Omega 127 (2024) 103099

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

Assessing the performance of banks through an improved sigma-mu multicriteria analysis approach

Silvia Angilella [a], Michalis Doumpos [b], Maria Rosaria Pappalardo [a,*], Constantin Zopounidis [b]

[a] Department of Economics and Business, University of Catania, Corso Italia 55, 95129 Catania, Italy
[b] Technical University of Crete, School of Production Engineering and Management, Financial Engineering Laboratory, University Campus, 73100 Chania, Greece

ARTICLE INFO

Keywords:
Multicriteria decision analysis
Composite indicators
Banking
ESG criteria

ABSTRACT

The σ-μ efficiency methodology has been recently introduced for multicriteria evaluation problems, based on the framework of Stochastic Multi-Attribute Acceptability analysis (SMAA), to address the uncertainty in the performance of a set of decision alternatives. The methodology builds, iteratively, a set of Pareto-Koopmans efficiency frontiers, which are used to assess the alternatives with respect to their expected performance and its variability, measured across different scenarios for the weights of the evaluation criteria. This paper presents an improved algorithmic implementation of this methodology that provides results that are consistent with the Pareto dominance relation between the alternatives. The proposed approach is employed to evaluate the performance of a sample of European banks which participated in the European stress tests conducted by the European Banking Authority, over the last five years available (2017–2021). The performance and efficiency of the banks is analyzed using financial criteria along with environmental, social, and governance (ESG) factors. Results from comprehensive and disaggregated analysis reveal performance disparities among banks in financial and ESG factors, highlighting the influence of country-specific green policies and individual bank practices. Valuable for the banking sector and regulators, the findings help identify operational inefficiencies and propose areas for performance enhancement, operational improvement, and innovation, with a focus on green practices.

1. Introduction

Banking performance and efficiency have been central research topics in the literature, investigating different aspects of a bank, such as financial, social and environmental factors or also measuring its impacts in terms of systemic risk, as part of a financial network. On the basis of the multi-faceted nature of banks' operations, it is easy to understand the complexity behind the evaluation framework of bank performance and efficiency. Within this context, an important issue involves the specification of appropriate models for aggregating the different points of view that should be considered. To this end, composite indicators (CIs) come into play. In recent years, CIs have been considered as useful tools in different contexts, as evident by the plethora of such indicators released by global organizations, such as the World Bank, the European Commission, the Organization for Economic Cooperation and Development (OECD), etc.

CIs have also been a research theme for many scholars. From a mathematical point of view, CIs give rise to different issues: the choice of the aggregation function, the structure of the criteria and the procedures for defining their weights, the robustness of the results, and the participation of eventual stakeholders (see Beliakov et al. [1] and Greco et al. [2] for details on methodological issues of CIs and OECD [3] for a useful guide for practitioners on this topic).

Within this research stream, a recent methodology for constructing CIs, called σ-μ efficiency analysis, has been introduced by Greco et al. [4], combining tools and models from the field of multi-criteria decision aid (MCDA), with ideas stemming from efficiency measurement. The methodology starts with evaluating a set of units (i.e., decision alternatives) through a MCDA model (e.g., a weighted sum) taking into account multiple evaluation criteria. The evaluation is based on a simulation process that considers the whole set of admissible criteria weights under the framework of Stochastic Multi-Attribute Acceptability analysis (SMAA, Lahdelma et al. [5]). The simulation process provides estimates about the expected (mean, μ) performance of the alternatives and its variability (σ) over the different weighting scenarios used in the simulation. An alternative is considered efficient, compared to its peers, if it achieves high performance with low variability. The analysis is based on the concept of Pareto-Koopmans efficiency, under

* Corresponding author.
E-mail addresses: silvia.angilella@unict.it (S. Angilella), mdoumpos@tuc.gr (M. Doumpos), maria.pappalardo1@unict.it (M.R. Pappalardo), kzopounidis@tuc.gr (C. Zopounidis).

https://doi.org/10.1016/j.omega.2024.103099
Received 2 October 2023; Accepted 13 April 2024
Available online 24 April 2024

which an alternative a_i is efficient if there is no convex combination C of the remaining units, such that $\mu_C \geq \mu_i$ and $\sigma_C \leq \sigma_i$, with at least one of these inequalities being strict. In Greco et al. [4], it is argued that if an alternative is quite far from the σ-μ Pareto-Koopmans efficiency frontier (PKF), then it is not meaningful to compare it to the peers of that frontier. Instead, benchmarks for these remote peers should be those much more closer. For this reason, the appealing and interesting idea of σ-μ efficiency methodology, inspired by the study of Seiford and Zhu [6], is to construct a set of PKFs through an iterative process. Each of these frontiers define a local efficiency score for the alternatives, which are then aggregated (i.e., averaged over the set of PKFs) to obtain the overall evaluation results.

In this paper, we note that the standard iterative implementation of σ-μ efficiency analysis approach, may lead to results that are inconsistent with the Pareto dominance relation among the alternatives in the σ-μ plane. To address this limitation an improved iterative algorithmic procedure is developed. The proposed procedure is employed to evaluate bank performance through financial and non-financial criteria related to environmental, social, and governance factors (ESG, Un [7]). The combination of financial and ESG criteria was also considered by Ishizaka et al. [8] for describing the performance of US banks through a multicriteria clustering approach based on the integration of SMAA with the PROMETHEE method. In contrast to this clustering process, the approach developed in this study enables the derivation of a performance score for the banks that allows the ranking of the banks. For the purposes of this application, a sample of banks that participated in the European stress tests of the European Banking Authority (EBA) is considered over the period 2017–2021. The empirical results highlight the discrepancies that exist between the performance of banks on financial and ESG factors.

The rest of the paper is organized as follows: Section 2 presents a brief literature review on CIs and efficiency analysis. Section 3 presents the details of the proposed multicriteria methodology, together with an illustrative example. Section 4 is devoted to the application of the proposed methodology on the evaluation of banks that have participated in the EU-wide stress tests, including the description of the data and the multicriteria evaluation results as well as a comparison to the outputs of a well known efficiency model. Finally, Section 5 concludes the paper and discusses some practical implications of the analysis and a few directions for future research.

2. Literature review

In recent decades, extensive research has been conducted on the topic of banking performance and efficiency, primarily due to its role in facilitating investment allocation, offering insights into a country's financial strength, and thereby aiding in the prevention of economic crises [9]. As suggested by Berger and Humphrey [10], measuring bank efficiency provides valuable understanding at both the micro and macro levels of analysis. Nevertheless, given the broad range of activities undertaken by banks, such as their role as intermediaries, liquidity providers, information channels, risk managers and drivers of innovations, bank performance has become a subject of extensive discussion among academics.

The literature can be described through the identification of two major research streams. The first focuses on efficiency measurement through an input–output analysis approach. In this context, bank efficiency typically pertains to two main dimensions: technical efficiency (TE) and allocative efficiency (AE), which are challenging to disentangle [11]. TE focuses on achieving the optimal output using the least amount of inputs and costs, while AE deals with the effective allocation of various inputs to produce a diverse range of outputs [12].

The assessment of overall economic efficiency (EE), obtained by multiplying TE and AE, is typically based on parametric or non-parametric approaches. The former relies on a particular functional form and entails estimating an economic function that establishes the relationship between the outputs of a decision-making unit (DMU) and its input factors. Non-parametric approaches, on the other hand, adopt an optimization-based approach without assuming a specific production (input–output) function.

On the parametric side, stochastic frontier analysis (SFA, Aigner et al. [13]) has been the most popular approach for bank efficiency analysis. SFA has the advantage of distinguishing between inefficiency arising from random shocks, by those resulting from the technical inefficiency of the firm and controlling for measurement errors and other random effects. However, it has the limitation of imposing a specific functional form of production or cost function that, if incorrectly specified, could lead to significantly biased estimated efficiency. Major real-world applications and methodological contributions have used the SFA approach for estimating banking efficiency [14,15]. Among these, some have estimated cost and profit functions [16–19], while some others a production function [20]. With respect to the specific functional form, the translog [15,21,22] and the Fourier flexible form [14] have been mainly employed within the banking industry. Regarding non-parametric methods, data envelopment analysis (DEA) and free disposal hull (FDH) have been widely popular methods. DEA relies on linear programming (LP) formulations to construct a piece-wise linear frontier that encompasses the observed data points, thus resulting in a convex production possibilities set. Furthermore, it assesses the relative efficiency of each DMU by considering its distance from the efficiency frontier [23]. In contrast, FDH assumes free disposability of inputs and outputs rather than convexity, thus leading to larger estimates of average efficiency than DEA [24].

A significant part of the literature has utilized various DEA models to evaluate the efficiency of banks [25–28]. Additionally, numerous studies have been conducted to compare the efficiency outcomes obtained by applying DEA and SFA methodologies in the banking industry. However, the findings of these studies show inconsistent results in terms of the congruence and correlation of efficiency outcomes between the two methodologies [29–31]. More specifically, researchers employing DEA models found greater efficiency scores compared to stochastic models because of the substantial and favorable fluctuation of the random factor (see the recent literature review of Aiello and Bonanno 32 for a thorough assessment of the factors influencing the diversity of bank efficiency).

Nevertheless, it should be noted that approaches for evaluating the efficiency of banks in an input–output context focus on financial aspects (e.g., cost efficiency, allocative efficiency, profit efficiency, production efficiency, etc.). While efficiency estimates are useful for analyzing the operation of banks, the performance of banking institutions cannot be fully captured by an input–output analysis; it requires the adoption of a broader perspective of the risks and challenges that banks face. Thus, another large part of the literature has considered a different context that goes beyond the derivation of efficiency estimates. In this area, two sub-streams can be identified, the first focusing the development of models for evaluating bank failures and the second focusing on the construction of CIs for assessing the performance of banks.

The analysis of the risk of failure of the banks is typically addressed in a predictive modeling context. While statistical models are widely used in the finance literature, approaches stemming from operational research (OR) and artificial intelligence (AI) have also attracted significant interest [33], mainly due to their superior predictive performance. As far as OR methodologies are concerned, some indicative applications include the use of MCDA models [34,35], stochastic approaches [36], as well as network-based models [37], and simulation [38]. Among AI approaches, one can mention methodologies based on neural networks [39], support vector machines [40], fuzzy models [41], ensembles [42,43], as well as technologies such as natural language processing and deep learning [44].

The last part of the literature, which is more closely related to our study, involves the evaluation of bank performance by constructing CIs, which allow for comprehensive analyses to be performed combining both expert judgments as well as data-driven approaches. The use of CIs is motivated by the recognition that individual components alone cannot fully capture the multidimensional nature of overall performance. CIs facilitate comprehensive assessments, enabling judgments on the overall system performance with a more holistic approach [45]. The development of a CI involves two pivotal stages: the weighting process and the aggregation procedure, where assigned weights are applied to individual indicators, ultimately determining the robustness of the CI [46]. Two widely employed approaches for weighting and aggregating CIs are MCDA and DEA. Notably, the weighting aspect within MCDA models tends to be more contentious than DEA, as it requires experts and analysts well-versed in economic priorities and theoretical foundations [47]. However, the involvement of experts in the performance assessment of banks serves to highlight the diverse perspectives of stakeholders, including customers, investors, public administration, and regulators, thus enhancing the overall reliability of the analysis. Various CIs have been proposed by academics to evaluate the performance of banks. Some pioneering studies focus solely on a financial-centered perspective for assessment. A notable illustration of this is the Bank Overall Financial Strength Index (BOFSI), developed by Doumpos et al. [48]. This index integrates five financial criteria aligned with the categories of the CAMEL framework. Over time, CIs have evolved to adopt a more systematic approach and have begun incorporating additional aspects, such as ESG factors. A recent example of this is the Banking Sustainability Performance Index (BSPI), elaborated by Rebai et al. [49], which is employed to evaluate and rank according to their sustainability degree, whereas Gulati et al. [50] employed a DEA-based approach to evaluate the performance of banks with respect to their corporate governance practices.. A broader perspective was adopted by Gaganis et al. [51], who introduced the CISEF index to assess the social, environmental, and financial performance of corporate clients, considering a broad range of internal and external stakeholders. To address conflicting preferences among shareholders, the CISEF index incorporates the σ-μ efficiency analysis introduced by Greco et al. [4].

Table 1 in the supplementary material summarizes the models described in this section to assess the performance and efficiency of banks. Following the last of the above streams of the literature, this study builds on the grounds of the works of Greco et al. [4],Gaganis et al. [51], and presents a new improved implementation of the σ-μ efficiency analysis approach, which ensures the derivation of results that are consistent with standard MCDA principles. The σ-μ analysis methodology fits well the context of bank performance evaluation, because it combines elements from judgmental MCDA approaches with a data-driven process based on the principles of DEA. The proposed approach for implementing the σ-μ analysis is applied to the evaluation of the performance of European banks, based on financial and non-financial factors, as well as taking into consideration the economic uncertainty of the external environment.

3. Methodological framework

Having reviewed the relevant literature on methodologies for the construction of CIs in the area of banking, this section presents the proposed approach, which combines elements from the areas of MCDA and efficiency analysis. The presentation starts with an outline of the σ-μ efficiency assessment framework (Section 3.1), and then proceeds in Section 3.2 with the description of the proposed methodology, which overcomes the limitations of the basic σ-μ efficiency analysis procedure. An illustrative example is given in Section 3.3.

3.1. An outline of σ-μ efficiency analysis

Let $A = \{a_1, a_2, \dots, a_m\}$ be a set of alternatives to be evaluated on the basis of a set of n criteria $G = \{g_1, g_2, \dots, g_n\}$ with $\mathcal{I} = \{1, 2, \dots, m\}$ being the set of indices of the alternatives. Each alternative $a_i \in A$ is described by the vector of its evaluations on the criteria set G, i.e., $a_i = [g_1(a_i), g_2(a_i), \dots, g_n(a_i)]$, where $g_h(a_i)$ indicates the evaluation of alternative a_i on criterion g_h. The criteria are aggregated into a single performance score for each alternative using a CI. Several aggregation functions have been analyzed in the literature (see Beliakov et al. [1] and references therein). The most simple model used in many papers is the weighted average, which is defined as follows for any $a_i \in A$:

$$CI(a_i, \mathbf{w}) = \sum_{h=1}^{n} g_h(a_i) w_h, \tag{1}$$

where $\mathbf{w}$ is the vector with the criteria weight, such that $\mathbf{w} \in W = \{(w_1, w_2, \dots, w_n) \in \mathbb{R}^n : w_h \geq 0 \text{ and } \sum_{h=1}^{n} w_h = 1\}$.

There are several methodological issues related to the construction of CIs, with the specification of the criteria weights being the most crucial point (see OECD [3]). On the one hand, a well-defined set of weights could be representative of a specific decision-maker (DM). On the other hand, the whole space of weight vectors W should be taken into account if multiple set of stakeholders' preferences have to be incorporated as is often the case in many contexts, such as financial, economic, or environmental management decisions. To this end, in this study a simulation analysis based on the framework of Stochastic Multiobjective Acceptability Analysis (SMAA, Lahdelma and Salminen [52]) is implemented, taking into account several stakeholders' points of view.

The SMAA methodology enables the computation of the probability that an alternative is ranked first, second, and so on (rank acceptability index) as well as the identification of the preferences of a typical DM that correspond to the best evaluation for an alternative a_i (central weight vector). As observed in Gaganis et al. [51], SMAA indices are not suitable for CIs, especially when the number of alternatives is quite high, because the estimated probabilities tend to become very small. Moreover, since the nature of a CI is essentially cardinal, it is more useful to have a metric giving the magnitude of the performance of an alternative. To this end, stemming from the Markowitz mean–variance analysis framework [53], Greco et al. [4] proposed to synthesize the performance of an alternative with its mean value μ and standard deviation σ.

More specifically, under the SMAA framework, the mean value and the standard deviation of the CI for each alternative $a_i \in A$, are computed using a random sampling of S vectors of criteria weights, uniformly distributed over the unit simplex. A typical choice for the number of simulated weights vectors is to set $S = 10\,000$ (Lahdelma et al. [5]).

Considering the sampled vectors of criteria weights, the mean (μ_i) and standard deviation (σ_i) of the CI for each alternative $a_i \in A$ are defined as follows:

$$\mu_i = \frac{1}{S}\sum_{r=1}^{S} CI(a_i, \mathbf{w}_r), \qquad \sigma_i = \sqrt{\frac{1}{S}\sum_{r=1}^{S} [CI(a_i, \mathbf{w}_r) - \mu_i]^2}. \tag{2}$$

Larger values of μ_i denote better performance and lower values of σ_i denote a more stable performance adhering to a large portion of stakeholders.

Based on these two estimates, one can define a σ-μ Pareto dominance relation D for any pair of alternatives $\{a_i, a_j\}$, as follows:

$$a_i \, D \, a_j \Leftrightarrow (\mu_i \geq \mu_j \wedge \sigma_i < \sigma_j) \vee (\mu_i > \mu_j \wedge \sigma_i \leq \sigma_j). \tag{3}$$

Under this definition, alternative a_i dominates a_j if and only if:

(i) the mean performance score of a_i is at least as high as the one of a_j, and
(ii) the standard deviation (i.e., the uncertainty/variability) of the performance of a_i is at most equal to that of a_j,

with at least one of the inequalities involved in these conditions being strict.

A concept stricter than the σ-μ Pareto efficiency is given by σ-μ Pareto-Koopmans efficiency, under which an alternative a_i is efficient if there is no convex combination $x = (\mu_x, \sigma_x)$ of the remaining alternatives, with

$$\mu_x = \sum_{j \neq i} \lambda_j \mu_j \quad \text{and} \quad \sigma_x = \sum_{j \neq i} \lambda_j \sigma_j,$$

such that $x\, D\, a_i$, where $\lambda_1, \ldots, \lambda_m \geq 0$ and $\lambda_1 + \cdots + \lambda_m = 1$.

The σ-μ Pareto-Koopmans efficiency of an alternative a_i is tested with the following LP problem:

$$\max \delta_i \quad \text{s.t.} \quad \begin{cases} \mu_i \alpha - \sigma_i \beta \geq \mu_j \alpha - \sigma_j \beta + \delta_i \;\forall\; j \neq i \\ \alpha + \beta = 1 \\ \alpha, \beta \geq 0, \delta_i \in \mathbb{R} \end{cases} \tag{4}$$

If there exists a pair (α, β) with a non-negative δ_i which verifies the constraints of LP (4), then a_i is efficient in the Pareto-Koopmans sense. The signs of the coefficients α and $-\beta$ are in accordance with the Markowitz's mean–variance theory, under which μ_i and σ_i have to be, respectively, maximized and minimized.

The value of δ_i can be interpreted as a measure of efficiency, with negative values indicating inefficiency and positive values corresponding to greater efficiency.

In Greco et al. [4], a concept of local efficiency was introduced, since in many cases an alternative is quite far from the σ-μ PKF, so it is not meaningful to compare it with alternatives on this frontier. For this reason, in Greco et al. [4] a sequence of σ-μ PKFs has been taken into consideration, following the idea originally introduced by Seiford and Zhu [6]. In this process, a sequence of σ-μ PKFs is obtained, denoted by $F_1, F_2, \ldots, F_p$. Each frontier F_k at level k consists of alternatives that are efficient (in terms of the Pareto-Koopmans efficiency concept) compared to the rest of the alternatives, excluding those belonging to "higher" (lower level) frontiers, i.e. to $\mathcal{F}_{k-1} = \{F_1 \cup \cdots \cup F_{k-1}\}$.

Similarly to the concept of global efficiency of an alternative a_i evaluated with the LP problem (4), a local $\sigma - \mu$ Pareto-Koopmans efficiency score δ_{ik} at level k, can be defined by solving the following LP:

$$\max \delta_{ik} \quad \text{s.t.} \quad \begin{cases} \mu_i \alpha - \sigma_i \beta \geq \mu_j \alpha - \sigma_j \beta + \delta_{ik} \;\forall\; j \neq i, j \in \mathcal{P}_k \\ \alpha + \beta = 1 \\ \alpha, \beta \geq 0 \end{cases} \tag{5}$$

where $\mathcal{P}_k = \mathcal{I} \setminus \mathcal{F}_{k-1}$ is the set of peers for evaluating the local efficiency of the alternatives with respect to frontier k. Henceforth, the maximum value of δ_{ik} derived from the solution of this LP, will be denoted as δ^*_{ik}. The local frontier F_k consists of alternatives not belonging to $\mathcal{F}_{k-1}$, such that $\delta^*_{ik} \geq 0$.

The sum of the local efficiency scores $\delta^*_{i1}, \delta^*_{i2}, \ldots$ for each alternative is used to derive its global efficiency score s_i, i.e., $s_i = \sum_k \delta^*_{ik}$. Since s_i can be also negative, a normalized score $\bar{s}_i \in [0, 1]$ can be computed for any alternative $a_i \in A$, as follows:

$$\bar{s}_i = \frac{s_i - \min_i s_i}{\max_i s_i - \min_i s_i}. \tag{6}$$

To illustrate the process, a numerical example is presented below. The example is also used to demonstrate a shortcoming of the above implementation of the σ-μ efficiency analysis approach, as the results may violate the dominance relation between the alternatives.

An example

To illustrate the process of σ-μ efficiency analysis, let us consider the following example with five alternatives evaluated on μ and σ as shown in Table 1.

The procedure starts with the derivation of the first frontier F_1. To this end, problem (5) is solved for all alternatives, setting $\mathcal{P}_1 = \{1, \ldots, 5\}$, for all $i = 1, \ldots, 5$. The results for the resulting local efficiency scores are shown on the left upper chart of Fig. 1. According to these results, frontier F_1 consists of alternatives a_1 and a_4, which have positive local efficiency scores ($\delta^*_{11} = 0.002$, $\delta^*_{41} = 0.003$). In the second iteration, the identification of the second frontier, F_2, is based on the solution of problem (5) for all alternatives, setting $\mathcal{P}_2 = \{2, 3, 5\}$ (i.e., the peer set consists of the alternatives not yet assigned to an efficiency frontier). The results for the new local efficiency scores are shown on the upper right chart of Fig. 1. Among the alternatives not yet assigned to an efficiency frontier, only a_2 has a positive local efficiency score ($\delta^*_{22} = 0.01$). Therefore, $F_2 = \{a_2\}$. In the final iteration, problem (5) is solved again for all alternatives, using the peer set defined by the remaining two alternatives $\mathcal{P}_3 = \{3, 5\}$. According to the results shown in the bottom chart of Fig. 1, among the two unassigned alternatives, a_3 has a positive local efficiency score. Therefore, it is assigned to the third local efficiency frontier, i.e., $F_3 = \{a_3\}$, whereas the remaining alternative a_5 is assigned to the last (worst) frontier F_4, and the procedure ends.

Table 1
Alternatives with σ-μ evaluations

	a_1	a_2	a_3	a_4	a_5
μ	0.723	0.721	0.714	0.665	0.709
σ	0.024	0.024	0.034	0.021	0.076

Table 2
Efficiency scores according to the σ-μ efficiency analysis procedure for the example.

Alternative	δ^*_{i1}	δ^*_{i2}	δ^*_{i3}	s_i	$\bar{s}_i$
a_1	0.0020	0.0020	0.0100	0.0140	0.7895
a_2	−0.0001	0.0100	0.0100	0.0199	0.8930
a_3	−0.0090	−0.0070	0.0420	0.0260	1.0000
a_4	0.0030	0.0030	0.0130	0.0190	0.8772
a_5	−0.0140	−0.0120	−0.0050	−0.0310	0.0000

Table 2 summarizes the results, including the local efficiency scores $(\delta_{i1}, \delta_{i2}, \delta_{i3})$, the global scores ($s_i$), and the normalized global scores ($\bar{s}_i$). According to the results, the alternatives are ranked as $a_3 \succ a_2 \succ a_4 \succ a_1 \succ a_5$, where $\succ$ denotes the preference relation. However, as evident from Fig. 1, alternative a_3 is dominated by a_1 and a_2, thus indicating that the results are inconsistent with the dominance relations among the alternatives.

The inconsistency in the results arises due to the way the peer set $\mathcal{P}_k$ is defined when evaluating the alternatives. For instance, at level $k = 2$, alternative a_2 is compared against alternatives a_3 and a_5, whereas at the same level, alternative a_1 is compared against a_2, a_3, and a_5. Given the closeness (i.e., similarity) of a_1 to a_2, this leads to a lower local efficiency score for a_1 than a_2 ($\delta^*_{12} = 0.002 < \delta^*_{22} = 0.01$). In order to address this problem, in the next section we propose an improved implementation of the iterative process of σ-μ efficiency analysis.

3.2. Determining the σ-μ PKFs considering the pareto dominance relation

To overcome the problem illustrated in the example of the previous sub-section, the iterative procedure used in the σ-μ efficiency analysis is extended to consider the Pareto dominance relations among the alternatives. The process is implemented through algorithm 1.

The key idea underlying the proposed algorithm, is that if, for instance, an alternative B is benchmarked (i.e., evaluated) against some other alternatives C, D, and E, then any alternative A that dominates B should also be compared against C, D, and E. Using the same set of peers to evaluate both A and B, ensures that the efficiency score of the dominating alternative A will not be lower than the one of the dominated alternative B. If A and B are not compared against the same set of peers (e.g., B is compared to C, D, E, whereas A is compared to B, C, D, E), then it may happen that the efficiency score of A will be lower than the one of B, even though A dominates B. This reasoning should also consider the possibility that except for alternative B, alternative A may also dominate another alternative among those considered in the evaluation, e.g., alternative C. So, the evaluation of A should be

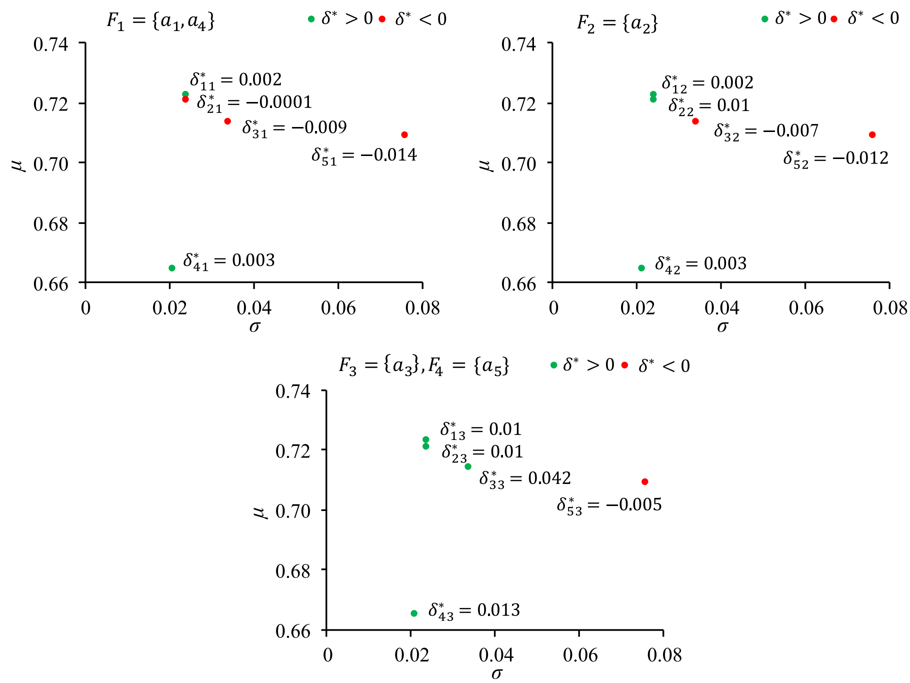

Fig. 1. σ-μ frontiers for the alternatives in the example and local efficiency scores.

repeated to compare it against the set of peers that is used to evaluate alternative C. Given the two ways to evaluate A (through B and C), the best (i.e., maximum) of the two resulting scores should be used to ensure that A will not be judged as being inferior to either B or C.

Having in mind this informal description of the proposed process, at each iteration k of the standard implementation of the σ-μ analysis that was described in the previous sub-section, the efficiency front F_k is obtained by solving LP (5) in lines 14–15 of algorithm 1 for all alternatives not previously assigned to a higher-level frontier (i.e., for all alternatives in $\mathcal{I} \setminus \mathcal{F}_{k-1}$). In the proposed implementation, these steps are only applied to alternatives from higher-level frontiers that do not dominate any of the remaining alternatives (see the condition in line 13). For alternatives not assigned to a higher-level frontier that dominate at least one of the remaining alternatives, the steps in lines 6–12 are implemented. These steps operate following the reasoning described above, so that when evaluating the performance of an alternative a_i, against its peers from $\mathcal{I} \setminus \mathcal{F}_{k-1}$ at level k, all alternatives that dominate a_i are also compared against the exact same set of peers. Thus, this part of the algorithm (i.e., lines 6–12) ensures that the obtained performance scores will be consistent with the dominance relations between the alternatives.

As explained above, an alternative a_ℓ that dominates a_i may also dominate other alternatives at level k, thus leading to multiple local efficiency scores, denoted by $\Delta^*_{i\ell k}$, each corresponding to the evaluation of a_ℓ based on the set of peers defined by the dominated alternative a_i. The final performance score for such comparisons examined at level k is defined by the maximum of all the different results that are obtained at this level (see line 12). The maximum is selected in order to ensure that the results will be consistent with the dominance relation.

3.3. Illustration of the process

To illustrate the proposed methodology, let us consider again the example of the previous sub-section with the data of the five alternatives in Table 1. To facilitate the presentation Table 3 presents the dominance relations between the alternatives. Each entry (i, j) of this table indicates whether the dominance relation $a_i \, D \, a_j$ holds or not. For

Algorithm 1 Derivation of the $\sigma - \mu$ efficiency scores

Input: Mean performance and standard deviation (μ_i, σ_i) for a set $\mathcal{I}$ of m alternatives
Output: Global and local efficiency scores $(\delta_i, \delta^*_{ik})$
1: $F_0 \leftarrow \emptyset$
2: Solve LP (5) for all $i \in \mathcal{I}$
3: $F_1 \leftarrow \{i \in \mathcal{I} : \delta^*_{i1} \geq 0\}$
4: $k \leftarrow 2$
5: **while** $|\mathcal{F}_{k-1}| < m - 1$ **do**
6: $\quad \Delta^*_{i\ell k} \leftarrow \infty, \forall i, \ell \in \{1, \dots, m\}$
7: $\quad$ **for all** $i \in \mathcal{I} \setminus \mathcal{F}_{k-1}$ **do**
8: $\quad\quad \mathcal{P}_{ik} \leftarrow \mathcal{I} \setminus \{i \cup \mathcal{F}_{k-1}\}$
9: $\quad\quad$ Specify the set $\mathcal{D}_i$ of alternatives dominating a_i: $\mathcal{D}_i \leftarrow \{j \in \mathcal{F}_{k-1} : a_j D a_i\}$
10: $\quad\quad$ If $\mathcal{D}_i \neq \emptyset$, use $\mathcal{P}_{ik}$ to solve (5) and obtain efficiency scores $\Delta^*_{i\ell k}, \forall \ell \in \{\mathcal{D}_i \cup i\}$
11: $\quad$ **end for**
12: $\quad \delta^*_{ik} \leftarrow \max_\ell \Delta^*_{\ell ik}, \forall i \mid \Delta^*_{\ell ik} < \infty$
13: $\quad$ **for all** $i \in \mathcal{F}_{k-1}$ such that $\nexists \, a_j \in \{\mathcal{I} \setminus \mathcal{F}_{k-1}\} \mid a_i D a_j$ **do**
14: $\quad\quad \mathcal{P}_k \leftarrow \mathcal{I} \setminus \mathcal{F}_{k-1}$
15: $\quad\quad$ Solve LP (5) to obtain the efficiency score δ^*_{ik}
16: $\quad$ **end for**
17: $\quad F_k \leftarrow \{i \in \mathcal{I} \setminus \mathcal{F}_{k-1} : \delta^*_{ik} \geq 0\}$
18: $\quad k \leftarrow k + 1$
19: **end while**
20: $\delta_i \leftarrow \sum_k \delta^*_{ik}, \forall i$

Table 3
Dominance relations between the alternatives of the example.

	a_1	a_2	a_3	a_4	a_5
a_1	–	D	D	–	D
a_2	–	–	D	–	D
a_3	–	–	–	–	D
a_4	–	–	–	–	–
a_5	–	–	–	–	–

instance, a_1 dominates a_2, a_3 and a_5. The dominance relations are also apparent by observing Fig. 1 that was used in the previous sub-section.

The first iteration (level $k = 1$) is identical to the one of the standard procedure of the σ-μ analysis, as explained in the example of the previous sub-section. More specifically, LP (5) is solved for all alternatives and the first set of local efficiency scores are obtained: $\delta^*_{11} = 0.002$, $\delta^*_{21} = -0.0001$, $\delta^*_{31} = -0.009$, $\delta^*_{41} = 0.003$, $\delta^*_{51} = -0.014$. These are the same local scores to the ones previously reported in column δ^*_{i1} of Table 2. The cases with non-negative scores are assigned to the first frontier, i.e., $F_1 = \{a_1, a_4\}$.

At the second iteration (i.e., level $k = 2$), all alternatives (including those from F_1) are compared against a_2, a_3, and a_5, which have not been assigned to a previous frontier. For the non-dominated alternative a_4, which was previously assigned to the first frontier F_1, lines 13–16 of Algorithm 1 are followed. Accordingly, this alternative is compared to $\{a_2, a_3, a_5\}$ and the following LP is solved, which yields the optimal solution $\delta^*_{42} = 0.003$:

$$\max \delta_{42} \quad \text{s.t.} \quad \begin{cases} \mu_4\alpha - \sigma_4\beta \geq \mu_2\alpha - \sigma_2\beta + \delta_{42} \\ \mu_4\alpha - \sigma_4\beta \geq \mu_3\alpha - \sigma_3\beta + \delta_{42} \\ \mu_4\alpha - \sigma_4\beta \geq \mu_5\alpha - \sigma_5\beta + \delta_{42} \\ \alpha + \beta = 1 \\ \alpha, \beta \geq 0, \delta_{42} \in \mathbb{R} \end{cases} \tag{7}$$

Given that a_4 does not dominate any of the remaining alternatives $\{a_2, a_3, a_5\}$, the solution of the above LP suffices for the assessment of the local efficiency of this alternative at the current level $k = 2$. This does not apply, however, to a_1, which dominates all three of the remaining alternatives. Thus, the local efficiency of a_1 should be considered taking into account $\{a_2, a_3, a_5\}$, as discussed in the process below, according to lines 7–12 of Algorithm 1.

More specifically, a_2 is compared against $\{a_3, a_5\}$, i.e., $\mathcal{P}_{22} = \{a_3, a_5\}$ in line 8 of Algorithm 1. The same set of peers is also used to evaluate the performance of a_1, because $a_1 D a_2$ (line 9). The resulting local efficiency scores for a_1 and a_2 are $\Delta^*_{212} = \Delta^*_{222} = 0.01$, where the first index corresponds to alternative a_2 that defines the set of peers, the second index corresponds to the alternative being evaluated, and the third index indicates the current level ($k = 2$).

Continuing the process, again at level $k = 2$, alternative a_3 is compared against $\{a_2, a_5\}$, i.e., $\mathcal{P}_{32} = \{a_2, a_5\}$ in line 8 of algorithm 1. From the alternatives previously assigned to frontier F_1, alternative a_1 dominates a_3 ($\mathcal{D}_3 = \{a_1\}$ in line 9). Therefore, the evaluation of the performance of a_1 should be repeated, this time using the same set of peers as those employed for a_3, i.e., $\{a_2, a_5\}$. The obtained efficiencies are $\Delta^*_{332} = -0.007$ and $\Delta^*_{312} = 0.002$.

At this point, there are two assessments for the local efficiency of a_1 at level $k = 2$. The first $\Delta^*_{212} = 0.01$ is taken using the set of peers $\{a_3, a_5\}$ that are defined by a_2, whereas the second is $\Delta^*_{312} = 0.002$ through the set of peers $\{a_2, a_5\}$ that are defined by a_3. The same process should also be applied to assess the local efficiency of a_5 at level $k = 2$, with the set of peers being $\mathcal{P}_{52} = \{a_2, a_3\}$. Given that a_1 also dominates a_5 ($\mathcal{D}_5 = \{a_1\}$ in line 9), a third evaluation of a_1 should be obtained, by comparing it against $\{a_2, a_3\}$. The obtained results are $\Delta^*_{552} = -0.012$ and $\Delta^*_{512} = 0.002$.

A summary of the above results at level $k = 2$ is presented in the left part of Table 4. The peers' set column indicates the set of peers against which each alternative is compared at level $k = 2$. Alternatives a_2, a_3, and a_5, which were not assigned in the first frontier F_1, are compared against the sets of peers $\{a_3, a_5\}$ (for a_2), $\{a_2, a_5\}$ (for a_3), and $\{a_2, a_3\}$ (for a_5). These comparisons lead to one evaluation for each of a_2, a_3, a_5, which is shown in column δ^*_{i2} (for simplicity the $\Delta^*_{i\ell 2}$ scores are omitted for these three alternatives as they coincide with the given δ^*_{i2} values). Alternative a_4, which belongs to the first frontier F_1, also has a unique set of peers, namely $\{a_2, a_3, a_5\}$, because it does not dominate any the remaining alternatives. On the other hand, alternative a_1, which was also assigned to the first frontier, dominates all of the remaining alternatives $\{a_2, a_3, a_5\}$. Thus, its evaluation at level $k = 2$ is based on three different sets of peers, each corresponding to one of the alternatives that are dominated by a_1:

(1) the set $\{a_3, a_5\}$, which is used to evaluate a_2,
(2) the set $\{a_2, a_5\}$, which is used to evaluate a_3, and
(3) the set $\{a_2, a_3\}$, which is used to evaluate a_5.

Table 4
Summary of comparisons and results at levels $k = 2$ and $k = 3$.

Alternative	Level $k = 2$			Level $k = 3$		
	Peers' set	$\Delta^*_{i\ell 2}$	δ^*_{i2}	Peers' set	$\Delta^*_{i\ell 3}$	δ^*_{i3}
a_1	$\{a_3, a_5\}$	0.010	0.010	$\{a_3\}$	0.052	0.052
	$\{a_2, a_5\}$	0.002		$\{a_5\}$	0.010	
	$\{a_2, a_3\}$	0.002				
a_2	$\{a_3, a_5\}$	–	0.010	$\{a_3\}$	0.052	0.052
				$\{a_5\}$	0.010	
a_3	$\{a_2, a_5\}$	–	−0.007	$\{a_5\}$	0.042	0.042
a_4	$\{a_2, a_3, a_5\}$	–	0.003	$\{a_3, a_5\}$	0.013	0.013
a_5	$\{a_2, a_3\}$	–	−0.012	$\{a_3\}$	−0.005	−0.005

Table 5
Local and global efficiencies scores.

	δ^*_{i1}	δ^*_{i2}	δ^*_{i3}	s_i	$\bar{s}_i$
a_1	0.0020	0.0100	0.0520	0.0640	1.000
a_2	−0.0001	0.0100	0.0520	0.0619	0.978
a_3	−0.0090	−0.0070	0.0420	0.0260	0.600
a_4	0.0030	0.0030	0.0130	0.0190	0.526
a_5	−0.0140	−0.0120	−0.0050	−0.0310	0.000

Each of these comparisons yields a different local efficiency score for a_1, as shown in column $\Delta^*_{i\ell 2}$. More specifically, the first comparison yields a score of $\Delta^*_{212} = 0.01$ for a_1, the second yields a score of $\Delta^*_{312} = 0.002$ and the last one yields a score of $\Delta^*_{512} = 0.002$. The maximum of these scores is the local efficiency score of a_1 that is shown in column δ^*_{i2}.

From the local efficiency scores δ^*_{ik} at level $k = 2$, it is evident that a_1, a_2, and a_4 have positive scores. Given that a_1 and a_4 have already been assigned to F_1, only a_2 is assigned to the second frontier F_2. Thus, after completing the second round of the process, there are two frontiers, $F_1 = \{a_1, a_4\}$ and $F_2 = \{a_2\}$. The process continues at level $k = 3$ with the remaining alternatives a_3 and a_5. The summary of comparisons and results is shown in the right part of Table 4, whereas Table 5 presents the final results, including all local efficiencies ($\delta^*_{i1}, \delta^*_{i2}, \delta^*_{i3}$), the global efficiency scores (s_i), and the normalized scores ($\bar{s}_i$). It is evident, that the proposed algorithmic implementation of the σ-μ efficiency analysis yields performance scores that respect the dominance conditions between the alternatives.

4. Application

This section presents the application of the proposed methodology to the evaluation of the performance of European banks. We start with the description of the data and then proceed with the presentation and discussion of the results. The results have been obtained in two settings: (i) by considering the comprehensive set of evaluation criteria (Section 4.2), and (ii) through a "disaggregated" analysis that distinguishes between the performance of the banks on financial and ESG criteria (Section 4.3).

4.1. Data description

The empirical analysis carried out in this paper relies on a sample of 76 European banks that have participated to the 2023 EU-wide stress test conducted by the EBA to assess the resilience of EU banks to severe

Table 6
Sample of banks from the 2023 EU-wide stress test of EBA.

Bank	Country	Bank	Country
BNP Paribas	France	Intesa Sanpaolo	Italy
Groupe Crédit Agricole	France	Mediobanca	Italy
Société Générale	France	UniCredit	Italy
Commerzbank	Germany	ABN AMRO Bank	Netherlands
Deutsche Bank	Germany	ING Groep	Netherlands
Deutsche Pfandbriefbank	Germany	Banco Bilbao Vizcaya Argentaria	Spain
Alpha Services & Holdings	Greece	Banco de Sabadell	Spain
Eurobank Ergasias Services	Greece	Banco Santander	Spain
National Bank of Greece	Greece	Bankinter	Spain
AIB Group	Ireland	CaixaBank	Spain
Bank of Ireland Group	Ireland	Unicaja Banco	Spain
Banca Monte dei Paschi di Siena	Italy	Skandinaviska Enskilda Banken	Sweden
Banco BPM	Italy	Svenska Handelsbanken	Sweden
BPER Banca	Italy	Swedbank	Sweden

adverse shocks (such as credit, market, counterparty, and operational risk, EBA [54]).

To evaluate the performance of the banks, we combine financial and non-financial information about the operation of the banks, as well as information about the external environment at the host country in which each bank operates. This information was collected from three sources:

- The BankFocus database of Moody's Analytics was used to obtain data on financial variables that capture bank-specific vulnerabilities. The selection of the financial criteria was based on the CAMELS framework [55], which is a widely accepted context for analyzing the financial soundness of banks, on six key dimensions. These refer to Capital adequacy, Asset quality, Management effectiveness, Earnings power, Liquidity, and Sensitivity to market risks.
- The Refinitiv database of Thomson Reuters was employed to obtain non-financial data about the exposure of each bank on environmental, social, and governance (ESG) risk factors, in accordance with the Sustainable Development Goals (SDGs) of Agenda 2030 (see Refinitiv [56] for the description of the ESG score methodology). Over the past decade, ESG has become a major point of interest among practitioners and researchers in the banking sector [57]. For European banks, this is clearly demonstrated by the introduction of climate stress tests by the EBA, whose aim is to analyze the resilience of the EU banking system to the potential impact of climate risks. The consideration of ESG criteria in the analysis allows the investigation of the impact they have on the performance and soundness of the banks.
- Regarding the external environment, the Economic Policy Uncertainty index (EPU_index)[1] is used as an indicator of economic policy uncertainty at the host country of each bank. The index is compiled using data collected from articles regarding policy uncertainty from the leading newspapers. Higher values of the index represent stronger uncertainty and higher banking risks [58].

Because of missing ESG and EPU data for some banks and countries in the sample, our final dataset consists of 28 European banks from eight countries, evaluated on CAMELS, ESG, and EPU criteria over the last five years available, i.e. 2017–2021. However, since values on the EPU_index are collected monthly, its annual median value has been used. Furthermore, the linear interpolation method has been employed in the final sample to estimate the unknown values for some specific missing years in CAMELS and ESG variables.

The full list of banks is reported in Table 6, while the selected CAMELS, ESG and EPU criteria are described in Table 7. The same table indicates the preference direction of the criteria, distinguishing between criteria that are positively (max) or negatively (min) associated with bank performance. For instance, capital adequacy has a positive association with the performance of banks, because it decreases the risk of failure. On the other hand, the cost to income ratio is negatively associated with the performance of banks, because it indicates the cost burden of banks in relation to their income (i.e, revenue). Table A.1 in the Appendix provides the full data matrix.

Moreover, the presence of outliers in the considered sample has been detected with the interquartile range method (IRQ), which consists in verifying one of the following inequalities:

$$g_h(a_i) < Q1 - 1.5(Q3 - Q1) \quad \text{or} \quad g_h(a_i) > Q3 + 1.5(Q3 - Q1), \tag{8}$$

where $g_h(a_i)$ is the evaluation of bank a_i on criterion g_h, while $Q1$ and $Q3$ represent, respectively, the first and the third quartiles for the distribution on criterion g_h.

Finally, once data were trimmed to the maximum and minimum values that are not outliers, they have been normalized on a [0, 1] scale through the equations shown below, according to whether a criterion is positively or negatively associated with the bank performance.

$$\bar{g_h}(a_i) = \frac{g_h(a_i) - \min_h}{\max_h - \min_h} \quad \text{or} \quad \bar{g_h}(a_i) = \frac{\max_h - g_h(a_i)}{\max_h - \min_h}, \tag{9}$$

where $\bar{g_h}(a_i)$ is the normalized value, and $\min_h$ and $\max_h$ are, respectively, the minimum and maximum values of criterion g_h.

Table A.2 in the Appendix presents the trimmed data[2], while Table 8 provides the main summary statistics of the European banks on the evaluation criteria per year.

4.2. Comprehensive evaluation results

The application of the proposed methodology leads to a set of σ-μ PKFs presented in Fig. 2 for the latest year, i.e. 2021, and in the supplementary material for all years, (2017–2021). The 2021 results comprise 11 frontiers for all the considered years, except for 2019 (10 frontiers).

The local (δ^*_{ik}) and the normalized global efficiency scores ($\bar{s}_i$), have been computed for each bank of the sample by year, and the comprehensive results are provided in the supplementary material.

In most of the years, the Swedish banks are ranked in the top three positions, followed by Dutch and some Spanish banks (Banco Santander, Bankinter, CaixaBank). Intermediate positions are filled by most of Italian (Unicredit, Intesa Sanpaolo, Mediobanca, Banco BPM) and Irish banks. The lowest rank positions are assigned to German (mainly in the most recent years), French, Greek, and some Italian banks (BPER Banca, Banco BPM, Banca Monte dei Paschi di Siena) in the previous years (2017–2019).

It is worth noting that the global efficiency scores ($\bar{s}_i$) and consequently σ-μ rank positions vary over time. To capture this variation

[1] https://www.policyuncertainty.com

[2] The data trimmed to the maximum and minimum values are denoted, respectively, with asterisk and bold.

Table 7
Description of the selected CAMELS, ESG, and EPU criteria and their preference direction.

Variable Acronyms	Variable name	Description	Unit of measure	Source	Preference direction
C	Capital adequacy	Total capital adequacy ratio	%	BankFocus	Max
A	Asset quality	Loan loss reserves/Gross customer loan & advances	%	BankFocus	Min
M	Management soundness	Cost to income (Efficiency) ratio	%	BankFocus	Min
E	Earnings	ROE : Net income/Total equity	%	BankFocus	Max
L	Liquidity	Basel Liquidity Coverage Ratio (LCR)	%	BankFocus	Max
S	Sensitivity to market risk	Trading revenues/ Operating revenues	%	BankFocus	Min
E_score	Environmental pillar score	Aggregate score combining: - Emission score - Innovation score - Resource use score	[0–100]	Refinitiv ESG scores	Max
S_score	Social pillar score	Aggregate score of combining: - Community score - Human rights score - Product responsibility score - Workforce score	[0–100]	Refinitiv ESG scores	Max
G_score	Governance pillar score	Aggregate score combining: - Corporate social responsibility strategy score - Management score - Shareholders score	[0–100]	Refinitiv ESG scores	Max
EPU_index	Economic Policy Uncertainty index	Monthly index based on the relative frequency of each country's leading newspaper articles, including terms pertaining to the economy (E), policy (P), and uncertainty (U).	yearly median of monthly frequency	www.policyuncertainty.com	Min

Table 8
Summary statistics of CAMELS, ESG and EPU_index criteria in the sample period (2017–2021).

Year	Statistics	C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
2017	Mean	17.84	3.62	64.00	6.15	145.38	3.06	71.50	74.46	62.68	124.20
	St. Dev.	3.25	2.62	13.29	5.05	44.32	4.47	22.28	12.83	21.18	63.83
	Min	13.30	0.18	39.59	−5.58	70.32	−6.60	23	46	19	70.15
	Max	24.30	7.23	93.77	13.39	245	10.87	98	98	93	272.37
2018	Mean	16.76	3.28	64.78	6.87	148.08	0.44	71.07	74.79	65.61	127.88
	St. Dev.	3.32	2.51	13.83	4.39	38.29	3.02	22.96	11.49	20.29	47.50
	Min	12.40	0.17	39.61	−0.89	70.32	−6.60	23	52	18	62.68
	Max	24.30	7.23	93.61	15.55	245	10.62	97	97	95	235.21
2019	Mean	17.44	2.80	66.40	5.58	151.76	1.80	74.86	76.50	68.21	139.72
	St. Dev.	2.95	2.21	14.71	4.82	32.90	5.52	21.45	10.42	17.68	53.14
	Min	13.73	0.20	44.34	−5.58	97	−6.60	23	56	32	86.93
	Max	24.30	7.23	99.14	12.96	245	10.87	96	96	95	257.09
2020	Mean	18.34	2.75	65.97	1.51	174.88	1.84	76.75	77.00	73.79	181.54
	St. Dev.	2.94	1.93	15.23	4.82	34.61	5.41	21.16	10.09	15.08	68.71
	Min	13.50	0.15	42.07	−5.58	123.68	−6.60	23	54	26	73.12
	Max	24.30	7.23	99.14	9.16	245	10.87	97	96	93	283.95
2021	Mean	18.50	2.32	65.94	8.06	181.83	3.68	79.21	79.32	74.50	151.58
	St. Dev.	2.79	1.39	16.22	5.09	35.86	4.68	20.79	9.48	17.46	72.77
	Min	13.80	0.13	39.59	−5.58	129	−4.50	23	52	20	64.45
	Max	23.30	5.66	99.14	17.60	245	10.87	100	96	94	283.95

over the years, the heatmap representation of $\overline{s}_i$ and the Kendall's τ coefficient are employed and their corresponding results are presented, respectively, in Fig. 3 and Table 9. More specifically, the heatmap in Fig. 3 indicates, with different colors, the $\overline{s}_i$ values by bank-year. Dark blue squares denote the highest $\overline{s}_i$ values and thus the highest (better) rank positions, while the lighter squares indicate the lower $\overline{s}_i$ values and therefore the worst rank positions. Variations of rank positions over the years are clearly observed for Italian banks (Intesa Sanpaolo, Banco BPM, BPER Banca) and some Spanish (Banco Santander), Greek (National Bank of Greece), German (Commerzbank), and French banks (BNP Paribas). Instead, Swedish, Irish, Dutch, and some Spanish (Bankinter, Banco Bilbao Vizcaya Argentaria) and French banks (Groupe Crédit Agricole, Société Générale) maintain almost identical σ-μ rank positions over the years (or alternatively show no significant variations of σ-μ rank positions over the years).

The Kendall's τ coefficient in Table 9, represents the degree of concordance between columns of ranked data. From this table, a very strong correlation emerges between the rank data of 2017 and 2018 (0.7619), as well as between the results of 2017 and 2019 (0.7460). Instead, the weaker correlations between the most recent rank data of 2021, and those of the previous years, i.e. 2018 (0.4815) and

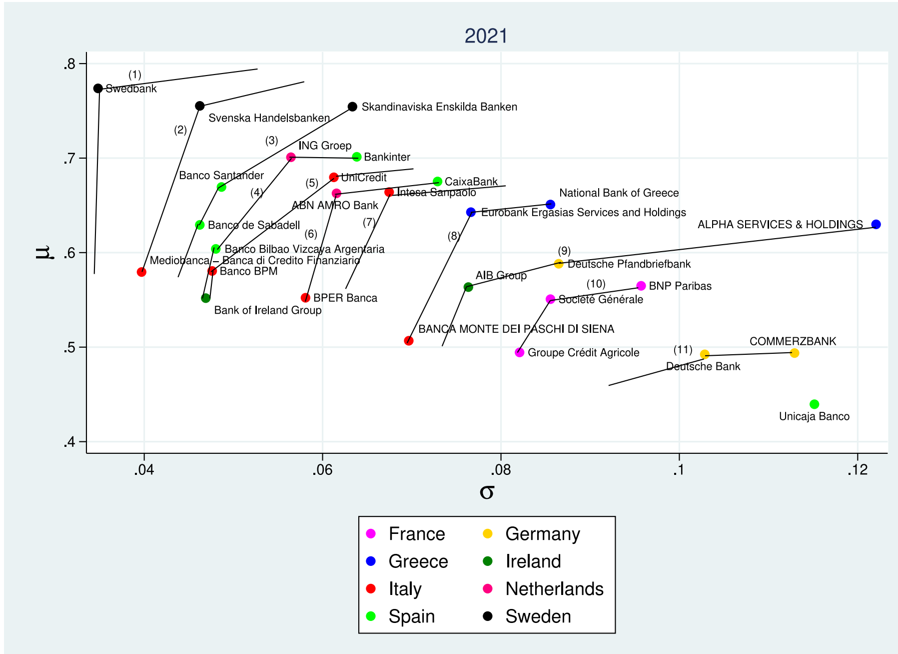

Fig. 2. Family of σ-μ PKFs by country with respect to the whole set of criteria in 2021.

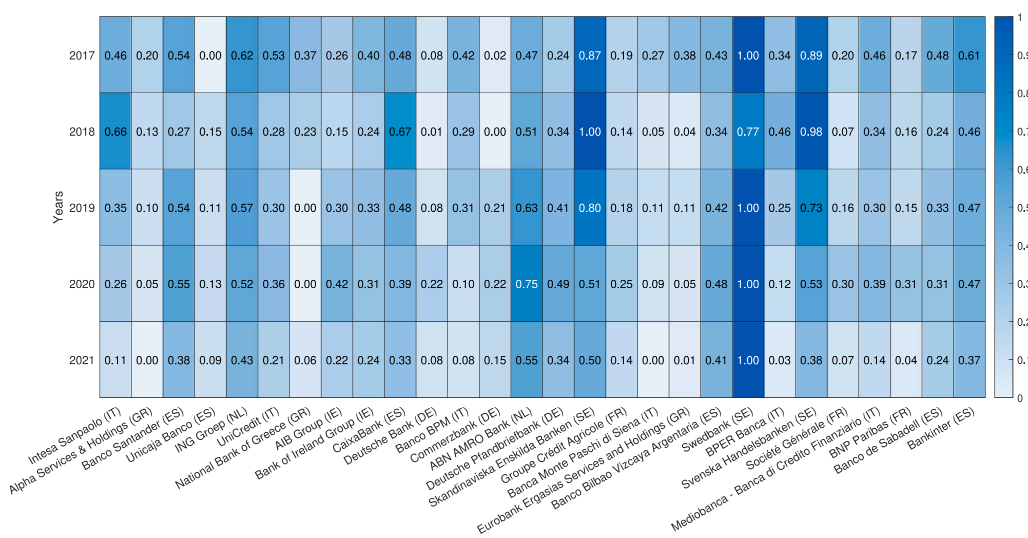

Fig. 3. Heatmap of the global efficiency scores ($\bar{s}_i$).

Table 9
Kendall's τ coefficient of the rankings obtained with the global efficiency scores ($\bar{s}_i$).

	2020	2019	2018	2017
2021	0.6190	0.6455	0.4815	0.4974
2020		0.6667	0.5026	0.5185
2019			0.6878	0.7460
2018				0.7619

2017 (0.4974), suggest significant changes in CAMELS, ESG or EPU evaluation criteria over time.

4.3. Disaggregated analysis

In this section, the performance of the banks is evaluated by considering, separately, the financial and non-financial characteristics of the banks, i.e., CAMELS and ESG criteria, respectively. The aim is to make a comparison in terms of σ-μ PKFs with respect to financial and non-financial criteria (i.e., CAMELS and ESG) and highlight possible discrepancies between these evaluations.

In particular, Figs. 4 and 5 depict the families of σ-μ frontiers in 2021 with regard to CAMELS and ESG criteria, respectively, while the set of σ-μ PKFs for all the considered years is provided in the supplementary material. Concerning the CAMELS criteria, 11 frontiers have been identified for each year, except for 2020 and 2017 (9 frontiers). Similarly, 11 frontiers have also been detected for the ESG criteria in most years, except for 2021 (13 frontiers) and 2017 (9 frontiers).

Some interesting results emerge after observing the apparent differences between the top-ranked banks in the two evaluation dimensions (i.e., CAMELS and ESG). For instance, in most of the years, Swedish banks are ranked in the top six positions with respect to their financial soundness (i.e., CAMELS criteria), along with some German (Deutsche Pfandbriefbank), Italian (Mediobanca), Dutch (ABN AMRO Bank), and Spanish banks (Bankinter). However, these banks have

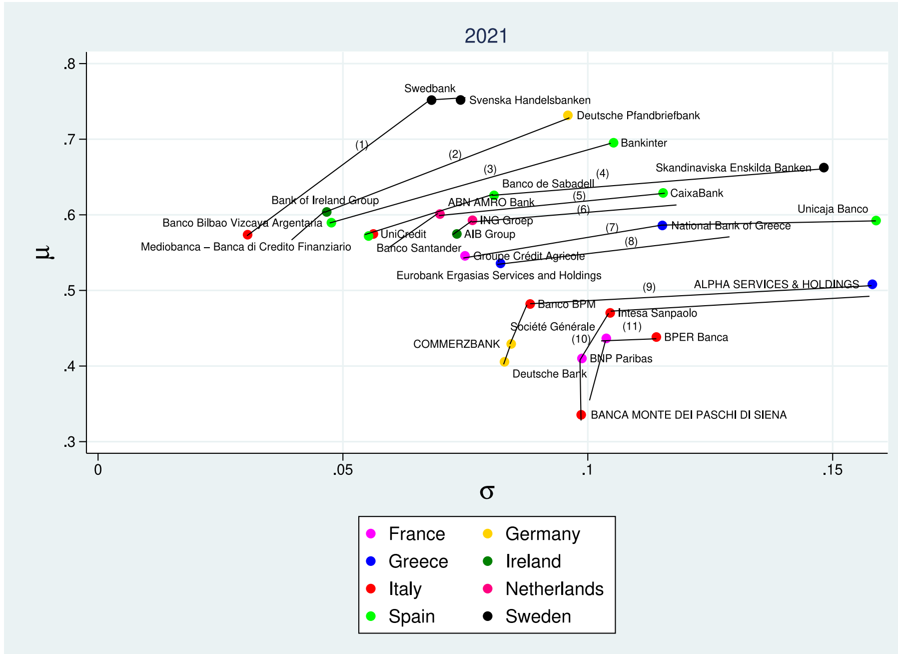

Fig. 4. A set of σ-μ PKFs by country with respect to CAMELS criteria in 2021.

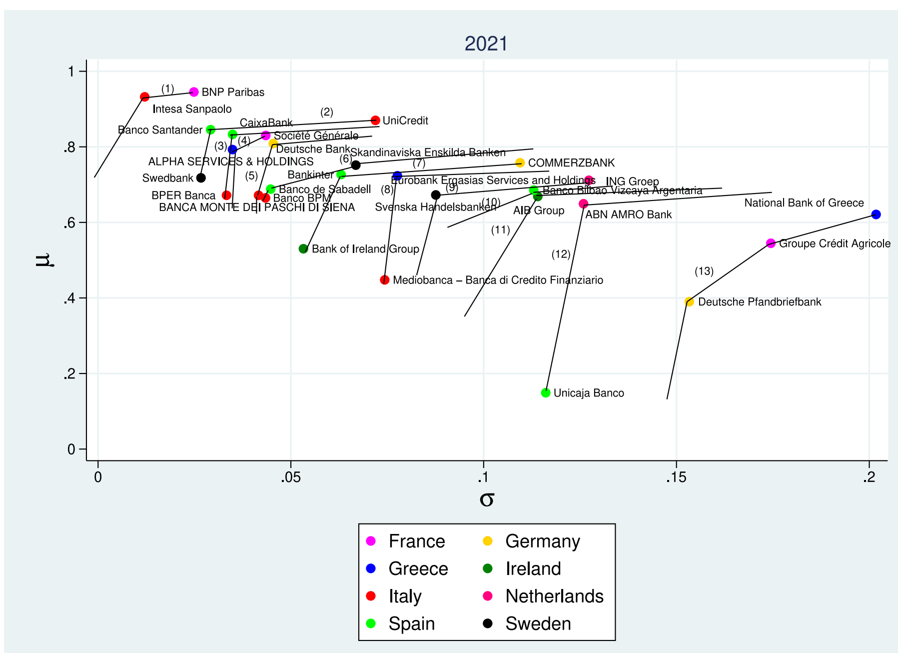

Fig. 5. A set of σ-μ PKFs by country with respect to ESG criteria in 2021.

below average rank positions with respect to ESG criteria. Indeed, Deutsche Pfandbriefbank and Mediobanca, are the most striking examples of inconsistent results between CAMELS (first eight positions) and ESG criteria (last four positions) in most of the years.

Conversely, most of the French banks (BNP Paribas, Société Générale), along with an Italian (Unicredit) Spanish (Banco Santander) and German bank (Deutsche Bank), reach the highest rank positions with respect to ESG criteria, despite their low financial performance in terms of the CAMELS criteria.

Such discrepancies in the performance of the banks could be justified by the implementation of specific green domestic policies in certain countries and the financing of low-carbon transition policies by each bank. A research carried out by ShareAction, a British association dedicated to promoting greener and more sustainable finance over the years, drew up a ranking of the 20 largest European banks based on their implementation of policies for climate-related risks and low-carbon transition. Their results corroborate our findings, indicating that French banks are the best in terms of green practices, due to their high investments in non-polluting companies [59]. For instance, BNP Paribas, acknowledged as the leader in the field, has decided to stop financing gas and oil from shale and bituminous sands, pipelines, drilling in the Arctic, and the tobacco industry. Meanwhile, Société Générale, defined as challenger, has committed to delivering 120 billion euros in financing for the energy transition between 2019 and

Table 10
Kendall's τ coefficient of the rankings on CAMELS and ESG criteria obtained with the global efficiency scores ($\bar{s}_i$).

		ESG				
		2021	2020	2019	2018	2017
CAMELS	2021	−0.1852	−0.2275	−0.1323	−0.1376	−0.0423
	2020		−0.1323	−0.1534	−0.2434	−0.1587
	2019			−0.1529	−0.1376	−0.0847
	2018				−0.0899	−0.0053
	2017					−0.0847

2023. ShareAction attributes part of the merit to the French law on the Energy Transition for Green Growth, specifically, which mandates that anyone handling investments must submit a formal report detailing their environmental impact.

Generally, apart from the French banks, few European credit institutions have implemented action plans to address climate risks and guide investors toward green finance. For instance, according to the ShareAction report, the Spanish Banco Santander and the Dutch ABN AMRO, have not yet defined a proper oil and natural gas policy, even though they plan to do it soon. Nonetheless, some progress is being made in managing climate-related risks and opportunities.

To strengthen the aforementioned discussion, Table 10 computes the Kendall's τ coefficient between the financial and non-financial ranking (i.e., CAMELS versus ESG) of the banks over the years. It is worth noticing that the correlation between the most recent year (2021) and the older year (2017) is quite small in absolute terms because ESG policies were more overlooked by banks than in recent times. In addition, the negative correlation between CAMELS and ESG criteria corroborates the discrepancies found between the performance of banks on financial and ESG factors, where the highest-ranked positions with respect to CAMELS criteria correspond to the worst ranking evaluations on ESG criteria, and vice versa. Therefore, the results indicate a trade-off between the financial soundness of the banks and their non-financial performance on ESG factors, at least in the short term [60,61].

4.4. Comparison to the benefit-of-the doubt efficiency-based model

In order to compare the results of the proposed approach to another similar methodology, this sub-section presents the results of the benefit-of-the-doubt (BoD) technique [62]. Similarly to the σ-μ analysis, the BoD also relies on ideas from efficiency measurement and it has been extensively used for performance evaluation and the development of CIs in various fields (see, Walheer [63] and the references therein).

As it was noted in the literature review section, one of the principal issues in the building of CIs is related to the procedure for weighting the criteria. Indeed, the choice of the weights is one of the most crucial aspects as it entails subjectivity, thus raising concerns especially when several different stakeholders are involved in the evaluation process. Under the BoD framework, the determination of weights is data-driven giving more importance to criteria that describe the strengths of an alternative over its peers.

More specifically, in the context of the BoD methodology, the performance score $CI(a_i, \mathbf{w})$ of an alternative a_i is obtained through a weighted average model, in which the weight vector $\mathbf{w}$ is specified through the solution of the following LP problem:

$$\begin{aligned} \max_{\mathbf{w}} \quad & CI(a_i, \mathbf{w}) = \sum_{h=1}^{n} w_h \tilde{g_h}(a_i) \\ \text{s.t.} \quad & \sum_{h=1}^{n} w_h \tilde{g_h}(a_j) \leq 1, \qquad \forall j = 1, 2, \ldots, m \\ & LB \leq \frac{w_h \tilde{g_h}(a_i)}{\sum_{h=1}^{n} w_h \tilde{g_h}(a_i)} \leq UB, \quad \forall h = 1, 2, \ldots, n \end{aligned} \tag{10}$$

The first constraint defines the scale of the CI, such that the best performing alternative will have a CI equal to 1. The second set of inequalities introduce bounds for the weights of the criteria. In the present analysis we employ "pie-share" constraints that ensure that none of the indicators will be excluded during the evaluation process [62]. LB and UB represent, respectively, a lower and upper bound on the weights (in relative form). For instance, a LB= 0.1 means that no criteria will have a weight lower than 10% of the total weight.

The BoD model has been applied to the sample of banks across all the considered years, covering the entire set of criteria for comparison purposes. The bounds on the (relative) weights of the criteria are set such that LB=0.025 and UB=1, implying that the contribution of a criterion to the performance score of each alternative will be equal to 2.5%. This specification was made to enhance the similarity of the BoD model with our SMAA implementation.

The results, including the performance scores and the rankings from the BoD and the σ-μ models, are presented in Table 11, with a specific focus on the most recent year (i.e., 2021). It is evident that the results of the BoD model exhibit lower discriminating power, with eight banks placed at the top position and seven others having scores higher than 0.9. In contrast, the evaluations derived through the proposed implementation of the σ-μ efficiency analysis, discriminate the banks better, as the scores span the whole range from 0 to 1.

Moreover, an examination of the rankings derived by the scores of the two models highlights some notable disparities, and highlights the refinements obtained with the σ-μ analysis as opposed to another widely used approach. More specifically, the largest ranking difference, in absolute terms, is observed for National Bank of Greece (a 16-point difference between the two rankings), followed by Eurobank Ergasias Services and Holdings (15-point difference), Groupe Crédit Agricole (13-point difference), AIB Group (13-point difference), Unicaja Banco (10-point difference), and BNP Paribas (10-point difference). All these banks are ranked in low positions according to the results of the σ-μ model, whereas BoD assigns them in much better positions. In contrast, banks ranked as the best (Swedbank, Svenska Handelsbanken, Skandinaviska Enskilda Banken, ING Groep) or intermediate (Intesa Sanpaolo, Bank of Ireland Group, Banca Monte Dei Paschi di Siena, Deutsche Pfandbriefbank), have similar rankings according to the two models. Detailed comparisons spanning all years of the analysis (2017–2021) are provided in the supplementary material.

5. Conclusions

The aim of this study was to extend the σ-μ efficiency analysis methodology proposed by Greco et al. [4], through the incorporation of the Pareto dominance relation among alternatives. To address this issue, we introduced an iterative algorithmic approach to compute the global efficiency scores of alternatives and to identify a set of σ-μ PKFs, leading to evaluation results that are consistent with the dominance relation.

The proposed approach was employed to evaluate the performance of European banks that participated in the 2023 EU-wide stress test, along financial and ESG criteria. The results showed that Swedish banks are ranked in the top three positions in most years, followed by Dutch and some Spanish banks. Meanwhile, most Italian and Irish banks presented intermediate positions, with the worst being held by German, French, and Greek banks. Top-ranked banks have performed consistently well over the years, whereas banks with intermediate or worse rank positions showed some variability over time. Moreover, the comparison between the evaluation obtained based solely on financial criteria on the one hand, and ESG criteria on the other, showed notable differences. These variations may be attributed to the implementation of specific green domestic policies in certain countries and the policies followed by each bank with respect to ESG factors.

The results of the model are valuable for the banking sector and various regulatory bodies, including the EBA, the Basel Committee on Banking Supervision (BCBS), and the Financial Stability Board (FSB),

Table 11
Comparison of the scores and rankings between the BoD model and σ-μ efficiency analysis for the sample of EU banks in 2021.

Banks	BoD score	σ-μ score	BoD rank	σ-μ rank
Intesa Sanpaolo	0.786	0.463	13	11
Alpha Services & Holdings	0.735	0.200	14	22
Banco Santander	0.987	0.542	2	6
Unicaja Banco	0.232	0.000	18	28
ING Groep	0.962	0.623	4	4
UniCredit	1.000	0.533	1	7
National Bank of Greece	1.000	0.367	1	17
AIB Group	0.909	0.257	7	20
Bank of Ireland Group	0.827	0.398	12	15
CaixaBank	1.000	0.483	1	8
Deutsche Bank	0.155	0.076	20	26
Banco BPM	0.908	0.420	8	14
Commerzbank	0.152	0.020	21	27
ABN AMRO Bank	0.961	0.474	5	10
Deutsche Pfandbriefbank	0.157	0.242	19	21
Skandinaviska Enskilda Banken	0.950	0.874	6	3
Groupe Crédit Agricole	0.830	0.195	11	24
Banca Monte Dei Paschi Di Siena	0.671	0.269	17	19
Eurobank Ergasias Services and Holdings	1.000	0.377	1	16
Banco Bilbao Vizcaya Argentaria	0.980	0.433	3	13
Swedbank	1.000	1.000	1	1
BPER Banca	0.833	0.340	10	18
Svenska Handelsbanken	1.000	0.892	1	2
Société Générale	0.704	0.197	16	23
Mediobanca	0.852	0.457	9	12
BNP Paribas	0.729	0.168	15	25
Banco de Sabadell	1.000	0.478	1	9
Bankinter	1.000	0.609	1	5

entailing diverse practical implications. For example, they can support the identification of operational inefficiencies and areas where banks can improve their performance, enhance their operations, and encourage improvements and innovation within the banking sector. The CAMELS framework, concentrating on aspects like capital adequacy, asset quality, and sensitivity to market risk, offers insights into a bank's risk management practices, enabling regulatory bodies to ensure effective risk management for overall financial stability. Moreover, the inclusion of ESG criteria and the EPU index in the evaluation process encourages sustainable practices and provides insights into how banks are adapting to external economic conditions, leading to policy adjustments for stability. These transparent and more rigorous assessments enhance market confidence and investor trust, attract investments, and foster overall market stability. Regulatory bodies can use these findings to formulate and adjust policies based on efficiency analysis results. For instance, if certain banks consistently underperform in key areas, regulators may implement targeted policies to address specific challenges and improve overall industry performance. Conversely, if certain banks perform well, regulatory bodies can encourage competition by fostering an environment where banks strive to improve efficiency, ultimately leading to better services and products for consumers. Hence, regular efficiency analysis, as proposed in this work, could serve as an early warning system for potential issues within the banking sector, allowing regulatory bodies to intervene promptly if inefficiencies or risks are detected, preventing larger systemic problems. It is worth noting that the incorporation of the Pareto dominance relation into σ-μ efficiency analysis introduces both strengths and weaknesses. On the positive side, this integration enhances the decision-making process by excelling in identifying non-dominated solutions. It provides a clear set of superior banks across various dimensions and contributes to maintaining diversity among efficient banks, facilitating the identification of the optimum.

However, certain weaknesses exist, notably the sensitivity to outliers, which can significantly skew the outcomes both before and after calculating σ and μ. To address this sensitivity, appropriate data preprocessing procedures were used in the present study, before computing σ and μ, as detailed in Section 4.1. Additionally, potential subjectivity bias in composite indicator construction, especially in weighting or aggregation methods, was mitigated through a sampling procedure for criteria weights involving 10,000 iterations. This approach considers a broad range of decision-making preferences and helps reduce bias towards dimensions most affected by extreme alternatives.

To validate our results, we compared the σ-μ efficiency analysis with the BoD model, a widely used data-driven weighting technique. Unlike the BoD model, which tends to generate a flattened final ranking with several ties at the top position, our model excels in establishing a precise one-to-one correspondence between positions and alternatives, yielding more refined results.

Finally, we envisage some potential directions for future research, regarding the area of decision-making in the domain of banking and finance, as well as methodological issues for the construction of CIs. Regarding financial decision-making and the role of non-financial factors such as ESG, it is worth noting that recent studies (Billio et al. [64], Del Vitto et al. [65]), have noted the lack of common metrics among rating agencies regarding the definition of ESG factors (characteristics, attributes, and standards). This heterogeneity severely questions the reliability of existing ESG ratings as measures of corporate sustainable performance. Therefore it is important to further investigate the results obtained in this paper with data on the ESG scores provided by other rating agencies.

In addition, to address the growing interest in the recent literature concerning systemically important banks, a further future direction could be to investigate the role of ESG factors as additional criteria in predicting bank financial distress [66] using a combination of MCDA methods, such as the hybrid classification methodology proposed by Angilella and Pappalardo [67].

Regarding the construction of CIs, one unresolved issue pertains to the selection of an appropriate MCDA method to aggregate the underlying criteria. In the literature different aggregation methods can be found, beyond the simple additive weighting model. Among others, one can mention non-linear value function models and relational approaches (e.g., outranking models, the TOPSIS method, etc.). Therefore, a comparative analysis of different MCDA aggregation methods could provide insights into their strengths and weaknesses for specific areas where the construction of CIs is important.

Table A.1
Performance matrix.

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
			2017	17.90	7.23	57.39	12.99	82.90	0.58	89	90	46	70.15
			2018	16.50	5.69	61.62	7.49	108.30	−3.05	88	94	48	106.21
a_1	Intesa Sanpaolo	Italy	2019	17.00	4.25	60.19	7.35	133.70	−7.52	89	93	42	113.63
			2020	19.20	3.31	66.06	4.95	159.10	−7.88	97	93	82	164.69
			2021	18.90	2.83	64.94	6.31	184.50	15.91	97	94	92	103.85
			2017	18.40	23.48	50.01	0.22	10.80	3.28	59	64	78	91.76
			2018	14.00	23.06	44.68	0.65	57.20	0.32	83	74	68	94.03
a_2	Alpha Services & Holdings	Greece	2019	14.90	19.20	49.41	1.24	103.60	0.80	85	74	71	86.93
			2020	16.00	19.70	43.87	1.24	150.00	−0.83	79	72	64	73.12
			2021	13.80	5.66	−341.12	−47.80	196.40	−4.50	81	88	84	64.45
			2017	14.48	2.77	56.47	7.68	142.00	2.50	87	92	88	105.76
			2018	14.77	2.58	54.34	8.68	158.00	3.22	84	90	95	111.32
a_3	Banco Santander	Spain	2019	15.02	2.32	56.63	7.33	147.00	2.59	85	90	93	131.52
			2020	16.16	2.54	55.33	−8.44	168.00	6.98	92	90	88	180.49
			2021	16.41	2.34	55.09	9.95	163.00	2.26	91	87	88	147.92
			2017	13.30	4.31	83.15	3.55	701.00	0.17	0	66	44	105.76
			2018	13.70	3.65	85.87	3.89	468.00	0.05	5	69	38	111.32
a_4	Unicaja Banco	Spain	2019	15.40	2.63	99.14	4.34	319.00	0.07	13	72	32	131.52
			2020	16.60	2.86	70.48	1.94	310.00	−0.02	14	70	26	180.49
			2021	15.80	2.38	109.40	17.60	307.00	0.91	15	68	20	147.92
			2017	19.14	0.78	55.86	9.76	114.00	3.89	88	76	64	78.56
			2018	18.44	0.75	58.84	9.30	123.00	4.17	87	75	61	62.67
a_5	ING Groep	Netherlands	2019	19.09	0.74	57.06	8.93	127.00	6.12	86	71	74	87.35
			2020	20.09	0.95	60.88	4.60	137.00	5.36	84	70	81	131.92
			2021	21.01	0.83	60.89	8.97	139.00	4.68	86	68	90	91.27

(*continued on next page*)

Moreover, it could be interesting to investigate the extension of the μ-σ analysis framework to consider additional parameters of the distributions of the alternatives' evaluations, such as skewness and kurtosis. The consideration of advanced simulation procedures, such as Markov Chain Monte Carlo approaches [68], could also enhance the efficiency of the SMAA process. Such approaches could be employed to extend the simulation process in order to cover different decision modeling parameters for the evaluation of the alternatives, instead of focusing solely on scenarios about the weights of the criteria. Scenarios for the decision environment (e.g., economic conditions) would also be analyzed to consider the uncertainty not only with respect to the specification of the evaluation model, but also with respect to the external conditions that affect the decision outcomes. Extending further this line of research, it would be interesting to examine the construction of CIs in a dynamic framework to benchmark and monitor the performance of the alternatives over different time periods. This, for instance, could be particularly useful in the development of an augmented credit rating, incorporating both financial and ESG aspects, for entities like Small and Medium Enterprises (SMEs). Indeed, these latter, despite playing a pivotal role in driving growth and fostering innovation in the modern economy, suffer from the absence of suitable credit rating models for assessing accurately their creditworthiness. This challenge arises due to their limited historical financial data on which the traditional credit rating systems' are mainly based.

Finally, the analysis of the coherence and validity of CIs should be further explored through the development of formal procedures, measures, and further empirical analysis [69].

CRediT authorship contribution statement

Silvia Angilella: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Michalis Doumpos:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Maria Rosaria Pappalardo:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Constantin Zopounidis:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: none

Data availability

Data will be made available on request.

Acknowledgments

This study (for the first and third authors) was funded by the European Union - NextGenerationEU, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP E63C22002120006). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

Appendix A

See the Excel file Supplementarymaterial for:

- the table summarizing the models discussed in the literature review section and evaluating the performance and efficiency of banks;
- results on $\sigma-\mu$ PKFs with respect to all criteria; CAMELS criteria; and ESG criteria for all considered years (i.e.2017–2021);
- comparison of results, in terms of scores and rankings, between the BoD model and the $\sigma-\mu$ efficiency analysis for all considered years.

Table A.1 (*continued*).

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_6	UniCredit	Italy	2017	18.10	6.70	70.15	9.61	185.29	2.92	83	80	73	70.15
			2018	15.80	5.44	65.31	7.57	164.14	1.22	83	89	86	106.21
			2019	17.69	4.30	71.24	5.65	143.00	9.97	84	89	85	113.63
			2020	20.02	3.62	76.29	−4.63	171.00	2.51	94	86	86	164.69
			2021	19.68	3.00	72.18	2.53	182.00	−2.48	100	83	87	103.85
a_7	National Bank of Greece	Greece	2017	17.00	21.25	59.89	−5.58	189.30	−9.95	76	78	87	91.76
			2018	12.86	23.90	74.63	−0.89	198.30	−2.35	10	63	87	94.03
			2019	13.73	16.47	64.72	−4.49	207.30	4.85	26	88	95	86.93
			2020	13.76	9.15	42.07	0.69	232.20	0.58	31	87	93	73.12
			2021	15.54	5.16	56.43	15.06	242.00	4.71	36	86	91	64.45
a_8	AIB Group	Ireland	2017	19.00	5.28	60.88	8.18	129.00	2.87	48	66	39	191.86
			2018	19.10	3.24	63.39	7.88	128.00	−0.14	63	70	72	155.27
			2019	20.50	1.99	82.18	2.56	157.00	−3.96	61	65	62	158.21
			2020	20.80	4.22	78.34	−5.52	193.00	−2.06	59	74	74	244.97
			2021	21.90	3.23	84.32	4.72	203.00	0.13	57	83	86	233.39
a_9	Bank of Ireland Group	Ireland	2017	17.90	3.01	72.10	7.16	136.00	5.44	46	70	37	191.86
			2018	17.20	2.18	74.02	6.72	136.00	1.94	45	68	53	155.27
			2019	17.40	1.62	71.47	4.29	138.00	4.09	49	73	60	158.21
			2020	18.00	2.84	86.19	−7.35	152.83	0.96	50	70	48	244.97
			2021	21.40	2.50	66.40	9.30	181.00	3.66	56	76	65	233.39
a_{10}	CaixaBank	Spain	2017	15.70	3.05	72.16	6.92	202.00	0.45	67	93	73	105.76
			2018	15.30	2.55	67.61	8.60	200.00	0.90	66	92	55	111.32
			2019	15.70	2.07	75.32	6.79	179.00	2.02	86	89	92	131.52
			2020	18.10	2.31	63.28	5.46	276.00	1.38	83	88	90	180.49
			2021	17.90	2.34	87.63	14.76	336.00	1.32	85	89	88	147.92
a_{11}	Deutsche Bank	Germany	2017	18.40	0.96	93.77	−1.00	140.00	10.51	98	85	69	167.46
			2018	17.50	1.03	93.61	0.47	140.40	0.20	97	84	71	176.79
			2019	17.40	0.93	108.52	−8.20	141.20	0.83	96	87	80	202.74
			2020	17.30	1.12	88.88	0.98	144.80	9.21	96	85	80	308.51
			2021	17.70	1.02	85.07	3.66	133.10	7.29	97	86	75	283.95
a_{12}	Banco BPM	Italy	2017	15.20	9.81	73.32	5.23	125.61	1.17	65	67	40	70.15
			2018	12.40	8.46	77.51	−0.67	154.13	−0.60	69	64	62	106.21
			2019	15.50	4.37	72.22	6.57	165.03	−6.21	71	75	60	113.63
			2020	17.70	3.75	81.52	0.14	191.00	−0.64	76	79	74	164.69
			2021	18.40	2.89	71.98	4.34	209.00	8.26	84	76	69	103.85
a_{13}	Commerzbank	Germany	2017	17.50	1.37	90.64	0.74	138.60	−0.94	83	74	69	167.46
			2018	15.90	0.91	82.22	3.28	135.66	2.02	79	72	63	176.79
			2019	16.40	0.85	80.74	2.25	132.72	2.94	88	70	65	202.74
			2020	17.70	1.21	92.18	−10.01	135.70	0.27	91	71	79	308.51
			2021	18.40	1.09	93.14	1.19	145.10	3.35	93	72	88	283.95
a_{14}	ABN AMRO Bank	Netherlands	2017	26.60	0.89	60.03	13.08	104.00	3.11	80	74	65	78.56
			2018	27.30	0.83	59.40	11.00	119.00	1.92	88	65	72	62.67
			2019	25.90	0.90	62.12	9.53	134.00	0.99	89	66	53	87.35
			2020	23.90	1.36	67.43	−0.21	149.00	2.03	90	65	75	131.92
			2021	22.40	0.93	77.11	5.61	168.00	1.88	93	65	73	91.27
a_{15}	Deutsche Pfandbriefbank	Germany	2017	22.20	0.18	57.81	6.37	187.00	−1.01	41	48	81	167.46
			2018	24.90	0.26	51.89	5.50	212.00	−0.42	42	52	71	176.79
			2019	20.40	0.30	48.04	5.53	182.00	−3.92	43	56	61	202.74
			2020	21.40	0.60	47.34	3.67	279.00	−1.33	46	54	68	308.51
			2021	22.40	0.86	45.35	6.66	227.00	4.06	49	52	75	283.95
a_{16}	Skandinaviska Enskilda Banken	Sweden	2017	24.20	0.31	48.06	10.29	145.00	13.67	84	84	55	95.46
			2018	22.20	0.34	47.99	15.55	147.00	10.62	93	80	44	107.19
			2019	23.30	0.42	45.88	12.96	218.00	13.42	92	79	58	103.33
			2020	25.10	0.57	47.83	9.16	163.00	10.68	90	76	91	111.79
			2021	23.10	0.46	42.26	13.16	145.00	9.51	92	77	80	101.93
a_{17}	Groupe Crédit Agricole	France	2017	17.40	2.72	65.20	6.52	137.00	6.88	94	74	50	272.37
			2018	17.20	2.52	63.79	7.67	133.00	−0.23	93	71	42	235.20
			2019	16.90	2.28	63.28	7.70	131.60	19.20	95	76	56	257.08
			2020	18.97	2.31	62.88	4.41	157.30	13.79	95	73	54	313
			2021	17.46	1.90	61.00	8.90	183.00	6.46	96	74	32	244.37
a_{18}	Banca Monte Dei Paschi di Siena	Italy	2017	14.97	26.36	83.85	−33.57	199.50	3.10	79	70	19	70.15
			2018	12.80	9.71	89.97	3.10	190.20	−2.19	57	71	18	106.21
			2019	14.70	6.43	86.15	−12.48	152.40	−17.58	57	79	41	113.63
			2020	13.50	2.19	106.54	−29.22	196.70	−14.15	76	78	71	164.69
			2021	14.56	2.84	84.85	5.01	172.70	10.87	76	77	77	103.85

(*continued on next page*)

Table A.1 (*continued*).

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_{19}	Eurobank Ergasias Services and Holdings	Greece	2017	15.00	21.45	48.25	1.61	43.64	7.44	62	72	64	91.76
			2018	14.00	19.54	51.46	1.84	70.32	1.19	62	70	64	94.03
			2019	17.10	15.97	55.71	1.90	97.00	−2.42	63	72	75	86.93
			2020	14.60	8.50	46.78	−23.09	123.68	−17.31	62	84	72	73.12
			2021	15.30	4.58	47.69	5.80	152.24	−1.83	67	89	80	64.45
a_{20}	Banco Bilbao Vizcaya Argentaria	Spain	2017	15.30	3.27	56.27	8.90	128.00	0.03	87	90	93	105.76
			2018	15.50	3.15	54.83	11.59	127.00	3.01	84	86	88	111.32
			2019	15.42	3.14	54.62	7.89	128.00	1.92	83	81	80	131.52
			2020	15.92	3.75	53.44	4.10	149.00	3.55	93	81	73	180.49
			2021	16.98	3.37	55.03	11.52	165.00	0.63	95	84	51	147.92
a_{21}	Swedbank	Sweden	2017	30.70	0.26	39.59	13.39	173.00	2.47	83	75	89	95.46
			2018	21.50	0.38	39.61	14.23	144.00	1.50	82	70	91	107.19
			2019	21.80	0.41	44.34	12.71	182.00	2.76	83	69	86	103.33
			2020	21.00	0.49	54.47	7.90	174.00	1.97	81	66	72	111.79
			2021	22.40	0.30	45.40	11.91	163.00	0.83	86	81	72	101.93
a_{22}	BPER Banca	Italy	2017	16.47	11.49	69.24	3.09	113.70	1.90	27	63	28	70.15
			2018	14.93	7.09	71.82	9.10	154.30	−2.74	31	69	36	106.21
			2019	14.94	5.21	79.33	7.45	158.90	−7.00	81	69	64	113.63
			2020	19.30	3.27	71.53	4.14	200.10	−7.88	78	77	65	164.69
			2021	16.16	2.98	80.14	8.14	215.10	10.25	75	85	66	103.85
a_{23}	Svenska Handelsbanken	Sweden	2017	28.30	0.25	45.56	10.64	133.00	4.28	74	71	34	95.46
			2018	21.00	0.17	46.11	12.20	146.00	−0.10	71	74	38	107.19
			2019	23.20	0.20	48.06	10.59	147.00	−0.13	87	78	43	103.33
			2020	24.30	0.15	52.08	9.09	150.00	0.11	93	76	62	111.79
			2021	23.30	0.13	47.47	10.75	152.00	0.00	94	76	61	101.93
a_{24}	Société Générale	France	2017	17.00	2.87	73.94	5.36	140.00	38.00	89	91	82	272.37
			2018	16.50	2.51	71.14	7.31	129.00	0.82	97	90	83	235.20
			2019	18.90	2.31	71.85	5.75	119.00	63.09	96	88	82	257.08
			2020	18.92	2.45	75.58	0.29	149.00	13.33	94	88	83	313.00
			2021	18.69	2.12	68.18	8.94	129.00	37.13	96	84	84	244.37
a_{25}	Mediobanca	Italy	2017	17.11	3.91	60.41	8.08	245.00	−11.47	23	46	53	70.15
			2018	18.11	3.63	55.12	8.92	186.00	−6.60	36	57	74	106.21
			2019	17.28	3.35	57.25	8.35	143.00	4.34	37	57	59	113.63
			2020	18.67	3.28	56.60	6.17	165.00	−2.28	39	58	57	164.69
			2021	18.78	3.32	55.43	7.29	158.00	0.96	46	71	63	103.85
a_{26}	BNP Paribas	France	2017	14.60	3.54	68.95	7.66	121.00	16.41	96	98	89	272.37
			2018	15.00	3.05	71.93	7.57	132.00	−3.48	95	97	88	235.20
			2019	15.50	2.56	70.27	7.67	125.00	26.33	95	96	90	257.08
			2020	16.31	2.58	68.23	6.32	154.00	9.61	94	96	93	313.00
			2021	16.40	2.38	67.33	8.06	143.00	14.57	95	96	94	244.37
a_{27}	Banco de Sabadell	Spain	2017	16.02	2.56	61.68	6.09	133.00	4.06	86	62	57	105.76
			2018	14.13	2.35	72.38	2.53	151.00	0.59	86	73	75	111.32
			2019	15.00	1.95	69.86	5.50	158.00	−2.07	84	71	58	131.52
			2020	15.91	2.02	73.52	0.02	198.00	7.86	83	75	76	180.49
			2021	17.70	2.06	71.61	3.50	221.41	−3.49	86	77	71	147.92
a_{28}	Bankinter	Spain	2017	14.32	1.62	57.41	11.37	141.20	1.10	85	66	89	105.76
			2018	14.29	1.31	62.62	11.23	144.20	0.62	83	65	94	111.32
			2019	13.94	1.18	63.02	11.01	153.70	1.05	82	69	93	131.52
			2020	15.02	1.43	65.02	5.27	198.10	0.26	80	74	89	180.49
			2021	15.39	1.47	60.75	25.63	228.60	0.83	82	77	84	147.92

Data Source: EBA, BankFocus, ESG Refinitiv, EPU_index.

Table A.2
Performance matrix with the outperformers trimmed to the lower and the upper bound.

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_1	Intesa Sanpaolo	Italy	2017	17.90	7.23	57.39	12.99	82.90	0.58	89	90	46	70.15
			2018	16.50	5.69	61.62	7.49	108.30	−3.05	88	94	48	106.21
			2019	17.00	4.25	60.19	7.35	133.70	**−6.60**	89	93	42	113.63
			2020	19.20	3.31	66.06	4.95	159.10	**−6.60**	97	93	82	164.69
			2021	18.90	2.83	64.94	6.31	184.50	10.87*	97	94	92	103.85

(*continued on next page*)

Table A.2 (*continued*).

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_2	Alpha Services & Holdings	Greece	2017	18.40	7.23*	50.01	0.22	**70.32**	3.28	59	64	78	91.76
			2018	14.00	7.23*	44.68	0.65	**70.32**	0.32	83	74	68	94.03
			2019	14.90	7.23*	49.41	1.24	103.60	0.80	85	74	71	86.93
			2020	16.00	7.23*	43.87	1.24	150.00	−0.83	79	72	64	73.12
			2021	13.80	5.66	**39.59**	**−5.58**	196.40	−4.50	81	88	84	64.45
a_3	Banco Santander	Spain	2017	14.48	2.77	56.47	7.68	142.00	2.50	87	92	88	105.76
			2018	14.77	2.58	54.34	8.68	158.00	3.22	84	90	95	111.32
			2019	15.02	2.32	56.63	7.33	147.00	2.59	85	90	93	131.52
			2020	16.16	2.54	55.33	**−5.58**	168.00	6.98	92	90	88	180.49
			2021	16.41	2.34	55.09	9.95	163.00	2.26	91	87	88	147.92
a_4	Unicaja Banco	Spain	2017	13.30	4.31	83.15	3.55	245*	0.17	**23**	66	44	105.76
			2018	13.70	3.65	85.87	3.89	245*	0.05	**23**	69	38	111.32
			2019	15.40	2.63	99.14	4.34	245*	0.07	**23**	72	32	131.52
			2020	16.60	2.86	70.48	1.94	245*	−0.02	**23**	70	26	180.49
			2021	15.80	2.38	99.14*	17.60	245*	0.91	**23**	68	20	147.92
a_5	ING Groep	Netherlands	2017	19.14	0.78	55.86	9.76	114.00	3.89	88	76	64	78.56
			2018	18.44	0.75	58.84	9.30	123.00	4.17	87	75	61	62.67
			2019	19.09	0.74	57.06	8.93	127.00	6.12	86	71	74	87.35
			2020	20.09	0.95	60.88	4.60	137.00	5.36	84	70	81	131.92
			2021	21.01	0.83	60.89	8.97	139.00	4.68	86	68	90	91.27
a_6	UniCredit	Italy	2017	18.10	6.70	70.15	9.61	185.29	2.92	83	80	73	70.15
			2018	15.80	5.44	65.31	7.57	164.14	1.22	83	89	86	106.21
			2019	17.69	4.30	71.24	5.65	143.00	9.97	84	89	85	113.63
			2020	20.02	3.62	76.29	−4.63	171.00	2.51	94	86	86	164.69
			2021	19.68	3.00	72.18	2.53	182.00	−2.48	100	83	87	103.85
a_7	National Bank of Greece	Greece	2017	17.00	7.23*	59.89	−5.58	189.30	**−6.60**	76	78	87	91.76
			2018	12.86	7.23*	74.63	−0.89	198.30	−2.35	**23**	63	87	94.03
			2019	13.73	7.23*	64.72	−4.49	207.30	4.85	26	88	95	86.93
			2020	13.76	7.23*	42.07	0.69	232.20	0.58	31	87	93	73.12
			2021	15.54	5.16	56.43	15.06	242.00	4.71	36	86	91	64.45
a_8	AIB Group	Ireland	2017	19.00	5.28	60.88	8.18	129.00	2.87	48	66	39	191.86
			2018	19.10	3.24	63.39	7.88	128.00	−0.14	63	70	72	155.27
			2019	20.50	1.99	82.18	2.56	157.00	−3.96	61	65	62	158.21
			2020	20.80	4.22	78.34	−5.52	193.00	−2.06	59	74	74	244.97
			2021	21.90	3.23	84.32	4.72	203.00	0.13	57	83	86	233.39
a_9	Bank of Ireland Group	Ireland	2017	17.90	3.01	72.10	7.16	136.00	5.44	46	70	37	191.86
			2018	17.20	2.18	74.02	6.72	136.00	1.94	45	68	53	155.27
			2019	17.40	1.62	71.47	4.29	138.00	4.09	49	73	60	158.21
			2020	18.00	2.84	86.19	**−5.58**	152.83	0.96	50	70	48	244.97
			2021	21.40	2.50	66.40	9.30	181.00	3.66	56	76	65	233.39
a_{10}	CaixaBank	Spain	2017	15.70	3.05	72.16	6.92	202.00	0.45	67	93	73	105.76
			2018	15.30	2.55	67.61	8.60	200.00	0.90	66	92	55	111.32
			2019	15.70	2.07	75.32	6.79	179.00	2.02	86	89	92	131.52
			2020	18.10	2.31	63.28	5.46	245*	1.38	83	88	90	180.49
			2021	17.90	2.34	87.63	14.76	245*	1.32	85	89	88	147.92
a_{11}	Deutsche Bank	Germany	2017	18.40	0.96	93.77	−1.00	140.00	10.51	98	85	69	167.46
			2018	17.50	1.03	93.61	0.47	140.40	0.20	97	84	71	176.79
			2019	17.40	0.93	99.14*	**−5.58**	141.20	0.83	96	87	80	202.74
			2020	17.30	1.12	88.88	0.98	144.80	9.21	96	85	80	283.95*
			2021	17.70	1.02	85.07	3.66	133.10	7.29	97	86	75	283.95
a_{12}	Banco BPM	Italy	2017	15.20	7.23*	73.32	5.23	125.61	1.17	65	67	40	70.15
			2018	12.40	7.23*	77.51	−0.67	154.13	−0.60	69	64	62	106.21
			2019	15.50	4.37	72.22	6.57	165.03	−6.21	71	75	60	113.63
			2020	17.70	3.75	81.52	0.14	191.00	−0.64	76	79	74	164.69
			2021	18.40	2.89	71.98	4.34	209.00	8.26	84	76	69	103.85
a_{13}	Commerzbank	Germany	2017	17.50	1.37	90.64	0.74	138.60	−0.94	83	74	69	167.46
			2018	15.90	0.91	82.22	3.28	135.66	2.02	79	72	63	176.79
			2019	16.40	0.85	80.74	2.25	132.72	2.94	88	70	65	202.74
			2020	17.70	1.21	92.18	**−5.58**	135.70	0.27	91	71	79	283.95*
			2021	18.40	1.09	93.14	1.19	145.10	3.35	93	72	88	283.95

(*continued on next page*)

Table A.2 (*continued*).

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_{14}	ABN AMRO Bank	Netherlands	2017	24.3*	0.89	60.03	13.08	104.00	3.11	80	74	65	78.56
			2018	24.3*	0.83	59.40	11.00	119.00	1.92	88	65	72	62.67
			2019	24.3*	0.90	62.12	9.53	134.00	0.99	89	66	53	87.35
			2020	23.90	1.36	67.43	−0.21	149.00	2.03	90	65	75	131.92
			2021	22.40	0.93	77.11	5.61	168.00	1.88	93	65	73	91.27
a_{15}	Deutsche Pfandbriefbank	Germany	2017	22.20	0.18	57.81	6.37	187.00	−1.01	41	48	81	167.46
			2018	24.3*	0.26	51.89	5.50	212.00	−0.42	42	52	71	176.79
			2019	20.40	0.30	48.04	5.53	182.00	−3.92	43	56	61	202.74
			2020	21.40	0.60	47.34	3.67	245*	−1.33	46	54	68	283.95*
			2021	22.40	0.86	45.35	6.66	227.00	4.06	49	52	75	283.95
a_{16}	Skandinaviska Enskilda Banken	Sweden	2017	24.20	0.31	48.06	10.29	145.00	10.87*	84	84	55	95.46
			2018	22.20	0.34	47.99	15.55	147.00	10.62	93	80	44	107.19
			2019	23.30	0.42	45.88	12.96	218.00	10.87*	92	79	58	103.33
			2020	24.3*	0.57	47.83	9.16	163.00	10.68	90	76	91	111.79
			2021	23.10	0.46	42.26	13.16	145.00	9.51	92	77	80	101.93
a_{17}	Groupe Crédit Agricole	France	2017	17.40	2.72	65.20	6.52	137.00	6.88	94	74	50	272.37
			2018	17.20	2.52	63.79	7.67	133.00	−0.23	93	71	42	235.20
			2019	16.90	2.28	63.28	7.70	131.60	10.87*	95	76	56	257.08
			2020	18.97	2.31	62.88	4.41	157.30	10.87*	95	73	54	283.95*
			2021	17.46	1.90	61.00	8.90	183.00	6.46	96	74	32	244.37
a_{18}	Banca Monte Dei Paschi di Siena	Italy	2017	14.97	7.23*	83.85	**−5.58**	199.50	3.10	79	70	19	70.15
			2018	12.80	7.23*	89.97	3.10	190.20	−2.19	57	71	18	106.21
			2019	14.70	6.43	86.15	**−5.58**	152.40	**−6.60**	57	79	41	113.63
			2020	13.50	2.19	99.14*	**−5.58**	196.70	**−6.60**	76	78	71	164.69
			2021	14.56	2.84	84.85	5.01	172.70	10.87	76	77	77	103.85
a_{19}	Eurobank Ergasias Services and Holdings	Greece	2017	15.00	7.23*	48.25	1.61	**70.32**	7.44	62	72	64	91.76
			2018	14.00	7.23*	51.46	1.84	70.32	1.19	62	70	64	94.03
			2019	17.10	7.23*	55.71	1.90	97.00	−2.42	63	72	75	86.93
			2020	14.60	7.23*	46.78	**−5.58**	123.68	**−6.60**	62	84	72	73.12
			2021	15.30	4.58	47.69	5.80	152.24	−1.83	67	89	80	64.45
a_{20}	Banco Bilbao Vizcaya Argentaria	Spain	2017	15.30	3.27	56.27	8.90	128.00	0.03	87	90	93	105.76
			2018	15.50	3.15	54.83	11.59	127.00	3.01	84	86	88	111.32
			2019	15.42	3.14	54.62	7.89	128.00	1.92	83	81	80	131.52
			2020	15.92	3.75	53.44	4.10	149.00	3.55	93	81	73	180.49
			2021	16.98	3.37	55.03	11.52	165.00	0.63	95	84	51	147.92
a_{21}	Swedbank	Sweden	2017	24.3*	0.26	39.59	13.39	173.00	2.47	83	75	89	95.46
			2018	21.50	0.38	39.61	14.23	144.00	1.50	82	70	91	107.19
			2019	21.80	0.41	44.34	12.71	182.00	2.76	83	69	86	103.33
			2020	21.00	0.49	54.47	7.90	174.00	1.97	81	66	72	111.79
			2021	22.40	0.30	45.40	11.91	163.00	0.83	86	81	72	101.93
a_{22}	BPER Banca	Italy	2017	16.47	7.23*	69.24	3.09	113.70	1.90	27	63	28	70.15
			2018	14.93	7.09	71.82	9.10	154.30	−2.74	31	69	36	106.21
			2019	14.94	5.21	79.33	7.45	158.90	**−6.60**	81	69	64	113.63
			2020	19.30	3.27	71.53	4.14	200.10	**−6.60**	78	77	65	164.69
			2021	16.16	2.98	80.14	8.14	215.10	10.25	75	85	66	103.85
a_{23}	Svenska Handelsbanken	Sweden	2017	24.3*	0.25	45.56	10.64	133.00	4.28	74	71	34	95.46
			2018	21.00	0.17	46.11	12.20	146.00	−0.10	71	74	38	107.19
			2019	23.20	0.20	48.06	10.59	147.00	−0.13	87	78	43	103.33
			2020	24.30	0.15	52.08	9.09	150.00	0.11	93	76	62	111.79
			2021	23.30	0.13	47.47	10.75	152.00	0.00	94	76	61	101.93
a_{24}	Société Générale	France	2017	17.00	2.87	73.94	5.36	140.00	10.87*	89	91	82	272.37
			2018	16.50	2.51	71.14	7.31	129.00	0.82	97	90	83	235.20
			2019	18.90	2.31	71.85	5.75	119.00	10.87*	96	88	82	257.08
			2020	18.92	2.45	75.58	0.29	149.00	10.87*	94	88	83	283.95*
			2021	18.69	2.12	68.18	8.94	129.00	10.87*	96	84	84	244.37
a_{25}	Mediobanca	Italy	2017	17.11	3.91	60.41	8.08	245.00	**−6.60**	23	**48**	53	70.15
			2018	18.11	3.63	55.12	8.92	186.00	−6.60	36	57	74	106.21
			2019	17.28	3.35	57.25	8.35	143.00	4.34	37	57	59	113.63
			2020	18.67	3.28	56.60	6.17	165.00	−2.28	39	58	57	164.69
			2021	18.78	3.32	55.43	7.29	158.00	0.96	46	71	63	103.85
a_{26}	BNP Paribas	France	2017	14.60	3.54	68.95	7.66	121.00	10.87*	96	98	89	272.37
			2018	15.00	3.05	71.93	7.57	132.00	−3.48	95	97	88	235.20
			2019	15.50	2.56	70.27	7.67	125.00	10.87*	95	96	90	257.08
			2020	16.31	2.58	68.23	6.32	154.00	9.61	94	96	93	283.95*
			2021	16.40	2.38	67.33	8.06	143.00	10.87*	95	96	94	244.37

(*continued on next page*)

Table A.2 (*continued*).

	Alternatives			Criteria									
	Bank	Country	Year	CAMELS						ESG			EPU
				C	A	M	E	L	S	E_score	S_score	G_score	EPU_index
a_{27}	Banco de Sabadell	Spain	2017	16.02	2.56	61.68	6.09	133.00	4.06	86	62	57	105.76
			2018	14.13	2.35	72.38	2.53	151.00	0.59	86	73	75	111.32
			2019	15.00	1.95	69.86	5.50	158.00	−2.07	84	71	58	131.52
			2020	15.91	2.02	73.52	0.02	198.00	7.86	83	75	76	180.49
			2021	17.70	2.06	71.61	3.50	221.41	−3.49	86	77	71	147.92
a_{28}	Bankinter	Spain	2017	14.32	1.62	57.41	11.37	141.20	1.10	85	66	89	105.76
			2018	14.29	1.31	62.62	11.23	144.20	0.62	83	65	94	111.32
			2019	13.94	1.18	63.02	11.01	153.70	1.05	82	69	93	131.52
			2020	15.02	1.43	65.02	5.27	198.10	0.26	80	74	89	180.49
			2021	15.39	1.47	60.75	17.60*	228.60	0.83	82	77	84	147.92
			MIN	12.40	0.13	39.59	−5.58	70.32	−6.60	23	48	18	62.67
			MAX	24.30	7.23	99.14	17.60	245.00	10.87	100	98	95	283.95

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.omega.2024.103099.

References

[1] Beliakov G, Pradera A, Calvo T. Aggregation functions: A guide for practitioners. In: Studies in fuzziness and soft computing, vol. 221, Berlin Heidelberg: Springer; 2007.

[2] Greco S, Ishizaka A, Tasiou M, Torrisi G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Soc Indic Res 2019;141(1):61–94.

[3] OECD. Handbook on constructing composite indicators: Methodology and user guide. Paris: OECD Publishing; 2008.

[4] Greco S, Ishizaka A, Tasiou M, Torrisi G. Sigma-Mu efficiency analysis: A methodology for evaluating units through composite indicators. European J Oper Res 2019;278(3):942–60.

[5] Lahdelma R, Hokkanen J, Salminen P. SMAA - Stochastic multiobjective acceptability analysis. European J Oper Res 1998;106(1):137–43.

[6] Seiford LM, Zhu J. Context-dependent data envelopment analysis—Measuring attractiveness and progress. Omega 2003;31(5):397–408.

[7] Un. Who cares wins: Connecting financial markets to a changing world. Tech. Report, United Nations Global Compact; 2004.

[8] Ishizaka A, Lokman B, Tasiou M. A stochastic multi-criteria divisive hierarchical clustering algorithm. Omega 2021;103:102370.

[9] Tecles PL, Tabak BM. Determinants of bank efficiency: The case of Brazil. European J Oper Res 2010;207(3):1587–98.

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

[11] Brissimis SN, Delis MD, Tsionas EG. Technical and allocative efficiency in European banking. European J Oper Res 2010;204(1):153–63.

[12] Farrell MJ. The measurement of productive efficiency. J Roy Statist Soc Ser A 1957;120(3):253–81.

[13] Aigner D, Lovell CK, Schmidt P. Formulation and estimation of stochastic frontier production function models. J Econom 1977;6(1):21–37.

[14] Kasman A, Yildirim C. Cost and profit efficiencies in transition banking: The case of new EU members. Appl Econ 2006;38(9):1079–90.

[15] Williams J. Efficiency and market power in latin American banking. J Financ Stab 2012;8(4):263–76.

[16] Badunenko O, Kumbhakar SC, Lozano-Vivas A. Achieving a sustainable cost-efficient business model in banking: The case of European commercial banks. European J Oper Res 2021;293(2):773–85.

[17] Kraft E, Hofler R, Payne J. Privatization, foreign bank entry and bank efficiency in Croatia: A Fourier-flexible function stochastic cost frontier analysis. Appl Econ 2006;38(17):2075–88.

[18] Lensink R, Meesters A, Naaborg I. Bank efficiency and foreign ownership: Do good institutions matter? J Bank Financ 2008;32(5):834–44.

[19] Delis M, Iosifidi M, Tsionas MG. Endogenous bank risk and efficiency. European J Oper Res 2017;260(1):376–87.

[20] Almanidis P, Karakaplan M, Kutlu L. A dynamic stochastic frontier model with threshold effects: U.S. bank size and efficiency. J Prod Anal 2019;52:69–84.

[21] Tsionas MG. Optimal combinations of stochastic frontier and data envelopment analysis models. European J Oper Res 2021;294(2):790–800.

[22] Poghosyan T, Kumbhakar SC. Heterogeneity of technological regimes and banking efficiency in former socialist economies. J Prod Anal 2010;33:19–31.

[23] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European J Oper Res 1978;2(6):429–44.

[24] Tulkens H. On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. J Prod Anal 1993;4(1–2):183–210.

[25] Kaffash S, Marra M. Data envelopment analysis in financial services: A citations network analysis of banks, insurance companies and money market funds. Ann Oper Res 2017;253(1):307–44.

[26] Hammami H, Ngo T, Tripe D, Vo D-T. Ranking with a Euclidean common set of weights in data envelopment analysis: With application to the Eurozone banking sector. Ann Oper Res 2022;311(2):675–94.

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

[28] Lozano-Vivas A, Pastor JT. Do performance and environmental conditions act as barriers for cross-border banking in Europe? Omega 2010;38(5):275–82.

[29] Delis MD, Koutsomanoli-Fillipaki A, Staikouras CK, Katerina G. Evaluating cost and profit efficiency: A comparison of parametric and nonparametric methodologies. Appl Financial Econ 2009;19(3):191–202.

[30] Weill L. Measuring cost efficiency in European banking: A comparison of frontier techniques. J Prod Anal 2004;21:133–52.

[31] Ferrier GD, Lovell CK. Measuring cost efficiency in banking: Econometric and linear programming evidence. J Econometrics 1990;46(1–2):229–45.

[32] Aiello F, Bonanno G. On the sources of heterogeneity in banking efficiency literature. J Econ Surv 2018;32(1):194–225.

[33] Doumpos M, Zopounidis C, Gounopoulos D, Platanakis E, Zhang W. Operational research and artificial intelligence methods in banking. European J Oper Res 2023;306(1):1–16.

[34] Tsagkarakis M-P, Doumpos M, Pasiouras F. Capital shortfall: A multicriteria decision support system for the identification of weak banks. Decis Support Syst 2021;145:113526.

[35] Manthoulis G, Doumpos M, Zopounidis C, Galariotis E. An ordinal classification framework for bank failure prediction: Methodology and empirical evidence for US banks. European J Oper Res 2020;282(2):786–801.

[36] Calabrese R, Osmetti SA. A new approach to measure systemic risk: A bivariate copula model for dependent censored data. European J Oper Res 2019;279(3):1053–64.

[37] Paltalidis N, Gounopoulos D, Kizys R, Koutelidakis Y. Transmission channels of systemic risk and contagion in the European financial network. J Bank Financ 2015;61:S36–52, Global Trends in Banking, Regulations, and Financial Markets.

[38] Ye W, Zhou Y, Chen P, Wu B. A simulation-based method for estimating systemic risk measures. European J Oper Res 2024;313(1):312–24.

[39] Lagasio V, Pampurini F, Pezzola A, Quaranta AG. Assessing bank default determinants via machine learning. Inform Sci 2022;618:87–97.

[40] Gogas P, Papadimitriou T, Agrapetidou A. Forecasting bank failures and stress testing: A machine learning approach. Int J Forecast 2018;34(3):440–55.

[41] Behbood V, Lu J, Zhang G, Pedrycz W. Multistep fuzzy bridged refinement domain adaptation algorithm and its application to bank failure prediction. IEEE Trans Fuzzy Syst 2015;23(6):1917–35.

[42] Climent F, Momparler A, Carmona P. Anticipating bank distress in the Eurozone: An extreme gradient boosting approach. J Bus Res 2019;101:885–96.

[43] Petropoulos A, Siakoulis V, Stavroulakis E, Vlachogiannakis NE. Predicting bank insolvencies using machine learning techniques. Int J Forecast 2020;36(3):1092–113.

[44] Rönnqvist S, Sarlin P. Bank distress in the news: Describing events through deep learning. Neurocomputing 2017;264:57–70.

[45] Smith P. Developing composite indicators for assessing health system efficiency. In: Smith P, editor. Measuring up: Improving health system performance in OECD countries. Paris: OECD; 2002, p. 295–316.

[46] Saisana M, Saltelli A, Tarantola S. Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. J Roy Statist Soc Ser A 2005;168(2):307–23.

[47] Hope C, Parker J, Peake S. A pilot environmental index for the UK in the 1980s. Energy Policy 1992;20(4):335–43.

[48] Doumpos M, Gaganis C, Pasiouras F. Bank diversification and overall financial strength: International evidence. Financ Mark, Inst Instrum 2016;25(3):169–213.

[49] Rebai S, Azaiez MN, Saïdane D. A multi-attribute utility model for generating a sustainability index in the banking sector. J Clean Prod 2016;113:835–49.

[50] Gulati R, Kattumuri R, Kumar S. A non-parametric index of corporate governance in the banking industry: An application to Indian data. Socio-Econ Plan Sci 2020;70:100702.

[51] Gaganis C, Pasiouras F, Tasiou M, Zopounidis C. CISEF: A composite index of social, environmental and financial performance. European J Oper Res 2021;291(1):394–409.

[52] Lahdelma R, Salminen P. SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. Oper Res 2001;49(3):444–54.

[53] Markowitz H. Portfolio selection. J Finance 1952;7(1):77–91.

[54] EBA. 2023 EU-Wide stress test - Methodological note. Technical report, European Banking Authority; 2022, https://t.ly/f5BbB.

[55] Sahajwala R, Van den Bergh P. Supervisory risk assessment and early warning systems. Technical report 4, Basel, Switzerland: Bank of International Settlements; 2000.

[56] Refinitiv. Environmental, social and governance scores, brochure on refinitiv website, May 2022. 2022, https://t.ly/kzD_c. [Accessed 25 September 2023].

[57] Bătae O, Dragomir V, Feleagă L. The relationship between environmental, social, and financial performance in the banking sector: A European study. J Clean Prod 2021;290:125791.

[58] Baker SR, Bloom N, Davis SJ. Measuring economic policy uncertainty. Q J Econ 2016;131(4):1593–636.

[59] ShareAction. Banking on a low-carbon future II – A ranking of the 20 largest European banks' responses to climate change. 2020, Available online at: https://t.ly/auILQ. [Accessed 25 September 2023].

[60] Di Tommaso C, Thornton J. Do ESG scores effect bank risk taking and value? Evidence from European banks. Corp Soc Responsib Environ Manag 2020;27(5):2286–98.

[61] Menicucci E, Paolucci G. ESG dimensions and bank performance: An empirical investigation in Italy. Corp Gov: Int J Bus Soc 2023;23(3):563–86.

[62] Cherchye L, Moesen W, Rogge N, Puyenbroeck TV. An introduction to 'benefit of the doubt' composite indicators. Soc Indic Res 2007;82:111–45.

[63] Walheer B. A sequential benefit-of-the-doubt composite indicator. European J Oper Res 2024;(in press).

[64] Billio M, Costola M, Hristova I, Latino C, Pelizzon L. Inside the ESG ratings:(Dis) agreement and performance. Corp Soc Responsib Environ Manag 2021;28(5):1426–45.

[65] Del Vitto A, Marazzina D, Stocco D. ESG ratings explainability through machine learning techniques. Ann Oper Res 2023;1–30.

[66] Citterio A, King T. The role of environmental, social, and governance (ESG) in predicting bank financial distress. Finance Research Letters 2023;51:103411.

[67] Angilella S, Pappalardo MR. Assessment of a failure prediction model in the European energy sector: A multicriteria discrimination approach with a PROMETHEE based classification. Expert Syst Appl 2021;184:115513.

[68] Tsionas MG. Multi-objective optimization using statistical models. European J Oper Res 2019;276(1):364–78, URL https://www.sciencedirect.com/science/article/pii/S0377221718311160.

[69] Paruolo P, Saisana M, Saltelli A. Ratings and rankings: Voodoo or science? J R Stat Soc Ser A: Stat Soc 2013;176(3):609–34.