



Incorporating causal modeling into data envelopment analysis for performance evaluation

Hirofumi Fukuyama¹ · Mike Tsionas^{2,3,4} · Yong Tan⁵

Accepted: 20 June 2023
© The Author(s) 2023

Abstract

The risk factors in banking have been considered an undesirable carryover variable by the literature. Methodologically, we consider the risk factor using loan loss reserves as a desirable carryover input with dynamic characteristics, which provides a new framework in the dynamic network Data Envelopment Analysis (DEA) modelling. We substantiate our formulation and results using novel techniques for causal modelling to ensure that our dynamic network model admits a causal interpretation. Finally, we empirically examine the impact of risk from various economic sectors on efficiency. Our results show that the inefficiencies were volatile in Chinese banking over the period 2013–2020, and we further find that the state-owned banks experienced the highest levels of inefficiency and volatility. The findings report that credit risk derived from the agricultural sector and the Water Conservancy, Environment and Public Facilities management sector decreases bank efficiency, while credit risk derived from the wholesale and retail sector improves bank efficiency. The results of our innovative causal modelling show that our pioneering modelling on the role of loan loss reserves is valid. In addition, from an empirical perspective, our second-stage analysis regarding the impact of risk derived from different economic sectors on bank efficiency can be applied to other banking systems worldwide because of our successful validation from causal modelling. Our attempt to incorporate causal inference into DEA can be generalized to future studies of using DEA for performance evaluation.

Keywords Data envelopment analysis · Dynamic inefficiency · Causal modelling · Two-stage network · Chinese banks

✉ Yong Tan
y.tan9@bradford.ac.uk
Hirofumi Fukuyama
fukuyama@fukuoka-u.ac.jp

¹ Department of Business Management, Faculty of Commerce, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 8140180, Japan

² Montpellier Business School, 2300 Avenue des Moulins, 34080 Montpellier, France

³ Lancaster University Management School, Lancaster LA1 4YX, UK

⁴ Montpellier Research in Management (MRM; EA 4557) University Montpellier, Montpellier, France

⁵ School of Management, University of Bradford, Bradford, West Yorkshire BD7 1DP, UK

1 Introduction

The operational research in bank efficiency has undergone rapid development. Lots of advanced non-parametric methods and parametric methods have been proposed and used in estimating efficiency in the banking industry. These methods include, but are not limited, to: Bayesian stochastic frontier analysis; network Data Envelopment Analysis; quantile regression; hyperbolic measurement; fuzzy super-efficiency; hybrid minimax reference point-DEA; conditional directional distance approach; satisficing DEA approach; and parallel production frontier approach, among others. In general, Bayesian stochastic frontier analysis and quantile regression analysis belong to the parametric methods, the rest of the above-mentioned methods are the examples of the non-parametric methods. The Bayesian stochastic frontier analysis specifies the functional form in estimating efficiency; therefore, the main disadvantage of this method lies in the fact that the efficiency results would be sensitive to different functional forms. The quantile regression analysis estimates the production process for benchmark banks located at top conditional quantiles, however, it suffers from the similar disadvantage. The hyperbolic measurement has a higher ability to model undesirable outputs, and the fuzzy super-efficiency model benefits from the advantage of being able to provide a better ranking among banks. The conditional directional distance approach can estimate bank efficiency while accounting for the time effect. The hybrid minimax reference point-DEA can measure efficiency from the perspective that each bank branch is regarded to be unique by itself, and different branches can have different sizes and targets for specific market segments. The satisficing DEA approach can incorporate risk factors as inputs in the bank efficiency analysis, while the parallel frontiers provide consistent measures of technical changes for all Decision-Making Units (DMUs) at different periods. The Malmquist Productivity Indices (MPIs) and technical changes obtained in this way have the property of circularity. Finally, the network DEA model benefits from the advantage of being able to open the black box and better identify the source of inefficiency in the production process. It would be able to consider different aspects of the banking production process carefully and comprehensively through incorporating carryover variables in designing the production process (Zhu, 2022), which is also one of the advantages of the non-parametric methods over the parametric ones. In addition to the estimation of bank efficiency, DEA has been used to evaluate merger gains in the banking industry (Amin & Boamah, 2020). Not only in the banking sector, but DEA has been proposed and applied to different economic sectors or different types of DMUs from different perspectives, this includes the proposal of a hybrid DEA-Machine learning approach for predicting performance of micro, small and medium enterprises (Boubaker et al., 2023); the estimation of managerial ability of listed firms (Dalwali et al., 2023); the development of DEA in supplier selection (Dutta et al., 2022); the evaluation of healthcare system's efficiency under DEA and compromise programming DEA (Habib & Shawan, 2020; Lozano et al., 2020; Mourad et al., 2021; Rouyendegh et al., 2019). For more detailed studies regarding DEA and its applications in operations and data analytics, please refer to Chen et al. (2019).

Few attempts have been made to consider the "carryover" characteristics in the banking production process using DEA. Among the DEA studies incorporating carryover variables in the production process, two groups can be classified: one of them treats assets as one of the variables with the carryover characteristics, and the other group treats "credit risk" as the carryover variable. As can be seen in Sect. 2, the theoretical literature review, previous studies have used two credit risk indicators as the carryover variables: non-performing loans and loan loss reserves. Both of these indicators have been treated as the carryover variables with an

undesirable characteristic. This is mainly due to the consideration that non-performing loans are the type of loans that are lent by the banks, but the borrowers do not pay back to the banks when the loans become due. Therefore, non-performing loans are a type of bank risk that banks try to minimize. In terms of loan loss reserves, they have a close linkage with non-performing loans by the fact that larger volumes of loan loss reserves will be made if there is an expectation that there would be an increase in the volumes of non-performing loans. Thus, lower levels of loan loss reserves are pursued by the banks in their operation.

Our study considers credit risk in a different way. Although non-performing loans are an indicator of instability, banks take different measures to absorb this negative shock. A certain amount of loan loss provisions is set aside by the banks every year, and loan loss reserves are the accumulation of loan loss provisions over several years (Monokroussos et al., 2017). Therefore, from this perspective, setting aside higher volumes of loan loss reserves can be understood as a way to improve bank stability. The relationship between loan loss reserves and the risk of bank failure is further confirmed by Bushman and Williams (2012) as well as Ng and Roychowdhury (2014).

Not only can the desirability of loan loss reserves be reflected by their impact on credit risk, but loan loss reserves are also found to have a positive impact on bank earnings (Beaver & Engel, 1996). Therefore, our proposal of treating loan loss reserves with a desirable characteristic can be validated by the literature, while we build on the literature by incorporating it into the production process under a dynamic nature.

As an important component in the banking production process, the role played by deposits has been a debatable issue among academic researchers. Berger and Humphrey (1992) provided a systematic and comprehensive discussion related to the role of deposits from different perspectives. The asset approach in defining bank outputs argues that bank deposits and other liabilities should be treated as bank inputs due to the consideration that they are the raw materials in generating loans and other banking assets. In comparison, the user cost approach argues that the treatment of deposits as inputs or outputs depends on their net contribution to bank revenue. If the financial cost of deposits is higher than the opportunity cost, they will be considered inputs. If the opportunity cost is higher, deposits should be treated as outputs.

Thirdly, the value-added approach argues that all bank liabilities and assets have some output characteristics, and there is no clear boundary in defining the difference between inputs and outputs. Finally, explicit revenue in banking argues that banks accumulate implicit revenues because of the payment of below-market interest rates. This means that substantial service outputs are generated by deposits. However, because the explicit revenue generated by deposits is very small, there is still a high level of controversy as to whether to treat deposits as inputs or outputs.

Although several rounds of banking reforms in China aimed to improve competitive conditions and bank stability (Tan, 2016), the non-performing loan ratios in the Chinese banking industry consistently increased from 2013 to 2019, with the highest ratio of 1.9%. In comparison, other countries, including those in Asia and Europe, had much lower non-performing loan ratios. For example, Singapore had non-performing loan ratios of no more than 1.4% over the same period. The United Kingdom experienced a significant drop in the non-performing loan ratio from 2013 to 2015, and afterwards, although the ratio suffered from some volatility, it remained no more than 1.2%. Finally, South Korea consistently kept its non-performing loan ratios no higher than 0.6%. These statistics and relevant comparisons show that the Chinese banking industry has higher levels of instability. As Tan and Floros (2013) argue, improving bank efficiency will improve bank stability. Therefore, investigating efficiency in the Chinese banking context is essential.

Over the past few years, several studies have examined efficiency in the Chinese banking context (2022b; Antunes et al., 2022; Fukuyama & Tan, 2021a, 2021b; Fukuyama et al., 2022a; Tan & Tsionas, 2022; Tan et al., 2021), and relevant methodological contributions have been made. However, little attempt has yet been made to clearly and explicitly discuss the concept and importance of loan loss reserves in the banking production process, particularly the dynamic nature of this variable, and no effect has yet been given to correctly incorporate this essential factor in the modeling framework. In addition, the empirical banking studies place too much emphasis on the bank-specific determinants of efficiency, with no consideration given to examining the interaction between the banking industry and the operational performance of various economic sectors. Considering the nature of banking services, which cover loan services to various industries, examining the impact of non-performing loans derived from different economic sectors on bank performance will not only significantly contribute to the bank literature but will also provide more tailored policy implications in terms of credit allocations to different economic sectors.

The current study contributes to the literature as follows: (1) Loan loss reserves, rather than being regarded as an undesirable input or a carryover output variable, are treated as a "desirable input" with carryover characteristics because they provide stability to bank operations. We propose that loan loss reserves are produced at the first stage along with the intermediate products of loans and securities investment. Loan loss reserves affect the production at the second stage of the production process in the next period, not the current one. (2) In addition to considering the role played by loan loss reserves under a network DEA model, we engage in a second-phase analysis examining the influence of industry-level risk on efficiency in the Chinese banking sector. No previous studies in banking have addressed this issue. We justify our method by providing a causal interpretation within a dynamic network framework.

Establishing causal relationships without experimental data is known to be difficult (Pearl, 2009; Peters et al., 2013, 2017; Pfister et al., 2019). For causal models, it is well-known that changes in the environment or active interventions in the covariates should not affect their prediction properties since the effect of confounding variables has been considered. Without correctly establishing a causal relationship, the results associated with technical efficiency would be spurious and misleading. Our paper proposes a new technique that facilitates formal statistical testing for the causal interpretation of a given non-parametric model with a dynamic framework. The results of this causal modelling show that our proposed framework, considering the carryover characteristics of loan loss reserves in the banking production process, is valid. This indicates that our theoretical and innovative proposal regarding the dynamic role played by loan loss reserves as the desirable inputs can be supported. In addition, the causal interpretation for our second-stage regression analysis is also accepted and shows that our modelling, as well as the estimated results, are reasonable, reliable, and robust. Specifically, although no study has empirically investigated the impact of risk derived from different economic sectors on bank efficiency, this cannot necessarily lead to an examination of this issue without proper validation. In addition to our theoretical arguments, our causal modelling framework supports our innovative proposal on the investigation of this topic. In summary, besides the purpose of robustness check and validation, facilitated by our innovative causal modelling techniques, more generally, our study provides a pioneering example of including the operational research method and statistical modelling together to address the banking issue from both the perspectives of efficiency analysis and empirical banking. For all the efficiency estimation studies in the future, more efforts should be given to the validation of the relevant proposed modelling framework in the production process. For the empirical

banking analysis, in addition to relevant econometrics techniques, our statistical causal modelling framework is recommended to be incorporated into the analysis for validation. We set out the remainder of the current paper in the following way: Sect. 2 reviews the literature. Section 3 presents our research design. The empirical results and discussion are provided in Sect. 4. Finally, the concluding remarks are given in Sect. 5.

2 Literature review

2.1 Contextual background

A country's economy is composed of the public and private sectors. The public sector enjoys an advantage over the private sector in terms of receiving stronger government support, especially in investment and financial needs. In contrast, a lack of funding is one of the main obstacles faced by the private sector, which hinders its investment activities. The private sector is characterized by a large number of small and medium-sized enterprises in different economic sectors. According to Tan (2017), these enterprises provided 70% of employment and contributed 60% of China's GDP in 2012, but only received 30% of bank loans. Based on data from the World Bank Open Data between 2013 and 2020, the volume of domestic credits provided to the private sector by banks in China has generally increased, with the volume of credits provided in 2020 being 1.82 times the country's GDP.

The lack of funding provided by banks to the private sector is mainly due to the small and medium-sized enterprises' lack of capital and the uncertainty in their operations. This presents a level of risk when granting credits to these enterprises, resulting in non-performing loans that have a significant impact not only on banks but also on the operation and development of other economic sectors. The risk derived from the volume of non-performing loans is a historical issue in the Chinese banking industry, which the Chinese government has addressed through various measures, including capital injection, write-off of non-performing loans, establishment of four asset management companies, and the China Banking Regulatory Commission (Tan, 2016). However, academic scholars have not investigated the source of non-performing loans, particularly the distribution of non-performing loans across different economic sectors and their impact on banks and the economy.

Apart from the issue of non-performing loans, the Chinese government has implemented relevant measures to enhance bank performance and competitiveness. The current Chinese banking industry structure includes banks with different ownership types, among which state-owned commercial banks, joint-stock commercial banks, and city commercial banks are the three largest ownership groups in terms of total assets. Competition among banks of the same ownership type is expected to have a positive impact on performance. The Chinese government has also introduced foreign strategic investors and allowed initial public offerings of domestic Chinese banks, among other measures (Tan, 2014).

2.2 Theoretical literature review

Originally, the static DEA model was proposed by Ferrier (1994) to investigate and compare the efficiency levels between proprietary and cooperative firms. This work can be regarded as one example of how DEA can be applied specifically to estimate the efficiency level of firms with different ownership types and/or characteristics. Shephard and Färe (1980) developed dynamic production theory by presenting an axiomatic analysis that allows for the

inventories of standard inputs, intermediate products, and the final outputs. Following this study, Hackman (1990) extended this analysis and defined a dynamic production system as a network of production activities that produce the final outputs.

Regarding Hackman's analysis, Färe and Grosskopf (1996a, 1996b) presented a dynamic DEA framework that explicitly connects a sequence of single-period technologies and allows the producers to either produce the final outputs or carry over some outputs to adjust production in the subsequent period. In Färe and Grosskopf (1996a, 1996b)'s framework, network DEA can be thought of as the formulation of intertemporal production that allows each time-specific activity to represent the same production unit in different time periods. Their framework can be regarded as the foundation of dynamic network DEA since some final outputs or intermediate products in period t are carried over to period $t + 1$ as inventories.

One study related to dynamic DEA other than that of Färe and Grosskopf (1996a, 1996b) is Sengupta (1995), who developed the adjustment cost-based analysis by seeking to decide the optimal levels of inputs over a time horizon and estimate the optimal inputs to calculate the level of overall efficiency. Another study is the cost function-based work of Nemoto and Goto (2003), who provided optimal control-theoretic models for examining the dynamic DEA efficiency of public entities. Tone and Tsutsui (2014) proposed a slacks-based dynamic network DEA approach that allows carryovers and classifies various kinds of production variables.

Applying the idea of Färe and Grosskopf (1996a, 1996b)'s theoretical approach into bank efficiency measurement, Fukuyama and Weber (2015) allow a bank to decrease or increase the current production of loans and securities investments by controlling the amount of carryovers in a future period. More recently, using the trade-off relationship (Eq. 4), Fukuyama et al., (2022a, 2022b) examined the Chinese bank efficiency with a sequential structure and a behavioral causal analysis.

While we acknowledge Tone and Tsutsui (2014) as a contribution to DEA, their approach differs from that of Färe and Grosskopf (1996a, 1996b) in the sense that they do not have a constraint related to Eq. (2). Therefore, our dynamic network model is directly in line with Färe and Grosskopf (1996a, 1996b). Related to dynamic DEA, there are two useful survey papers: One is Fallah-Fini et al. (2014), who identified five primary characteristics: (i) production delays; (ii) inventories; (iii) quasi-fixed of physical capital leading to embodied technical change; (iv) adjustment costs; and (v) learning models dealing with disembodied technical change. Another survey article on dynamic efficiency is Mariz et al., (2018), which focused on intermediate and carryover variables in dynamic DEA. According to Mariz et al., (2018), the intertemporal variables are: (i) intertemporal inputs and outputs; (ii) storable inputs; (iii) quasi-fixed inputs; (iv) cost of adjustments; (v) lagged productive effects; and (vi) carryovers. Our approach considers (v) and (vi).

A dynamic model under a two-stage framework was proposed by Fukuyama and Weber (2015). The proposed model considered undesirable inputs, carryover inputs, and carryover outputs in addition to the original inputs, intermediate output, and outputs with desirable and undesirable characteristics. More specifically, labor, capital, and equity of the current year were used as the initial inputs, together with the undesirable input (non-performing loans) of the previous year ($t-1$), to produce two intermediate products in year t , including (1) deposits, and (2) other raised funds. These two intermediate outputs produced in year t at stage 1 were used as inputs, together with two carryover inputs in year $t-1$ (loans and securities), to produce three different types of final outputs in stage 2. These are desirable outputs in year t (loans and securities), carryover outputs in year t (carryover assets), and an undesirable output in year t (non-performing loans).

Another study by Bansal et al. (2022) proposed the use of dynamic Luenberger Productivity Indices and applied them to the Indian banking industry. The dynamic nature of the modelling framework was reflected by the inclusion of two desirable carryovers, namely unused assets and net profits, from period t to period $t + 1$. Non-performing loans were considered as the undesirable carryover input from period t to $t + 1$, while the study did not account for the dynamic nature of loan loss provisions and instead treated them as a desirable input. Efforts have been made to investigate the role of loan loss reserves as a carryover output in the banking production process (Chao et al., 2015).

Moreover, Fukuyama et al., (2023) proposed a dynamic DEA behavioural model and applied it to the Chinese banking industry. The model not only considered the dual role of intermediate and final outputs in a dynamic manner, but also treated loan loss reserves as the carryover generated at the second stage of the production process in time period t and used it as the carryover input in the first-stage production in time period $t + 1$. In this study, we differentiate ourselves from Fukuyama et al., (2023) by more explicitly considering the desirability and positive role played by loan loss reserves in the banking production process. Instead of considering the carryover input of loan loss reserves in the first-stage production in time period $t + 1$, we treat loan loss reserves in period t as a carryover variable in the first stage of production, along with the intermediate products generated in time period $t + 1$, to generate the final outputs. This modelling framework is designed to consider that the generation of final outputs (income) depends on loans, which are supported by loan loss reserves.

Fukuyama et al., (2023) not only contributed to the DEA banking literature by proposing an innovative behavioural model, but they were also pioneers in proposing a causal analysis to validate the framework of the production process. Their work provided a good example for future DEA studies as the proposal of inputs and outputs in the production process is the basis for the implementation of the model. The choice of inputs and outputs is typically based on relevant economic theory, but due to different types of observational data used for different sectors of the economy, there could be a causal interpretation problem. The proposal and estimation of a causal model would be able to solve this issue and provide validity to the DEA model proposed. In the causal modelling framework, the study tested that, conditional on income, equity, and loan loss reserves in the previous period, no factor affects personal expenses, total deposits, and fixed assets. Personal expenses, total deposits, and fixed assets are the only factors influencing loans and securities investment. Loan loss reserves in the current period may be caused by themselves in the previous period, and income and equity are only affected by loans and securities investment. Table 1 provides a summary of the reviewed studies above.

In order to highlight the importance of incorporating carryover variables in network DEA, we also review relevant studies that adopted network DEA in the banking industry without considering the carryover variable in the production process and make a comparison to the results of the above-reviewed literature. Regarding the Japanese banking industry, Fukuyama and Matousek (2018) estimated Nerlove's inefficiency between 2001 and 2013 under a two-stage network DEA. Labor and capital were used as the original inputs in the first stage to generate the intermediate product (deposits). Deposits were then further used as inputs in the second stage to produce loans, non-performing loans, and securities investment. The results obtained are completely opposite to Fukuyama and Weber (2015), which shows that the level of inefficiency between 2007 and 2009 is higher than the one between 2008 and 2010.

Regarding the Canadian Credit unions, Dia et al., (2020, 2022) used a three-stage network DEA to estimate efficiency between 2007 and 2017. The first stage of production used total assets, employee compensation, and other operating expenses to generate deposits, which

Table 1 Review of previous efficiency studies

Study	Aim	Period investigated	Inputs and outputs	DEA model	Results
Fukuyama and Weber (2015)	Measuring Japanese bank performance	2006–2010	Inputs: labor, physical capital, and equity capital Intermediates: deposits and other raised funds Final outputs: loans, securities investments and non-performing loans Carryover: carryover assets from securities	Two-stage dynamic network DEA approach	There is a deterioration in the efficiency level in the period 2008–2010, compared to the one of 2007–2009

Table 1 (continued)

Study	Aim	Period investigated	Inputs and outputs	DEA model	Results
Bansal et al. (2022)	Assessing productivity performance of Indian banks	2010–2017	Desirable inputs of division 1 Labor Fixed assets Quasi-fixed input of division 1 Equity Link between divisions 1 and 2 Deposits Desirable input of division 2 Inter-bank borrowings Undesirable input of division 2 Non-performing loans (NPLs) at period t-1 Desirable carryovers of division 2 Unused assets at period t-1 Net profit at period t-1 Links between divisions 2 and 3 Performing loans Investments Undesirable output of division 2 Non-performing loans (NPLs) at period t Desirable input of division 3 Loan loss provisions Outputs of division 3 Net-interest income Non-interest income	three-stage dynamic network DEA	The dynamic sequential Malmquist-Luenberger productivity index is higher than the dynamic Malmquist-Luenberger productivity index for the overall sample

Table 1 (continued)

Study	Aim	Period investigated	Inputs and outputs	DEA model	Results
Chao et al. (2015)	Measuring the efficiency of Taiwanese banks	2005–2011	Inputs (stage 1): operating cost and capital utilization expense Intermediates: Customer capital, Human capital, Process capital and Innovation capital Stage 2 outputs: Loans, Investments and Write-offs Carryover in stage 2: Nonperforming loans Stage 3 output: revenue Carryover in stage 3: loan loss reserves	dynamic network slack-based measure DEA model	The capability for intellectual capital creation starts to improve from 2008 to 2010. The profitability of the banks has been deteriorating since the 2008 financial crisis
Fukuyama et al. (2023)	Measuring the performance in the Chinese banking industry	2010–2018	Dual-role inputs produced at the first stage: personal expenses, fixed assets and total deposits Second-stage Intermediate products used at the first stage in period $t + 1$: gross loans and total securities investment Third-stage Dual-role outputs as a carryover in period t and used at the first stage 1 as a carryover input in period $t + 1$: income and total equity Pure carryover generated at the second stage and used as a carryover input at the first stage in period $t + 1$: loan loss reserves	Dynamic network DEA with a sequential structure	The inefficiency level is around 0.14, although a slight volatility has been observed. The highest efficiency is dominated by state-owned banks. Although foreign banks are less efficient than joint-stock banks, they are more efficient than city banks. Rural banks have the highest inefficiency

were further used as inputs to generate loans and securities in the second stage. Finally, the revenue generation stage generated interest income and non-interest income. The results show that there had been a sharp decline in the level of efficiency over the examined period, which is mainly attributed to the managerial issues at the revenue generation stage. Regarding the Chinese banking industry, Fukuyama and Tan (2022a, 2022b) designed a three-stage DEA framework to estimate efficiency without considering the carryover characteristics in the production process. Labor and other operating expenses were used to produce intangibles, which were the reflection of innovation activities. Labor, as the share inputs, together with fixed assets and deposits, were used to generate securities investment, good loans, and bad loans. The securities investment and good loans are related to the production of income in the final stage, and bad loans are related to the production of equity capital. The results show that there is a much stronger volatility in the inefficiency level in the Chinese banking industry, compared to the results reported by Fukuyama et al., (2023), which considered the carryover characteristics in the production process. It is understandable that the different results obtained here are related to the data sample selected as well as the variables adopted, but whether or not incorporating the carryover variables in network DEA does matter for the results. As Zhou et al. (2019) argue, neglecting carryovers can overestimate bank performance.

2.3 Empirical literature review and hypothesis development

Thanks to the development of operational research, risk can now be incorporated in the evaluation of efficiency when addressing its influence on banking (Fukuyama & Matousek, 2018). Another stream of studies tests the impact of risk on efficiency using various econometric techniques, such as the three-stage least square estimator, Granger-causality test, bootstrapped truncated regression, fractional logit regression, and Tobit regression, among others (Tan & Anchor, 2017; Tan & Floros, 2013, 2018). All empirical banking studies focus on risk at the bank level. Tan and Floros (2013) find a positive relationship between bank risk and bank efficiency using a sample of Chinese commercial banks facilitated by the three-stage least square estimator. Tan and Floros (2018) use the Granger-causality test to investigate the relationship between risk and efficiency in the Chinese banking industry and find a significant and positive impact of credit risk on bank efficiency. Tan and Anchor (2017) use the bootstrapped truncated regression and fractional logit regression to investigate the impact of credit risk on bank efficiency in China but do not find any robust impact.

In addition to empirical investigations and evidence, the bad luck hypothesis (Berger & DeYoung, 1997) states that an increase in problem loans is primarily attributable to external events rather than managers' skills or their appetite for risk. Higher risk increases costs and managerial efforts, leading to a decline in bank efficiency. These types of bank-specific risks are crucial for bank stability in China, where the banking industry provides credits to every sector of the economy. This raises important questions: (1) what is the level of credit risk from different economic sectors? (2) What would be the impact of credit risk from various economic sectors on efficiency in banking? Addressing these questions would help regulatory authorities in China generate concrete and important policies.

As argued by the bad luck hypothesis, the negative impact of risk on efficiency is primarily due to the increase in bank cost, while allocating credits to a specific economic sector comes with a level of risk. However, we must also consider various other factors when assessing the impact of industry-level risk on bank efficiency, such as the importance and size of the economic sector, the characteristics of a specific sector, the economic sector's relationship with the banking sector, and the type of credits allocated. For example, for sectors that play a

crucial role in the economy and have a high demand for banking credits, there could be cost reduction for the banks derived from economies of scale. On the other hand, if a sector is less developed and carries a higher level of risk, it could lead to an increase in bank cost, resulting in a negative impact on bank efficiency. Furthermore, if an economic sector is closely related to the banking sector, there would be a possibility that they could help reduce the level of risk by undertaking credit allocation to companies with higher levels of risk. Lastly, the type of credits allocated to a specific economic sector matters for the relationship between risk and efficiency. For instance, green credit, allocated to environmentally friendly projects, has a higher level of risk, and there is a negative impact of this type of loan on bank efficiency (Galan & Tan, 2022). Based on these arguments, we propose the following hypotheses:

H1 For the economic sector that plays an important role in the economy and demands a large amount of banking credits, banks may have a potential cost reduction advantage, and there may be a positive impact of risk from this sector on bank efficiency.

H2 If a specific economic sector has a higher level of risk, there will be a negative impact of risk from this sector on bank efficiency.

H3 For the economic sector that has a close relationship with the banking sector and plays a similar role, the impact of risk on bank efficiency will depend on the level of risk brought to the banks by this sector and the level of risk reduction made by this sector for the banking sector. If the former is higher than the latter, there will be a negative impact of risk on bank efficiency. Otherwise, if the latter is higher, there will be a positive impact of risk on bank efficiency.

H4 The impact of risk on bank efficiency from a specific economic sector will depend on the type of credits allocated. Specifically, for the economic sector receiving green credits from banks, risk would have a negative impact on bank efficiency.

3 Research design

3.1 Data description

The current paper collects data from 64 Chinese banks that operated between 2013 and 2020. The sample includes five different bank types: 22 foreign banks, 21 city banks, 7 rural banks, 9 joint-stock banks, and 5 state-owned banks. The data was collected from the Fitch Connect database, which provides financial data for over 30,000 banks worldwide. Three original inputs were used: (1) labor, (2) deposits, and (3) physical capital (fixed assets), to produce two intermediate products: loans and securities investments. We argue that loan loss reserves are produced together with the intermediate products of loans and securities investment, but will not directly affect the next stage of production. These two intermediate products are used as inputs in the next stage of the production process, together with loan loss reserves from the previous period, to generate income. One special characteristic of this paper is considering the role played by loan loss reserves in the production process. The loan loss reserves are produced in the first stage in year t , and we treat them as desirable inputs in year $t + 1$, together with loans and securities investment, to produce income. The desirable feature of loan loss reserves is attributed to the fact that they provide bank stability. In general, we follow the intermediation approach in selecting inputs and outputs (Sealey and Lindley, 1977). While

building on the proposal of the intermediation approach, we extend the model into a multi-period, two-stage dynamic DEA model. In our model, we used physical capital as one input because it includes land, buildings and facilities, and computer equipment, which can be depreciated. Hence, physical capital can be regarded as the equivalent item of fixed assets. In contrast, one of the important components of total assets is loans, which have been widely used as outputs. Our use of physical capital as one of the inputs follows Tan et al. (2021).

3.2 Study models and variable measurement

3.2.1 Dynamic network bank production

A two-stage network technology is developed in this section. Let “T” represent the transposition operator. For time periods $t = 0, 1, \dots, \tau$, bank $j = 1, \dots, J^t$ converts three standard inputs including deposits, labour and physical capital, $\mathbf{x}^t = (x_1^t, x_2^t, x_3^t)^T \in \mathbb{R}_+^3$, in Stage 1 to produce the two intermediate outputs of loans and securities investments, $\mathbf{z}^t = (z_1^t, z_2^t)^T \in \mathbb{R}_+^2$, which are then transformed in Stage 2 to produce a single desirable output of income, $y^t \in \mathbb{R}_+$, and a scalar undesirable output of loan loss reserve, $c^t \in \mathbb{R}_+$. The loan loss reserve is carried over to the next period. That is, this carryover variable c^t links two different periods in two ways: (a) the reserve for loan losses produced in Stage 1 is chosen by optimally deciding the configuration of good outputs and the carryover variable; and (b) the loan loss reserve in period $t-1$ (c^{t-1}) affects production in Stage 2 production in period t in the sense that such reserve produced in the previous period is used in Stage 2 in the current period.

Therefore, we assume that managers decide the trade-offs between the final desirable outputs and the carryover between current and future periods. Greater amounts of loans z_1^t and securities investments z_2^t as well as the loans loss reserve in the previous period c^{t-1} are desirable inputs because such activities create larger amounts of income y^t . On the other hand, larger amounts of z_1^t, z_2^t lead to the contraction of the amount of c^t , given the fixed amounts of other assets-related items, $(other\ assets)^t$ on the asset side of the balance sheet. To explain the relationship among z_1^t, z_2^t and c^t , we write the total assets, $(total\ assets)^t$, on the asset side of the balance sheet as follows:

$$(total\ assets)^t = z_1^t + z_2^t + (other\ assets)^t \tag{1}$$

We treat the loan loss reserve as a positive carryover variable. While the loan loss reserve is reported as a negative value on the asset side of the balance sheet, is a value greater than zero in this study. Hence, $(total\ assets)^t$ is expressed as

$$(total\ assets)^t = (nominal\ assets)^t - c^t \tag{2}$$

where $(nominal\ assets)^t$ is the nominal assets in time t . The nominal total assets distinguished from the item called total assets in the balance sheet as shown in (2), $(nominal\ assets)^t$ is the sum of the total assets reported as the one in the balance sheet and the loan loss reserve (carryover variable). Combining Eq. (1) and Eq. (2) yields

$$\underbrace{(nominal\ assets)^t - c^t}_{(total\ assets)^t} = z_1^t + z_2^t + (other\ assets)^t \tag{3}$$

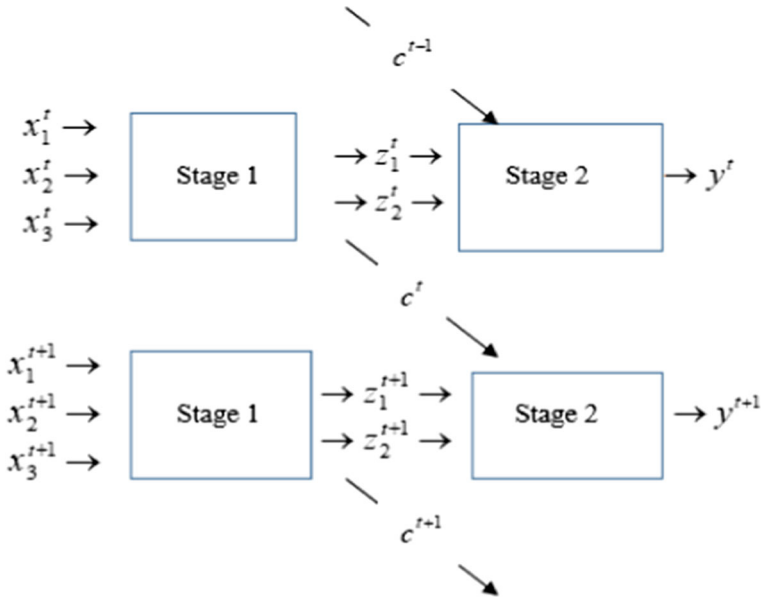


Fig. 1 Dynamic network structure

Consequently, we can write¹

$$(nominal\ assets)^t - (other\ assets)^t = z_1^t + z_2^t + c^t \tag{4}$$

Assume that bank managers make decisions on the amounts of z_1^t, z_2^t and c^t , while keeping $(nominal\ assets)^t$ fixed. Now suppose that the bank managers would like to increase c^t . To do so, the total assets item in the balance sheet must be decreased in view of Eq. (2) and the sum of z_1^t, z_2^t , and $(other\ assets)^t$ must be decreased according to Eq. (1). If the decrease in $(total\ assets)^t$ is totally made by decreasing $(other\ assets)^t$ then Eq. (4) states that it is possible to increase c^t without changing z_1^t and z_2^t . It should be noted that z_1^t and z_2^t can be changed as long as Eq. (4) when c^t is increased. Figure 1 show the bank production structure.

To develop a two-stage network framework, we add “j” in the subscripts of production factors for DMUj as $x_j^t, z_j^t, y_j^t, c_j^{t-1}$ and c_j^t to represent the jth decision-making unit, DMU_j^t. Then, the technology of stage 1 in period t is defined by

$$T^{1t} = \left\{ \begin{pmatrix} \mathbf{x}^t \\ \mathbf{z}^t \\ c^t \end{pmatrix} \in \mathbb{R}_+^{3+2+1} \mid \mathbf{x}^t \text{ can produce } \begin{pmatrix} \mathbf{z}^t \\ c^t \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} \mathbf{x}^t \\ \mathbf{z}^t \\ c^t \end{pmatrix} \in \mathbb{R}_+^{3+2+1} \mid \begin{matrix} x_1^t \geq \sum_j x_{1j}^t \lambda_j^t, x_2^t \geq \sum_j x_{2j}^t \lambda_j^t, x_3^t \geq \sum_j x_{3j}^t \lambda_j^t, z_1^t \leq \sum_j z_{1j}^t \lambda_j^t, \\ z_2^t \leq \sum_j z_{2j}^t \lambda_j^t, z_1^t + z_2^t + c^t \leq \sum_j (z_{1j}^t + z_{2j}^t + c_j^t) \lambda_j^t, \lambda_j^t \geq 0, j = 1, \dots, J \end{matrix} \right\} \tag{5}$$

¹ Equation (4), which modifies the accounting identify-based formulation used by Fukuyama and Weber (2013), clarifies why we use the carryover variable c^t as well as z_1^t and z_2^t , in the carryover constraints in our dynamic-network DEA analysis given in what follows.

and the Stage 2 technology in the same period is defined by

$$T^{2t} = \left\{ \begin{pmatrix} c^{t-1} \\ z^t \\ y^t \end{pmatrix} \in \mathbb{R}_+^{1+2+1} \mid \begin{pmatrix} c^{t-1} \\ z^t \end{pmatrix} \text{ can produce } y^t \right\}$$

$$= \left\{ \begin{pmatrix} c^{t-1} \\ z^t \\ y^t \end{pmatrix} \in \mathbb{R}_+^{1+2+1} \mid \begin{matrix} c^{t-1} \geq \sum_j^{J'} c_j^{t-1} \lambda_j^{2t}, z_1^t \geq \sum_j^{J'} z_{1j}^t \lambda_j^{2t}, z_2^t \geq \sum_j^{J'} z_{2j}^t \lambda_j^{2t}, \\ y^t \leq \sum_j^{J'} y_j^t \lambda_j^{2t}, \lambda_j^{2t} \geq 0, j = 1, \dots, J' \end{matrix} \right\} \quad (6)$$

where $\lambda_j^{1,t}$ and $\lambda_j^{2,t}$ are intensity variables related to the bank j 's Stage 1 and Stage 2 technologies, respectively, at time t . In Eq. (5), a pair of period technologies are connected by the following constraint

$$z_1^t + z_2^t + c^t \leq \sum_j^{J'} (z_{1j}^t + z_{2j}^t + c_j^t) \lambda_j^{1t} \quad (7)$$

and Eq. (7) implements a dynamic DEA structure given in Eq. (3). We utilise Eq. (7) in view of the accounting practice that the loan loss reserve, a proxy of the carryover variable, is reported as a nonpositive number on the asset side of the bank balance sheet. In Stage 1, banks combine inputs \mathbf{x}^t to produce intermediate products z^t and carryover c^t . We believe the constraint (7) reflects the current regulatory situation in Chinese banking and other environmental conditions; consequently, the number of carryovers is not completely chosen freely. Bank managers, therefore, have some latitude on the choice of carryovers.

In Stage 2 the bank converts the earlier period carryover c^{t-1} and z^t to produce final output y^t . Connecting (5) and (6), the network technology in period t is described as:

$$NT^t = \left\{ \begin{pmatrix} \mathbf{x}^t \\ z^t \\ c^t \\ c^{t-1} \\ y^t \end{pmatrix} \in \mathbb{R}_+^{3+2+1+1+1} \mid \begin{pmatrix} \mathbf{x}^t \\ z^t \\ c^t \end{pmatrix} \in T^1 \& \begin{pmatrix} c^{t-1} \\ z^t \\ y^t \end{pmatrix} \in T^2 \right\} \quad (8)$$

Explicitly incorporating Eq. (3), we construct a nonparametric DEA technology in period t as follows:

$$NT^t = \left\{ \begin{pmatrix} \mathbf{x}^t \\ z^t \\ c^t \\ c^{t-1} \\ y^t \end{pmatrix} \in \mathbb{R}_+^{3+2+1+1+1} \mid \begin{matrix} x_1^t \geq \sum_j^{J'} x_{1j}^t \lambda_j^{1t}, x_2^t \geq \sum_j^{J'} x_{2j}^t \lambda_j^{1t}, x_3^t \geq \sum_j^{J'} x_{3j}^t \lambda_j^{1t}, z_1^t \leq \sum_j^{J'} z_{1j}^t \lambda_j^{1t}, \\ z_2^t \leq \sum_j^{J'} z_{2j}^t \lambda_j^{1t}, z_1^t + z_2^t + c^t \leq \sum_j^{J'} (z_{1j}^t + z_{2j}^t + c_j^t) \lambda_j^{1t}, \\ c^{t-1} \geq \sum_j^{J'} c_j^{t-1} \lambda_j^{2t}, z_1^t \geq \sum_j^{J'} z_{1j}^t \lambda_j^{2t}, z_2^t \geq \sum_j^{J'} z_{2j}^t \lambda_j^{2t}, \\ y^t \leq \sum_j^{J'} y_j^t \lambda_j^{2t}, \lambda_j^{2t} \geq 0, j = 1, \dots, J' \end{matrix} \right\} \quad (9)$$

Then a τ -period dynamic network technology is constructed using $NT^t (t = 1, \dots, \tau)$ as follows:

$$NT^t = \left\{ \left(\begin{array}{c} (\mathbf{x}^1, \dots, \mathbf{x}^\tau)^\top \\ (\mathbf{z}^1, \dots, \mathbf{z}^\tau)^\top \\ (c^1, \dots, c^\tau)^\top \\ (y^1, \dots, y^\tau)^\top \end{array} \right) \in \mathbb{R}_+^{3\tau+2\tau+(\tau+1)+\tau} \mid \begin{array}{c} (\mathbf{x}^1 \\ z^1 \\ c^0 \\ c^1 \\ y^1 \end{array} \right) \in NT^1, \dots, \left(\begin{array}{c} \mathbf{x}^\tau \\ z^\tau \\ c^\tau \\ c^\tau \\ y^\tau \end{array} \right) \in NT^\tau \right\} \quad (10)$$

Let $\mathbf{g}^t = (\mathbf{g}^{x,t}, g_3^{y,t})^\top \in \mathbb{R}_+^{3+1}$ be a period-dependent direction vector utilised to scale exogenous inputs and outputs to the efficient frontier.

Using dynamic network technology (10), we define the dynamic τ -period network directional distance function for bank o as:

$$DNIneff = \max \alpha^1 + \alpha^2 + \dots + \alpha^\tau$$

subject to:

$$\begin{aligned} x_{1o}^t - \alpha^t g_1^{x,t} &\geq \sum_j x_{1j}^t \lambda_j^t, x_{2o}^t - \alpha^t g_1^{x,t} \geq \sum_j x_{2j}^t \lambda_j^t, x_{3o}^t - \alpha^t g_1^{x,t} \geq \sum_j x_{3j}^t \lambda_j^t, \\ z_1^t &\leq \sum_j z_{1j}^t \lambda_j^t, z_2^t \leq \sum_j z_{2j}^t \lambda_j^t, z_1^t + z_2^t + c^t \leq \sum_j (z_{1j}^t + z_{2j}^t + c_j^t) \lambda_j^t, \\ c^{t-1} &\geq \sum_j c_j^{t-1} \lambda_j^t, z_1^t \geq \sum_j z_{1j}^t \lambda_j^t, z_2^t \geq \sum_j z_{2j}^t \lambda_j^t, y_o^t + \alpha^t g_3^{y,t} \leq \sum_j y_j^t \lambda_j^t, \\ \lambda_j^t &\geq 0, \lambda_j^{2t} \geq 0, z_1^t \geq 0, z_2^t \geq 0, c^{t-1} \geq 0, \alpha^t \geq 0, (\forall j = 1, \dots, J^t, \forall t = 1, \dots, \tau) \end{aligned} \quad (11)$$

where $DNIneff$ stands for dynamic network inefficiency.

Färe and Grosskopf (1996a) strove to utilise the relationship between final outputs and carryover outputs to model a dynamic DEA problem. In a banking context, Fukuyama and Weber (2013, 2015, 2017a, 2017b) and Yu et al. (2021) for example implemented a dynamic production structure by considering the final outputs and the carryover simultaneously. The standard Malmquist indexes (Caves et al., 1982a, 1982b); Färe et al. (1994) and Luenberger productivity (Chambers, 2002) indicators that do not utilise dynamic structures such as (8), may be thought of as static dynamic models. A black-box version of the static directional distance function was developed by Chambers et al., (1996, 1998). The original version provides the maximum reduction in inputs and simultaneous expansion in outputs given a production technology by using a single scaling factor. The dynamic network directional distance function utilises τ scaling factors $\alpha^1, \alpha^2, \dots, \alpha^\tau$.

Before concluding this section, let us mention why we adopted the current dynamic framework which adopts the directional distance function (Chambers et al., 1986, 1989) rather than the slack-based measure (Lozano, 2016; Moreno & Lozano, 2018; Tone, 2001; Tone & Tsutsui, 2010; Fukuyama & Weber, 2010). The directional distance function is only consistent with the weakly efficient frontier, although the slack-based measure projects all decision-making units to the strongly efficient frontier. Despite this fact, we have chosen the directional distance function as the basis because, in a static setting, the variable and the constant returns to scale directional distance functions completely characterises the production technologies consistent to the constant and the variable returns to scale production possibility sets due to Charnes et al., (1978) and Banker et al. (1984), respectively. By contrast, the reference technology derived from slack-based measures is not equivalent to the resultant constant or the variable returns to scale technology (as was stated for example by Fukuyama, Matousek

& Tzeremes, 2021). The *DNIneff* model (11) shows how the current carryovers not only are identified (i.e., Eq. (3) or Eq. (4)) but also are carried over to the next period (i.e., Eq. (7)). In other words, *DNIneff* considers the final outputs and carryovers jointly in a dynamic decision-making framework. In contrast, Tone and Tsutsui's (2010) dynamic slacks-based framework does not directly incorporate such a dynamic role² of carryovers (see the conclusion section of Tone et al., 2019 for example). We think it is of great significance to consider the carryovers across time periods in productivity analysis, although we acknowledge Tone and Tsutsui's (2010) contribution which helped increase many dynamic DEA based publications. Moreover, while the individual bank-based carryover specification by Fukuyama and Weber (2013, 2015) considers various levels of time deposits, our specification does not use time deposits partly due to data limitations and hence it is quite useful for dynamic bank production analysis in the situations of data limitations. In view of these considerations, the directional distance function is used in this study.

In this study, we have adopted network DEA because of its strengths as follows: (1) various returns to scale properties (i.e. variable returns to scale and constant returns to scale) and the dynamic network structure can be implemented simultaneously as a simple optimization problem by adding or deleting the constraint on the sum of the intensity variables; (2) our dynamic network DEA imposes the nature of the returns to scale with respect to the intensity variables, but not the functional form. The imposition of a functional form by the parametric model does not only require the addition of the parameter constraints, but also make the estimation framework difficult. For instance, if the multi-output setting is represented by the standard Cobb–Douglas production function, the convexity of the production possibility set will not be allowed. On the other hand, if the multi-output setting is represented by a more flexible functional model, such as a translog or quadratic model, the imposition of a dynamic network structure and various returns to scale may make the estimation complicated. It should be noted that implementing noise in the parametric model may sometimes relatively easy. So, to cope with this limitation of the dynamic network DEA, we check its validity by means of causal modelling similar to Fukuyama et al., (2023).

Although the return to scale is usually considered in a static framework, this study defines the technology with a carryover variable with respect to previous and current periods. Therefore, our technology is somehow different from the standard production technology, but each time-period technology N^t given in Eq. (9) exhibits constant returns to scale in the following sense. $N^t = \delta N^t$ for all $\delta > 0$. Note that if T is the production possibility set based on Charnes et al. (1978), then T exhibits constant returns to scale, i.e., $T = \delta T$ for all $\delta > 0$. Note that constant returns to scale is considered to be a socially desirable scale. Our model employs the input orientation because the Chinese banks are too big and should be down sized (Fukuyama & Tan, 2021a, 2021b; Wilson & Zhao, 2022).

3.2.2 Causal interpretation

In general, we have a joint probability distribution in the form $p(x^t, z^t, c^t, c^{t-1}, y^t)$ which can be written as

$$p(x^t, z^t, c^t, c^{t-1}, y^t) = p(y^t | x^t, z^t, c^t, c^{t-1}) p(z^t | x^t, c^t, c^{t-1}) p(c^t | x^t, c^{t-1}) p(x^t | c^{t-1}) p(c^{t-1}). \quad (12)$$

² Färe and Grosskopf (1996b) discuss the necessity of using some kinds of dynamic constraints in general dynamic modelling.

P represents probability distribution, $|$ stands for “conditional on”, therefore, $p(y^t|x^t, z^t, c^t, c^{t-1})$ means conditional on x^t, z^t, c^t, c^{t-1} , the probability distribution of y^t .

Under restriction (2), the general specification (12) is represented as follows:

$$p(x^t, z^t, c^t, c^{t-1}, y^t) = p(y^t|z^t)p(z^t|c^t)p(c^t|x^t)p(x^t). \quad (13)$$

which means the following claim:

1. $p(y^t|x^t, z^t, c^t, c^{t-1})$ depends only on z^t .
2. $p(z^t|x^t, c^t, c^{t-1})$ depends only on x^t .
3. $p(c^t|x^t, c^{t-1})$ does not depend on c^{t-1} .
4. $p(x^t|c^{t-1})$ does not depend on c^{t-1} .

Alternatively, based on the production process we designed in our dynamic network DEA, in the causal analysis, we try to test (1) securities and loans in the current period are the only factors affecting net income in the current period; (2) labour, capital and deposits in the current period are the only factors affecting the production of loans and securities in the current period; (3) loan loss reserves in the current period are not affected by the ones in the previous period; (4) labour, capital and deposit in the current period are not affected by loan loss reserves in the previous period.

The decomposition (13) may or may not be supported by the data; this is particularly a cause for concern when the sample has the annual data, however the panel dataset is characterised in our sample. In order to deal with the curse of dimensionality, we approximate the probability distributions to test the decomposition. Let $\hat{u}_{(k)}$ and $F_{(k)}$ be, respectively, an inefficiency and the machine-learning approximation dependent on parameter $\theta_{(k)}$, which are parameters for adjusting the values $\hat{u}_{(k)}$. Let

$$\hat{u}_{(k)} = F_{(k)}(y^t, c^t, c^{t-1}, z^t, x^t; \theta_{(k)}), k \in \{1, 2, 3, 4\}, \quad (14)$$

represent the four machine-learning approximations of probability distributions obtained from the DEA (in)efficiency scores. Figure 1 indicates the following:

- (A) $F_{(1)}$ depends only on z^t but not c^t, c^{t-1}, x^t .
- (B) $F_{(2)}$ depends only on x^t but not c^{t-1}, z^t or anything else.
- (C) $F_{(3)}$ depends only on c^{t-1} but not x^t or anything else.
- (D) $F_{(4)}$ does not depend on anything else.

Without experimental data, it is not easy to establish the causal relationships (Pearl, 2009; Peters et al., 2013, 2017 and Pfister et al., 2019). However, the prediction properties are not influenced by the changes in the environment or interventions in the covariates due to the fact that the confounding variables are taken into consideration in the causal models. An example of confounding is that c^{t-1} influences both z^t and c^t and a spurious relationship between z^t and c^t occurs. If the confounding effect of c^{t-1} is ignored, then such an approach provides a misleading result.

Hence, we are interested in finding a reasonable causal relation, so we turn to the deconfounding approach established in a general setting. The deconfounding method is proposed by Wang and Blei (2019), in which a set of common factors are used to represent the relationship between certain variables:

$$Y_{it} = \delta_i f_t + \xi_{it}, \quad (15)$$

where δ_i represents a $d \times 1$ vector of factor loadings for bank³ $i = 1, \dots, J$, time t dependent variable f_t is the single common factor, and ξ_{it} is a $d \times 1$ error term vector. The common

³ In this paper, we have a balanced panel dataset and hence $J = J^t$ for all t .

factor f_t at t is constant over i and factor loadings δ_i of bank i are constant over at t . Here, upper case letter Y_{it} represents a potential outcome from the population, assessed at its single cause f_t , which is in the data set.

A number of procedures can be used to extract the common factor from the data and the nonlinear models are recommended by Wang and Blei (2019); one example would be the quadratic factor model as below:

$$Y_{it} = \delta_{(I)}f_t + \delta_{(II)}f_t^2 + \xi_{it}, \tag{16}$$

where $\delta_{(I)}$ and $\delta_{(II)}$ are $d \times 1$ vector of factor loadings. There are a few additional studies discussing the alternative methods for coping with deconfounding (Chernozhukov et al., 2017; Schneeweiss et al., 2009; McCaffrey et al., 2004; and Lee et al., 2010).

A dynamic factor model with the following form is proposed due to the consideration that high persistency is the characteristic of the missing environmental variables:

$$\begin{aligned} Y_{it} &= \varphi(f_t; \delta_{(I)}, \delta_{(II)}) + \xi_{it}, \\ f_t &= \vartheta f_{t-1} + \xi_{f,it}, \end{aligned} \tag{17}$$

where Y_{it} is a vector function represented by $\varphi(f_t; \delta_{(1)}, \delta_{(2)})$, the autoregressive parameter is expressed as ϑ , and the error term is denoted by $\xi_{f,it}$ with a zero mean and normalised variance to one.

In addition, a mean vector zero and diagonal covariance matrix Σ are the properties of a multivariate normal distribution $\xi_{it} \sim \mathcal{N}_G(0, \Sigma)$. Any persistent missing variables which are not observed by the researcher but are important in defining the operational context and environment could be picked up by the Dynamic Factor Model (DFM). For a set of variables under consideration, a proper causal interpretation can be facilitated by a proper deconfounding (Wang and Blei, 2009). Put in alternative words, a proper deconfounding indicates:

- (a) Models pass the causal interpretation test.
- (b) In the context of cross-validating or hold out samples, the predictive ability of DEA models is reasonable.

Using the deconfounding method explained above, this paper estimate $\delta_{(I)}$ and $\delta_{(II)}$ in Eq. (17). In our empirical analysis, the sample distribution of Y_{it} , the realization of which is y_{it} , depends on the estimates of components of $\delta_{(I)}$ and $\delta_{(II)}$ for each bank $i = 1, \dots, J$ at time t .

3.2.3 Algorithm for implementing the causal structure

The finite mixture of multivariate normal (FMMN) distributions is used to approximate the underlying distributions, through which the restrictions (12)-(13) are tested. Let $\mathbf{W}^t = (\mathbf{c}^t, \mathbf{z}^t, \mathbf{x}^t, \mathbf{c}^{t-1})$. Noting \mathbf{c}^{t-1} is the auxiliary variable and applying the standard Markov chain Monte Carlo (MCMC) methods typically used in Bayesian modelling, we test $p(\mathbf{y}^t | \mathbf{x}^t, \mathbf{z}^t, \mathbf{c}^t)$ as follows.

Step 1. An M variate normal distribution is used by approximate $p(\mathbf{y}^t | \mathbf{W}^t)$, in which there are G components. Hence, the number of outputs is represented by M , we can write

$$p(\mathbf{y}^t | \mathbf{W}^t) = \sum_{g=1}^G \omega_g f_M^{\mathcal{N}}(\mathbf{y}^t; \mu_g, \Sigma_g). \tag{18}$$

In Eq. (18) the mixing probabilities are represented by ϖ_g , an $M \times 1$ vector of means is denoted by μ_g , an $M \times M$ covariance matrix is expressed by Σ_g , and the density of an M -variate normal distribution is proxied by $f_M^N(\mathbf{y}^t; \mu, \Sigma)$, μ and Σ stand for the mean vector and covariance matrix, respectively. The following are functions of \mathbf{W}^t for mixing probabilities, means, and covariances:

$$\mu_g = \mathbf{W}^t \theta_{\mu,g} \quad \forall g = 1, \dots, G, \quad (19)$$

$$\varpi_g = \exp\left\{\mathbf{W}^t \theta_{\varpi,g}\right\} \quad \forall g = 1, \dots, G, \quad (20)$$

$$C_{kh,g} = \exp\left\{\mathbf{W}^t \theta_{\Sigma,kh,g}\right\} \quad \forall g = 1, \dots, G, k \leq h = 1, \dots, M, \quad (21)$$

A lower triangular matrix is represented by $C_g = [C_{kh,g}]$ so that $C_g C_g' = \Sigma_g$, $\theta_{\mu,g}$, and vectors of parameters are denoted by $\theta_{\varpi,g}$, $\theta_{\Sigma,kh,g}$ (jointly called $\theta \in \mathbb{R}^d$). From this perspective, \mathbf{W}^t is included as a component in the functions of all mixture parameters, in which case the tight approximations to any conditional distributions can be yielded when there is an increase in G (Geweke & Keane, 2007; Norets, 2010; Norets & Pelenis, 2012).

Step 2. The parameters related to \mathbf{x}^t , \mathbf{c}^t , as represented by θ are tested to make sure that they are jointly zero. Equivalently, in order for the flexible FMNM to be fitted to the distribution $p(\mathbf{y}^t | \mathbf{W}^t)$, the same methodology as the one adopted in Step 1 is used.

Step 3. Compute the p -values of the three models in Steps 1 and 2.

4 Empirical results and discussion

4.1 Descriptive statistics

The statistics of the production factors are described in Table 2. Upon examining the standard deviation of the variables, it can be observed that Chinese banks in the sample have the largest difference in terms of the volumes of deposits, followed by the outputs loans and securities. This can be attributed to the fact that the Chinese banking industry has a variety of different ownership types and the size of operation among them is substantially different. The difference in the volumes of income generated among the banks in the sample is smaller than the difference in the volumes of outputs generated, indicating that larger banks should focus more on effective and efficient cost management. The smallest difference among the production factors of the banks in the sample is observed for labor. This difference can be mainly attributed to the fact that different numbers of workers are employed for different banks based on the size of the operation. Additionally, the difference can also be explained by the fact that there is a large gap in salary/wage levels between employees and management personnel.

4.2 Pairwise correlation and multicollinearity test

We conducted pairwise correlation analysis using the Pearson correlation coefficient to test the relationship between inefficiency and each of the independent variables. Additionally, this correlation coefficient was used to check for the potential issue of multicollinearity among the independent variables. Table 3 presents the results. Two important findings from the table are: (1) the results indicate that risk derived from the agricultural sector and risk from the financial services sector have a negative impact on bank inefficiency, while risk derived

Table 2 Descriptive statistics of inputs and outputs in the production process

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
Original inputs					
Labour	576	9863.064383	23,483.11032	16.29	122,487
Physical capital	576	11,906.55497	31,127.86549	0.9	155,480
Deposits	576	1,824,193.274	4,284,368.08	477.6	27,099,725
Intermediate products					
loans	576	1,131,314.502	2,776,283.476	57.1	17,822,565
Securities investment	576	558,562.3203	1,262,141.57	4.79	7,948,361
Carryover variable					
Loan loss reserves	576	33,372.71087	86,249.93115	0.09712	613,191.747
Final output					
Income	576	189,982.9417	408,621.3724	51.28205128	2,391,901.463

NB: all variables are in the unit of million RMB

Table 3 Pearson correlation coefficients

	Inefficiency	Agricultural sector	Wholesale sector	Water Conservancy, Environment and Public Facilities	Financial services sector
Inefficiency	1				
Agricultural sector	- 0.0051	1			
Wholesale sector	0.0022	0.4936	1		
Water Conservancy, Environment and Public Facilities	0.0073	0.4333	0.5582	1	
Financial services sector	- 0.0090	0.6008	0.6638	0.0530	1

Inefficiency is derived/estimated from our proposed dynamic two-stage network DEA model, while agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors

from the wholesale sector and Water Conservancy, Environment and Public Facilities have a positive impact. However, in our regression analysis using either bootstrapped truncated regression or Tobit regression, the signs and significance of the independent variables may differ from those obtained from the Pearson correlation coefficient; (2) the coefficients among the independent variables are below 0.8, indicating that the variables used in our analysis are not affected by multicollinearity issues (Tan & Floros, 2012).

4.3 Results of the study models

The directions are equal to the average values of labour, physical capital, deposits and income for all sample periods. Therefore, we use the same direction vector for all sample years. Take the first sample bank (Agricultural Bank of China) as an example, the dynamic network inefficiency of this bank is equal to

$$\begin{aligned} DNIneff &= \alpha^{2013} + \alpha^{2014} + \alpha^{2015} + \alpha^{2016} + \alpha^{2017} + \alpha^{2018} + \alpha^{2019} + \alpha^{2020} \\ &= 3.872067 + 3.292366 + 3.517703 + 4.540133 + 4.237997 + 4.185962 + 3.769933 + 3.704772 \\ &= 31.12093 \end{aligned}$$

Figure 2 presents the results on the level of inefficiency, which shows fluctuations during the period from 2013 to 2020. These findings are in line with Antunes et al. (2021), who reported a volatile efficiency trend in Chinese banking after 2014. Notably, the figure illustrates a significant increase in inefficiency in 2016. This was due to the government's removal of restrictions on deposit and loan interest rates after the completion of interest rate liberalization in 2015, which intensified competition in the banking sector. The resulting deterioration in efficiency is consistent with Tan and Floros (2019). The substantial increase in the inefficiency level in 2016 was not revealed by Liu et al., (2020a, 2020b), which proposed a two-stage DEA model based on the meta-frontier boundary and intermediate outputs goal setting to estimate bank efficiency in China. We further notice that there is an improvement in

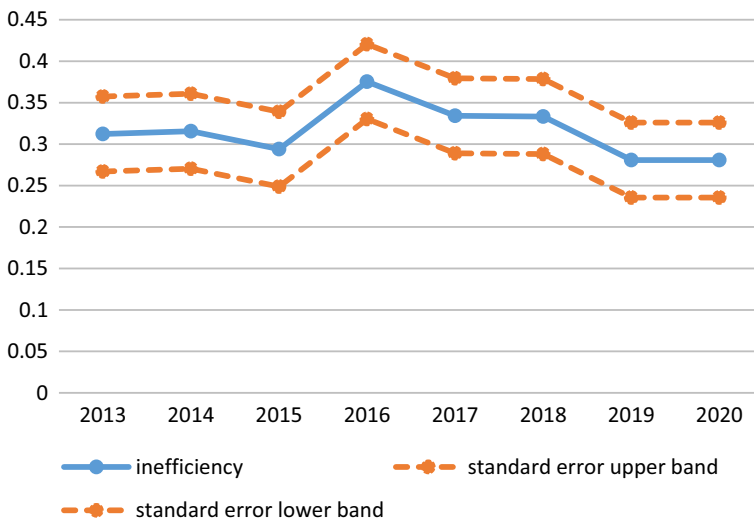


Fig. 2 Inefficiency scores between 2013 and 2020

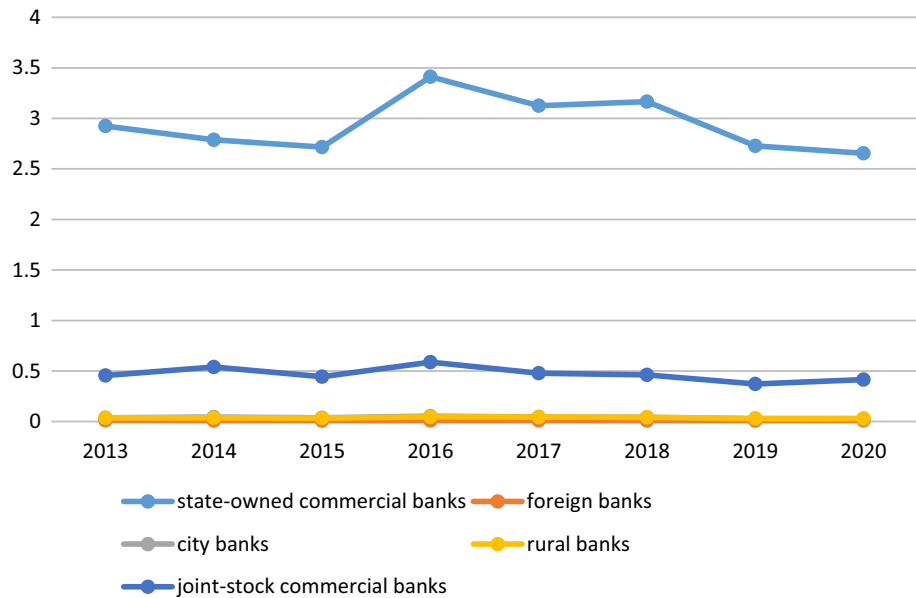


Fig. 3 Inefficiency scores of different bank types

the efficiency level after 2016 (i.e., decrease in the inefficiency level); this can be attributed to the fact that, in 2017, the Chinese banking industry deepened the integration between traditional banking and internet financial platforms, and the use of financial technology in traditional banking has been further embedded in the operation. The use of financial technology in the banking operation reduces bank cost, which is reflected in the improvement in bank efficiency.

Figure 3 shows that the highest inefficiency and inefficiency volatility were experienced by the state-owned banks. This is in contrast with An et al. (2015), who proposed a slacks-based two-stage DEA model to estimate bank efficiency in China and reported that the level of efficiency of state-owned banks was higher than that of the joint-stock banks. Another study by Liu et al., (2020a, 2020b) reported the similar results regarding the superior performance of state-owned banks compared to joint-stock banks and city banks. This is mainly attributed to the fact that they proposed a two-stage meta-frontier DEA network models for measuring the level of efficiency in the Chinese banking industry and the sample covered in their research is smaller compared to ours because foreign banks were not considered by them. We further notice that the highest efficiency was possessed by foreign banks. The figure shows that stable inefficiencies were characterised in foreign, city and rural banks and, compared to state-owned banks and joint-stock banks, they perform significantly better. This is in contrast with the findings of Shih et al. (2007). The highest and lowest efficiency occupied by state-owned and foreign banks is in accordance with Berger et al. (2009).

In order to show the accuracy of our proposed innovative model, we compare the inefficiency scores generated from the model with two alternative models, including the one which does not treat loan loss reserves as a carryover variable as well as the one derived from the single network directional distance function.⁴ Figure 4 shows the results. As we can see from

⁴ Please see Appendix A1 and A2 for detail.

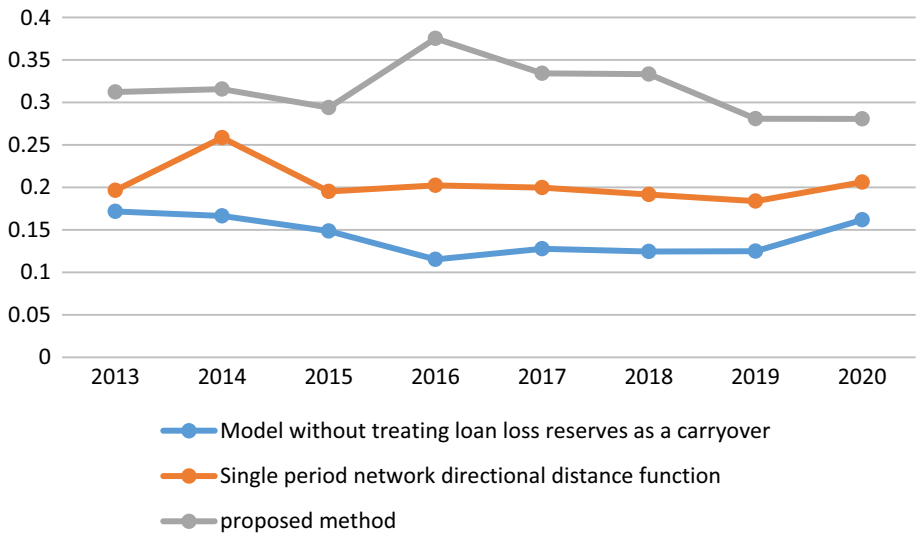


Fig. 4 Comparison of inefficiency scores among the proposed model, the model without treating Loan Loss Reserves as a carryover and single period network directional distance function

the figure, the model which does not treat loan loss reserves as a carryover variable and the one from the single network directional distance function generate lower inefficiency scores compared to the ones from our proposed approach. In other words, the efficiency scores generated by these two alternative models are inflated. In addition, it is noticed that our proposed method produces inefficiency scores with a higher degree of volatility over the period.

In the second-phase regression analysis, we investigated the impact of credit risk⁵ derived from four economic sectors on bank efficiency, including the agricultural sector, the wholesale sector, the Water Conservancy, Environment, and Public Facilities management sector, as well as the Financial Services sector under the bootstrapped truncated regression. This is another contribution we make. All the empirical banking studies concentrated on the influence of risk at the bank-level; no attempt has been made to examine the impact of credit risk at the industry-level on bank efficiency. We collected the data regarding the non-performing loan ratios of 25 different economic sectors over the period 2013–2020 from the annual financial statements published by the China Banking Regulatory Commission. The detailed figures can be found in Table 4.⁶

We can see from the table that the level of credit risk derived from allocating credits to international organizations is quite low compared to that of other economic sectors, with non-performing loan ratios for most years being 0. On the other hand, the data clearly indicates that

⁵ The credit risk investigated in our second-stage analysis is different from the key variable loan loss reserve we used in our first-stage modelling framework. Credit risk in our second-stage analysis mainly means the volumes of non-performing loans obtained by the banks from different industrial sectors, as a proportion of gross loans. While loan loss reserves in our first-stage modelling framework mainly indicate the accumulation of loan loss provisions.

⁶ We would like to thank the anonymous referee regarding the comment of controlling the year-effect and bank-specific effect in the regression analysis. Therefore, for both the bootstrapped truncated regression as well as the Tobit regression, we create dummy variables for the years as well as the banks used in our sample and include them as additional independent variables in our regression analysis, through which we would be able to generate more accurate and robust results by controlling the year-specific and bank-specific effects.

Table 4 Non-performing loan ratios of different economic sectors in China over the period 2010–2017

Industry/NPLR	NPLR2013	NPLR2014	NPLR2015	NPLR2016	NPLR2017	NPLR2018	NPLR2019	NPLR2020
Agriculture Forestry, Animal Husbandry and Fishing	2.27	2.64	3.54	3.57	4.4	4.51	4.48	4.55
Mining	0.31	1.04	2.33	3.57	3.7	3.91	3.83	3.76
Manufacturing	1.79	2.42	3.35	3.85	4.2	4.56	4.32	4.49
Production and Supply of Electricity, Gas and Water	0.51	0.34	0.37	0.35	0.5	0.65	0.71	0.68
Construction	0.5	0.72	1.39	1.67	1.8	1.96	1.87	1.78
Wholesale and Retail Trade	2.16	3.05	4.25	4.68	4.7	5.1	4.8	4.9
Transport, Storage and Post	0.68	0.52	0.58	0.54	0.7	0.83	0.88	0.9
Lodging and Catering Services	1.27	1.47	2.26	2.68	3.4	2.99	3.15	3.5
Information Transmission, Computer Services and Software	0.95	1.15	1.06	0.79	1.1	0.95	1.07	0.84
Financial Services	0.12	0.21	0.19	0.13	0.3	0.38	0.26	0.33
Real Estate	0.48	0.5	0.81	1.04	1.1	0.97	1.03	1.09
Leasing and Business Services	0.29	0.33	0.53	0.52	0.6	0.68	0.59	0.66
Scientific Research, Technical Services	0.68	0.66	0.8	1.12	0.8	0.75	0.88	0.95
Management of Water Conservancy, Environment, and Public Facilities	0.11	0.08	0.12	0.16	0.1	0.13	0.12	0.15
Services to Households, Maintenance and Other Services	1.05	1.43	2.07	2.55	2.6	2.58	2.62	2.6
Education	0.89	0.58	0.46	0.42	0.5	0.53	0.58	0.55
Health and Social Work	0.18	0.11	0.12	0.16	0.3	0.25	0.46	0.48
Culture, Sports and Entertainment	0.57	0.65	0.82	0.87	1.3	1.41	1.1	1.26

Table 4 (continued)

Industry/NPLR	NPLR2013	NPLR2014	NPLR2015	NPLR2016	NPLR2017	NPLR2018	NPLR2019	NPLR2020
Public Management, Social Security and Social Organisation	0.32	0.25	0.2	0.19	0.1	0.28	0.26	0.35
International Organisations	0	3.8	0	0	0	0	0	0
Personal Loans (excluding personal business loans)	0.53	0.62	0.79	0.76	0.7	0.73	0.69	0.72
Credit Card	1.28	1.49	1.84	1.9	1.6	1.7	1.79	1.82
Automobiles	2.04	2.07	2.15	2.29	1.5	1.8	1.93	2.11
Home Mortgage Loans	0.26	0.29	0.39	0.36	0.3	0.32	0.35	0.38
Other Personal Loans	0.93	1.19	1.59	1.73	1.4	1.38	1.45	1.57

NB: NPLR represents non-performing loan ratio. The data is from the annual reports of China Banking Regulatory Commission as well as the banks' annual financial statements. The NPLR is calculated as the ratio between non-performing loans and gross loans and it is in the unit of %

the level of credit risk derived from allocating credits to the agricultural sector is very high. Our aim is to investigate the impact of credit risk at the industry-level on bank efficiency and examine whether the impact of risk at the industry-level on bank efficiency differs between sectors with a higher risk level and those with a lower level of risk. To achieve this, we first need to select economic sectors with low and high risk. We compute the average non-performing loan ratios for each economic sector and compare the values. We select the two economic sectors with the highest average values as the target sample for sectors with a high level of risk, while the two economic sectors with the lowest average values are selected as the target sample for sectors with a low level of risk. After the calculation, we select the agricultural sector and the wholesale sector as the sectors with a high level of risk, with average non-performing loan ratios of more than 3.7 and 4.2, respectively. The Water Conservancy, Environment and Public Facilities management sector and the Financial Services sector have the lowest levels of credit risk, with average non-performing loan ratios of 0.12125 and 0.24, respectively. Therefore, our second-stage model to investigate the impact of industry-level risk on bank efficiency can be expressed as follows:

$$DNIneff_{it} = \beta_0 + \beta_1 Agriculture_t + \beta_2 Wholesale_t + \beta_3 WCEPF_t + \beta_4 Financial_t + \varepsilon_{it}$$

where $DNIneff_{it} = \alpha_i^t$ is the dynamic network inefficiency derived from the first-stage analysis,⁷ *Agriculture*, *Wholesale*, *WCEPF* and *Financial* stand for four different industries, namely, the agricultural industry, the wholesale sector, the Water Conservancy, Environment and Public Facilities Management sector and the financial services sector, i and t represent bank and year, respectively. The results regarding the impact of credit risk derived from these four economic sectors on efficiency are presented in Table 5. We find that credit risk derived from the agricultural sector decreases efficiency. However, we further find that credit risk derived from the wholesale and retail sector improves bank efficiency. We also find that credit risk from the Water Conservancy, Environment and Public Facilities sector increases bank inefficiency. Looking at the coefficients and probability values of the agricultural sector and the Water Conservancy, Environment and Public Facilities sector, we can find out that the latter is less significant but with a bigger size of the coefficient, which means that the latter has a bigger impact on the inefficiency level (i.e. one unit of increase in the risk level results in a larger degree of decline in the efficiency level compared to the agricultural sector). Finally, we find that credit risk derived from allocating credits to the financial services sector is not significant.

4.4 Validity analyses and robustness tests

Tobit regression⁸ is used to double check the accuracy of our results; the findings are provided in Table 6. The results confirm that credit risk from the agricultural sector decreases bank efficiency, while credit risk from the wholesale and retail sector increases bank efficiency. Our findings confirm that credit risk from the Water Conservancy, Environment, and Public Facilities sector has a larger negative impact on bank efficiency than the agricultural sector, as evidenced by a significantly larger coefficient. Finally, bank efficiency is not affected by credit risk derived from the financial services sector. We use the methods in the previous section and the Bayes factors to decide whether the restricted versions in (12)–(13) are valid. Our results are reported in Table 7.

⁷ The parameter α_i^t is the dynamic network inefficiency of bank i at year t in Eq. (11).

⁸ Based on the Monte Carlo experiments, Banker et al. (2019) argued that the ordinary least squares estimator and the Tobit regression model perform better than the bootstrap estimator in some cases.

Table 5 Empirical results (bootstrapped truncated regression)

Dependent variable	Observed coefficient	Bootstrap standard error	Z	P > Z	Normal-based (95% confidence interval)	
Inefficiency						
Independent variables						
Agricultural sector	0.253	0.05	2.22	0.08	0.012	0.218
Wholesale sector	- 0.077	0.011	- 2.88	0.005	- 0.238	- 0.031
Water Conservancy, Environment and Public Facilities	0.81	0.144	1.88	0.049	- 0.046	0.617
Financial services sector	- 0.55	0.348	- 0.79	0.44	- 0.818	0.397
Constant	0.146	0.039	4.55	0.000	0.11	0.265
Year effect	Controlled					
Bank-specific effect	Controlled					
/sigma	0.233	0.018	18.18	0.000	0.131	0.214
Log likelihood	228.85					
Number of observations	512					
Wald chi (2)	295.69***					

NB: *, **, *** represent statistically significant at 10%, 5% and 1%, respectively. Inefficiency is derived/estimated from our proposed dynamic two-stage network DEA model, while agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors

The odds against the restrictions of the causal structure A and B, as reflected by the results in Table 7, are overwhelming, which indicates that without deconfounding, there would be an invalid causal structure in the dataset. This is within our expectation due to the fact that some important variables may be missed in the data generation process. Quadratic deconfounding in Eq. (16) provides the marginal results. Finally, quadratic (dynamic) deconfounding in Eq. (17) provides the results which are largely consistent with a causal structure. Furthermore, we can test the causal interpretation of the truncated regression results in Tables 5 and 6 (bootstrapped truncated regression, and Tobit regression, respectively). Since specifications A, B, and C are rejected, we focus on specification D and we re-estimate the specifications in Tables 5 and 6 by adding the dynamic factor and its square into the specification. Additionally, in this context, we can test the same specification (viz. bootstrapped truncated regression, and Tobit regression) for causality. The results are reported in Table 8.

As the two specifications (the model without considering loan loss reserves as well as the one with a single period directional distance function) are rejected, they do not admit a causal interpretation, so the respective efficiency evaluation models should not be used. To evaluate requirements (b) and (c) (viz. the predictive ability of the DEA models in cross-validating or hold-out samples should be “reasonable”, and the predictive ability of the DEA models in cross-validating or hold-out samples remains invariant under active interventions

Table 6 Empirical results (Tobit regression)

Dependent variable	Observed coefficient	Bootstrap standard error	Z	P > Z	Normal-based (95% confidence interval)	
Inefficiency						
Independent variables						
Agricultural sector	0.088	0.058	1.90	0.052	- 0.009	0.183
Wholesale sector	- 0.071	0.044	- 2.55	0.013	- 0.179	- 0.035
Water Conservancy, Environment and Public Facilities	0.42	0.177	2.99	0.016	0.158	0.695
Financial services sector	- 0.11	0.261	- 0.25	0.899	- 0.554	0.482
Constant	0.16	0.031	4.66	0.000	0.082	0.209
Year effect	Controlled					
Bank-specific effect	Controlled					
Log likelihood	241.36					
Number of observations	512					
LR Chi (2)	249.33***					
Pseudo R2	0.0238					

NB: *, **, *** represent statistically significant at 10%, 5% and 1%, respectively. Inefficiency is derived/estimated from our proposed dynamic two-stage network DEA model, while agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors

Table 7 *p*-values

Restrictions	No deconfounding	Linear deconfounding in Eq. (15)	Quadratic deconfounding in Eq. (16)	Quadratic (Dynamic) deconfounding in Eq. (17)
A	0.000	0.014	0.082	0.171
B	0.0001	0.0012	0.154	0.182
C	0.000	0.0032	0.082	0.087
D	0.000	0.000	0.057	0.144

into the covariates of the model), we proceed as follows: For the specifications in Tables 3 and 4, we compute a pseudo-R-squared as the squared correlations between the actual and the predicted values. The requirement (b) would hold provided the pseudo-R-squared in cross-validating or hold-out samples does not differ significantly relative to the baseline pseudo-R-squared in the baseline specification (that uses all the data); we call this ratio τ .

Table 8 Deconfounding of models (model without considering Loan Loss Reserves and single period network directional distance function)

	Bootstrapped truncated regression	Tobit
Loan Loss Reserves are modelled with a carryover characteristic	0.226	0.186
Loan Loss Reserves are excluded from the model	0.000	0.000
Single period network directional distance function	0.004	0.001
<i>Specification D by sectors</i>		
Agricultural sector	0.317	0.310
Wholesale sector	0.215	0.217
Water Conservancy, Environment and Public Facilities	0.185	0.185
Financial services sector	0.215	0.214
Agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors		
<i>Specification in Tables 3 and 4 with the dynamic factor and its square</i>		
Agricultural sector	0.995	0.997
Wholesale sector	0.990	0.986
Water Conservancy, Environment and Public Facilities	0.988	0.992
Financial services sector	0.990	0.990
Agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors		
Equality of coefficients in alternative estimation sub-samples (<i>p</i> -values)		
Agricultural sector	0.225	0.220
Wholesale sector	0.157	0.152
Water Conservancy, Environment and Public Facilities	0.282	0.285
Financial services sector	0.316	0.315
Agricultural sector, wholesale sector, Water Conservancy, Environment and public facilities as well as financial services sector represent the non-performing loans ratios of the Chinese banking industry as a whole from allocating credits to these four different economic sectors		

Based on the τ values reported in Table 8, which are very close to 1, we see that in alternative sub-samples, the behaviour of the truncated regression and the one of the Tobit regression are not dramatically different. For this purpose, we generated 500 randomly selected sub-samples from the original data and split them into an estimation sample consisting of 80% observations and a cross-validation sample of approximately 20%.

Another test in this context is to examine whether the parameter estimates change when we consider alternative estimation samples similar to the above. The relevant *p*-values are reported in Table 8. The relevant test is a χ^2 -statistic of the form $\chi^2 = d'V^{-1}d$, where $d = b_1 - b_2$, b_1 and b_2 correspond to the parameter estimates from the truncated/Tobit

Table 9 The distribution variance of inefficiency scores

Year	<i>p</i> -value (Mann–Whitney)	<i>p</i> -value (Kruskal–Wallis)	Spearman rank correlation (sig.)
(2013–2014)	0.7896	0.9413	0.9934 (0.000)
(2014–2015)	0.7822		0.9896 (0.000)
(2015–2016)	0.3793		0.9828 (0.000)
(2016–2017)	0.8600		0.9834 (0.000)
(2017–2018)	0.9886		0.9924 (0.000)
(2018–2019)	0.3806		0.9267 (0.000)
(2019–2020)	0.9184		0.9673 (0.000)

regression in the whole sample, b_2 is the average of the parameter estimates from the truncated/Tobit regression in alternative sub-samples, and $V = V_1 + V_2$, where V_1 and V_2 are the corresponding covariance matrices. The empirical results show that both of the two regressions (truncated and Tobit) admit a causal interpretation in our context. In addition to testing the internal validity of the adopted model, following Habib and Mourad (2022) and Habib and Kayani (2022), we also verified the external validity of the adopted model by examining the consistency of the results over time (Table 9). The Mann–Whitney U test showed no statistically significant variance in the inefficiency score distribution for the study years 2013–2014 ($p = 0.7896$), 2014–2015 ($p = 0.7822$), 2015–2016 ($p = 0.3793$), 2016–2017 ($p = 0.8600$), 2017–2018 ($p = 0.9886$), 2018–2019 ($p = 0.3806$), and 2019–2020 ($p = 0.9084$). The Kruskal–Wallis H test also revealed no statistically significant variation in the efficiency score distribution over the study period ($p = 0.9413$). The Spearman rank correlation between each year was highly significant as well. Therefore, the general distribution of inefficiency scores and the rate of efficient DMUs do not appear to change significantly from one period to another, and the banks ranked as efficient mostly remain the same from one period to another.

4.5 Discussions

After reviewing the theoretical literature on efficiency estimation in the banking context, a number of studies have incorporated the risk factor in the production process. We contribute to the literature in bank efficiency analysis by proposing a dynamic two-stage network DEA analysis, in which loan loss reserves, one of the main risk indicators in banking, have been treated as a carryover input variable with a desirable characteristic. In the second-stage empirical analysis, existing literature studies mainly focus on the influence of bank-level risk on bank efficiency. We contribute to the empirical banking literature by focusing on the impact of industry-level risk on bank efficiency.

Our findings indicate that the Chinese banks have undergone ups and downs with regard to the inefficiency levels; the deterioration and improvement in the efficiency level are mainly attributed to the change in competitive conditions and the use of financial technology in banking operations. The impact of competition on efficiency has been documented in the competition-efficiency and competition-inefficiency hypotheses (Konara et al., 2019), while the influence of financial technology on bank efficiency is supported by Ataullah et al. (2004).

Our results are in contrast with Fukuyama and Tan (2022a, 2022b), which reported an improvement in efficiency in the Chinese banking industry between 2013 and 2016. This difference in findings is mainly attributed to the fact that two different methods were adopted. State-owned banks experienced the highest inefficiency and inefficiency volatility, while foreign banks possessed the highest efficiency. The lowest efficiency of state-owned commercial banks is mainly attributed to the strong support they received from the government. In addition, the complicated and bureaucratic procedure of this bank ownership type leads to significant waste of resources, which eventually drags down the level of efficiency. In comparison, foreign banks have a more open, transparent, and simpler operational pattern. Furthermore, foreign banks have more advanced technology in banking operations and risk management, resulting in a more optimal allocation of resources and an improvement in the level of efficiency. Joint-stock banks were found to be less efficient than city and rural banks. This is different from Antunes et al. (2022), which reported much higher volatility in the level of efficiencies for all ownership types, except for state-owned banks. This is mainly attributed to the fact that Antunes et al. (2022) used a basic DEA model, in which the multi-stage production process and the carryover characteristics of the variables are not considered.

In the second-stage analysis, we further evaluate the impact of risk from four industries on efficiency in banking, including the agricultural sector, the wholesale and retail trade sector, the Water Conservancy, Environment and Public Facilities management sector, and the Financial Services sector. The first two have the highest average non-performing loan ratios among all economic sectors, while the latter two have the lowest average non-performing loan ratios. The bootstrapped truncated regression shows that credit risk derived from the agricultural sector decreases bank efficiency. This negative impact of the agricultural industry on bank performance is evidenced and supported by Osei-Assibey and Asenso (2015), who argued that the agricultural sector is highly risky, and banks may be forced to scale back their lending activities in this sector. In China, there has been a specific type of loan tailored to agriculture called "sannong." More specifically, this type of loan focuses on farmers, rural areas, and agricultural-related activities. The characteristics of this type of loan are that the volume is small, and the allocation of this type of loan has a dispersed distribution, which increases bank costs.

We find that credit risk derived from the wholesale and retail sector improves bank efficiency. The positive impact of the wholesale and retail sector on bank efficiency can be explained by the fact that there is a very important role played by this economic sector. As argued by Su et al. (2022), the proportion of economic value added in China from the wholesale and retail trades sector increased from 1.11% in 1992 to 7.81% in 2017. The benefits derived from economies of scale for allocating credits to this economic sector are more than the losses derived from the volumes of non-performing loans. The positive and significant impact of the wholesale sector and retail sector on bank efficiency can also be explained by the fact that this specific economic sector has a significant and positive externality effect. In other words, the development of this sector will significantly improve the development of other economic sectors such as the manufacturing sector, transportation sector, agricultural sector, among others. The benefits banks receive from the development of these economic sectors are more than enough to offset the potential losses from the non-performing loans of the wholesale and retail sector, resulting in an overall improvement in bank efficiency.

We report that credit risk derived from the Water Conservancy, Environment, and Public Facilities management sector decreases bank efficiency. Danye (2020) argued that in China, the credits allocated to the Water Conservancy, Environment, and public facilities management sector focus on green loans, while green credits come with higher risk levels, which

further result in a decline in bank efficiency (Galan & Tan, 2022). In the Chinese context, one of the main issues in terms of green loans is this type of loans allocated to specific types of businesses is still underdeveloped, the businesses applying for green loans are short of information disclosure, the resulted increase in the problem of asymmetric information will increase bank risk. In addition, the Chinese banking industry now implemented different levels of interest rate charged to normal loans and green loans, with the latter receiving more beneficial policies from the perspective of a lower level of interest rate. Considering that the resources available for the banks are fixed, the volumes of green loans allocated by the banks not only will receive lower levels of income, but also will reduce the volume of “normal loans” that could be allocated by the banks. Due to the fact that the normal loans are the main source of bank income and profits, this would have a significant and negative impact on bank performance, at least in the short term.

Finally, bank efficiency is not affected by credit risk derived from the financial services sector. Li (2022) argued that there is a dual aim for the financial services sector, one is to provide financial services and the other is profit maximization. Because non-banking financial institutions can reduce the level of bank risk by providing credits to individuals or companies with higher levels of risk, the increase in the level of risk for the banks, from the perspective of higher volumes of non-performing loans from this sector, is compensated by the reduction in bank risk, facilitated by credit allocation from the non-banking financial institutions. Comparing the coefficients between the agricultural sector and the Water Conservancy, Environment and Public Facilities management sector, we can conclude that the latter has a bigger impact, as reflected by the larger coefficient.

In order to ensure the robustness of our first-stage efficiency results, as well as the second-stage regression analysis, we have tested the internal and external validity of our proposed models. We tested the causal structures of our proposed network DEA by examining different models, including the proposed model, the model without consideration of loan loss reserves, and the single-period network directional distance function. We confirmed the validity of our proposed approach, which includes loan loss reserves with a carryover characteristic in the production process. Finally, we looked at the equality of coefficients in alternative subsamples. The results show that the parameters do not differ, which guarantees the validity of our second-stage regression analysis.

Regarding the external validity of our proposed models, we used Mann–Whitney U test, Kruskal–Wallis H test, and the correlation of Spearman rank. The results show no statistically significant variation in the efficiency score distribution over the study, and the rate of efficient banks did not change significantly from one period to another. We further find that the banks ranked as efficient remained mostly harmonious from one period to another. Our current study moves the current body of knowledge forward. Therefore, for future studies investigating relevant relationships using econometric analysis, the interval validity of the econometric techniques adopted can be confirmed by relevant statistical causal analysis. For future efficiency studies incorporating a two-stage analysis, the validity of the proposed methods should be checked internally and externally.

5 Concluding remarks

5.1 Research conclusions

The current study significantly contributes to the subject area of empirical banking and operational research in banking by proposing a two-stage dynamic DEA model that considers the desirability of loan loss reserves in the banking production process with a carryover role. We validate our proposed production process using causal inference analysis. Empirically, we estimate the impact of industry-level risk on bank efficiency in a second-stage analysis facilitated by bootstrapped truncated regression and Tobit regression. The findings based on the deconfounding of models show that both of the two regressions (truncated and Tobit) admit a causal interpretation in our context, which guarantees the interval validity of our model and results. We also check the external validity of our results through the Mann–Whitney U test and Kruskal–Wallis H test. The findings show that our results are robust and consistent.

Our first-stage efficiency analysis shows that the inefficiencies were volatile in Chinese banking over the period 2013–2020. We further find that the state-owned banks experienced the highest levels of inefficiency and volatility, while foreign banks have the highest level of efficiency among different ownership types. We also find that although the efficiency level of city and rural banks is lower than that of foreign banks, it is still higher than that of joint-stock commercial banks. Our second-stage analysis shows that credit risk derived from the agricultural sector and the Water Conservancy, Environment and Public Facilities management sector decreases bank efficiency, while credit risk derived from the wholesale and retail sector improves bank efficiency.

5.2 Policy implications

Our results provide important and interesting policy recommendations: (1) additional efforts should be made to carefully check the projects related to credits allocated to the agricultural sector, as enhanced risk monitoring, checking, and management will reduce credit risk derived from this economic sector and further improve bank efficiency; (2) the volumes of credits allocated to the wholesale and retail sector should be increased; (3) additional efforts should also be made to focus on closely monitoring, checking, and managing the credits allocated to the Water Conservancy, Environment, and Public Facilities management sector; (4) as indicated by our results, we recommend that special efforts should be made to reduce credit risk derived from the economic sectors with lower levels of credit risk, because this will significantly increase bank efficiency compared to the economic sectors with higher credit risk.

Our results provide important and interesting policy recommendations: (1) additional efforts should be made to carefully check the projects related to credits allocated to the agricultural sector, as enhanced risk monitoring, checking, and management will reduce credit risk derived from this economic sector and further improve bank efficiency; (2) the volumes of credits allocated to the wholesale and retail sector should be increased; (3) additional efforts should also be made to focus on closely monitoring, checking, and managing the credits allocated to the Water Conservancy, Environment, and Public Facilities management sector; (4) as indicated by our results, we recommend that special efforts should be made to reduce credit risk derived from the economic sectors with lower levels of credit risk, because this will significantly increase bank efficiency compared to the economic sectors with higher credit risk.

From a theoretical perspective, the benefits enjoyed by large banks from economies of scale and scope do not play a decisive role in improving performance in the Chinese banking industry. Instead, a simpler, more transparent, and straightforward operating procedure, as well as advanced techniques/technology employed in operations and risk management, are of vital importance in improving bank efficiency in China. The literature has well-documented that risk has an impact on bank performance/efficiency, while no discussion has yet been initiated to investigate the level of risk derived from different economic sectors on bank efficiency. Our study provides new discoveries regarding the impact of risk from different economic sectors on bank efficiency. Based on our findings, we argue from a theoretical perspective that risk derived from economic sectors does not necessarily be harmful to bank efficiency, although there is a negative impact on bank stability. In particular, we have the novel finding that risk derived from the wholesale and retail sector benefits bank efficiency. This implies that for economic sectors that can generate positive externality to the development of other economic sectors, the benefits received by the banks can cover the potential losses from non-performing loans allocated to these sectors. At an overall level, various DEA models established from measuring performance of different economic sectors from different perspectives are short of theoretical support in terms of the validity of the proposed models. We provide a pioneering practice of supplementing and supporting the test of validity by using causal inference analysis. In addition, the econometric analysis, which was widely practiced by previous literature studies through relevant econometric techniques, lacks robustness and accuracy. We provide an example for future studies to include both econometric analysis and statistical analysis for robustness checking.

5.3 Limitations and future directions

Although the current study makes innovations from the modelling perspective as well as the empirical perspective, it still suffers from few limitations and the future research can be directed mainly in the following areas: (1) the robustness of the results can be further checked by considering the role played by the materiality of investment in the influence of industry-level risk on bank efficiency. In other words, a specific economic sector may have a higher risk, therefore, banks will have low exposure in that sector. The investigation of this issue would further provide and confirm the validity of our results; (2) The current study investigates the influence of the selected four specific economic sectors, however, there is a potential issue that a bank or some banks in the sample may not lend to one of more of the selected sectors. We tried to address this issue in our study, however, the annual financial statements provided by the banks do not disclose enough information about this. Thus, further studies could collect more data through an alternative channel and confirm that all the banks in the sample provided credits to all the selected economic sectors. This will further strengthen the accuracy and robustness of our results. Shen et al. (2022) proposed an interesting nonparametric approach to estimate the shadow pricing of non-performing loans, which can be regarded as the measurement of risk performance in the banking sector. Future studies could use our dataset, together with the nonparametric approach proposed, to evaluate the shadowing pricing of loan loss reserves. This would help check and obtain the robustness regarding the level of risk performance in the Chinese banking sector. Finally, Bădin et al. (2019) argued that conditional efficiency measures are needed when the production process does not depend only on the inputs and outputs but may be influenced by external factors and/or environmental variables. Therefore, another area of future research could further build

on the existing dynamic two-stage DEA model by incorporating the environmental variables in the modelling framework.

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/>.

Appendix A1: Model without treating loan loss reserves as a carryover

For comparison purposes, we formulate an alternative model, *DNIneff2*, to our proposed model (11) by.

- deleting the carryover constraints associated with loan loss reserves, and treating the loan loss reserves as the final output at Stage 1.
- Replacing $c^{t-1} \geq \sum_j c_j^{t-1} \lambda_j^{2t}$ with $c_o^{t-1} \geq \sum_j c_j^{t-1} \lambda_j^{2t}$ and hence $c^{t-1} \geq 0$ is deleted.

Formally,

$$DNIneff2 = \max \alpha^1 + \alpha^2 + \dots + \alpha^\tau$$

subject to:

$$\begin{aligned} x_{1o}^t - \alpha^t g_1^{x,t} &\geq \sum_j x_{1j}^t \lambda_j^{1t}, x_{2o}^t - \alpha^t g_2^{x,t} \geq \sum_j x_{2j}^t \lambda_j^{1t}, x_{3o}^t - \alpha^t g_3^{x,t} \geq \sum_j x_{3j}^t \lambda_j^{1t}, \\ z_1^t &\leq \sum_j z_{1j}^t \lambda_j^{1t}, z_2^t \leq \sum_j z_{2j}^t \lambda_j^{1t}, c_o^t + \alpha^t g^{c,t} \leq \sum_j c_j^t \lambda_j^{1t}, \\ c_o^{t-1} &\geq \sum_j c_j^{t-1} \lambda_j^{2t}, z_1^t \geq \sum_j z_{1j}^t \lambda_j^{2t}, z_2^t \geq \sum_j z_{2j}^t \lambda_j^{2t}, y_o^t + \alpha^t g^{y,t} \leq \sum_j y_j^t \lambda_j^{2t}, \\ \lambda_j^{1t} &\geq 0, \lambda_j^{2t} \geq 0, z_1^t \geq 0, z_2^t \geq 0, \alpha^t \geq 0, (\forall j = 1, \dots, J^t, \forall t = 1, \dots, \tau) \end{aligned} \quad (A1)$$

Appendix A2: Single period network directional distance function in period t

$$NDD^t = \max \alpha^t$$

subject to:

$$\begin{aligned} x_{1o}^t - \alpha^t g_1^{x,t} &\geq \sum_j x_{1j}^t \lambda_j^{1t}, x_{2o}^t - \alpha^t g_2^{x,t} \geq \sum_j x_{2j}^t \lambda_j^{1t}, x_{3o}^t - \alpha^t g_3^{x,t} \geq \sum_j x_{3j}^t \lambda_j^{1t}, \\ z_1^t &\leq \sum_j z_{1j}^t \lambda_j^{1t}, z_2^t \leq \sum_j z_{2j}^t \lambda_j^{1t}, c_o^t \leq \sum_j c_j^t \lambda_j^{1t} \\ c_o^{t-1} &\geq \sum_j c_j^{t-1} \lambda_j^{2t}, z_1^t \geq \sum_j z_{1j}^t \lambda_j^{2t}, z_2^t \geq \sum_j z_{2j}^t \lambda_j^{2t}, y_o^t + \alpha^t g^{y,t} \leq \sum_j y_j^t \lambda_j^{2t}, \\ \lambda_j^{1t} &\geq 0, \lambda_j^{2t} \geq 0, z_1^t \geq 0, z_2^t \geq 0, \alpha^t \geq 0, (\forall j = 1, \dots, J^t, \forall t = 1, \dots, \tau) \end{aligned} \quad (A2)$$

NB: Eq. (A2) corresponds to Eq. (A1).

References

- Amin, G. R., & Boamah, M. I. (2020). A new inverse DEA cost efficiency model for estimating potential merger gains: A case of Canadian banks. *Annals of Operations Research*, 295, 21–36.
- 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, 13–35.
- Antunes, J., Hadi-Vencheh, A., Jamshidi, A., Tan, Y., & Wanke, P. (2022). Bank efficiency estimation in China: DEA-RENNA approach. *Annals of Operations Research*, 315, 1373–1398. <https://doi.org/10.1007/s10479-021-04111-2>
- Ataullah, A., Cockerill, T., & Le, H. (2004). Financial liberalization and bank efficiency: A comparative analysis of India and Pakistan. *Applied Economics*, 36, 1915–1924.
- Bădin, L., Daraio, C., & Simar, L. (2019). A bootstrap approach for bandwidth selection in estimating conditional efficiency measures. *European Journal of Operational Research*, 277, 784–797.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092.
- Banker, R. D., Natarajan, R., & Zhang, D. (2019). Two-stage estimation of the impact of contextual variables in stochastic frontier production function models using data envelopment analysis: Second stage OLS versus bootstrap approaches. *European Journal of Operational Research*, 278(2), 368–384. <https://doi.org/10.1016/j.ejor.2018.10.050>
- Bansal, P., Kumar, S., Mehra, A., & Gulati, R. (2022). Developing two dynamic Malmquist-Luenberger productivity indices: An illustrated application for assessing productivity performance of Indian banks. *Omega*, 107, 102538. <https://doi.org/10.1016/j.omega.2021.102538>
- Beaver, W. H., & Engel, E. E. (1996). Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics*, 22(1–3), 177–206. [https://doi.org/10.1016/S0165-4101\(96\)00428-4](https://doi.org/10.1016/S0165-4101(96)00428-4)
- Berger, A. N., & DeYoung, R. (1997). Problem loans and cost efficiency in commercial banking. *Journal of Banking and Finance*, 21, 849–870.
- Berger, A. N., Hasan, I., & Zhou, M. (2009). Bank ownership and efficiency in China: What will happen in the world's largest nation? *Journal of Banking and Finance*, 33(1), 113–130. <https://doi.org/10.1016/j.jbankfin.2007.05.016>
- Berger, A., & Humphrey, D. (1992). Measurement and efficiency issues in commercial banking. In Z. Griliches (Ed.), *Output measurement in service sectors, National Bureau of Economic Research, Studies in Income and Wealth* (Vol. 56, pp. 245–300). University of Chicago Press.
- Boubaker, S., Le, T. D. Q., Ngo, T., & Manita, R. (2023). Predicting the performance of MSMEs: A hybrid DEA-machine learning approach. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-023-05230-8>
- Bushman, R. M., & Williams, C. D. (2012). Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. *Journal of Accounting and Economics*, 54(1), 1–18. <https://doi.org/10.1016/j.jacceco.2012.04.002>
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982a). Multilateral comparisons of output, input, and productivity using superlative index numbers. *The Economic Journal*, 92(365), 73–86. <https://doi.org/10.2307/2232257>
- Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982b). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50(6), 1393–1414. <https://doi.org/10.2307/1913388>
- Chambers, R. G. (2002). Exact nonradial input, output, and productivity measurement. *Economic Theory*, 20(4), 751–765. <https://doi.org/10.1007/s001990100231>
- Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, 70(2), 407–419. <https://doi.org/10.1006/jeth.1996.0096>
- Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions and Nerlovian efficiency. *Journal of Optimization Theory and Applications*, 98, 351–364. <https://doi.org/10.1023/A:1022637501082>
- 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(S1), S133–S151. <https://doi.org/10.1080/1540496X.2014.998887>
- Chen, Y., Cook, W. D., & Lim, K. (2019). Preface: DEA and its applications in operations and data analytics. *Annals of Operations Research*, 278, 1–4.

- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2017). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*, 21(1), C1–C68. <https://doi.org/10.1111/ectj.12097>
- Dalwai, T., Habib, A. M., Mohammadi, S. S., & Hussainey, K. (2023). Does managerial ability and auditor report readability affect corporate liquidity and cost of debt? *Asian Review of Accounting*. <https://doi.org/10.1108/ARA-06-2022-0151>
- Danye, H. (2020). Research on the impact of green credit on profitability of commercial banks in China. *E3S Web of Conferences*, 214, 03008.
- Dia, M., Takouda, P. M., & Gomohammadi, A. (2020). Correction to: Assessing the performance of Canadian credit unions using a three-stage network bootstrap DEA. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-020-03628-2>
- Dia, M., Takouda, P. M., & Gomohammadi, A. (2022). Assessing the performance of Canadian credit unions using a three-stage network bootstrap DEA. *Annals of Operations Research*, 311, 641–673.
- Dutta, P., Jaikumar, B., & Arora, M. S. (2022). Applications of data envelopment analysis in supplier selection between 2000 and 2020: A literature review. *Annals of Operations Research*, 315, 1399–1454.
- Fallah-Fini, S., Triantis, K., & Johnson, A. L. (2014). Reviewing the literature on non-parametric dynamic efficiency measurement: State-of-the-art. *Journal of Productivity Analysis*, 41(1), 51–67.
- Färe, R., & Grosskopf, S. (1996a). Productivity and intermediate products: A frontier approach. *Economics Letters*, 50(1), 65–70. [https://doi.org/10.1016/0165-1765\(95\)00729-6](https://doi.org/10.1016/0165-1765(95)00729-6)
- Färe, R., & Grosskopf, S. (1996b). *Intertemporal production frontiers: With dynamic DEA*. Kluwer.
- Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review*, 84(1), 66–83.
- Ferrier, G.D. (1994). Ownership type, property rights, and relative efficiency. In: Data envelopment analysis: theory, methodology, and applications. Springer, Dordrecht. https://doi.org/10.1007/978-94-011-0637-5_14.
- Fukuyama, H., & Matousek, R. (2018). Nerlovian revenue inefficiency in a bank production context: Evidence from Shinkin Banks. *European Journal of Operational Research*, 271(1), 317–330. <https://doi.org/10.1016/j.ejor.2018.05.008>
- Fukuyama, H., Matousek, R., & Tzeremes, N. (2021). Minimum distance efficiency measure in bank production: a directional slack inefficiency approach. *Journal of the Operational Research Society*. <https://doi.org/10.1080/01605682.2021.1943020>
- Fukuyama, H., & Tan, Y. (2021a). Corporate social behaviour: Is it good for efficiency in the Chinese banking industry? *Annals of Operations Research*, 306, 383–413.
- Fukuyama, H., & Tan, Y. (2021b). Investigating into the dual role of loan loss reserves in banking production process. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-021-04365-w>
- Fukuyama, H., & Tan, Y. (2022a). 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*, 27(1), 953–974. <https://doi.org/10.1002/ijfe.2185>
- Fukuyama, H., & Tan, Y. (2022b). Implementing strategic disposability for performance evaluation: Innovation, stability, profitability and corporate social responsibility in Chinese banking. *European Journal of Operational Research*, 296, 652–668.
- Fukuyama, H., Tsionas, M., & Tan, Y. (2023). Dynamic network data envelopment analysis with a sequential structure and behavioural-causal analysis: Application to the Chinese banking industry. *European Journal of Operational Research*, 307, 1360–1373. <https://doi.org/10.1016/j.ejor.2022.09.028>
- Fukuyama, H., & Weber, W. L. (2010). A slacks-based inefficiency measure for a two-stage system with bad outputs. *Omega*, 38(5), 239–410. <https://doi.org/10.1016/j.omega.2009.10.006>
- Fukuyama, H., & Weber, W. L. (2013). A dynamic network DEA model with an application to Japanese Shinkin banks. In P. Pasiouras (Ed.), *Efficiency and productivity growth: Modelling in the financial services industry* (pp. 193–213). Wiley.
- Fukuyama, H., & Weber, W. L. (2015). Measuring Japanese bank performance: A dynamic network DEA approach. *Journal of Productivity Analysis*, 44, 249–264. <https://doi.org/10.1007/s11223-014-0403-1>
- Fukuyama, H., & Weber, W. L. (2017a). Measuring Bank Performance with a dynamic network Luenberger indicator. *Annals of Operations Research*, 250, 85–104. <https://doi.org/10.1007/s10479-015-1922-5>
- Fukuyama, H., & Weber, W. L. (2017b). Japanese bank productivity, 2007–2012: A dynamic network approach. *Pacific Economic Review*, 22(4), 649–676. <https://doi.org/10.1111/1468-0106.12199>
- Galan, J. E., & Tan, Y. (2022). Green light for green credit? Evidence from its impact on bank efficiency. *International Journal of Finance and Economics*. <https://doi.org/10.1002/ijfe.2697>
- Geweke, J., & Keane, M. (2007). Smoothly mixing regressions. *Journal of Econometrics*, 138(1), 252–290. <https://doi.org/10.1016/j.jeconom.2006.05.022>

- Habib, A. M., & Kayani, U. N. (2022). Does the efficiency of working capital management affect a firm's financial distress? Evidence from UAE. *Corporate Governance*, 22, 1567–1586.
- Habib, A. M., & Mourad, N. (2022). Analyzing the efficiency of working capital management: A new approach based on DEA-malmquist technology. *Operations Research Forum*, 3, 32.
- Habib, A. M., & Shahwan, T. M. (2020). Measuring the operational and financial efficiency using a Malmquist data envelopment analysis: a case of Egyptian hospitals. *Benchmarking: an International Journal*, 27, 2521–2536.
- Hackman, S. T. (1990). An axiomatic framework of dynamic production. *Journal of Productivity Analysis*, 1, 309–324.
- 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.
- Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, 29(3), 337–346. <https://doi.org/10.1002/sim.3782>
- Li, Z. (2022). The influence of economic institution on finance sector credit allocation in China. *Economic Research-Ekonomska Istraživanja*, 35, 728–745.
- Liu, X., Sun, J., Yang, F., & Wu, J. (2020b). How ownership structure affects 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. (2020a). DEA considering technological heterogeneity and intermediate output target setting: The performance analysis of Chinese commercial banks. *Annals of Operations Research*, 291, 605–626.
- Lozano, S. (2016). Slacks-based inefficiency approach for general networks with bad outputs: An application to the banking sector. *Omega*, 60, 73–84. <https://doi.org/10.1016/j.omega.2015.02.012>
- Lozano, S., Soltani, N., & Dehnohalaji, A. (2020). A compromise programming approach for target setting in DEA. *Annals of Operations Research*, 288, 363–390.
- Mariz, F. B. A. R., Almeida, M. R., & Aloise, D. (2018). A review of Dynamic Data Envelopment Analysis: state of the art and applications. *International Transactions in Operational Research*, 25, 469–505.
- McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, 9(4), 403–425. <https://doi.org/10.1037/1082-989X.9.4.403>
- Monokroussos, P., Thomakos, D., Alexopoulos, T. A., & Tsioli, E. L. (2017). The determinants of loan loss provisions: An analysis of the greek banking system in light of the sovereign debt crisis. In P. Monokroussos & C. Gortsos (Eds.), *Non-performing loans and resolving private sector insolvency* (pp. 181–225). Palgrave Macmillan.
- Moreno, P., & Lozano, S. (2018). Super SBI Dynamic Network DEA approach to measuring efficiency in the provision of public services. *International Transactions in Operational Research*, 25(2), 715–735. <https://doi.org/10.1111/itor.12257>
- Mourad, N., Habib, A. M., & Tharwat, A. (2021). Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis. *Decision Science Letters*, 10, 301–310.
- Nemoto, J., & Goto, M. (2003). Measurement of dynamic efficiency in production: An application of data envelopment analysis to Japanese electric utilities. *Journal of Productivity Analysis*, 19, 191–210.
- Ng, J., & Roychowdhury, S. (2014). Do loan loss reserves behave like capital? Evidence from recent bank failures. *Review of Accounting Studies*, 19(3), 1234–1279. <https://doi.org/10.1007/s11142-014-9281-z>
- Norets, A. (2010). Approximation of conditional densities by smooth mixtures of regressions. *Annals of Statistics*, 38(3), 1733–1766. <https://doi.org/10.1214/09-AOS765>
- Norets, A., & Pelenis, J. (2012). Bayesian modelling of joint and conditional distributions. *Journal of Econometrics*, 168(2), 332–346. <https://doi.org/10.1016/j.jeconom.2012.02.001>
- Osei-Assibey, E., & Asenso, J. K. (2015). "Regulatory capital and its effect on credit growth, non-performing loans and bank efficiency: Evidence from Ghana. *Journal of Financial Economic Policy*, 7, 401–420.
- Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge University Press.
- Peters, J., Janzing, D., & Schölkopf, B. (2013). Causal inference on time series using structural equation models. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (pp. 585–592). Curran Associates.
- Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of causal inference: Foundations and learning algorithms*. MIT Press.
- Pfister, N., Bühlmann, P., & Peters, J. (2019). Invariant causal prediction for sequential data. *Journal of the American Statistical Association*, 114(527), 1264–1276. <https://doi.org/10.1080/01621459.2018.1491403>
- 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.

- Schneeweiss, S., Rassen, J. A., Glynn, R. J., Avorn, J., Mogun, H., & Brookhart, M. A. (2009). High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology*, 20(4), 512–522. <https://doi.org/10.1097/EDE.0b013e3181a663cc>
- Sengupta, J. K. (1995). *Dynamics of data envelopment analysis: Theory of systems efficiency*. Kluwer Academic.
- Shen, Z., Li, J., Vardanyan, M., & Wang, B. (2022). Nonparametric shadow pricing of non-performing loans: A study of the Chinese banking sector. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-022-05088-2>
- Shephard, R. W., and Fare, R. (1980). Dynamic Theory of Production correspondences. *Mathematical Systems in Economics*, Vol. 54. New York: Oelgeschlager, Gunn and Hain.
- Shih, V., Zhang, Q., & Liu, M. (2007). Comparing the performance of Chinese banks: A principal component analysis. *China Economic Review*, 18(1), 15–34. <https://doi.org/10.1016/j.chieco.2006.11.001>
- Su, Y., Zou, Z., Ma, X., & Ji, J. (2022). Understanding the relationships between the development of the construction sector, carbon emissions, and economic growth in China: Supply-chain level analysis based on the structural production layer difference approach. *Sustainable Production and Consumption*, 29, 730–743.
- Tan, Y. (2014). *Performance, risk, and competition in the Chinese banking industry*. Chandos Publishing.
- Tan, Y. (2016). *Investigating the performance of chinese banks: efficiency and risk features*. Palgrave Macmillan.
- Tan, Y. (2017). The impacts of competition and shadow banking on profitability: Evidence from the Chinese banking industry. *The North American Journal of Economics and Finance*, 42, 89–106.
- Tan, Y., & Anchor, J. (2017). The impacts of risk-taking behaviour and competition on technical efficiency: Evidence from the Chinese banking industry. *Research in International Business and Finance*, 41, 90–104.
- Tan, Y., & Floros, C. (2012). Bank profitability and inflation: The case of China. *Journal of Economic Studies*, 39, 675–696.
- 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(2), 144–161. <https://doi.org/10.1504/IJBAAF.2019.099424>
- Tan, Y., & Tsionas, M. (2022). Modelling sustainability efficiency in banking. *International Journal of Finance and Economics*, 27(3), 3754–3772. <https://doi.org/10.1002/ijfe.2349>
- Tan, Y., Wanke, P., Antunes, J., & Emrouznejad, A. (2021). Unveiling endogeneity between competition and efficiency in Chinese banks: A two-stage network DEA and regression analysis. *Annals of Operations Research*, 306, 131–171. <https://doi.org/10.1007/s10479-021-04104-1>
- Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3), 498–509. [https://doi.org/10.1016/S0377-2217\(99\)00407-5](https://doi.org/10.1016/S0377-2217(99)00407-5)
- Tone, K., Kweh, Q. L., Lu, W. M., & Ting, I. W. K. (2019). Modeling investments in the dynamic network performance of insurance companies. *Omega*, 88, 237–247. <https://doi.org/10.1016/j.omega.2018.09.005>
- Tone, K., & Tsutsui, M. (2010). Dynamic DEA: A slacks-based measure approach. *Omega*, 38(3–4), 145–156.
- Wang, Y., & Blei, D. M. (2019). The Blessings of Multiple Causes. *Journal of the American Statistical Association*, 114(528), 1574–1596. <https://doi.org/10.1080/01621459.2019.1686987>
- Wilson, P. W., & Zhao, S. (2022). Investigating the performance of Chinese banks over 2007–2014. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-022-04925-8>
- Yu, M. M., Lin, C. I., Chen, K. C., & Chen, L. H. (2021). Measuring Taiwanese bank performance: A two-system dynamic network data envelopment analysis approach. *Omega*, 98, 102145. <https://doi.org/10.1016/j.omega.2019.102145>
- 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.
- Zhu, J. (2022). DEA under big data: Data enabled analytics and network data envelopment analysis. *Annals of Operations Research*, 309, 761–783.