition and cost efficiency in the Chinese banking industry. *International Journal of Banking, Accounting and Finance*, 10, 144–161.
- Tan, Y., Floros, C., & Anchor, J. (2017). The profitability of Chinese banks: Impacts of risk, competition and efficiency. *Review of Accounting and Finance*, 16, 86–105.
- Thanassoulis, E., Kortelainen, M., Johnes, G., & Johnes, J. (2011). Cost and efficiency of higher education institutions in England: A DEA analysis. *Journal of the Operational Research Society*, 62, 1282–1297.
- Tone, K., & Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega*, 38, 145–156.
- Walheer, B., & He, M. (2020). Technical efficiency and technology gap of the manufacturing industry in China: Does firm ownership matter? *World Development*, 127, 104769. <https://doi.org/10.1016/j.worlddev.2019.104769>
- Wang, K., Huang, W., Wu, J., & Liu, Y. (2014). Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. *Omega*, 44, 5–20.
- Wang, Q., Zhao, Z., Zhou, P., & Zhou, D. (2013). Energy efficiency and production technology heterogeneity in China: A meta-frontier DEA approach. *Economic Modelling*, 35, 283–289.
- Wang, R., & Feng, Y. (2020). Research on China's agricultural carbon emission efficiency evaluation and regional differentiation based on DEA and Theil models. *International Journal of Environmental Science and Technology*. <https://doi.org/10.1007/s13762-020-02903-w>
- Wang, X., Ding, H., & Liu, L. (2019). Eco-efficiency measurement of industrial sectors in China: A hybrid super-efficiency DEA analysis. *Journal of Cleaner Production*, 229, 53–64.
- Wanke, P., Barros, C. P., & Emrouznejad, A. (2016). Assessing productive efficiency of banks using integrated fuzzy-DEA and bootstrapping: A case of Mozambican banks. *European Journal of Operational Research*, 249, 378–389.
- Wanke, P., Tan, Y., Antunes, J., & Hadi-Vencheh, A. (2020). Business environment drivers and technical efficiency in the Chinese energy industry: A robust Bayesian stochastic frontier analysis. *Computers and Industrial Engineering*, 144, 106487. <https://doi.org/10.1016/j.cie.2020.106487>
- Wu, D. S., Yang, Z. J., & Liang, L. A. (2006). Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank. *Expert Systems with Applications*, 31, 108–115.
- Wu, Y., & Goh, M. (2010). Container port efficiency In emerging and more advanced markets. *Transportation Research Part E: Logistics and Transportation Review*, 46, 1030–1042.
- Yin, P., Chu, J., Wu, J., Ding, J., Yang, M., & Wang, Y. (2020). A DEA-based two-stage network approach for hotel performance analysis: An internal cooperation perspective. *Omega*, 93, 102035. <https://doi.org/10.1016/j.omega.2019.02.004>
- Yu, M. (2010). Assessment of airport performance using the SBM-NDEA model. *Omega*, 38, 440–452.
- Yu, M. M., & Chen, L. (2020). Evaluation of efficiency and technological bias of tourist hotels by a meta-frontier DEA model. *Journal of the Operational Research Society*, 71, 718–732.
- Zha, Y., Liang, N., Wu, M., & Bian, Y. (2016). Efficiency evaluation of banks in China: A dynamic two-stage slacks-based measure approach. *Omega*, 60, 60–72.
- Zhong, W., Yuan, W., Li, S. X., & Huang, Z. (2011). The performance evaluation of regional R&D investment in China: An application of DEA based on the first official China economic census data. *Omega*, 39, 447–455.
- Zhou, P., Ang, B. W., & Wang, H. (2012). Energy and CO2 emission performance in electricity generation: A non-radial directional distance function approach. *European Journal of Operational Research*, 221, 625–635.

- Zhou, X., Xu, Z., Chai, J., Yao, L., Wang, S., & Lev, B. (2019). Efficiency evaluation for banking systems under uncertainty: A multi-period three-stage DEA model. *Omega*, 85, 68–82.
- Zhou, Y., Xing, X., Fang, K., Liang, D., & Xu, C. (2013). Environmental efficiency analysis of power industry in China based on an entropy SBM model. *Energy Policy*, 57, 68–75.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.