A unified framework for nonperforming loan modeling in bank production: An application of Data Envelopment Analysis

Hirofumi Fukuyama[1], Roman Matousek[2*,3], Nickolaos G. Tzeremes[4]

[1]Faculty of Commerce, Fukuoka University, 8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan. Tel: +81-92-871-6631 (ext. 4402), Fax: +81-92-864-2938. E-mail: fukuyama@fukuoka-u.ac.jp

[2*]Corresponding author, School of Business and Management, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom. E-mail address: r.matousek@qmul.ac.uk. (corresponding author)

[3]Business School, Shandong Normal University, 88 East Wenhua Road, Lixia District, Jinan, China.

[4]University of Thessaly, Department of Economics, 28th October 78, 38333, Volos, Greece. E-mail address: bus9nt@econ.uth.gr.

Abstract:

The aim of this paper is to conduct a comparative analysis of three environmental approaches in the context of a bank production framework, considering the presence of nonperforming loans (NPLs). Specifically, we examine banks' inefficiency levels using the "by-production technology," "joint-weak disposable technology," and "material balanced technology." To ensure comparability within a directional slack inefficiency framework, we propose a two-step procedure. The study is based on a sample of 379 prominent banks operating in the United States from 2003 to 2017. Our findings reveal that the material balance and by-production technologies result in estimated inefficiency measures with higher sensitivity compared to the estimator utilizing the joint-weak disposable technology. Additionally, we identify distinct properties among the estimators, emphasizing their unique characteristics for modeling nonperforming loans. Finally, our paper sheds light on the differences between the three estimators in relation to banks' inefficiency levels, considering the incorporation of nonperforming loans in the production process.

Keywords: Data envelopment analysis; Banking; By-production technology; Joint-weak disposable technology; Material balanced technology; Nonperforming loans.

1. Introduction

Efficiency measurement in the banking sector has been extensively studied for over three decades, particularly using data envelopment analysis (DEA) [8,10,11]. However, a persistent challenge in this field is the incorporation of non-performing loans (NPLs) into the bank's production function [21]. According to Drake [9] the inclusion of NPLs when modeling bank production process via DEA is essential for a more realistic assessment of bank efficiency[1].

Traditionally, DEA studies are based on environmental and natural resource economics to model NPLs within the bank production framework. In this context, NPLs are considered as an undesirable output, akin to pollutants, while net loans and securities are regarded as desirable outputs. The exploration of undesirable outputs within the DEA framework dates back to the late 1980s, with seminal works by Färe *et al*. [13] examining economic activities involving by-products. Following these pioneering studies, further contributions by Chung *et al*. [6] and Färe *et al*. [14] introduced the concept of weak disposability (WD) within the production possibility set. This assumption implies that the reduction of undesirable outputs is only achievable if desirable outputs are proportionally reduced, without any jointness between the two. While the WD assumption is widely adopted in modeling undesirable outputs [11], scholars have cautioned about the application of WD under variable returns to scale (VRS) and non-decreasing returns to scale (NDRS) assumptions [12, 33].

These potential shortcomings, however, have been tackled by the works of Kuosmanen [30], Kuosmanen and Podinovski [32] and Kuosmanen and Matin [31]. Based on the WD assumption, the studies provided an alternative setting when modeling undesirable outputs in the production process. Specifically, they provide a formulation of WD that allows for non-uniform abatement factors and

[1] For intriguing applications and modeling approaches related to the subject matter, please refer to the studies conducted by Mehdiloo and Podinovski [38], Xie *et al.* [50], and Jin *et al*. [29].

preserves the linear structure of the model, asserting that the conventional specification of WD under the VRS assumption leads to underestimation of the abatement possibilities.

At this point it must be emphasized that different approaches have been proposed to address the modeling of undesirable outputs, including transformations that convert undesirable outputs into desirable outputs [5, 45, 47]. Some studies have even considered undesirable outputs as inputs in the production process [26, 36, 37, 41]. However, these methodological treatments have faced criticism for deviating from the standard axioms of the physical law of production theory and for disregarding the direct links between desirable and undesirable outputs with input factors [12].

Conversely, Førsund [15], Murty *et al*. [39], and Dakpo *et al*. [7] have proposed approaches that consider undesirable outputs as freely disposable inputs or having weak disposability with null-jointness. These approaches avoid trade-offs among inputs, desirable outputs, and undesirable outputs. To address the challenges associated with modeling undesirable outputs, Murty *et al*. [39] introduced a by-production approach based on the costly disposability (CD) assumption. This framework treats undesirable outputs as a separate sub-technology, operating independently from the desirable output sub-technology. By considering the CD assumption, the model accounts for the distinct nature of undesirable outputs and their impact on overall estimated production efficiency.

Another notable methodological framework, the material balance approach, draws from the laws of thermodynamics and has been explored by Rødseth [44]. Under this approach, weak-G disposability of inputs and outputs is assumed, adhering to the summing-up condition, which mandates that increases in desirable outputs resulting from increased input usage or reduced desirable outputs must correspond to increases in undesirable outputs during disposal. Furthermore, an extension of the weak-G disposability has been introduced by Hampf and Rødseth [28] based on the notion of generalized weak G-disposability assumption.

When considering the modeling of NPLs within the context of bank efficiency under the assumption of weak-G disposability from a mathematical programming perspective, it assumes that the technology and production possibilities vary for each individual bank. However, such an outcome may not reflect a realistic scenario especially in cases when attempting to assess the sources of differing efficiency levels. These efficiency variations could be due to disparities in input-output combinations or disparities in the technology set.

In contrast, the concept of generalized weak G-disposable technology, as introduced by Hampf and Rødseth [28], provides a solution to this potential shortcoming. The adoption of weak G-disposable technology, takes into account the input qualities of all banks and forms a convex hull of the combined quality-specific technology sets. Consequently, the adopted generalized weak G-disposable technology encompasses a convex meta-technology that incorporates both the technical and quality differences among banks [28, p. 613]. Building upon the work of Hampf and Rødseth [28], our approach incorporates Material Balance (MB) based frontier models tailored to our specific context. These models utilize material flow coefficients, which, in our context, serve as representations of banks' Non-Performing Loans (NPLs) factors. These factors play a crucial role in calculating the risk associated with banks, stemming from the transformation of their inputs into desired outputs[2].

These NPLs factors essentially encapsulate the distribution of loan quality issued by banks. When we refer to 'loan quality,' we are alluding to the standards and criteria applied by bank managers when approving specific loans. These stringent selection processes aim to safeguard the bank against potential losses by minimizing or, ideally, avoiding the creation of NPLs [20, 22].

[2]These coefficients capture the distribution of loan quality, allowing for the modeling of input heterogeneity based on banks' quality characteristics. Importantly, following Hampf and Rødseth [28], these factors are not considered additional inputs, avoiding dimensionality-related issues in the banks' production process. The variability in coefficients across banks introduces diversity in the technology set's properties among assessed banks.

It's worth noting that these NPLs factors are incorporated to account for variations in banks' input heterogeneity, which in turn captures banks' distinct quality characteristics. Importantly, as emphasized by Hampf and Rødseth [28], these factors are not considered additional inputs in our model. Therefore, they do not directly influence or enter into the banks' production process. This approach mitigates any issues related to the dimensionality of the model. Furthermore, given that these coefficients vary among different banks, the properties of the technology set also exhibit variability across the banks under evaluation. This variance reflects the inherent diversity in how banks manage and assess their loan portfolios, contributing to a more nuanced and accurate evaluation of each institution's performance.

In addition, it must be mentioned that the adopted model, based on Hampf and Rødseth [28], serves as an extension of the MB approach initially introduced by Rødseth [44] and later refined by Hampf and Rødseth [27]. This model takes into consideration heterogeneous risk factors. In Rødseth's [44] work, the MB condition was applied within the framework of weak G-disposability, which is related to the concept of weak disposability. Under the weak G-disposability axiom, risk factors are treated as constants when defining a bank's technology. This assumption implies that loan quality remains constant across the evaluated banks. However, one might argue that this assumption could be unrealistic, as it depends on the specific conditions under which each bank operates. Therefore, we propose that, instead of adhering solely to the weak G-disposability axiom, it would be more appropriate in the context of bank production to adopt the concept of generalized weak G-disposability. This approach provides a more suitable framework for modeling a bank's production function.

Several studies have examined the modeling of banks' production function and have treated NPLs differently when estimating banks' performance. In the case of the Japanese banking sector, Fukuyama and Weber [23, 24] utilized the directional distance function (DDF) and a directional network slack-based inefficiency model, assuming weak disposability, to measure the efficiency of Japanese banks. Similarly, Barros *et al*. [2] used weak disposability to model NPLs in a non-radial directional distance function when investigating the performance of the Japanese banking sector. On the other hand, Akther *et al*. [1] applied a directional network slack-based inefficiency model, assuming weak disposability, to model NPLs in the Bangladesh banking industry. It seems that the most prevalent approach to modeling NPLs in the bank production process is the weak disposability assumption. Specifically, Wang *et al.* [49] employed an additive two-stage DEA model to analyze the Chinese banking sector, while Fukuyama and Weber [24] used a dynamic network efficiency model for the Japanese banking industry. Zhu *et al*. [52] used a non-radial Luenberger productivity estimator, and Lozano [35] used a network slacks-based inefficiency model to analyze the Chinese banking sector, both assuming weak disposability when modeling NPLs as bad outputs. It is evident that the weak disposability assumption is the dominant modeling approach for bad outputs in various DEA settings. This includes studies by Fukuyama and Matousek [18], Fukuyama and Weber [25], Fukuyama and Matousek [19], Partovi and Matousek [40], Yu *et al*. [51], and Fukuyama *et al*. [20] among others. However, only a few studies have utilized the costly disposability property introduced by Murty *et al.* [39] to model NPLs [17, 21], and Fukuyama *et al*. [22] is the only study that employed the material balanced approach [28, 42, 43, 44] to model the bank production process.

Our paper makes a significant contribution to the existing literature on DEA bank efficiency measurement by providing a unified framework[3] that models banks' production process in the presence of NPLs (Non-Performing Loans). Specifically, we propose a methodological framework based on the joint weak disposable technology referred to as "**K**," which was introduced by Kuosmanen [30], Kuosmanen and Podinovski [32], and Kuosmanen and Matin [31]. Additionally, we employ the by-production technology, denoted as "**MRL**," proposed by Murty *et al*. [39], and the material balanced technology,

[3] The term "unified framework" refers to a framework that integrates and combines various DEA models into a coherent system for the purpose of comparing how these models handle the negative or "bad" outputs in the production processes of banks. Essentially, it means that different DEA models are being brought together into one comprehensive framework to facilitate a comparative analysis.

denoted as "**HR**," which utilizes the G-weak disposability assumption introduced by Hampf and Rødseth [28], in order to accurately model banks' production process when NPLs are present.

Furthermore, we apply all three approaches to a sample of large U.S. banks to compare and analyze the estimated levels of bank performance using the three different estimators. To the best of our knowledge, this study is the first to compare the most widely used methodological framework, weak disposability, for modeling NPLs against two new and promising approaches: by-production and material balance technology. Interestingly, the material balanced approach, despite its popularity in environmental economics, has not received attention in the context of bank efficiency measurement with NPLs, even though it possesses several advantages over the weak disposability assumption.[4] As a result, to our knowledge is one of the few studies to apply the material balance approach in order to measure banks' production process.

The structure of our paper is as follows: Section 2 presents the variables and general modeling setting, while Section 3 describes a unified framework that allows for the comparison and application of the three approaches. In Section 4, we present our empirical findings, and finally, the concluding section summarizes our paper.

2. Description of Data and Model Specification

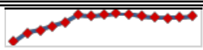
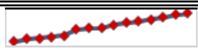
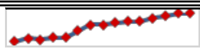
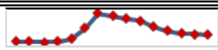
Our study uses a balanced panel consisting of 379 prominent banks in the United States, covering the period from 2003 to 2017 (5,685 observations). To gather the necessary data, we extracted information from the quarterly balance sheets of all banks in our sample, utilizing the Consolidated Report of Condition and Income, commonly referred to as Call Reports. These reports were obtained from the

[4] The MB framework presents two main advantages. Firstly, it provides the flexibility to modify the axioms of the neo-classical production model, ensuring consistency between the economic model and the materials balance principle. Secondly, it allows for the incorporation of heterogeneity in banks' input quality, as discussed by Hampf and Rødseth [27, 28].

Federal Financial Institutions Examination Council (FFIEC), specifically the FFIEC031 reporting forms[5], which encompass banks with both domestic and foreign offices.[6] Within our research, we examine different variables involved in the production process of banks, namely total fixed assets (reported in the FFIEC031 reporting form as: "premises and fixed assets") , the number of employees (reported in the FFIEC031 reporting form as: "number of full-time equivalent employees at end of current period"), total deposits (reported in the FFIEC031 reporting form as: "Deposits: Total (sum of items 1 through 6)"), non-performing loans- NPLs (reported in the FFIEC031 reporting form as: "Total loans and lease finance receivables: Non-accrual"), total securities (reported in the FFIEC031 reporting form as: "Total held-to-maturity securities + Total available-for-sale securities") and total loans (reported in the FFIEC031 reporting form as: "Total loans and leases"). Descriptive statistics for the variables used in our analysis are presented in Table 1. The trend lines presented in Table 1 demonstrate that, on average, banks have observed an upward trend in their total fixed assets, total deposits, total loans, and total securities over the analyzed period. However, starting from 2012, there has been a decline in the number of employees and NPL levels. This finding indicates that banks have made efforts to control costs, especially in the aftermath of the Global Financial Crisis, by managing their NPL levels and reducing expenses through a decrease in employee numbers.

[5] A sample of a FFIEC031 reporting form can be found at: https://www.ffiec.gov/pdf/FFIEC_forms/FFIEC031_202309_f.pdf.
[6]Note that for our calculations banks' production variables are deflated in 2015 prices using the GDP deflator.

Table 1: Descriptive statistics

		Total Fixed Assets (in 000's USD)	Number of Employees	Total Deposits (in 000's USD)	Total Securities (in 000's USD)	Total Loans (in 000's USD)	Non-performing loans (in 000's USD)
	Trend						
2003	mean	13,164,931.92	2,986	5,778,576.10	2,023,983.78	6,728,772.29	69,158.18
	std	87,962,349.43	17,886	28,177,836.92	10,803,749.24	38,430,742.39	579,224.95
2004	mean	16,498,600.77	3,472	7,268,536.13	2,566,910.98	8,509,517.08	60,625.78
	std	114,981,281.24	21,750	40,118,167.63	15,809,849.37	50,191,081.80	486,562.02
2005	mean	17,887,603.77	3,613	7,876,051.06	2,461,568.03	9,299,157.81	65,163.14
	std	122,003,922.33	22,644	42,086,891.29	15,196,133.26	53,859,087.36	505,662.90
2006	mean	20,224,613.05	3,829	8,380,540.24	2,747,977.37	10,356,437.68	72,076.43
	std	143,521,046.73	24,591	44,796,019.66	17,799,766.99	61,776,130.42	543,456.21
2007	mean	23,067,888.15	4,024	9,016,600.36	2,752,909.14	11,967,909.63	144,819.41
	std	166,944,972.62	26,714	49,530,501.04	17,011,992.59	74,264,512.11	967,396.08
2008	mean	27,143,673.49	4,423	11,796,267.31	3,648,724.81	14,161,389.06	422,933.66
	std	189,422,590.64	28,215	69,893,029.40	23,560,069.82	87,373,424.55	2,906,732.78
2009	mean	27,460,357.36	4,359	12,431,164.02	4,517,574.91	13,263,535.43	800,509.40
	std	193,422,935.63	27,466	72,642,062.80	30,431,637.09	80,982,926.87	5,775,485.00
2010	mean	27,962,926.14	4,415	12,687,669.78	4,599,909.94	13,572,834.65	728,936.64
	std	198,450,871.79	28,009	74,814,692.56	30,567,846.95	84,369,724.09	5,368,496.30
2011	mean	28,606,499.48	4,505	14,004,739.75	4,893,315.60	13,807,268.56	659,856.41
	std	199,097,545.13	28,347	82,434,257.56	31,163,776.06	84,267,350.16	5,110,881.97
2012	mean	29,730,862.19	4,470	15,170,789.03	5,064,027.24	14,256,782.56	599,688.70
	std	205,863,974.37	27,711	89,622,835.17	32,615,808.75	85,481,251.50	4,841,352.29
2013	mean	30,162,688.35	4,357	15,852,292.88	5,038,836.62	14,502,866.42	448,292.25
	std	207,054,177.12	26,581	94,429,872.30	31,535,128.49	85,802,815.08	3,532,838.37
2014	mean	31,758,076.15	4,278	16,845,791.69	5,527,040.90	15,038,944.96	343,447.59
	std	214,573,919.31	25,630	98,884,818.05	34,601,999.63	85,678,932.20	2,631,154.90
2015	mean	31,962,068.97	4,235	17,767,858.33	5,813,343.53	15,843,146.23	282,473.03
	std	208,855,812.12	24,921	102,368,281.32	35,782,682.64	88,895,549.46	2,074,052.48
2016	mean	33,861,043.97	4,280	19,064,365.76	6,204,599.02	16,786,815.63	258,256.30
	std	218,889,117.19	24,789	109,052,191.03	38,304,508.02	92,746,824.66	1,765,501.34
2017	mean	35,088,118.02	4,301	19,787,369.20	6,254,238.49	17,580,713.75	224,392.99
	std	224,278,901.58	24,597	112,028,345.96	38,340,395.91	95,584,736.64	1,516,592.06

Moreover, Table 2, presents the pre-mentioned variables in the three applied specifications. Let $\mathbf{x} = (x_1, .., x_N)$ and $\mathbf{y} = (y_{1,...,} y_M)$ be two standard (or basic) inputs and outputs. In addition, we have additional variables: *NPL* (bad output) and its linked Total Loan variable. In this study, the linked Total Loan variable, which is responsible for the production of nonperforming loans, play two different roles depending upon the three different modeling approaches (*K*, *MRL*, and *HR*). The linked Total Loan variable is a good output y^{TL} in the *K* approach, but it is a linked-input x^{TL} in the *MRL* and *HR* approaches. Hence, y^{TL} and x^{TL} do not appear simultaneously for any of the approaches adopted in this study (see Table 2)[7].

Table 2: Model description

Modeling specification	**x**	**y**	x^{TL}	y^{TL}	NPL
HR (material balance technology) and ***MRL*** (by-production technology)	Total Fixed Assets, Number of Employees	Total Deposits, Total Securities	Total Loans		NPLs
K (joint weak disposable technology)	Total Fixed Assets, Total Deposits, Number of Employees	Total Securities		Total Loans	NPLs

3. Methodological framework

Let $(\mathbf{x}_j, \mathbf{y}_j, NPL_j, y_j^{TL})$ be observed values for bank $j = 1, \ldots, J$. The *K* production possibility set[8] is written as

[7] Note that in HR and MRL modeling approach, we consider total loans as inputs and total deposits as outputs. This selection of total loans and total deposits as primary factors in our model aligns with the value-added approach, as established by seminal works such as Berger *et al.* [4] and Berger and Humphrey [3]. In this approach, both liabilities and assets are regarded as pivotal elements contributing value to the functioning of a bank. As articulated by Sealy and Lindley [46, p.1254], the production process of banks can be conceptualized from a macroeconomic perspective, focusing on the banking system's overall impact on the economy. In this context, loans and investments are regarded as inputs, while deposits represent the output in the banking production framework.

[8] The *K* production possibility set is based on the joint weak disposability proposed by Färe and Grosskopf [12], which uses a uniform abatement factor for all banks. Kuosmanen [30] suggested to use un-uniform abatement factors across firms, which

$$T_{\mathrm{K}} = \left\{ \begin{pmatrix} \mathbf{x}, \\ \mathbf{y} \\ NPL \\ y^{TL} \end{pmatrix} \in \mathbb{R}_{+}^{N+M+1+1} \left| \begin{array}{c} \mathbf{x} \geq \sum_{j=1}^{J} \mathbf{x}_j(\delta_j + \psi_j) \\ \mathbf{y} \leq \sum_{j=1}^{J} \mathbf{y}_j(\delta_j + \psi_j) \\ y^{TL} \leq \sum_{j=1}^{J} y_j^{TL} \delta_j \\ \theta_{\mathrm{K}} \cdot NPL = \sum_{j=1}^{J} NPL_j \cdot \delta_j \\ \sum_j^J(\delta_j + \psi_j) = 1, \ \ \theta\text{: free} \\ \delta_j \geq 0, j = 1, \ldots, J; \ \ \psi_j \geq 0, \ \ j = 1, \ldots, J \end{array} \right. \right\} \tag{1}$$

where

$$\theta_{\mathrm{K}} = \min \left\{ \theta > 0 \left| \begin{array}{c} \mathbf{x} \geq \sum_{j=1}^{J} \mathbf{x}_j(\delta_j + \psi_j); \ \ \mathbf{y} \leq \sum_{j=1}^{J} \mathbf{y}_j(\delta_j + \psi_j) \\ y^{TL} \leq \sum_{j=1}^{J} y_j^{TL} \delta_j \\ -\sum_{j=1}^{J} NPL_j \cdot \delta_j + \theta \cdot NPL = 0 \\ \sum_{j=1}^{J}(\delta_j + \psi_j) = 1, \ \ \theta\text{: free} \\ \delta_j \geq 0, j = 1, \ldots, J; \ \ \psi_j \geq 0, \ \ j = 1, \ldots, J \end{array} \right. \right\} \tag{2}$$

The Kuosmanen technology has the following disposability properties[9]:

Strong disposability of **x**:

$$(\mathbf{x}, \mathbf{y}, y^{TL}, NPL) \in T_{\mathrm{K}} \text{ and } (\hat{\mathbf{x}}, \mathbf{y}, y^{TL}, NPL) \geq (\mathbf{x}, \mathbf{y}, y^{TL}, NPL) \ \Rightarrow \ (\hat{\mathbf{x}}, \mathbf{y}, y^{TL}, NPL) \in T_{\mathrm{K}}$$

Strong disposability of **y** and y^b:

$$(\mathbf{x}, \mathbf{y}, y^{TL}, NPL) \in T_{\mathrm{K}} \text{ and } (\mathbf{x}, \mathbf{y}, y^{TL}, NPL) \geq (\mathbf{x}, \hat{\mathbf{y}}, \hat{y}^{TL}, NPL) \ \Rightarrow \ (\mathbf{x}, \hat{\mathbf{y}}, \hat{y}^{TL}, NPL) \in T_{\mathrm{K}}$$

Joint weak disposability between NPL and y^{TL} :

$$(\mathbf{x}, \mathbf{y}, y^{TL}, NPL) \in T_{\mathrm{K}} \text{ and } 1 \geq \vartheta \geq 0 \ \Rightarrow \ (\mathbf{x}, \mathbf{y}, \vartheta y^{TL}, \vartheta \cdot NPL) \in T_{\mathrm{K}}$$

---

specification is adopted in this study. See Färe and Grosskopf [12] and Kuosmanen and Podinovski [32] for some interesting discussion on this issue.

[9]Throughout the paper, we assume the production technologies satisfy non-emptiness, closedness, output boundedness, no free lunch, convexity, bad output essentiality and input essentiality, in addition to the relevant disposability properties. Note that different approaches adopt different disposability properties as stated in this section.

Next we turn to two technologies: Murty *et al.*, [39] (MRL) technology and Hampf and Rodseth [28] (HR) technology. In these two technologies the loan variable is treated as an input x^{TL}, which is linked to NPL. The MRL technology, implementing by-production, consists of two sub-technologies, the technology is defined as the intersection of the two sub-technologies:

$$T_{\text{MRL}} = T^1 \bigcap T^2 = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^{N+M+1+1} \middle| \begin{array}{c} \mathbf{x} \geq \sum_j^J \mathbf{x}_j \lambda_j^1; \quad \mathbf{y} \leq \sum_{j=1}^J \mathrm{y}_j \lambda_j^1 \\ x^{TL} \geq \sum_j^J x_j^{TL} \lambda_j^1 \\ \sum_{j=1}^J \lambda_j^1 = 1; \quad \lambda_j^1 \geq 0, j = 1..,J \\ x^{TL} \leq \sum_j^J x_j^{TL} \lambda_j^2 \\ \theta_{\text{MRL}} \cdot NPL = \sum_{j=1}^J NPL_j \lambda_j^2 \\ \sum_{j=1}^J \lambda_j^2 = 1; \ \lambda_j^2 \geq 0, j = 1..,J \\ s^{x^{TL,2}} \geq 0 \end{array} \right\} \tag{3}$$

where

$$T^1 = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^{N+M+1+1} \middle| \begin{array}{c} \mathbf{x} \geq \sum_j^J \mathbf{x}_j \lambda_j^1; \quad \mathbf{y} \leq \sum_{j=1}^J \mathrm{y}_j \lambda_j^1 \\ x^{TL} \geq \sum_j^J x_j^{TL} \lambda_j^1 \\ \sum_{j=1}^J \lambda_j^1 = 1; \quad \lambda_j^1 \geq 0, j = 1..,J \end{array} \right\}$$

$$T^2 = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^{N+M+1+1} \middle| \begin{array}{c} x^{TL} \leq \sum_j^J x_j^{TL} \lambda_j^2 \\ \theta_{\text{MRL}} \cdot NPL = \sum_{j=1}^J NPL_j \lambda_j^2 \\ \sum_{j=1}^J \lambda_j^2 = 1; \ \lambda_j^2 \geq 0, j = 1..,J \end{array} \right\}$$

$$\theta_{\text{MRL}} = \min\{\theta| \ (\mathbf{x}, \mathbf{y}, x^{TL}, \theta \cdot NPL) \in T_{\text{MRL}}\} \tag{4}$$

The third approach, defined relative to the HR technology, considered in this paper is the one which is constructed under the assumption of different qualities across banks based on the following property:

$$s^{NPL} = \alpha_d \cdot s^{x^{TL}}$$

where α_d is the *NPL*-specific proportional factor and s^{NPL} and $s^{x^{TL}}$ are the slack constraints associated with the nonperforming loans and the nonperforming loans-generating input. Hampf and Rodseth [28] provide the *d*-quality specific technology:

$$T_{\mathrm{HR},d} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^{N+M+1+1} \middle| \begin{array}{c} \mathbf{x} \geq \sum_j^J \mathbf{x}_{dj}\lambda_{dj}\,;\;\; \mathbf{y} \leq \sum_{j=1}^J \mathbf{y}_{dj}\lambda_{dj}\;; \\ x^{TL} - s^{x^{TL}} = \sum_j^J x_{dj}^{TL}\lambda_{dj}\,; \\ NPL - s^{NPL} = \sum_{j=1}^J NPL_{dj}\lambda_{dj}\,; \\ s^{NPL} = \alpha_d \cdot s^{x^{TL}};\quad \sum_{j=1}^J \lambda_{dj} = 1; \\ \lambda_{dj} \geq 0\;\; (\forall d, \forall j);\;\; s^{x^{TL}} \geq 0;\quad s^{NPL} \geq 0 \end{array} \right\}$$

Considering $d = 1, \ldots, D$ different qualities, the *HR* technology is defined by the union of the *d*-quality sub-technologies

$$T_{\mathrm{HR}} = \text{convex} \bigcup_{d=1}^{D} T_{\mathrm{HR},d} \tag{5}$$

Applying (5) with appropriate adjustments, Hampf and Rodseth [28] presents the following *NPL*-x^{TL} oriented Farrell-like efficiency measurement (see Appendix for details):

$$\theta_{\mathrm{HR}} = \min\{\theta \mid (\mathbf{x}, \mathbf{y}, \theta \cdot x^{TL}, \theta \cdot NPL) \in T_{\mathrm{HR}}\}$$

$$= \min \left\{ \theta \middle| \begin{array}{c} \mathbf{x}_o \geq \sum_{j=1}^J \mathbf{x}_j \lambda_j;\quad \mathbf{y}_o \leq \sum_{j=1}^J \mathbf{y}_j \lambda_j\,; \\ \theta \cdot x^{TL} = \sum_{j=1}^J x_j^{TL} \lambda_j\,; \\ \theta \cdot NPL = \sum_{j=1}^J NPL_j \lambda_j \\ \sum_{j=1}^J \lambda_j = 1;\;\; \lambda_j \geq 0,\; j = 1, \ldots, J;\;\; \theta\text{: free} \end{array} \right\} \tag{6}$$

Regarding T_{MRL} and T_{HR}, we present below the following assumptions (A) of strong disposability:

Strong disposability of $\mathbf{x}$ for *h=MRL*, *HR*:

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_h \text{ and } (\hat{\mathbf{x}}, \mathbf{y}, x^{TL}, NPL) \geq (\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \;\Rightarrow\; (\hat{\mathbf{x}}, \mathbf{y}, x^{TL}, NPL) \in T_h \tag{A1}$$

Strong disposability of $\mathbf{y}$ for *h=MRL*, *HR*:

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_h \text{ and } (\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \geq (\mathbf{x}, \hat{\mathbf{y}}, x^{TL}, NPL) \;\Rightarrow\; (\mathbf{x}, \hat{\mathbf{y}}, x^{TL}, NPL) \in T_h \qquad \text{(A2)}$$

The following properties differ between T_{MRL} and T_{HR}.

Reverse disposability of x^{TL} for T_{MRL}:

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_{\text{MRL}} \text{ and } (\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \geq (\mathbf{x}, \mathbf{y}, \hat{x}^{TL}, NPL) \;\Rightarrow\; (\mathbf{x}, \mathbf{y}, \hat{x}^{TL}, NPL) \in T_{\text{MRL}} \text{ (A3)}$$

Reverse disposability of *NPL* for T_{MRL} :

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_{\text{MRL}} \text{ and } \left(\mathbf{x}, \mathbf{y}, x^{TL}, \widehat{NPL}\right) \geq (\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \;\Rightarrow\; \left(\mathbf{x}, \mathbf{y}, x^{TL}, \widehat{NPL}\right) \in T_{\text{MRL}} \text{ (A4)}$$

Joint weak disposability[10] between x^{TL} and NPL for T_{HR}:

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_{\text{HR}} \text{ and } 1 \geq \theta \geq 0 \;\Rightarrow\; (\mathbf{x}, \mathbf{y}, \theta x^{TL}, \theta \cdot NPL) \in T_{\text{HR}} \qquad \text{(A5)}$$

Moreover, T_{HR} satisfies the weak G-disposability[11]:

$$(\mathbf{x}, \mathbf{y}, x^{TL}, NPL) \in T_{\text{HR}} \text{ and } s^{NPL} = \alpha \cdot s^{x^{TL}} \;\Rightarrow\; \left(\mathbf{x}, \mathbf{y}, x^{TL} - s^{x^{TL}}, NPL - s^{x^{TL}}\right) \in T_{\text{HR}} \qquad \text{(A6)}$$

where $s^{x^{TL}}$ and s^{NPL} are associated to the slacks of x^{TL} and NPL, respectively.

We adopt a two-step procedure for each approach using a nonparametric (DEA) method by considering the importance of nonperforming loans in bank production with respect to the *K*, *MRL* and *HR* approaches. In our bank production setting, the relation between the *NPL* variable and its linked variable (output or input) is a key point that we wish to focus on, and hence we consider minimization of *NPL* in Step 1. Then we consider projection points on the standard inputs $\mathbf{x}$ and outputs $\mathbf{y}$ in Step 2.

[10] The joint disposability properties for T_{HR} and for T_{K} are different because the production variables are different.
[11] Here α not α_d because the quality of *NPLs* is not considered.

Step 1. Estimate equations (2), (4) and (6), i.e., obtain : $\theta_{\mathrm{K}}, \theta_{\mathrm{MRL}}$ and θ_{HR}.

Step 2. Obtain the following directional slack inefficiency measures with respect to $\mathbf{x}$ and outputs $\mathbf{y}$:

$$\beta_{\mathrm{K}} = \max\left\{\frac{1}{3}\left(\frac{s_1^x+s_2^x}{N}+\frac{s_1^y+s_2^y}{M}+\frac{s^{y^{TL}}}{1}\right)\middle|\begin{pmatrix}\mathbf{x}-\mathbf{s}^x;\quad \mathbf{y}-\mathbf{s}^y\\ \theta_{\mathrm{K}}\cdot y^{TL}+s^{y^{TL}};\quad \theta_{\mathrm{FGK}}\cdot NPL\\ \mathbf{s}^x\geq\mathbf{0};\ \mathbf{s}^y\geq\mathbf{0};\ s^{y^{TL}}\geq 0\end{pmatrix}\in T_{\mathrm{K}}\right\} \tag{7}$$

$$\beta_{\mathrm{MRL}} = \max\left\{\frac{1}{3}\left(\frac{s_1^x+s_2^x}{N}+\frac{s_1^y+s_2^y}{M}+\frac{s^{x^{TL}}}{1}\right)\middle|\begin{pmatrix}\mathbf{x}-\mathbf{s}^x;\ \mathbf{y}+\mathbf{s}^y\\ \theta_{\mathrm{MRL}}\cdot NPL;\quad \theta_{\mathrm{MRL}}\cdot x^{TL}-s^{x^{TL}}\\ \mathbf{s}^x\geq\mathbf{0},\ \mathbf{s}^y\geq\mathbf{0};\ s^{x^{TL}}\geq 0\end{pmatrix}\in T_{\mathrm{MRL}}\right\} \tag{8}$$

$$\beta_{\mathrm{HR}} = \max\left\{\frac{1}{2}\left(\frac{s_1^x+s_2^x}{N}+\frac{s_1^y+s_2^y}{M}\right)\middle|\begin{pmatrix}\mathbf{x}-\mathbf{s}^x;\ \mathbf{y}-\mathbf{s}^y;\\ \theta_{\mathrm{HR}}\cdot NPL;\ \theta_{\mathrm{HR}}\cdot x^{TL};\\ \mathbf{s}^x\geq\mathbf{0};\quad \mathbf{s}^y\geq\mathbf{0}\end{pmatrix}\in T_{\mathrm{HR}}\right\} \tag{9}$$

where the denominator value of the first ratio in each objective function indicates the number of kinds of production variables and the denominator value associated with slacks is the number of slack variables for each production variable. For example, there are three kinds of variables ($\mathbf{x}, \mathbf{y}$ and y^{TL}) and N variables of $\mathbf{x}$.

4. Empirical Findings

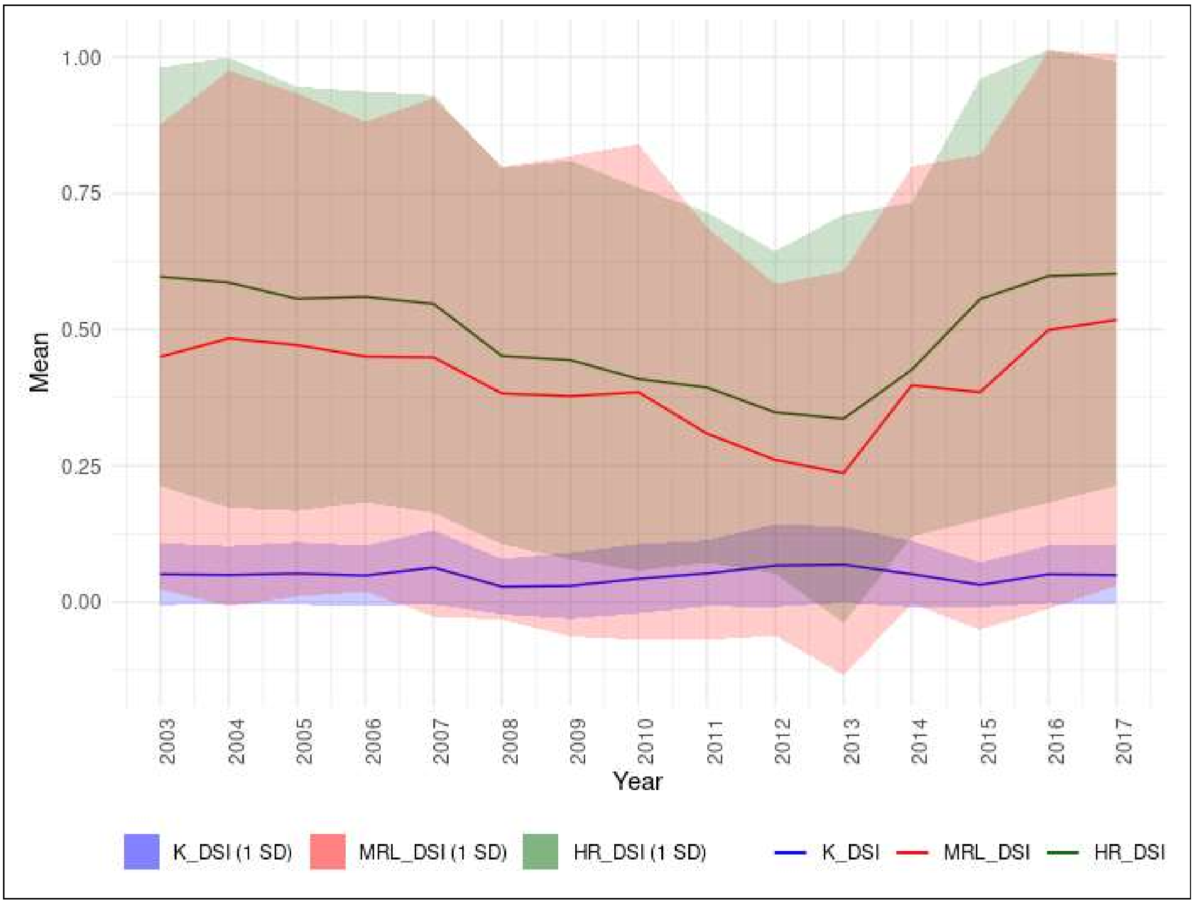
Figure 1 presents the diachronic per-year mean values of three estimators along with the associated uncertainties expressed as one standard deviation (1SD). The figure illustrates the directional distance slack inefficiency (DSI) measures under different specifications. The blue line represents the K_DSI estimator, which adopts the weak disposable technology based on Kuosmanen [30] and derives inefficiency estimates from equation 7. The red line represents the MRL_DSI estimator, which adopts the by-production technology based on Murty *et al.* [39] and derives inefficiency estimates from equation 8.

Lastly, the green line represents the HR_DSI estimator, which adopts the material balance technology by Hampf and Rodseth [28] and derives inefficiency estimates from equation 9.

Our sample comprises 5,685 observations over the entire period. Under the K_DSI estimator, 1463 observations are reported as fully efficient (i.e., zero inefficiency), whereas for the MRL_DSI estimator, 387 observations are reported as fully efficient. Similarly, under the HR_DSI estimator, 846 observations are reported as fully efficient. Figure 1 indicates that the mean inefficiency lines for MRL_DSI and HR_DSI exhibit a similar average trend, while the K_DSI estimator shows a distinct trend over the examined period. Notably, the K_DSI estimator demonstrates fewer fluctuations compared to the other two estimators, which exhibit higher standard deviation (SD) values. This discrepancy in fluctuations can be attributed to the MRL_DSI and HR_DSI estimators incorporating banks' heterogeneities, resulting in varying levels of inefficiency. Conversely, the K_DSI estimator lacks such properties, resulting in lower inefficiency fluctuation levels. Overall, the K_DSI estimator reports lower inefficiency levels compared to the MRL_DSI and HR_DSI estimators. Additionally, both the MRL_DSI and HR_DSI estimators show a decreasing trend in banks' inefficiency levels until 2013, followed by an increasing trend. In contrast, the K_DSI estimator does not exhibit this fluctuating inefficiency trend over the examined period. These differences in estimated efficiencies were expected due to the comparison of different models for banks' production processes. The MRL_DSI estimator, for instance, measures bank inefficiency using two sub-technologies: one for the treatment of desirable outputs and another for undesirable outputs (NPLs). This separation allows for distinct operations and impacts the inefficiency estimation. Similarly, the HR_DSI estimator employs the generalized weak G-disposability assumption, which incorporates banks' NPL factors as an outcome of transforming inputs to intended outputs. These factors represent the distribution of loan quality issued by banks and account for banks' input heterogeneity and quality characteristics.

Such specifications are not modeled under the K_DSI estimator, resulting in different inefficiency levels throughout the examined period.

Figure 1: Diachronic representation of banks' mean DSI levels



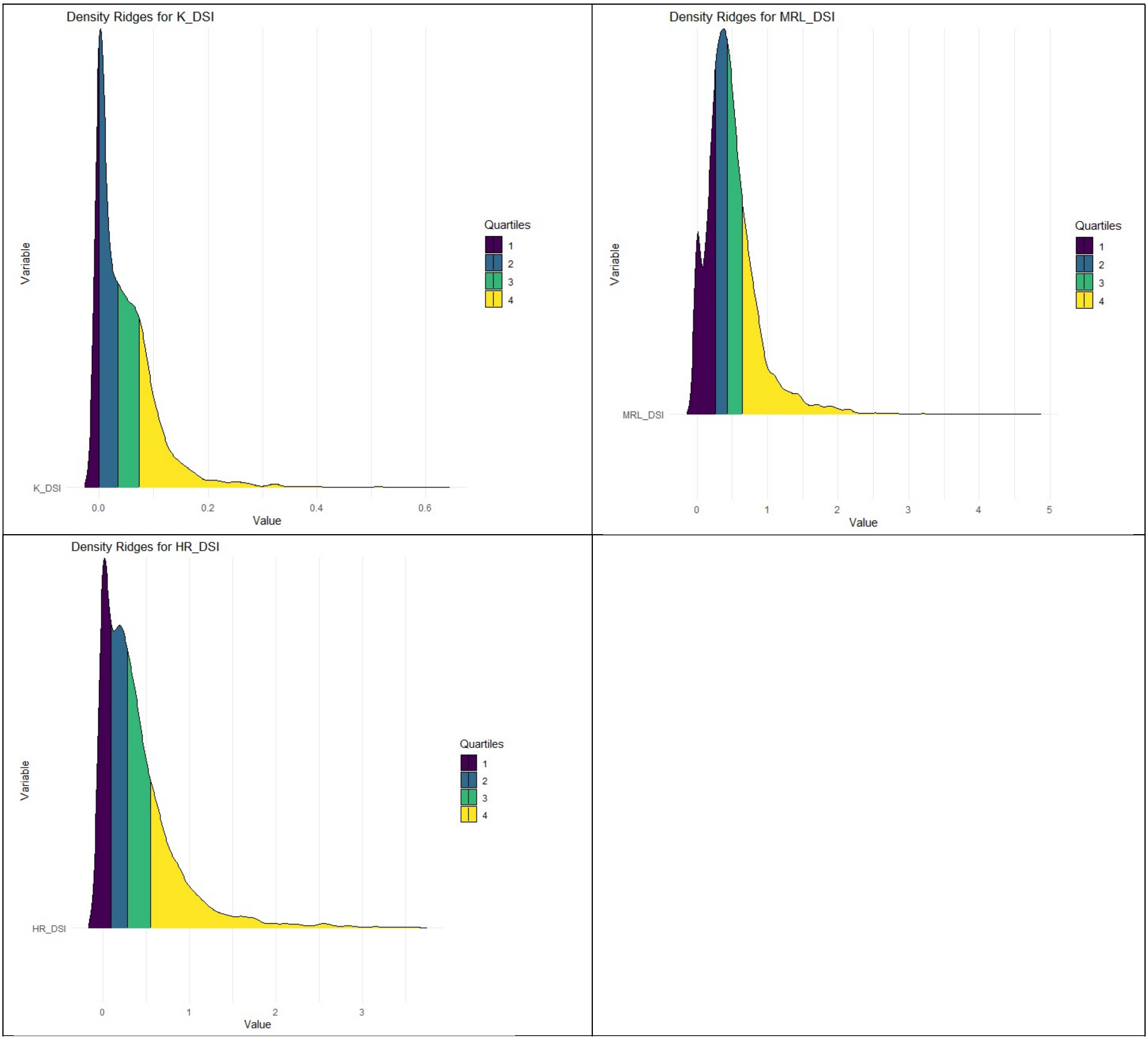
Note: The shaded boxes in various colors represent the uncertainty associated with the three estimated measures, expressed as one standard deviation (1SD). Additionally, the lines illustrate the annual mean values of the three estimators.

To examine the equality of the three inefficiency distributions, we employed the bootstrapped-based test proposed by Simar and Zelenyuk [48]. Their algorithm, based on Li's [34] work on density equality testing, compares distributions of efficiency scores estimated using DEA (Data Envelopment Analysis). Specifically, we applied "Algorithm I" from Simar and Zelenyuk [48] to assess the equality of inefficiency distributions for three pairs: "K_DSI against MRL_DSI," "K_DSI against HR_DSI," and "HR_DSI against MRL_DSI." For the pair "K_DSI against MRL_DSI," the bootstrapped-based test statistic yielded a value of 2209.311, with an estimated p-value of 0.000. This result indicates that we reject the null hypothesis of equality between the inefficiencies of the K_DSI and MRL_DSI estimators at a significance level of 0.1%. Similarly, when comparing the inefficiencies of the K_DSI and HR_DSI estimators, the test statistic was 14952.912, with an estimated p-value of 0.000. Consequently, we reject the null hypothesis of equality at the 0.1% level. Additionally, when examining the inefficiency distributions between the MRL_DSI and HR_DSI estimators, the bootstrapped-based test statistic was 118.8499, with an estimated p-value of 0.000. Thus, we once again reject the null hypothesis of equality between these two inefficiency distributions at the 0.1% level. In conclusion, all three estimators exhibit unequal inefficiency distributions.

Furthermore, Figure 2 complements the analysis by presenting density plots of quantile inefficiency estimates for the three DSI estimators. As previously observed, the weak disposability assumption employed by the K_DSI estimator results in lower inefficiency scores compared to the MRL_DSI and HR_DSI estimators. It is crucial to note that utilizing any of the DSI estimators requires subjective decisions regarding the technology used to model NPLs within the bank production framework. The K_DSI estimator, for instance, generates less fluctuated inefficiency estimates due to its inability to directly capture banks' heterogeneities or other NPL-related factors, which are accounted for by the MRL_DSI and HR_DSI estimators. Consequently, the inefficiencies obtained under the MRL_DSI and

HR_DSI estimators are higher, incorporating more uncertainty reflected in the estimated standard deviation values. Conversely, the MRL_DSI and HR_DSI estimators appear to be more sensitive in capturing structural and institutional changes in the banking industry, which subsequently affect banks' estimated inefficiency levels.

Figure 2: Density plots of the estimated banks' DSI levels



Note : The vertical axes denote the variable names, while the horizontal axis illustrates the values associated with each variable. The quartiles derived signify the segmentation of the data distribution into four intervals (quartiles), with each interval encompassing roughly 25% of the dataset. These quartile values are dynamically computed based on the filtered data, excluding extreme values.

5. Conclusions and policy implications

In this research study, we propose a unified framework for modeling banks' production process in the presence of non-performing loans (NPLs). Our analysis focuses on a sample of 379 prominent banks in the United States from 2003 to 2017. Firstly, we adopt the weak disposability assumption and incorporate heterogeneous factors across banks to model NPLs under the assumption of variable returns to scale [30, 31, 32]. Secondly, we employ the by-production approach, treating NPLs as a separate sub-technology operating independently from the desirable output sub-technology [39]. Lastly, we utilize the material balance approach to model banks' production process [28]. Notably, while Fukuyama *et al*. [22] have applied the material balance approach to banks' production processes, it has not been implemented in a directional slack inefficiency framework as developed in our study. Additionally, in our study, we assume the generalized weak-G disposability of inputs and outputs under the material balance approach, accounting for the heterogeneous quality of inputs in banks' production functions [28].

Furthermore, this study presents a unified methodological framework for the three approaches by employing a two-step procedure to make the approaches comparable. To compare the three approaches, we first minimize banks' NPL levels and then estimate the standard slack inefficiency model with respect to the standard banks' inputs and outputs. This two-step procedure is implemented to incorporate the generalized weak-G disposability assumption proposed by Hampf and Rodseth [28]. Specifically, in the first step, banks are ranked based on their radial efficiency measure, minimizing their NPL levels. In the second step, the maximum distance of only the input and output slacks ($\mathbf{s}^x$ and $\mathbf{s}^y$) is calculated.

Our empirical findings indicate that under the K_DSI estimator, the estimated bank inefficiency levels are lower compared to the inefficiency levels obtained from the MRL_DSI and HR_DSI estimators. Moreover, the MRL_DSI and HR_DSI estimators exhibit greater sensitivity compared to the K_DSI estimator, displaying higher fluctuations and increased uncertainty reflected in the estimated standard deviation values. Each model possesses unique attributes and can be applied in different situations where bank performance needs to be measured within distinct institutional and macro-micro bank industry environments.

In terms of analysing policy implications based on our empirical findings, we can conclude that the implementation of generalized weak G-disposable technology may be more suitable when modeling banks' production function with regards to non-performing loans (NPLs), particularly when closely monitoring the impact of NPL reduction policies on banks' efficiency levels in relation to factors contributing to NPL generation. Unlike the other two approaches, a crucial aspect of this approach is the consideration of weak-G disposability, specifically the "summing-up condition." As described by Hampf and Rødseth [27, p.142], in our MB modeling approach we assume zero abatement and, consequently, zero changes in the abatement of outputs (i.e., NPLs)[12]. In this context, the summation condition implies that the increases in the use of a bank's inputs and/or the reduction of the bank's good outputs must equal, the increases in NPLs when bank's inputs and outputs are disposed. Additionally, the application of generalized weak G-disposable technology encompasses all qualities of banks' inputs by forming the convex hull of the union of all quality-specific technology sets. Therefore, it includes all feasible production points within the quality-specific weakly G-disposable technologies, as well as their convex combinations. Consequently, it provides a convex meta-technology that accounts for both technical and quality differences.

Hampf and Rødseth [28, p.615] have highlighted that the implementation of the generalized weak G disposability when applying the material balance approach, in a sense, models both NPLs' generating

[12]Note that in our analysis we assume that $a = 0$ (i.e., no abatement activities). This is implied following formal derivations of the programming problems. However as demonstrated by Hampf and Rødseth [27] the $a > 0$ case can be adapted in order to allow for disposal of bank's abatement output.

inputs and NPLs as weakly disposable inputs. However, as in previous model (e.g., the $\boldsymbol{K}$ –model) when modeling NPLs and we assume a constant NPL factor for all observations, it implies an implausible trade-off. On the other hand, when we apply the generalized weak G-disposable technology assuming heterogeneity among banks' NPL generation factors, we allow for a meaningful trade-off. In such instances, a trade-off implies that when transitioning from a high NPL factor to a lower NPL factor by adjusting banks' input quality (i.e., total loans - x^{TL}), a larger quantity of this input (i.e. x^{TL}) is required to produce a desirable bank output.

It is crucial to note that the K_DSI estimator yields estimated levels of bank inefficiency that are significantly influenced by the implicit production function. For instance, it considers total deposits as inputs in the intermediation approach. On the other hand, the MRL_DSI and HR_DSI estimators treat total deposits as outputs in the value-added approach. These theoretical approaches establish the foundational framework for modeling the bank production process under the DEA framework, resulting in distinct efficiency estimates. Additionally, we acknowledge that the distribution properties of the analysed estimators are derived from the specific characteristics of the examined sample. However, it is crucial to conduct further investigations to achieve a comprehensive and conclusive understanding of the estimated inefficiency distributions. Despite these limitations, our study is the first to establish the applicability of the generalized weak-G disposability property within the directional slack inefficiency framework in the context of bank efficiency, in comparison with the other two established estimators. This opens up avenues for future research in modeling banks' production processes in the presence of NPLs.

Declarations Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

We would like to thank the Editor, the Associate Editor, and the anonymous reviewers for their valuable and constructive feedback. Any remaining errors in this work are solely the responsibility of the authors.

References

[1] Akther S, Fukuyama H, Weber WL. Estimating two-stage network slacks-based inefficiency: An application to Bangladesh banking. Omega 2013;41(1):88–96.

[2] Barros CP, Managi S, Matousek R. The technical efficiency of the Japanese banks: non-radial directional performance measurement with undesirable output. Omega 2012;40(1):1–8.

[3] Berger AN, Humphrey DB. Efficiency of financial institutions: International survey and directions for future research. Eur J Oper Res 1997;98(2):175–212.

[4] Berger AN, Hanweck GA, Humphrey DB. Competitive viability in banking: Scale, scope, and product mix economies. J Monet Econ 1987;20(3):501–520.

[5] Cherchye L, De Rock B, Walheer B. Multi-output efficiency with good and bad outputs. Eur J Oper Res 2015;240(3):872–881.

[6] Chung YH, Färe R, Grosskopf S. Productivity and undesirable outputs: a directional distance function approach. J Environ Manage 1997;51(3):229–240.

[7] Dakpo KH, Jeanneaux P, Latruffe L. Modelling pollution-generating technologies in performance benchmarking: Recent developments, limits and future prospects in the nonparametric framework. Eur J Oper Res 2016;250(2):347–359.

[8] Doumpos M, Zopounidis C, Gounopoulos D, Platanakis E, Zhang W. Operational Research and Artificial Intelligence Methods in Banking. Eur J Oper Res 2023;306:1–16.

[9] Drake L. Efficiency and productivity change in UK banking. Appl Financ Econ 2001;11(5):557–571.

[10] Emrouznejad A, Yang GL. A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. Socio-Econ Plann Sci 2018;61:4–8.

[11] Emrouznejad A, Parker BR, Tavares G. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. Socio-Econ Plann Sci 2008;42(3):151–157.

[12] Färe R, Grosskopf S. A comment on weak disposability in nonparametric production analysis. Am J Agric Econ 2009;91(2):535–538.

[13] Färe R, Grosskopf S, Lovell CK, Pasurka C. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Rev Econ Stat 1989;71(1):90–98.

[14] Färe R, Grosskopf S, Noh DW, Weber W. Characteristics of a polluting technology: theory and practice. J Econometrics 2005;126(2):469–492.

[15] Førsund FR. Pollution Modeling and Multiple-Output Production Theory, Discussion Paper #D-37/1998, Department of Economics and Social Sciences, Agricultural University of Norway 1998.

[16] Førsund FR. Multi-equation modelling of desirable and undesirable outputs satisfying the materials balance. Empir Econ 2018;54(1):67–99.

[17] Fukuyama H, Hashimoto A, Matousek R, Tzeremes NG. Analyzing bank “black boxes”: A two-stage Nerlovian profit inefficiency model. Expert Syst Appl 2021;183:115405.

[18] Fukuyama H, Matousek R. Modelling bank performance: A network DEA approach. Eur J Oper Res 2017;259(2):721–732.

[19] Fukuyama H, Matousek R. Nerlovian revenue inefficiency in a bank production context: evidence from Shinkin banks. Eur J Oper Res 2018;271(1):317–330.

[20] Fukuyama H, Matousek R, Tzeremes NG. A Nerlovian cost inefficiency two-stage DEA model for modeling banks' production process: Evidence from the Turkish banking system. Omega 2020;95:102198.

[21] Fukuyama H, Matousek R, Tzeremes NG. Bank production with nonperforming loans: A minimum distance directional slack inefficiency approach. Omega 2022;113:102706.

[22] Fukuyama H, Matousek R, Tzeremes NG. A material balance approach for modelling banks' production process with non-performing loans. J Oper Res Soc 2023; https://doi.org/10.1080/01605682.2023.2273460.

[23] Fukuyama H, Weber WL. Japanese banking inefficiency and shadow pricing. Math Comput Model 2008;48(11-12):1854–1867.

[24] Fukuyama H, Weber WL. A slacks-based inefficiency measure for a two-stage system with bad outputs. Omega 2010;38(5):398–409.

[25] Fukuyama H, Weber WL. Measuring Japanese bank performance: a dynamic network DEA approach. J Prod Anal 2015;44(3):249–264.

[26] Hailu A, Veeman TS. Non-parametric productivity analysis with undesirable outputs: an application to the Canadian pulp and paper industry. Am J Agric Econ 2001;83(3):605–616.

[27] Hampf B, Rødseth KL. Carbon dioxide emission standards for US power plants: An efficiency analysis perspective. Energy Econ 2015;50:140–153.

[28] Hampf B, Rødseth KL. Environmental efficiency measurement with heterogeneous input quality: A nonparametric analysis of US power plants. Energy Econ 2019;81:610–625.

[29] Jin F, Cai Y, Zhou L, Ding T. Regret-rejoice two-stage multiplicative DEA models-driven cross-efficiency evaluation with probabilistic linguistic information. Omega 2023;117:102839.

[30] Kuosmanen T. Weak disposability in nonparametric production analysis with undesirable outputs. Am J Agric Econ 2005;87(4):1077–1082.

[31] Kuosmanen T, Matin RK. Duality of weakly disposable technology. Omega 2011;39(5):504–512.

[32] Kuosmanen T, Podinovski V. Weak disposability in nonparametric production analysis: reply to Färe and Grosskopf. Am J Agric Econ 2009;91(2):539–545.

[33] Leleu H. Shadow pricing of undesirable outputs in nonparametric analysis. Eur J Oper Res 2013;231(2):474–480.

[34] Li Q. Nonparametric testing of closeness between two unknown distribution functions. Econometric Reviews 1996;15(3):261–274.

[35] Lozano S. Slacks-based inefficiency approach for general networks with bad outputs: An application to the banking sector. Omega 2016;60:73–84.

[36] Mahlberg B, Sahoo BK. Radial and non-radial decompositions of Luenberger productivity indicator with an illustrative application. Int J Prod Econ 2011;131(2):721–726.

[37] Mahlberg B, Luptacik M, Sahoo BK. Examining the drivers of total factor productivity change with an illustrative example of 14 EU countries. Ecol Econ 2011;72:60–69.

[38] Mehdiloo M, Podinovski VV. Selective strong and weak disposability in efficiency analysis. Eur J Oper Res 2019;276(3):1154–1169.

[39] Murty S, Russell RR, Levkoff SB. On modeling pollution-generating technologies. J Environ Econ Manage 2012;64(1):117–135.

[40] Partovi E, Matousek R. Bank efficiency and non-performing loans: Evidence from Turkey. Res Int Bus Financ 2019;48:287–309.

[41] Reinhard S, Lovell CK, Thijssen GJ. Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. Eur J Oper Res 2000;121(2):287–303.

[42] Rødseth KL. Capturing the least costly way of reducing pollution: a shadow price approach. Ecol Econ 2013;92:16–24.

[43] Rødseth KL. Efficiency measurement when producers control pollutants: A non-parametric approach. J Prod Anal 2014;42:211–223.

[44] Rødseth KL. Axioms of a polluting technology: a materials balance approach. Environ Resour Econ 2017;67(1):1–22.

[45] Scheel H. Undesirable outputs in efficiency valuations. Eur J Oper Res 2001;132(2):400–410.

[46] Sealey CW Jr, Lindley JT. Inputs, outputs, and a theory of production and cost at depository financial institutions. J Finance 1977;32(4):1251–1266.

[47] Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. Eur J Oper Res 2002;142(1):16–20.

[48] Simar L, Zelenyuk V. On testing equality of distributions of technical efficiency scores. Econometric Reviews 2006;25(4):497–522.

[49] Wang K, Huang W, Wu J, Liu YN. Efficiency measures of the Chinese commercial banking system using an additive two-stage DEA. Omega 2014;44:5–20.

[50] Xie Q, Xu Q, Chen L, Jin X, Li S, Li Y. Efficiency evaluation of China's listed commercial banks based on a multi-period leader-follower model. Omega 2022;110:102615.

[51] Yu A, Shao Y, You J, Wu M, Xu T. Estimations of operational efficiencies and potential income gains considering the credit risk for China’s banks. J Oper Res Soc 2019;70(12):2153–2168.

[52] Zhu N, Wang B, Wu Y. Productivity, efficiency, and non-performing loans in the Chinese banking industry. Soc Sci J 2015;52(4):468–480.

Appendix A: A procedure from Eq. (5) to Eq. (6).

In this appendix, we show how Eq. (5) leads to Eq. (6) by following Hampf and Rodseth (2019). Assuming there are d=1,…,D total loans qualities (heterogeneity of total loans across D different bank groups), d-quality-specific technology $T_{HR,d}$ is written as

$$T_{HR,d} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_{+}^{6} \left| \begin{array}{c} \mathbf{x} \geq \sum_{j=1}^{J_d} \mathbf{x}_{dj}\lambda_j \; ; \quad \mathbf{y} \leq \sum_{j=1}^{J_d} \mathbf{y}_{dj}\lambda_j \; ; \\ x^{TL} - s^{TL} = \sum_{j=1}^{J_d} x_{dj}^{TL}\lambda_j \; ; \; NPL - s^{NPL} = \sum_{j=1}^{J_d} NPL_{dj}\lambda_{dj} \; ; \\ s^{NPL} = \alpha_d s^{TL} \; ; \quad \sum_{j=1}^{J_d} \lambda_{dj} = 1; \\ \lambda_{dj} \geq 0 \; ; \quad s^{TL} \geq 0; \quad s^{NPL} \geq 0 \end{array} \right. \right\}$$

where J_d denotes the number of banks of loan quality d and α_d is the NPL-generating exogenous factor specific to group d. This d-specific technology can be written as

$$T_{HR,d} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_{+}^{6} \left| \begin{array}{c} \mathbf{x} \geq \sum_{j}^{J_d} \mathbf{x}_{dj}\lambda_{dj} \; ; \quad \mathbf{y} \leq \sum_{j=1}^{J_d} \mathbf{y}_{dj}\lambda_{dj} \; ; \\ x^{TL} = \sum_{j=1}^{J_d} x_{dj}^{TL}\lambda_{dj} \left(1 + \dfrac{s^{TL}}{\sum_{j}^{J_d} x_{dj}^{TL}\lambda_{dj}}\right) ; \\ NPL = \sum_{j=1}^{J_d} NPL_{dj}\lambda_{dj} \left(1 + \dfrac{\alpha_d s^{TL}}{\sum_{j=1}^{J_d} NPL_j\lambda_{dj}}\right); \\ \sum_{j=1}^{J_d} \lambda_{dj} = 1; \quad \lambda_{dj} \geq 0 \;\; (j = 1, \dots, J_d); \;\; s^{TL} \geq 0 \end{array} \right. \right\}$$

Since no single bank within group d can independently manage the relationship between non-performing loans and total loans, we put $NPL_{dj} = \alpha_d x_{dj}^{TL}$ and hence we have $NPL = \sum_{j=1}^{J_d} NPL_{dj}\lambda_{dj} = \alpha_d \sum_{j}^{J_d} x_{dj}^{TL}\lambda_{dj}$. Consequently, we have

$$T_{HR,d} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_{+}^{6} \left| \begin{array}{c} \mathbf{x} \geq \sum_{j}^{J_d} \mathbf{x}_{dj}\lambda_{dj} \; ; \quad \mathbf{y} \leq \sum_{j=1}^{J_d} \mathbf{y}_{dj}\lambda_{dj} \; ; \\ x^{TL} = \sum_{j=1}^{J_d} x_{dj}^{TL}\lambda_{dj}\, \rho; \\ NPL = \sum_{j=1}^{J_d} NPL_{dj}\lambda_{dj}\, \rho; \\ \sum_{j=1}^{J_d} \lambda_{dj} = 1; \quad \lambda_{dj} \geq 0 \;\; (j = 1, \dots, J_d); \; \rho \geq 1 \end{array} \right. \right\}$$

where $\rho = 1 + \frac{s^{TL}}{\sum_j^{J_d} x_{dj}^{TL}\lambda_{dj}}$. Since the generalized weak G-disposable technology T_{HR} is written as Eq. (5) as the convex hull, it can be written as

$$T_{\text{HR}} = \text{convex}\bigcup_{d=1}^{D} T_{\text{HR},d} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^6 \middle| \begin{array}{c} \mathbf{x} \geq \sum_{d=1}^{D}\sum_{j}^{J_d} \mathbf{x}_{dj}\lambda_{dj}\,; \quad \mathbf{y} \leq \sum_{d=1}^{D}\sum_{j=1}^{J_d} \mathbf{y}_{dj}\lambda_{dj} \\ x^{TL} = \sum_{d=1}^{D}\sum_{j=1}^{J_d} x_{dj}^{TL}\lambda_{dj}\,\rho\,; \\ NPL = \sum_{d=1}^{D}\sum_{j=1}^{J_d} NPL_{dj}\lambda_{dj}\,\rho\,; \; \sum_{d=1}^{D}\sum_{j=1}^{J_d}\lambda_{dj} = 1; \\ \lambda_{dj} \geq 0 \;\; (j = 1,\ldots,J_d; d = 1,\ldots,D); \quad \rho \geq 1 \end{array} \right\}$$

Considering $\sum_{d=1}^{D} J_d = J$, this technology can also be written as

$$T_{\text{HR}} = \left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \\ x^{TL} \\ NPL \end{pmatrix} \in \mathbb{R}_+^6 \middle| \begin{array}{c} \mathbf{x}_o \geq \sum_{j=1}^{J} \mathbf{x}_j\,\lambda_j; \quad \mathbf{y} \leq \sum_{j=1}^{J} \mathbf{y}_j\lambda_j\,; \\ x^{TL} = \sum_{j=1}^{J} x_j^{TL}\lambda_j\;\rho; \\ NPL = \sum_{j=1}^{J} NPL_j\lambda_{j;} \quad \rho \geq 1; \\ \sum_{j=1}^{J}\lambda_j = 1; \;\; \lambda_j \geq 0, \; j = 1,\ldots,J \end{array} \right\}$$

Relative to T_{HR}, the following nonlinear program can be constructed:

$$\min\left\{ \theta \middle| \begin{array}{c} \mathbf{x}_o \geq \sum_{j=1}^{J} \mathbf{x}_j\,\lambda_j; \quad \mathbf{y}_o \leq \sum_{j=1}^{J} \mathbf{y}_j\lambda_j\,; \\ \theta \cdot x^{TL} = \sum_{j=1}^{J} x_j^{TL}\lambda_j\,\rho; \\ \theta \cdot NPL = \sum_{j=1}^{J} NPL_j\lambda_j\,\rho \\ \sum_{j=1}^{J}\lambda_j = 1; \;\; \lambda_j \geq 0, \; j = 1,\ldots,J; \;\; \theta\text{: free} \end{array} \right\}$$

If ρ exceeds 1, the optimal value for θ would experience a further decrease in this minimization problem. So we set $\rho = 1$, leading to Eq. (6).