Ann Oper Res (2018) 266:551–587
https://doi.org/10.1007/s10479-018-2797-z

S.I.: ANALYTICAL MODELS FOR FINANCIAL MODELING AND RISK MANAGEMENT

Assessing efficiency profiles of UK commercial banks: a DEA analysis with regression-based feedback

Jamal Ouenniche[1] · **Skarleth Carrales**[1]

Published online: 5 March 2018

Abstract Data envelopment analysis (DEA) has witnessed increasing popularity in banking studies since 1985. In this paper, we propose a new DEA-based analysis framework with a regression-based feedback mechanism, where regression analysis provides DEA with feedback that informs about the relevance of the inputs and the outputs chosen by the analyst. Unlike previous studies, the DEA models used within the proposed framework could use both inputs and outputs, only inputs, or only outputs. So far, the UK banking sector remains relatively under researched despite its crucial importance to the UK economy. We use the proposed framework to address several research questions related to both the efficiency of the UK commercial banking sector and DEA analyses with and without regression-based feedback. Empirical results suggest that, on average, the commercial banks operating in the UK—whether domestic or foreign—are yet to achieve acceptable levels of overall technical efficiency, pure technical efficiency, and scale efficiency. On the other hand, DEA analyses with and without a linear regression-based feedback mechanism seem to provide consistent findings; however, in general DEA analyses without feedback tend to over- or under-estimate efficiency scores depending on the orientation of the analyses. Furthermore, in general, a linear regression-based feedback mechanism proves effective at improving discrimination in DEA analyses unless the initial choice of inputs and outputs is well informed.

Keywords Data envelopment analysis · Efficiency · UK commercial banks · DEA models without explicit inputs · DEA models without explicit outputs

1 Introduction

The banking sector plays a crucial socio-economic role at the regional, national and international levels. Banks are at the heart of financial systems in that they act as financial

✉ Jamal Ouenniche
Jamal.Ouenniche@ed.ac.uk

1 Business School, University of Edinburgh, 29 Buccleuch Place, Edinburgh EH8 9JS, UK

intermediaries; to be more specific, they borrow money by accepting deposits and issuing debt securities, and lend money both directly to their customers and indirectly through capital markets by investing in debt securities. Banks play an important role in money supply and the efficient allocation of financial resources in an economy. Banks make profits in exchange for their services including risk management. Nowadays, banks have a diversified portfolio of activities that range from personal, corporate and investment banking to trading of currency, commodities, and financial securities on stock markets. Because of the crucial importance of banking systems to the economy and the financial risks they face, banks are required to comply with both national and international regulations, and their performance is constantly monitored by both regulatory bodies and investors. In fact, poor performance often leads to distress which might lead to bankruptcy under some circumstances along with substantial financial, economic and social undesirable consequences.

In this paper, we assess the efficiency profiles of UK commercial banks. The UK banking system has specific distinctive features which distinguish it from other banking systems. In fact, the UK banking system is relatively big compared to the banking systems of other countries. Its size is the result of a combination of factors including its history, as the UK has been a financial centre since the eighteenth century. As a financial hub, the UK banking system offers the benefits of clustering such as higher productivity and wage. The robustness of the UK legal and regulatory structure along with the implicit government subsidy and its openness to trade and capital flow seem to provide attractive incentives and flexibility for foreign banks to do business in the UK and for domestic banks to do business abroad. As a result of some of these features, UK has the largest banking sector on a residency basis compared to US, Japan and the ten largest EU Economies with foreign banks on a residency basis, from 56 different countries, owning approximately 50% of the UK banking sector assets. In addition, nearly 1/5 of the global banking activity is booked in the UK. The contribution of foreign banks to the UK banking system and its economy is substantial as suggested by a growth from around 100% of nominal GDP in 1975 to around 450% of nominal GDP in 2013. This growth of 350% is due to the relatively large assets and liabilities account of foreign banks residing in the UK and representing more than four times the median figure for OECD countries. Last, but not least, the international nature of the UK banking system—foreign banks have a large operation in the UK and UK banks have a large operation abroad—along with the continuous reengineering of UK banking regulations enhances its banking system resilience. For more details on the features of the UK banking system, we refer the reader to the Bank of England publications (e.g., Davies et al. 2010; Bush et al. 2014; Burrows et al. 2015).

In this paper, we propose a revised methodological framework; namely, Data Envelopment Analysis (DEA) with a regression-based feedback mechanism along with new DEA models (i.e., DEA models without explicit inputs or outputs), and use it to assess the efficiency profiles of UK commercial banks. The proposed methodology is useful for variable selection especially when the lack of discrimination is a concern. It is used to address three research questions: (1) how do DEA analyses with and without a linear regression-based feedback mechanism compare? (2) how effective is a linear regression-based feedback mechanism in improving discrimination in DEA? and (3) when a feedback mechanism is used to inform the researcher or analyst about the relevance of the choices of inputs and outputs in a DEA analysis, how do radial models (e.g., CCR, BCC) and non-radial models (e.g., SBM) compare? From a practical perspective, we are questioning whether the efficiency determinants identified in previous studies (i.e., inputs and outputs in DEA analysis under the intermediation approach) are actually (empirically) contributing to efficiency or not and whether methodological choices (e.g., choice of DEA model to use, choice of metrics or proxies

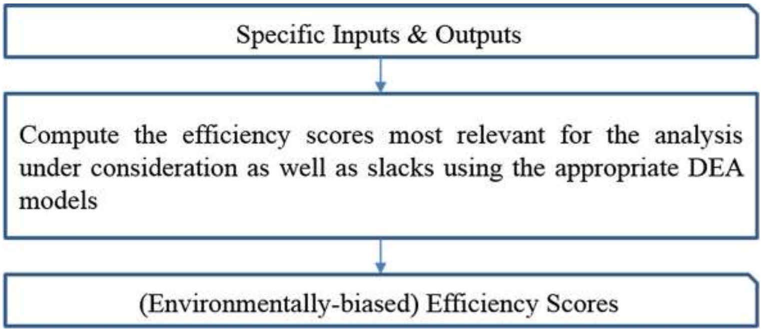

Fig. 1 Main steps of a single stage analysis

of performance criteria) have something to do with it. For the sake of completeness and update of analyses, we also address two conventional research questions: (4) are UK commercial banks managed efficiently? and (5) what are the drivers of UK Commercial Banks' efficiency? However, unlike previous contributions, which focus on the few largest UK commercial banks, these last two research questions are addressed for the whole UK commercial banking system. In our application, it turned out that the UK banking dataset we used requires and justifies the use of DEA models without explicit inputs or outputs when variable selection is informed by a feedback mechanism. Note that the feedback mechanism does not need to be regression-based.

The remainder of this paper is organised as follows. In Sect. 2, we classify the literature on efficiency assessment in banking according to several criteria and critically discuss some of the choices made in the literature. In Sect. 3, we propose a DEA-based sequential decision making process with regression-based feedback adjustment mechanisms along with new DEA models. In Sect. 4, we summarise our empirical investigation and its findings. Finally, Sect. 5 concludes the paper.

2 Landscape of research on efficiency assessment in banking

Research papers on efficiency assessment in banking could be classified into several categories depending on one's choice of the classification criterion. In this paper, we use three criteria to classify the literature on static DEA analyses; namely, type of analysis, type of approach, and country of focus.

With respect to the type of analysis, the literature could be divided into three categories. The first category of studies uses *Single Stage Analysis*—see Fig. 1 for a flow chart of a typical single stage analysis (e.g., Ferrier and Lovell 1990; Elyasiani and Mehdian 1992; Yue 1992; Grabowski et al. 1993; Fukuyama 1993; Zaim 1995; Pastor et al. 1997; Barr et al. 1993; Lozano-Vivas et al. 2002).

The second category of studies uses *Two-Stage Analysis* to overcome environment bias—see Fig. 2 for a flow chart of a typical two-stage analysis (e.g., Rangan et al. 1988; Elyasiani and Mehdian 1990; Aly et al. 1990; Favero and Papi 1995; Miller and Noulas 1996; Bhattacharyya et al. 1997; Chen 1998; Chu and Lim 1998; Barr et al. 1994; Barr and Siems 1997; Pasiouras 2008; Wanke and Barros 2014; Kwon and Lee 2015; Du et al. 2018). Note however that the efficiency scores obtained with a two-stage analysis would still be environmentally-biased, because the inputs and outputs used in the first stage are not adjusted for environment. In order to properly control for these environmental variables, one

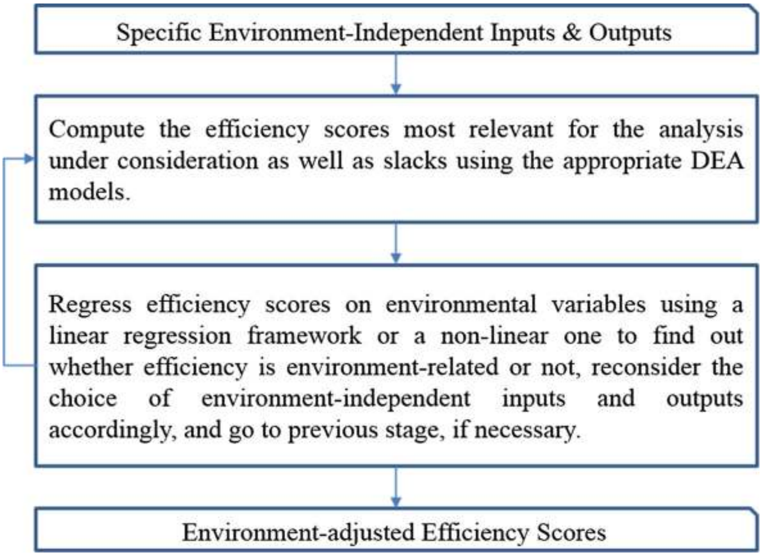

Fig. 2 Main steps of a two-stage analysis

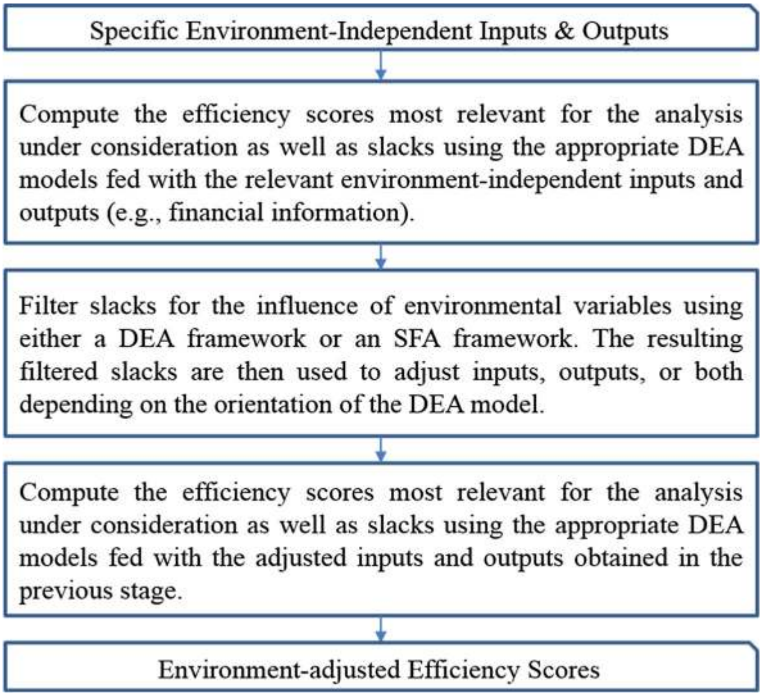

Fig. 3 Main steps of a three-stage analysis

could use a three-stage methodology. Finally the third category of studies uses *Three-Stage Analysis*—see Fig. 3 for a flow chart of a typical three-stage analysis (e.g., Pastor 2002; Drake et al. 2006; Liu and Tone 2008; Avkiran 2009; Liu 2018).

With respect to the type of assessment perspective, which drives the choices of inputs and outputs, we classify the literature into six categories; namely, the *intermediation approach* or perspective (e.g., Rangan et al. 1988; Ferrier and Lovell 1990; Charnes et al. 1990; Elyasiani and Mehdian 1990, 1992; Aly et al. 1990; Yue 1992; Grabowski et al. 1993; Fukuyama 1993; Zaim 1995; Favero and Papi 1995; Miller and Noulas 1996; Taylor et al. 1997; Chen 1998; Drake et al. 2006; Liu 2018), the *asset approach* (e.g., Favero and Papi 1995), the *production approach* (e.g., Drake et al. 2006; Liu and Tone 2008), the *value added approach* (e.g., Bhattacharyya et al. 1997; Pastor et al. 1997; Chu and Lim 1998; Pastor 2002; Das and Ghosh 2006), the *profit-oriented approach* (e.g., Berger and Mester 2003; Drake et al. 2006; Liu and Tone 2008), and the *user cost approach* (e.g., Hancock 1985a, b; Fixler and Zieschang 1992).

Recall that the intermediation approach considers banks as intermediation agents who collect funds and provide loans and other assets. The asset approach is a variant of the intermediation approach which consider banks as financial intermediaries between liability holders and those who receive bank funds. The production approach considers banks as production units that transform inputs into outputs, or producers of deposit accounts and loan services. Under the value added approach, the share of value added guides the choice of inputs and outputs. Under the profit approach, profit guides the choice of inputs and outputs. Finally, under the user cost approach, the net contribution to bank revenue determines the nature of inputs and outputs.

As to the country of focus, the literature could be divided into two main categories. The first category consists of *single country focused studies* and covers US Banks (Rangan et al. 1988; Ferrier and Lovell 1990; Elyasiani and Mehdian 1990, 1992; Aly et al. 1990; Yue 1992; Miller and Noulas 1996; Kwon and Lee 2015), UK Banks (Drake 2001; Webb 2003; Webb et al. 2010; Tanna et al. 2011), Italian Banks (Favero and Papi 1995), Turkish Banks (Zaim 1995; Kutlar et al. 2017), Japanese Banks (Fukuyama 1993; Liu and Tone 2008), Taiwanese Banks (Chen 1998; Liu 2018), Hong Kong Banks (Drake et al. 2006), Singaporean Banks (Chu and Lim 1998), Indian Banks (Bhattacharyya et al. 1997), Mozambique Banks (Wanke et al. 2016), and Korean Banks (Lee et al. 2017). The second category consists of *multi-country focused studies* and covers banks in several countries such as US, Australian, New Zealand, Austrian, Spanish, German, UK, Italian, Belgian, French, Danish, Luxembourg, Dutch, and Portuguese Banks (e.g., Pastor et al. 1997; Pastor 2002; Lozano-Vivas et al. 2002; Casu and Molyneux 2003; Pasiouras 2008; Avkiran 2009).

To conclude this section, it is worthy to mention that *single country focused studies* on banks using static DEA analyses (Drake 2001; Webb 2003; Webb et al. 2010; Tanna et al. 2011) focused exclusively on the few largest commercial banks in the UK, whereas this paper considers the whole UK commercial banking sector. We also would like to point out that other DEA methodologies have been used to assess the efficiency of banks; for example, Network DEA (e.g., Matthews 2013; Grigoroudis et al. 2013; Akther et al. 2013; Fukuyama and Matousek 2017; Gulati and Kumar 2017), Network DEA with undesirable variables (e.g., An et al. 2015; Liu et al. 2015), Dynamic DEA (e.g., Avkiran and Goto 2011; Fukuyama and Weber 2015, 2017), Dynamic Network DEA (e.g., Avkiran 2015; Chao et al. 2015; Fukuyama and Weber 2015, 2017; Zha et al. 2016; Wu et al. 2016; Fukuyama and Weber 2017b), Fuzzy DEA (e.g., Wang et al. 2014; Wanke et al. 2016; Hatami-Marbini et al. 2017), DEA with Bootstrapping (e.g., Ferrier and Hirschberg 1997), Fuzzy DEA with Bootstrapping (e.g., Wanke et al. 2016), and Stochastic DEA (e.g., Kao and Liu 2009). For a recent survey, we refer the reader to Kaffash and Marra (2017).

In the next section, we propose a DEA analysis with a regression-based feedback mechanism along with new DEA models to assess the efficiency profiles of banks, which we apply in the following section to the UK banking sector.

3 A DEA analysis with regression-based feedback mechanism

In this section, we shall describe the methodology and models we propose for assessing the efficiency profile of UK commercial banks. The proposed methodology is a sequential decision making process with a feedback adjustment mechanism; namely, a DEA-based analysis with a regression-based feedback mechanism.

DEA was first proposed by Charnes et al. (1978) as a frontier-based non-parametric approach to the relative performance evaluation of a set of n entities commonly referred to as decision making units (*DMUs*), where *DMUs* are viewed as production systems that make use of the same set of m inputs to produce the same set of s outputs. For each DMU, lot sizing decisions of both inputs and outputs are made by its management; that is, the quantity $x_{i,k}$ of input i ($i = 1, \ldots, m$) used by DMU_k ($k = 1, \ldots, n$) and the quantity $y_{r,k}$ of output r ($r = 1, \ldots, s$) produced by DMU_k ($k = 1, \ldots, n$). Unlike parametric methodologies, DEA does not require an explicit specification of the form of the production function, or equivalently the relationship between inputs and outputs. DEA is a mathematical programming-based methodology—for a detailed text on DEA, we refer the reader to Cooper et al. (2007).

In this paper, we are concerned with measuring overall technical efficiency, pure technical efficiency, and scale efficiency of UK commercial banks. Unlike previous studies, the particular features of UK banking data require additional types of DEA models. Therefore, we shall use both input- and output-oriented CCR models (Charnes et al. 1978); both input- and output-oriented BCC models (Banker et al. 1984); BCC models without explicit inputs, BCC-WEI, or without explicit outputs, BCC-WEO (Lovell and Pastor 1999); input-oriented, output-oriented, and non-oriented SBM models (Tone 2001); and SBM-WEI model (Liu et al. 2011) and SBM-WEO model that we propose. CCR and BCC models are described in Table 1, BCC models without explicit inputs or outputs are described in Table 2, SBM models are described in Table 3, and SBM models without explicit inputs or outputs are described in Table 4, where θ_k denotes the technical efficiency of DMU_k and measures the efficiency with which DMU_k transforms inputs into outputs, which reflects the quality of its management decisions, λ_j denotes the weight assigned to DMU_j in constructing the "ideal" benchmark of DMU_k; that is, its projection on the efficiency frontier, and $s^-_{i,k}$ and $s^+_{r,k}$ denote the slacks in input i and output r, respectively, which represent input excess and output shortfall. Recall that most DEA analyses make use of one or several inputs and one or several outputs; however, in some situations one might not have to use any inputs or any outputs—these situations or models are referred to as DEA models or analyses without explicit inputs or without explicit outputs. In a DEA analysis with a regression-based feedback mechanism one might have to discard all inputs or all outputs when regression analysis suggests that they do not drive or explain differences in efficiency profiles. However, in general, in DEA applications the use of DEA models without explicit inputs could be justified when one assumes that inputs are considered similar and equal for all DMUs as they operate, for example, in the same market (e.g., Halkos and Salamouris 2004). On the other hand, the use of DEA models without explicit outputs could be justified when one assumes that outputs are considered similar and

Table 1 CCR and BCC models

Formulation	Description
Objective function	
θ_k	θ_k is to be minimised or maximised depending on whether the analysis is input-oriented or output-oriented
Constraints	
$\sum_{j=1}^{n} \lambda_j x_{i,j} \leq \theta_k \cdot x_{i,k},\ \ i = 1, \ldots, m$ OR $\sum_{j=1}^{n} \lambda_j x_{i,j} \leq x_{i,k},\ \ i = 1, \ldots, m$	For each input i ($i = 1, \ldots, m$), the amount used by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should at most be equal to the amount used by DMU_k adjusted for the degree of technical efficiency of DMU_k or not depending on whether the analysis is input-oriented or not
$\sum_{j=1}^{n} \lambda_j y_{r,j} \geq \theta_k \cdot y_{r,k},\ \ r = 1, \ldots, s$ OR $\sum_{j=1}^{n} \lambda_j y_{r,j} \geq y_{r,k},\ \ r = 1, \ldots, s$	For each output r ($r = 1, \ldots, s$), the amount produced by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should be at least as large as the amount produced by DMU_k adjusted for the degree of technical efficiency of DMU_k or not depending on whether the analysis is output-oriented or not
$\sum_{j=1}^{n} \lambda_j = 1$	The technology is required to be convex in BCC models. This constraint is relaxed in CCR models
$\lambda_j \geq 0,\ \ j = 1, \ldots, n$ θ_k unrestricted	Other requirements including non-negativity

Table 2 BCC models without explicit inputs or outputs

Formulation	Description
Objective function	
θ_k	θ_k is to be minimised or maximised depending on whether the analysis is without explicit output or without explicit inputs
Constraints	
$\sum_{j=1}^{n} \lambda_j x_{i,j} \leq \theta_k \cdot x_{i,k},\ \ i = 1, \ldots, m$ OR $\sum_{j=1}^{n} \lambda_j y_{r,j} \geq \theta_k \cdot y_{r,k},\ \ r = 1, \ldots, s$	For each input i ($i = 1, \ldots, m$), the amount used by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should at most be equal to the amount used by DMU_k adjusted for the degree of technical efficiency of DMU_k, or for each output r ($r = 1, \ldots, s$), the amount produced by DMU_k's "ideal" benchmark should be at least as large as the amount produced by DMU_k adjusted for the degree of technical efficiency of DMU_k depending on whether the analysis is without explicit output or without explicit inputs
$\sum_{j=1}^{n} \lambda_j = 1$	The technology is convex
$\lambda_j \geq 0,\ \ j = 1, \ldots, n$ θ_k unrestricted	Other requirements including non-negativity

Table 3 SBM models

Formulation	Description
Objective function	
$\rho_k = \left(1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_{i,k}^-}{x_{i,k}}\right) \Big/ \left(1 + \frac{1}{s}\sum_{r=1}^{s}\frac{s_{r,k}^+}{y_{r,k}}\right)$ OR $\rho_k = 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_{i,k}^-}{x_{i,k}}$ OR $\rho_k = 1 \Big/ \left(1 + \frac{1}{s}\sum_{r=1}^{s}\frac{s_{r,k}^+}{y_{r,k}}\right)$	One of these ρ_k formulations is to be minimised depending on whether the analysis is non-oriented, input-oriented, or output-oriented
Constraints	
$\sum_{j=1}^{n}\lambda_j x_{i,j} + s_{i,k}^- = x_{i,k},\ i = 1, \ldots, m$	For each input i $(i = 1, \ldots, m)$, the amount used by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should at most be equal to the amount used by DMU_k
$\sum_{j=1}^{n}\lambda_j y_{r,j} - s_{r,k}^+ = y_{r,k},\ r = 1, \ldots, s$	For each output r $(r = 1, \ldots, s)$, the amount produced by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should be at least as large as the amount produced by DMU_k
$\sum_{j=1}^{n}\lambda_j = 1$	This constraint requires the technology to be convex; however, it could be relaxed
$\lambda_j \geq 0,\ j = 1, \ldots, n$ $s_{i,k}^- \geq 0,\ i = 1, \ldots, m$ $s_{r,k}^+ \geq 0,\ r = 1, \ldots, s$	Non-negativity requirements

equal for all DMUs as they operate, for example, under specific legislation or supply markets with fixed shares on which DMUs could not act upon in the short to medium term.

The flowchart of the proposed methodology is outlined in Fig. 4. Within this methodological framework, given a set of relevant environment-independent inputs and outputs specified by the analyst or researcher, DEA analysis with both inputs and outputs is first performed to compute the relevant efficiency scores for the analysis under consideration (e.g., overall technical efficiency, pure technical efficiency, scale efficiency) as well as slacks by solving the appropriate DEA models (e.g., CCR, BCC, SBM models).

For our banking application, inputs and outputs are supplied from banks' financial statements (i.e., balance sheet and income statement). These inputs and outputs are environment-independent because the study is performed on UK banks only, on one hand, and we do not test any specific event-related hypotheses, on the other hand. Then, the DEA scores are regressed on the initial inputs and output supplied by the analyst to find out whether they are statistically significant or not; that is, whether they drive the efficiency scores or not—any inputs or outputs which are not relevant (i.e., not statistically significant) are then discarded and the DEA analysis with both inputs and outputs is performed with a reduced set of inputs and outputs. When regression analysis suggests that none of the inputs or none of the outputs chosen by the analyst are relevant, DEA analysis without explicit inputs or without explicit outputs is performed using the relevant DEA models mentioned above. In sum, regression

Table 4 SBM models without explicit inputs or outputs

Formulation	Description
Objective function	
$\rho_k = 1 - \frac{1}{m}\sum_{i=1}^{m}\frac{s_{i,k}^{-}}{x_{i,k}}$ OR $\rho_k = 1/\left(1+\frac{1}{s}\sum_{r=1}^{s}\frac{s_{r,k}^{+}}{y_{r,k}}\right)$	One of these ρ_k formulations is to be minimised depending on whether the analysis is without explicit output or without explicit inputs
Constraints	
$\sum_{j=1}^{n}\lambda_j x_{i,j} + s_{i,k}^{-} = x_{i,k},\ i = 1,\ldots,m$ OR $\sum_{j=1}^{n}\lambda_j y_{r,j} - s_{r,k}^{+} = y_{r,k},\ r = 1,\ldots,s$	For each input i ($i = 1,\ldots,m$), the amount used by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should at most be equal to the amount used by DMU_k, or for each output r ($r = 1,\ldots,s$), the amount produced by DMU_k's "ideal" benchmark; i.e., its projection on the efficiency frontier, should be at least as large as the amount produced by DMU_k depending on whether the analysis is without explicit output or without explicit inputs
$\sum_{j=1}^{n}\lambda_j = 1$	This constraint requires the technology to be convex; however, it could be relaxed
$\lambda_j \geq 0,\ j = 1,\ldots,n$ $s_{i,k}^{-} \geq 0,\ i = 1,\ldots,m$ OR $s_{r,k}^{+} \geq 0,\ r = 1,\ldots,s$	The weights λ_j s are required to be non-negative as well as the relevant slacks depending on whether the analysis is without explicit output or without explicit inputs

analysis provides DEA with feedback that informs DEA about the relevance of the inputs and outputs chosen by the analyst.

Before we proceed with the application of the proposed DEA analysis with regression-based feedback, we hereafter position our contribution with respect to the literature on variable selection in DEA. So far, such literature could be divided into (1) Judgemental Screening or Expert Opinions such as Fuzzy Delphi Method (Arsad et al. 2017); (2) Statistical Tests and Bootstrapping (e.g., Banker 1996; Olson et al. 1980; Simar and Wilson 2001; Nataraja and Johnson 2011); (3) Dimensionality Reduction Techniques such as Principal Component Analysis (Ueda and Hoshiai 1997; Adler and Golany 2001, 2002; Cinca and Molinero 2004; Adler and Yazhemsky 2010; Nataraja and Johnson 2011); and (4) Variable Reduction Techniques such as Correlation Analysis and Variants (Nunamaker 1985; Jenkins and Anderson 2003; Eskelinen 2017; Adler and Yazhemsky 2010), Copula (Alpay and Akturk Hayat 2017), Efficiency Contribution Measure (Pastor 2002; Nataraja and Johnson 2011; Eskelinen 2017), Stepwise Procedures (Norman and Stoker 1991; Sigala et al. 2004; Wagner and Shimshak 2007; Subramanyam 2016; Sharma and Yu 2015), Akaike's Information Criterion rule (Li et al. 2017), Directional Technology Distance Function (Guarda et al. 2013), Regression Analysis (Lewin et al. 1982; Fanchon 2003; Ruggiero 2005; Luo et al. 2012; Golany and Roll 1989); Decision Tree Analysis (Lim 2008; Jain et al. 2016), and Genetic Algorithms (Madhanagopal and Chandrasekaran 2014). Our contribution falls into the subcategory of Regression Analysis; however, unlike previous contributions, ours use regression analysis within a feedback mechanism and allows for no-inputs or no-outputs situations.

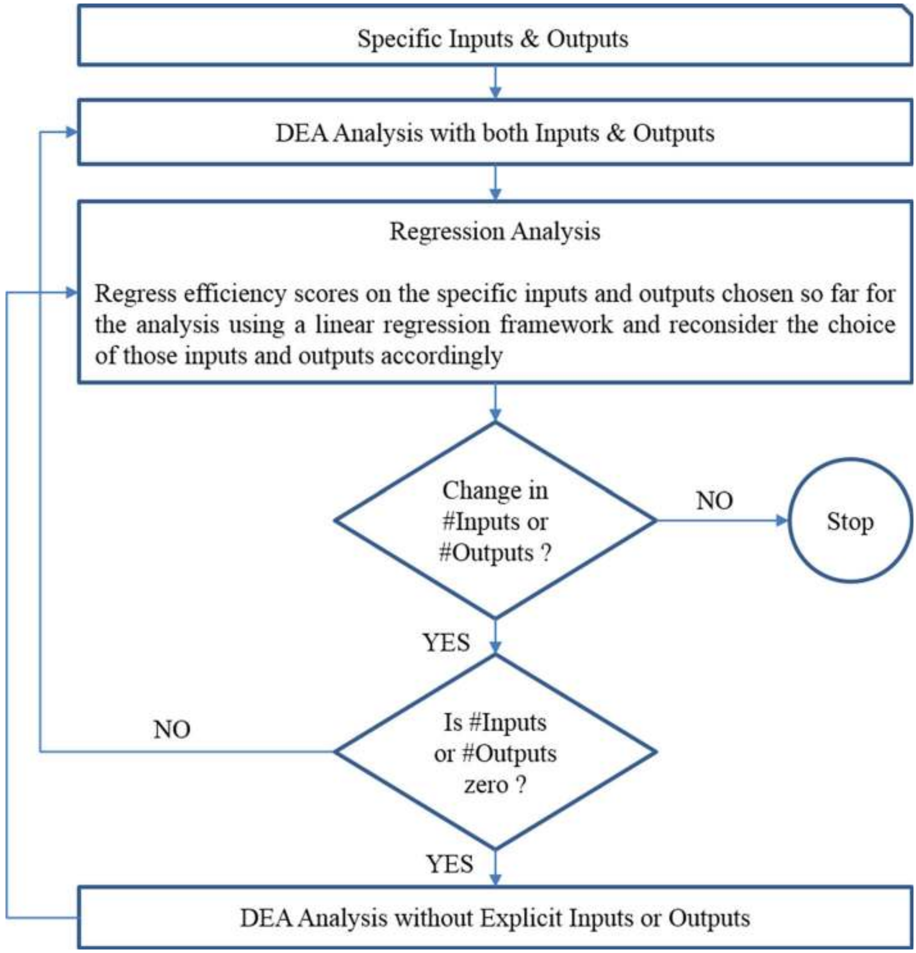

Fig. 4 DEA analysis with regression-based feedback

In the next section, we shall apply the proposed methodology to assess the efficiency profile of UK commercial banks.

4 Empirical study

In our empirical investigation, we used all UK commercial banks for which data is available from Bankscope, provided by Bureau van Dijk, over a period of 29 years; namely, 1987–2015. Our dataset includes 109 commercial banks and consists of a total number of 1171 bank-year observations or decision making units.

The choice of the inputs and outputs with which DEA models are fed is driven by the intermediation approach, where banks are considered as intermediation agents who collect funds and provide loans and other assets. For a discussion on the choice of inputs and outputs in banking applications, we refer the reader to Fethi and Pasiouras (2010). Our survey and classification of the inputs and outputs used in the literature (see Ouenniche et al. 2017) along with an analysis of the balance sheet and the income statement of UK commercial banks revealed that inputs are typically chosen based on resources, costs, or financial burden, whereas outputs are typically chosen based on bank's ability to provide financial services (i.e., Loans and Deposits), generate revenue (i.e., Income and Investments) and acquire more assets (i.e., Investments). However, our critical analysis of such choices suggests that some authors' choices—especially of inputs based on financial burden rather penalise the very means by which banks are able to perform their lending operations. Therefore, we selected inputs based only on resources (i.e., Labour as measured by Personnel Expenses—because the number of employees was not available for all UK banks; Capital as measured by Fixed Assets/Physical Capital or Equity/Financial Capital) and costs (i.e., Total Interest Expense; Total Expenses not including Personnel Expense). As to outputs, we selected them based on the ability of a bank to provide financial services (i.e., Gross Loans; Total Customer Deposits) and generate revenue (i.e., Total Income; Gross Interest and Dividend Income). We did not consider the ability of banks to acquire more assets or to make investments because small UK banks, which are part of our sample, are not quite involved in off-balance sheet activities. These chosen criteria could however be measured in different ways. In our empirical experiments, we used four setups or scenarios each corresponding to a different combination of measures—see Table 5 for details.

A snapshot of the 109 UK commercial banks in our dataset is summarised in Table 6 (see "Appendix"), where the figures are measured in millions of USD. Table 7 provides a snapshot of the leading UK commercial banks (see "Appendix"). Analysis of raw data on UK commercial banks in our dataset revealed that Pareto's Law holds; that is, eight leading banks (i.e., $8/109 \cong 7\%$ of UK commercial banks); namely, National Westminster Bank Plc—NatWest, The Royal bank of Scotland, Ulster bank, Lloyds bank, Bank of Scotland, Barclays, HSBC Bank Plc, and Standard Chartered Bank, together account for almost 87% of the stock of UK customer lending and deposits. In addition, as highlighted by some statistics on fixed assets, as a proxy for size (i.e., the first quartile of total assets in Table 7 is 400% bigger than the third quartile of total assets in Table 6); the UK commercial banks in our dataset, excluding the largest eight, are altogether smaller than the smallest bank of the largest ones. We also performed several analyses by size (e.g., total assets); market share (e.g., total customer deposits, gross loans), gross profitability (e.g., total income), operational expenses (e.g., personnel expenses), and origin (e.g., domestic, foreign)—see Table 8 in "Appendix". These analyses also support Pareto's Law. In addition, they highlight the importance of foreign banks in the UK; in fact, although foreign banks represent 38% of the total UK commercial banks as compared to 55% of domestic banks but the largest eight, their market share is bigger. Last, but not least, assuming that Personnel Expenses are a good proxy for the number of employees, the largest bank; namely, Barclays Bank Plc, employs about 50% of the labour used by all small domestic banks. We also investigated the UK commercial banks' ownership structure and found out that ownership structure is not a discriminatory feature, since 1 foreign bank in residency and 2 local banks are Limited Liability Corporations, 1 foreign bank in residency and 2 local banks are Mutual/Co-ops, and the remaining banks; that is, 39 foreign banks in residency and 64 local banks are Stock Corporations.

DEA analyses of the UK commercial banking sector, as represented by the 109 commercial banks in our dataset, are summarised as follows.

First, in input-oriented analyses (see Tables 9, 11 in "Appendix"), numerical results suggest that in the UK commercial banking system the combination of choices of measures of inputs matters; in other words, how resources and expenses are proxied as well as the combinations of these proxies matter for banks' levels of efficiencies. To be more specific, equity or financial

Table 5 Choices of measures of inputs and outputs for DEA analyses

Setup	Inputs					Output			
	Personnel expenses	Fixed assets	Equity	Total interest expense	Total expenses not including personnel expense	Gross loans	Total customer deposits	Gross interest and dividend income	Total income
1	x	x		x		x	x	x	
2	x	x			x	x	x		x
3	x		x	x		x	x	x	
4	x		x		x	x	x		x

capital (setups 4 and 3), as a proxy for resources, enhances on average overall technical efficiency (OTE) or CCR scores, overall technical efficiency adjusted for mix efficiency (adj-OTE) or SBM scores; pure technical efficiency (PTE) or BCC scores, and scale efficiency (SE) better than fixed assets or physical capital (setups 1 and 2); therefore, UK commercial banks are better at managing their equity or liquidity than their fixed assets, which is in line with the intermediation role of the banks. On the other hand, total expenses not including personnel expense (setups 4 and 2), as a proxy for expenses, seems to enhance on average OTE, adj-OTE, PTE, and SE better than total interest expense (setups 1 and 3). Judged on their use of inputs, on average, UK commercial banks fall short on overall technical efficiency, pure technical efficiency, and scale efficiency—see Tables 9 and 11. In fact, depending on the choice of measures of inputs across setups, average CCR scores vary between 0.3144 and 0.6119, average SBM scores (i.e., overall technical efficiency adjusted for mix efficiency) vary between 0.3577 and 0.5646, average BCC scores vary between 0.5132 and 0.6976, and average SE scores vary between 0.667 and 0.8796. In sum, the management of the UK commercial banking sector seems to be in need of further improvements. Commercial banks in the fourth quartile however seem to be scale efficient to a large extent; therefore, for these banks any further efficiency improvement efforts should be put on pure technical efficiency.

Second, most DEA analyses in banking have focused on input-oriented analyses, which is typically justified by the fact that bank managers have more control over the management of inputs than outputs. This is an arguable point of view as some outputs could be acted upon through better and more focused commercial strategies and marketing campaigns. In addition, in practice, the analysis of output-oriented DEA scores could provide important insight. Motivated by these concerns, we also performed output-oriented analyses of the UK commercial banks—see Tables 10 and 11 in “Appendix”. In output-oriented analyses (see Tables 10, 11), numerical results suggest that, in the UK commercial banking system, the choices of measures of outputs as well as the combinations of choices of measures of inputs matter; in other words, how income is proxied as well as the combinations of proxies of inputs matter for banks’ levels of efficiencies. To be more specific, regardless of the choice of inputs proxies, on average, OTE, PTE and SE are enhanced when total income (setups 2 and 4) is used as a proxy for income compared to gross interest and dividend income (setups 1 and 3). Consequently, on average, the management of UK commercial banks seem to be good at managing total income, but less so in generating gross interest and rewarding their shareholders through dividends. However, average adj-OTE figures are affected by both the choice of income proxies and the combinations of proxies of inputs; in fact, setup 4 enhances adj-OTE more than setup 3 followed by setup 2 then setup 1. Finally, in terms of scale efficiency, output-oriented results are in line with the input-oriented ones.

Third, regression feedback informs the analyst about the relevance of his or her choices of efficiency drivers (i.e., inputs and outputs). Our empirical analysis shows that taking account of regression feedback to revise DEA models always enhances discrimination and adjusts DEA scores downwards or upwards, depending on whether the DEA analysis is input-oriented or output-oriented—see Tables 9, 10, 11, 12, 13, 14, 15, 16 and 17 in “Appendix”. Note that, in the case of the UK commercial banks in our sample, the conclusions with respect to the efficiency profiles of banks remain the same. In sum, regardless of whether DEA analyses are performed with or without regression feedback, the UK commercial banking sector is in need of further efficiency improvements.

Fourth, in addition to enhancing discrimination amongst DMUs and adjusting their DEA scores, which in itself is a major issue in DEA applications, feedback reveals a completely new story on the actual drivers of a range of efficiency measures and exposes the importance of the choice of DEA models in estimating these measures. In the following paragraphs, we shall provide evidence of these claims.

In our empirical analysis, we used two types of regression feedback—see Tables 12, 13, 14, 15, 16 and 17 in “Appendix”. The first regression feedback—referred to as input focus regression analysis—involves regressing DEA scores on inputs. The second regression feedback—referred to as output focus regression analysis—involves regressing DEA scores on outputs. Depending on the statistical significance of inputs (respectively, outputs), some inputs (respectively, outputs) may have to be discarded and the DEA scores recomputed with a reduced set of inputs (respectively, outputs), if necessary. Note however that, in some cases, none of the inputs (respectively, outputs) proves to explain the behaviour of DEA scores in which case DEA models without explicit inputs (respectively, explicit outputs) would have to be solved—as illustrated by Setup 4 in output focus regression. So far, this case has not been encountered by or reported in previous studies, which has motivated the new methodological design in this research.

A summary of the statistically significant input and output drivers of efficiency is provided in Table 18, where Labour, as measured by Personnel expenses, seems to be the most consistent input driver of efficiency scores across all setups and DEA analyses, whereas the provision of financial services, as measured by Gross Loans, seems to be the most consistent output driver of efficiency scores across all setups and DEA analyses. The relevance of remaining drivers however depends on both the setups or combinations of drivers and the DEA analyses. Notice, however, that those setups (i.e., choices of combinations of drivers) that make the UK commercial banking sector look more efficient (e.g., Setup 4 without feedback) are the ones that are most affected by the regression feedback, on one hand, and those setups that lead to more conservative estimates of efficiency scores (e.g., Setup 1 without feedback) are less or not at all affected by the regression feedback, on the other hand. Therefore, the regression feedback serves as a correction mechanism in that it adjusts over- and under-estimated scores. These findings have important implications on the relevance of the choices of inputs and outputs and the combinations of their measures; in fact, they often tell the opposite story revealed by DEA analyses without regression feedback. For example, input-oriented DEA analyses without regression feedback suggested that UK commercial banks are better at managing their financial capital than their physical capital, which is in line with the intermediation role of the banks, but when feedback is incorporated the management of UK commercial banks does not seem to be doing such good job anymore in managing equity. In sum, the lessons to be learned could be summarised as follows. From the perspective of banks’ managers, DEA analyses without feedback make them look better, and most importantly it backs up their strategies of being intermediation agents in the economy. However, regulators and investors might be better off performing DEA analyses with feedback, alongside DEA analyses without feedback, to unveil different pictures.

Furthermore, with respect to the importance of the choice of DEA models in estimating efficiency measures, DEA analyses with input focus regression feedback provides a good example. In fact, empirical results suggest that, in some setups, DEA scores estimated by CCR and BCC models are not driven by the initial choice of inputs. For example, under Setup 2, CCR and BCC scores are only driven by Personnel Expenses. Interestingly, under the same setup, SBM scores are driven by Personnel Expenses, Fixed Assets (physical capital), and Total Expenses not including Personnel Expense. Further investigation of this fact revealed that the slacks associated with Fixed Assets, and Total Expenses not including Personnel Expense turn out to be important in magnitude, but ignored by radial measures of efficiency. SBM scores however take these slacks into account and thus avoid the elimination of Fixed Assets, and Total Expenses not including Personnel Expense through regression feedback.

In sum, ignoring slacks might result in the regression-based feedback suggesting that some efficiency determinants should be discarded when they should not. These findings suggest that, in practice, one should use slacks-based measures of efficiency instead of the conventional ones whenever possible, on one hand, and remind us of the importance for the DEA community to design new SBM based metrics to measure pure technical efficiency and scale efficiency, which are yet to be proposed, on the other hand.

Finally, our analysis of DEA scores of domestic and foreign banks suggests that their efficiency profiles are very similar regardless of which DEA models or regression analysis focus is used to estimate the scores—see, for example, Tables 19, 20, 21 and 22 in "Appendix" for illustration. Also, our analysis of DEA scores of large and smaller banks suggests that their efficiency profiles are very different. In fact, large banks are more overall technically efficient and pure technically efficient than the small ones, but the large ones seem to be less scale efficient than the small banks regardless of which DEA models or regression analysis focus is used to estimate the scores—see, for example, Tables 23, 24, 25 and 26 in "Appendix" for illustration.

In sum, our empirical analyses provided the following answers to our research questions. First, UK commercial banks need further efficiency improvements. Second, UK commercial banks' measures of efficiency seem to be driven by the inputs and outputs identified by researchers so far except when the combinations of measures and their interaction along with their slacks and the type of DEA models used for estimating efficiency scores come into play. Third, DEA analyses with and without a linear regression-based feedback mechanism seem to provide consistent findings in terms of inefficiency; however, compared to DEA analyses with feedback, in general DEA analyses without feedback tend to over- or underestimate efficiency scores depending on whether the analyses are input-oriented or output-oriented. Fourth, in general, a linear regression-based feedback mechanism proves effective at improving discrimination in DEA analyses unless the initial choice of inputs and outputs is well informed. Last, but not least, ignoring slacks might result in the regression-based feedback suggesting that some efficiency determinants should be discarded when they should not, which suggest that, in practice, one should use slacks-based measures of efficiency instead of the conventional ones whenever possible, on one hand, and remind us of the importance for the DEA community to design new SBM based metrics to measure pure technical efficiency and scale efficiency, which are yet to be proposed, on the other hand.

5 Conclusions

In this paper, we investigated the efficiency profiles of the UK commercial banking sector using a new DEA-based analysis framework with a regression-based feedback mechanism, where DEA models could use both inputs and outputs, only inputs, or only outputs. Note that the use of DEA models without explicit inputs or outputs is required when the regression-based feedback mechanism informs DEA analysis that all inputs or all outputs should be discarded, because they do not drive efficiency, which turned out to be the case in our empirical analysis of UK banking data. The proposed DEA analysis design was used to address several research questions related to both the UK commercial banking sector and DEA analyses with and without regression-based feedback—see Sect. 4 for details on our findings. Amongst these findings, it tuned out that performing DEA analyses with radial models such as CCR and BCC, which ignore slacks in computing technical efficiency scores, might result in the regression-based feedback suggesting that some efficiency drivers should be discarded when they should not. Therefore, we recommend that, in practice, one should use slacks-based measures of efficiency instead of the conventional ones whenever possible. These findings remind us of the importance for the DEA community to design new SBM based metrics to measure pure technical efficiency and scale efficiency, which are yet to be proposed.

Acknowledgements The authors would like to express their gratitude to two anonymous referees for their helpful comments and recommendations.

Appendix

See Tables 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 and 26.

Table 6 Statistics on all UK commercial banks in our dataset

Statistics	Inputs					Output			
	Personnel expenses	Fixed assets	Equity	Total interest expense	Total expenses not including personnel expense	Gross loans	Total customer deposits	Gross interest and dividend income	Total income
Minimum	145	3	496	17	176	155	2	156	592
1st quartile	3084	952	37,825	4430	7824	79,031	88,376	11,060	29,423
2nd quartile	8900	5184	184,532	27,658	45,300	559,289	566,231	57,864	151,137
3rd quartile	102,738	44,899	788,590	182,050	339,226	3,971,509	3,188,962	352,378	1,272,411
Maximum	20,018,117	28,031,677	104,117,263	57,559,609	63,550,393	1,113,372,106	887,561,640	73,422,162	164,071,334
Mean	587,769	523,085	4,136,704	1,279,699	1,907,944	42,454,571	36,800,174	2,295,624	6,341,943
SD	2,207,552	2,217,200	14,238,836	4,567,619	6,356,927	143,155,173	123,979,458	7,477,161	20,947,952

Table 7 Statistics on the largest UK commercial banks in our dataset

Statistics	Inputs					Output			
	Personnel expenses	Fixed assets	Equity	Total interest expense	Total expenses not including personnel expense	Gross loans	Total customer deposits	Gross interest and dividend income	Total income
Minimum	399,483	406,371	4,013,775	302,794	769,471	51,423,159	28,122,407	1,204,932	3,390,042
1st quartile	2,603,403	2,244,140	21,507,171	5,308,666	10,213,000	275,469,780	259,803,728	13,710,215	40,757,916
2nd quartile	5,818,473	4,407,448	33,460,279	10,110,862	16,524,428	404,934,564	372,874,000	20,017,169	57,496,490
3rd quartile	8,911,865	7,688,636	63,639,245	15,534,106	25,620,037	680,819,266	547,769,283	30,769,484	93,498,785
Maximum	20,018,117	28,031,677	104,117,263	57,559,609	63,550,393	1,113,372,106	887,561,640	73,422,162	164,071,334
Mean	6,378,980	5,842,691	42,820,658	12,205,320	18,872,460	441,836,080	393,482,462	22,820,758	64,658,760
SD	4,833,962	5,421,576	28,886,641	10,099,372	12,450,209	262,830,397	213,471,361	13,940,763	39,175,418

Table 8 Additional analyses of the UK commercial banking sector

	# Banks	Percentage (%)	Personnel expenses	Percentage (%)	Fixed assets	Percentage (%)	Gross loans	Percentage (%)	Total customer deposits	Percentage (%)	Total income
All commercial banks	109	100	688,278,073	100	612,533,104	100	49,714,302,196	100	43,093,003,273	100	7,426,415,455
5 Largest UK banking groups	8	7	606,003,076	88	555,055,662	91	41,974,427,620	84	37,380,833,843	87	6,142,582,203
Local banks	68	62	644,949,192	94	577,851,194	94	45,167,923,292	91	39,630,979,746	92	6,793,674,300
Foreign banks	41	38	43,328,882	6	34,681,910	6	4,546,378,903	9	3,462,023,527	8	632,741,155
Local banks–largest banks	60	55	38,946,115	6	22,795,532	4	3,193,495,672	6	2,250,145,903	5	651,092,097

Table 9 Summary statistics on input-oriented scores of overall technical, pure technical and scale efficiencies without regression feedback

	Statistics on CCR-IO scores				Statistics on BCC-IO scores				Statistics on SE-IO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0318	0.033	0.0357	0.0632	0.044	0.0381	0.0358	0.0637	0.0641	0.2525	0.041	0.263
1st quartile	0.1937	0.3091	0.2908	0.4773	0.2845	0.3815	0.3904	0.5557	0.4721	0.6363	0.6609	0.8195
2nd quartile	0.2612	0.3902	0.418	0.6057	0.4423	0.5129	0.5441	0.6894	0.689	0.8529	0.854	0.9431
3rd quartile	0.3729	0.4999	0.5875	0.7514	0.7094	0.7578	0.7944	0.867	0.895	0.9665	0.9674	0.9887
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.3144	0.4297	0.4512	0.6119	0.5132	0.5733	0.591	0.6976	0.667	0.7846	0.7847	0.8796
SD	0.1978	0.1829	0.2203	0.1967	0.2716	0.237	0.2502	0.1965	0.2466	0.1997	0.2189	0.1473

Table 10 Summary statistics on output-oriented scores of overall technical, pure technical and scale efficiencies without regression feedback

	Statistics on CCR-OO scores				Statistics on BCC-OO scores				Statistics on SE-OO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	1	1	1	1	1	1	1	1	1	1	1	1
1st quartile	2.6815	2.0002	1.7022	1.3308	1.3562	1.3233	1.2774	1.1791	1.0808	1.0266	1.0326	1.0078
2nd quartile	3.8284	2.5626	2.3925	1.6511	2.4235	2.026	1.9246	1.5049	1.3652	1.1287	1.1204	1.0455
3rd quartile	5.162	3.2357	3.4386	2.0952	3.6867	2.7762	2.8629	1.919	2.0277	1.4916	1.3582	1.1365
Maximum	31.4064	30.3305	28.02	15.822	30.698	20.6932	18.6209	11.8291	12.0431	4.2375	14.2522	2.2664
Mean	4.3804	2.7956	3.1031	1.8955	2.9497	2.2414	2.43	1.7138	1.7408	1.3365	1.3369	1.1152
SD	2.9099	1.6866	2.5788	1.0475	2.3493	1.4107	1.8062	0.9311	1.0381	0.4509	0.8369	0.1736

Table 11 Summary statistics on SBM efficiency scores without regression feedback

	Statistics on SBM-IO scores				Statistics on SBM-OO scores				Statistics on SBM scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0011	0.0022	0.0012	0.0023	0.0001	0.0001	0.0002	0.0002	0.0001	0.0001	0.0002	0.0002
1st quartile	0.157	0.2528	0.2765	0.4	0.0933	0.1177	0.1406	0.1546	0.0703	0.1069	0.1246	0.1404
2nd quartile	0.2511	0.3572	0.4168	0.5547	0.2271	0.2559	0.299	0.3406	0.1664	0.2247	0.2716	0.3219
3rd quartile	0.4587	0.5607	0.6651	0.7169	0.5096	0.533	0.5514	0.6255	0.3947	0.4615	0.5262	0.5947
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.3577	0.4373	0.4807	0.5646	0.3366	0.3572	0.3829	0.4104	0.2898	0.3269	0.364	0.3931
SD	0.2868	0.2586	0.2771	0.2417	0.3091	0.3053	0.3023	0.3068	0.3054	0.2981	0.3035	0.303

Table 12 Summary statistics on input-oriented scores of overall technical, pure technical and scale efficiencies with input focused regression feedback

	Statistics on CCR-IO scores				Statistics on BCC-IO scores				Statistics on SE-IO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0211	0.0085	0.0357	0.0085	0.044	0.0091	0.0358	0.0091	0.0372	0.0939	0.0410	0.0939
1st quartile	0.1429	0.0855	0.2908	0.0855	0.2845	0.1129	0.3904	0.1129	0.3322	0.3385	0.6612	0.3385
2nd quartile	0.2042	0.1095	0.418	0.1095	0.4423	0.1811	0.5441	0.1811	0.4852	0.7046	0.8541	0.7046
3rd quartile	0.2931	0.1451	0.5875	0.1451	0.7094	0.3455	0.7944	0.3455	0.7027	0.9524	0.9674	0.9524
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.2424	0.1264	0.4512	0.1264	0.5132	0.276	0.591	0.276	0.5247	0.6477	0.7847	0.6477
SD	0.1633	0.0085	0.2203	0.0772	0.2716	0.2371	0.2502	0.2371	0.2408	0.3040	0.2189	0.3040

Table 13 Summary statistics on output-oriented scores of overall technical, pure technical and scale efficiencies with input focused regression feedback

	Statistics on CCR-OO scores				Statistics on BCC-OO scores				Statistics on SE-OO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	1	1	1	1	1	1	1	1	1	1	1	1
1st quartile	4.1603	6.8929	1.7022	6.8929	1.3562	2.3027	1.4465	2.3027	1.6245	1.0380	1.0328	1.0380
2nd quartile	5.8167	9.1365	2.3925	9.1365	2.4235	4.8168	2.4248	4.8168	2.3409	1.7774	1.1205	1.7774
3rd quartile	8.2587	11.6993	3.4386	11.6993	3.6867	8.3935	4.1161	8.3935	3.6181	3.7435	1.3580	3.7435
Maximum	94.342	117.8434	28.02	117.8434	30.698	49.6054	41.4108	49.6054	21.4901	10.6462	14.2522	10.6462
Mean	7.4828	9.9778	3.1031	9.9778	2.9497	5.9386	3.4743	5.9386	0.3354	2.5006	1.3369	2.5006
SD	6.2755	7.2698	2.5788	7.2698	2.3493	4.8377	3.6333	4.8377	0.3089	1.7216	0.8369	1.7216

Table 14 Summary statistics on SBM efficiency scores with input focused regression feedback

	Statistics on SBM-IO scores				Statistics on SBM-OO scores				Statistics on SBM scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0011	0.0022	0.0012	0.0001	0.0002	0.0001	0.0002	0.0002	0.0002	0.0001	0.0002	0.0002
1st quartile	0.1602	0.2528	0.2765	0.0386	0.0884	0.1177	0.1406	0.1546	0.0704	0.1069	0.1246	0.1221
2nd quartile	0.2449	0.3572	0.4168	0.0782	0.188	0.2559	0.299	0.3406	0.1478	0.2247	0.2716	0.3085
3rd quartile	0.4734	0.5607	0.6651	0.2046	0.4104	0.533	0.5514	0.6255	0.3566	0.4615	0.5262	0.5708
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.343	0.4373	0.4807	0.1817	0.2913	0.3572	0.3829	0.4104	0.2618	0.3269	0.364	0.3737
SD	0.2677	0.2586	0.2771	0.2404	0.2799	0.3053	0.3023	0.3068	0.2773	0.2981	0.3035	0.2955

Table 15 Summary statistics on input-oriented scores of overall technical, pure technical and scale efficiencies with output focused regression feedback

	Statistics on CCR-IO scores				Statistics on BCC-IO scores				Statistics on SE-IO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0011	0.023	0.0011	0.016	0.0358	0.0283	0.0358	0.0304	0.0099	0.1863	0.0099	0.1628
1st quartile	0.1617	0.2632	0.1617	0.4441	0.3581	0.3388	0.3581	0.4404	0.3574	0.4997	0.3574	0.8085
2nd quartile	0.2664	0.3271	0.2664	0.5736	0.5251	0.4492	0.5251	0.5759	0.5805	0.6676	0.5805	0.9329
3rd quartile	0.3837	0.4101	0.3837	0.7234	0.762	0.7008	0.762	0.7487	0.7440	0.8438	0.7440	0.9875
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.2937	0.358	0.2937	0.5788	0.5627	0.5278	0.5627	0.5964	0.5570	0.6633	0.5570	0.8693
SD	0.1781	0.1533	0.1781	0.1991	0.258	0.2403	0.258	0.2042	0.2468	0.2027	0.2468	0.1559

Table 16 Summary statistics on output-oriented scores of overall technical, pure technical and scale efficiencies with output focused regression feedback

	Statistics on CCR-OO scores				Statistics on BCC-OO scores				Statistics on SE-OO scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	1	1	1	1	1	1	1	1	1	1	1	1
1st quartile	2.6815	2.4386	2.6062	1.6638	1.4829	1.8382	1.323	1.8382	1.1480	1.1283	1.2511	1.1283
2nd quartile	3.8284	3.0572	3.7533	2.1944	2.6466	3.0555	2.0381	3.0555	1.4734	1.4679	1.6070	1.4679
3rd quartile	5.162	3.7992	6.191	2.8296	4.027	3.9226	3.1517	3.9226	2.2064	1.9900	2.4887	1.9900
Maximum	31.4064	43.5696	947.6238	257.9347	47.3749	350.5737	47.3749	350.5737	12.0431	5.3689	132.4496	5.3689
Mean	4.3804	3.3063	7.665	2.7958	3.3081	3.5811	2.8136	3.5811	1.8552	1.6342	2.6234	1.6342
SD	2.9099	2.0551	41.5639	7.9682	3.0596	10.8695	2.7962	10.8695	1.1070	0.6150	6.0733	0.6150

Table 17 Summary statistics on SBM efficiency scores with output focused regression feedback

	Statistics on SBM-IO scores				Statistics on SBM-OO scores				Statistics on SBM scores			
	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4	Setup 1	Setup 2	Setup 3	Setup 4
Minimum	0.0003	0.0022	0.0002	0.0000	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0002	0.0000
1st quartile	0.0965	0.2528	0.1421	0.0062	0.1084	0.0815	0.1606	0.0875	0.0474	0.1069	0.1385	0.0640
2nd quartile	0.1873	0.3572	0.2727	0.0397	0.2448	0.1891	0.3672	0.2726	0.1164	0.2247	0.3173	0.2040
3rd quartile	0.3776	0.5607	0.487	0.1580	0.4845	0.4292	0.6017	0.4653	0.3141	0.4615	0.5625	0.3701
Maximum	1	1	1	1	1	1	1	1	1	1	1	1
Mean	0.2956	0.4373	0.3435	0.1203	0.3375	0.2975	0.4112	0.3265	0.2446	0.3269	0.3837	0.2737
SD	1.3633	0.2586	0.2702	0.1795	0.2953	0.2919	0.3024	0.2824	0.2951	0.2981	0.3019	0.2641

Table 18 Summary of drivers of efficiency scores after regression feedback

Inputs	Input-focus regression analysis						
	CCR-IO	BCC-IO	CCR-OO	BCC-OO	SBM-IO	SBM-OO	SBM
Setup 1							
Personnel expenses	X	X		X	X	X	X
Fixed assets		X	X	X	X	X	X
Total interest expense	X	X	X	X	X	X	X
Setup 2							
Personnel expenses	X	X	X	X	X	X	X
Fixed assets					X	X	X
Total expenses not including personnel expense					X	X	X
Setup 3							
Personnel expenses	X	X	X	X	X	X	X
Equity	X	X	X	X	X	X	X
Total interest expense	X	X	X		X	X	X
Setup 4							
Personnel expenses	X	X	X	X	X	X	X
Equity						X	X
Total expenses not including personnel expense						X	

Table 18 continued

Outputs	Output-focus regression analysis						
	CCR-IO	BCC-IO	CCR-OO	BCC-OO	SBM-IO	SBM-OO	SBM
Setup 1							
Gross loans	X	X	X	X	X	X	X
Total customer deposits	X	X	X	X	X	X	X
Gross interest and dividend income	X		X				
Setup 2							
Gross loans	X	X	X	X	X	X	X
Total customer deposits	X	X	X		X	X	X
Total income					X		X
Setup 3							
Gross loans	X	X	X	X	X	X	X
Total customer deposits		X		X		X	X
Gross interest and dividend income							
Setup 4							
Gross loans	X	X	X	X		X	X
Total customer deposits	X						
Total income							

Table 19 Summary of CCR-IO efficiency scores for domestic and foreign banks

	CCR-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Set up 4	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Minimum	0.0211	0.0318	0.0085	0.0189	0.0357	0.0381	0.0085	0.0189
1st quartile	0.1439	0.1334	0.0855	0.0855	0.3193	0.2357	0.0855	0.0855
2nd quartile	0.2094	0.1974	0.1084	0.1114	0.4568	0.3536	0.1084	0.1114
3rd quartile	0.2978	0.2779	0.1420	0.1588	0.6216	0.5317	0.1420	0.1588
Maximum	1	1	1	1	1	1	1	1
Mean	0.2526	0.2233	0.1213	0.1361	0.4860	0.3862	0.1213	0.1361
SD	0.1760	0.1345	0.0606	0.1007	0.2163	0.2132	0.0606	0.1007

Table 20 Summary of BCC-IO efficiency scores for domestic and foreign banks

	BCC-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Setup 4	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Minimum	0.0440	0.0869	0.0091	0.0214	0.0358	0.0778	0.0091	0.0214
1st quartile	0.2962	0.2704	0.1118	0.1179	0.4216	0.3455	0.1118	0.1179
2nd quartile	0.4563	0.4322	0.1787	0.1843	0.5990	0.4898	0.1787	0.1843
3rd quartile	0.7545	0.6343	0.3529	0.3303	0.8609	0.6890	0.3529	0.3303
Maximum	1	1	1	1	1	1	1	1
Mean	0.5295	0.4828	0.2811	0.2667	0.6252	0.5270	0.2811	0.2667
SD	0.2784	0.2561	0.2466	0.2182	0.2519	0.2343	0.2466	0.2182

Table 21 Summary of SE-IO efficiency scores for domestic and foreign banks

	SE-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Setup 4	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Minimum	0.1027	0.0372	0.0939	0.1627	0.1742	0.0410	0.0939	0.1627
1st quartile	0.3190	0.3618	0.3293	0.3637	0.6802	0.6347	0.3293	0.3637
2nd quartile	0.4759	0.4960	0.7037	0.7140	0.8671	0.8282	0.7037	0.7140
3rd quartile	0.7191	0.6735	0.9458	0.9616	0.9640	0.9726	0.9458	0.9616
Maximum	1	1	1	1	1	1	1	1
Mean	0.5270	0.5204	0.6390	0.6640	0.7998	0.7565	0.6390	0.6640
SD	0.2414	0.2398	0.3084	0.2955	0.1975	0.2521	0.3084	0.2955

Table 22 Summary of SBM-io efficiency scores for domestic and foreign banks

	SBM-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Setup 4	
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign
Minimum	0.0011	0.0024	0.0072	0.0022	0.0035	0.0012	0.0001	0.0009
1st quartile	0.1711	0.1272	0.2747	0.2239	0.3180	0.1955	0.0443	0.0272
2nd quartile	0.2800	0.2148	0.3773	0.3130	0.4583	0.3628	0.0820	0.0678
3rd quartile	0.5165	0.3848	0.6080	0.4890	0.7197	0.5651	0.2151	0.1722
Maximum	1	1	1	1	1	1	1	1
Mean	0.3858	0.3017	0.4651	0.3853	0.5237	0.4005	0.1941	0.1586
SD	0.2919	0.2674	0.2591	0.2499	0.2754	0.2623	0.2499	0.2197

Table 23 Summary of CCR-IO efficiency scores for large and small banks

	CCR-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Set up 4	
	Large	Small	Large	Small	Large	Small	Large	Small
Minimum	0.1300	0.0211	0.0830	0.0085	0.3554	0.0357	0.0830	0.0085
1st quartile	0.2020	0.1369	0.1205	0.0833	0.4724	0.2809	0.1205	0.0833
2nd quartile	0.2681	0.1986	0.1423	0.1066	0.5468	0.3990	0.1423	0.1066
3rd quartile	0.3100	0.2853	0.1728	0.1407	0.6769	0.5766	0.1728	0.1407
Maximum	0.6683	1.0001	0.3385	1.0000	1.0000	1.0000	0.3385	1.0000
Mean	0.2731	0.2397	0.1507	0.1243	0.5681	0.4409	0.1507	0.1243
SD	0.0979	0.1676	0.0467	0.0790	0.1292	0.2237	0.0467	0.0790

Table 24 Summary of BCC-IO efficiency scores for large and small banks

	BCC-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Set up 4	
	Large	Small	Large	Small	Large	Small	Large	Small
Minimum	0.5697	0.0440	0.4789	0.0091	0.6457	0.0358	0.4789	0.0091
1st quartile	0.8125	0.2750	0.6691	0.1085	0.9045	0.3753	0.6691	0.1085
2nd quartile	0.9531	0.4145	0.7886	0.1676	0.9905	0.5117	0.7886	0.1676
3rd quartile	1.0000	0.6304	0.9426	0.2806	1.0000	0.7357	0.9426	0.2806
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	0.8988	0.4791	0.7891	0.2306	0.9313	0.5609	0.7891	0.2306
SD	0.1211	0.2543	0.1627	0.1830	0.0989	0.2369	0.1627	0.1830

Table 25 Summary of SE-IO efficiency scores for large and small banks

	SE-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Set up 4	
	Large	Small	Large	Small	Large	Small	Large	Small
Minimum	0.1311	0.0372	0.0939	0.1228	0.3636	0.0410	0.0939	0.1228
1st quartile	0.2375	0.3670	0.1737	0.4111	0.5094	0.7003	0.1737	0.4111
2nd quartile	0.2818	0.5177	0.1854	0.7496	0.6040	0.8827	0.1854	0.7496
3rd quartile	0.3369	0.7360	0.1936	0.9615	0.7120	0.9721	0.1936	0.9615
Maximum	0.6683	1.0001	0.3385	1.0000	1.0000	1.0000	0.3385	1.0000
Mean	0.3041	0.5443	0.1907	0.6882	0.6124	0.8000	0.1907	0.6882
SD	0.1002	0.2399	0.0368	0.2835	0.1291	0.2188	0.0368	0.2835

Table 26 Summary of SBM-IO efficiency scores for large and small banks

	SBM-IO input-focus regression feedback							
	Setup 1		Setup 2		Setup 3		Set up 4	
	Large	Small	Large	Small	Large	Small	Large	Small
Minimum	0.4181	0.0011	0.5055	0.0022	0.5229	0.0012	0.3192	0.0001
1st quartile	0.5547	0.1486	0.6706	0.2425	0.7071	0.2690	0.6217	0.0354
2nd quartile	0.8138	0.2337	0.8843	0.3404	0.9272	0.3983	0.7432	0.0709
3rd quartile	1.0000	0.3873	1.0000	0.5017	1.0000	0.6097	0.8952	0.1405
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	0.7857	0.3185	0.8266	0.4028	0.8538	0.4477	0.7386	0.1324
SD	0.2068	0.2602	0.1750	0.2356	0.1598	0.2606	0.1814	0.1735

References

Adler, N., & Golany, B. (2001). Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. *European Journal of Operational Research, 132,* 260–273.

Adler, N., & Yazhemsky, E. (2010). Improving discrimination in data envelopment analysis: PCA–DEA or variable reduction. *European Journal of Operational Research, 202,* 273–284.

Akther, S., Fukuyama, H., & Weber, W. L. (2013). Estimating two-stage network slacks-based inefficiency: An application to Bangladesh banking. *Omega, 41*(1), 88–96.

Alpay, O., & Akturk Hayat, E. (2017). Copula approach to select input/output variables for DEA. *An International Journal of Optimization and Control: Theories & Applications, 7*(1), 28–34.

Aly, H. Y., Grabowski, R., & Pasurka, C. (1990). Technical, scale, and allocative efficiencies in US banking: An empirical investigation. *The Review of Economics and Statistics, 72*(2), 211–218.

An, Q., Chen, H., Wu, J., & Liang, L. (2015). Measuring slacks-based efficiency for commercial banks in China by using a two-stage DEA model with undesirable output. *Annals of Operations Research, 235*(1), 13–35.

Arsad, R., Abdullah, M. N., Alias, S., & Isa, Z. (2017). Selection input output by restriction using DEA models based on a fuzzy Delphi approach and expert information. *Journal of Physics: Conference Series, 892*(1), 012010.

Avkiran, N. K. (2009). Removing the impact of environment with units-invariant efficient frontier analysis: An illustrative case study with intertemporal panel data. *Omega, 37*(3), 535–544.

Avkiran, N. K. (2015). An illustration of dynamic network DEA in commercial banking including robustness tests. *Omega, 55,* 141–150.

Avkiran, N. K., & Goto, M. (2011). A tool for scrutinizing bank bailouts based on multi-period peer benchmarking. *Pacific-Basin Finance Journal, 19*(5), 447–469.

Banker, R. D. (1996). Hypothesis tests using data envelopment analysis. *The Journal of Productivity Analysis, 7,* 139–159.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30,* 1078–1092.

Barr, R. S., Seiford, L. M., & Siems, T. F. (1993). An envelopment-analysis approach to measuring the managerial efficiency of banks. *Annals of Operations Research, 45,* 1–19.

Barr, R. S., Seiford, L. M., & Siems, T. F. (1994). Forecasting bank failure: A non-parametric frontier estimation approach. *Louvain Economic Review, 60*(4), 417–429.

Barr, R. S., & Siems, T. F. (1997). Bank failure prediction using DEA to measure management quality. In R. S. Barr, R. V. Helgason & J. L. Kennington (Eds.), *Interfaces in computer science and operations research. Operations Research/Computer Science Interfaces Series* (Vol. 7). Boston, MA: Springer.

Berger, A. N., & Mester, L. J. (2003). Explaining the dramatic changes in performance of US banks: Technological change, deregulation, and dynamic changes in competition. *Journal of Financial Intermediation, 12,* 57–95.

Bhattacharyya, A., Lovell, C. K., & Sahay, P. (1997). The impact of liberalization on the productive efficiency of Indian commercial banks. *European Journal of Operational Research, 98,* 332–345.

Burrows, O., Low, K., & Cumming, F. (2015). Mapping the UK financial system. *Bank of England, Quarterly Bulletin, Q2,* 114–129.

Bush, O., Knott, S., & Peacock, C. (2014). Why is the UK banking system so big and is that a problem? *Bank of England, Quarterly Bulletin, Q4,* 385–395.

Casu, B., & Molyneux, P. (2003). A comparative study of efficiency in European banking. *Applied Economics, 35*(17), 1865–1876.

Chao, C. M., Yu, M. M., & Wu, H. N. (2015). An application of the dynamic network DEA model: The case of banks in Taiwan. *Emerging Markets Finance and Trade, 51*(sup1), S133–S151.

Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *Journal of econometrics, 46*(1–2), 73–91.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2,* 429–444.

Chen, T.-Y. (1998). A study of bank efficiency and ownership in Taiwan. *Applied Economics Letters, 5,* 613–616.

Chu, S. F., & Lim, G. H. (1998). Share performance and profit efficiency of banks in an oligopolistic market: Evidence from Singapore. *Journal of Multinational Financial Management, 8,* 155–168.

Cinca, C. S., & Molinero, C. M. (2004). Selecting DEA specifications and ranking units via PCA. *Journal of the Operational Research Society, 55,* 521–528.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software* (2nd ed.). Berlin: Springer.

Das, A., & Ghosh, S. (2006). Financial deregulation and efficiency: An empirical analysis of Indian banks during the post reform period. *Review of Financial Economics, 15*(3), 193–221.

Davies, R., Richardson, P., Katinaite, V., & Manning, M. (2010). Evolution of the UK banking system. *Bank of England, Quarterly Bulletin, Q4,* 321–332.

Drake, L. (2001). Efficiency and productivity change in UK banking. *Applied Financial Economics, 11,* 557–571.

Drake, L., Hall, M. J., & Simper, R. (2006). The impact of macroeconomic and regulatory factors on bank efficiency: A non-parametric analysis of Hong Kong's banking system. *Journal of Banking & Finance, 30,* 1443–1466.

Du, K., Worthington, A. C., & Zelenyuk, V. (2018). Data envelopment analysis, truncated regression and double-bootstrap for panel data with application to Chinese banking. *European Journal of Operational Research, 265*(2), 748–764.

Elyasiani, E., & Mehdian, S. M. (1990). A nonparametric approach to measurement of efficiency and technological change: The case of large US commercial banks. *Journal of Financial Services Research, 4,* 157–168.

Elyasiani, E., & Mehdian, S. (1992). Productive efficiency performance of minority and nonminority-owned banks: A nonparametric approach. *Journal of Banking & Finance, 16,* 933–948.

Eskelinen, J. (2017). Comparison of variable selection techniques for data envelopment analysis in a retail bank. *European Journal of Operational Research, 259*(2), 778–788.

Fanchon, P. (2003). Variable selection for dynamic measures of efficiency in the computer industry. *International Advances in Economic Research, 9*(3), 175–188.

Favero, C. A., & Papi, L. (1995). Technical efficiency and scale efficiency in the Italian banking sector: A non-parametric approach. *Applied Economics, 27,* 385–395.

Ferrier, G. D., & Hirschberg, J. G. (1997). Bootstrapping confidence intervals for linear programming efficiency scores: With an illustration using Italian banking data. *Journal of Productivity Analysis, 8*(1), 19–33.

Ferrier, G. D., & Lovell, C. K. (1990). Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of Econometrics, 46,* 229–245.

Fethi, M. D., & Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European Journal of Operational Research, 204,* 189–198.

Fixler, D. J., & Zieschang, K. D. (1992). User costs, shadow prices, and the real output of banks. In *Output measurement in the service sectors* (pp. 219–243). University of Chicago Press.

Fukuyama, H. (1993). Technical and scale efficiency of Japanese commerical banks: A non-parametric approach. *Applied Economics, 25,* 1101–1112.

Fukuyama, H., & Matousek, R. (2017). Modelling bank performance: A network DEA approach. *European Journal of Operational Research, 259*(2), 721–732.

Fukuyama, H., & Weber, W. L. (2015). Measuring Japanese bank performance: A dynamic network DEA approach. *Journal of Productivity Analysis, 44*(3), 249–264.

Fukuyama, H., & Weber, W. L. (2017). Measuring bank performance with a dynamic network Luenberger indicator. *Annals of Operations Research, 250*(1), 85–104.

Fukuyama, H., & Weber, W. L. (2017b). Japanese bank productivity, 2007–2012: A dynamic network approach. *Pacific Economic Review, 22*(4), 649–676.

Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega, 17*(3), 237–250.

Grabowski, R., Rangan, N., & Rezvanian, R. (1993). Organizational forms in banking: An empirical investigation of cost efficiency. *Journal of Banking & Finance, 17,* 531–538.

Grigoroudis, E., Tsitsiridi, E., & Zopounidis, C. (2013). Linking customer satisfaction, employee appraisal, and business performance: An evaluation methodology in the banking sector. *Annals of Operations Research, 205*(1), 5–27.

Guarda, P., Rouabah, A., & Vardanyan, M. (2013). Identifying bank outputs and inputs with a directional technology distance function. *Journal of Productivity Analysis, 40*(2), 185–195.

Gulati, R., & Kumar, S. (2017). Analysing banks' intermediation and operating efficiencies using the two-stage network DEA model: The case of India. *International Journal of Productivity and Performance Management, 66*(4), 500–516.

Halkos, G. E., & Salamouris, D. S. (2004). Efficiency measurement of the Greek commercial banks with the use of financial ratios: A data envelopment analysis approach. *Management Accounting Research, 15,* 201–224.

Hancock, D. (1985a). Bank profitability, interest rates, and monetary policy. *Journal of Money, Credit and Banking, 17*(2), 189–202.

Hancock, D. (1985b). The financial firm: Production with monetary and nonmonetary goods. *Journal of Political Economy, 93*(5), 859–880.

Hatami-Marbini, A., Agrell, P. J., Tavana, M., & Khoshnevis, P. (2017). A flexible cross-efficiency fuzzy data envelopment analysis model for sustainable sourcing. *Journal of Cleaner Production, 142*, 2761–2779.

Jain, R. K., Natarajan, R., & Ghosh, A. (2016). Decision tree analysis for selection of factors in DEA: An application to banks in India. *Global Business Review, 17*(5), 1162–1178.

Jenkins, L., & Anderson, M. (2003). A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research, 147*(1), 51–61.

Kaffash, S., & Marra, M. (2017). Data envelopment analysis in financial services: A citations network analysis of banks, insurance companies and money market funds. *Annals of Operations Research, 253*(1), 307–344.

Kao, C., & Liu, S. T. (2009). Stochastic data envelopment analysis in measuring the efficiency of Taiwan commercial banks. *European Journal of Operational Research, 196*(1), 312–322.

Kutlar, A., Kabasakal, A., & Ekici, M. S. (2017). Efficiency of commercial banks in Turkey and their comparison: Application of DEA with Tobit analysis. *International Journal of Mathematics in Operational Research, 10*(1), 84–103.

Kwon, H.-B., & Lee, J. (2015). Two-stage production modeling of large U.S. banks: A DEA-neural network approach. *Expert Systems with Applications, 42*(19), 6758–6766.

Lee, Y. J., Joo, S.-J., & Park, H. G. (2017). An application of data envelopment analysis for Korean banks with negative data. *Benchmarking: An International Journal, 24*(4), 1052–1064.

Lewin, A. Y., Morey, R. C., & Cook, T. J. (1982). Evaluating the administrative efficiency of courts. *Omega, 10*(4), 401–411.

Li, Y., Shi, X., Yang, M., & Liang, L. (2017). Variable selection in data envelopment analysis via Akaike's information criteria. *Annals of Operations Research, 253,* 453–476.

Lim, S. (2008). A decision tree-based method for selection of input–output factors in DEA. In *Proceedings of the 2008 international conference on artificial intelligence, ICAI 2008*, 14–17 July 2008.

Liu, H.-H. (2018). Applying three-stage DEA on the operational performance of foreign banks in Taiwan. *International Review of Applied Economics, 32*(1), 104–118.

Liu, J., & Tone, K. (2008). A multistage method to measure efficiency and its application to Japanese banking industry. *Socio-Economic Planning Sciences, 42,* 75–91.

Liu, W. B., Zhang, D. Q., Meng, W., Li, X. X., & Xu, F. (2011). A study of DEA models without explicit inputs. *Omega: The International Journal of Management Science, 39*(5), 472–480.

Liu, W., Zhou, Z., Ma, C., Liu, D., & Shen, W. (2015). Two-stage DEA models with undesirable input-intermediate-outputs. *Omega, 56,* 74–87.

Lovell, C. A. K., & Pastor, J. T. (1999). Radial DEA models without inputs or without outputs. *European Journal of Operational Research, 118,* 46–51.

Lozano-Vivas, A., Pastor, J. T., & Pastor, J. M. (2002). An efficiency comparison of European banking systems operating under different environmental conditions. *Journal of Productivity Analysis, 18*(1), 59–77.

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*(1), 1118–1123.

Madhanagopal, R., & Chandrasekaran, R. (2014). Selecting appropriate variables for DEA using genetic algorithm (GA) search procedure. *International Journal of Data Envelopment Analysis and Operations Research, 1*(2), 28–33.

Matthews, K. (2013). Risk management and managerial efficiency in Chinese banks: A network DEA framework. *Omega, 41*(2), 207–215.

Miller, S. M., & Noulas, A. G. (1996). The technical efficiency of large bank production. *Journal of Banking & Finance, 20,* 495–509.

Nataraja, N. R., & Johnson, A. L. (2011). Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research, 215,* 662–669.

Norman, M., & Stoker, B. (1991). *Data envelopment analysis: The assessment of performance*. London: Wiley.

Nunamaker, T. R. (1985). Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation. *Managerial and decision Economics, 6*(1), 50–58.

Olson, J. A., Schmidt, P., & Waldman, D. A. (1980). A Monte Carlo study of estimators of stochastic frontier production functions. *Journal of Econometrics, 13,* 67–82.

Ouenniche, J., Carrales, S., Fukuyama, H., & Tone, K. (2017). An account of DEA-based contributions in the banking sector. In K. Tone (Ed.), *Advances in DEA theory and applications: With examples in forecasting models*. Wiley series in operations research and management science. ISBN: 978-1-118-94562-9.

Pasiouras, F. (2008). 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, 30*(2), 187–223.

Pastor, J. M. (2002). Credit risk and efficiency in the European banking system: A three-stage analysis. *Applied Financial Economics, 12,* 895–911.
Pastor, J., Perez, F., & Quesada, J. (1997). Efficiency analysis in banking firms: An international comparison. *European Journal of Operational Research, 98,* 395–407.
Rangan, N., Grabowski, R., & Aly, H. Y. (1988). The technical efficiency of US banks. *Economics Letters, 28,* 169–175.
Ruggiero, J. (2005). Impact assessment of input omission on DEA. *International Journal of Information Technology and Decision Making, 4*(03), 359–368.
Sharma, M. J., & Yu, S. J. (2015). Stepwise regression data envelopment analysis for variable reduction. *Applied Mathematics and Computation, 253*, 126–134.
Sigala, M., Airey, D., Jones, P., & Lockwood, A. (2004). ICT paradox lost? A stepwise DEA methodology to evaluate technology investments in tourism settings. *Journal of Travel Research, 43*(2), 180–192.
Simar, L., & Wilson, P. W. (2001). Testing restrictions in nonparametric efficiency models. *Communications in Statistics-Simulation and Computation, 30*(1), 159–184.
Subramanyam, T. (2016). Selection of input–output variables in data envelopment analysis: Indian commercial banks. *International Journal of Computer & Mathematical Sciences, 5*(6), 2347–8527.
Tanna, S., Pasiouras, F., & Nnadi, M. (2011). The effect of board size and composition on the efficiency of UK banks. *International Journal of the Economics of Business, 18,* 441–462.
Taylor, W. M., Thompson, R. G., Thrall, R. M., & Dharmapala, P. (1997). DEA/AR efficiency and profitability of Mexican banks a total income model. *European Journal of Operational Research, 98*(2), 346–363.
Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research, 130,* 498–509.
Ueda, T., & Hoshiai, Y. (1997). Application of principal component analysis for parsimonious summarization of DEA inputs and/or outputs. *Journal of the Operations Research Society of Japan, 40*(4), 466–478.
Wagner, J. M., & Shimshak, D. G. (2007). Stepwise selection of variables in data envelopment analysis: Procedures and managerial perspectives. *European Journal of Operational Research, 180*(1), 57–67.
Wang, W. K., Lu, W. M., & Liu, P. Y. (2014). A fuzzy multi-objective two-stage DEA model for evaluating the performance of US bank holding companies. *Expert Systems with Applications, 41*(9), 4290–4297.
Wanke, P., & Barros, C. (2014). Two-stage DEA: An application to major Brazilian banks. *Expert Systems with Applications, 41*(5), 2337–2344.
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*(1), 378–389.
Webb, R. (2003). Levels of efficiency in UK retail banks: A DEA window analysis. *International Journal of the Economics of Business, 10,* 305–322.
Webb, R., Bryce, C., & Watson, D. (2010). The effect of building society demutualisation on levels of relative efficiency at large UK commercial banks. *Journal of Financial Regulation and Compliance, 18,* 333–355.
Wu, Y.-C., Ting, I. W. K., Lu, W.-M., Nourani, M., & Kweh, Q. L. (2016). The impact of earnings management on the performance of ASEAN banks. *Economic Modelling, 53,* 156–165.
Yue, P. (1992). Data envelopment analysis and commercial bank performance: A primer with applications to Missouri banks. *Federal Reserve Bank of St. Louis Review, 74,* 31–45.
Zaim, O. (1995). The effect of financial liberalization on the efficiency of Turkish commercial banks. *Applied Financial Economics, 5,* 257–264.
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.