European Journal of Operational Research 335 (2026) 868–887

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

Robust multiplier data envelopment analysis under data uncertainty: Correlated budgeted and order statistics models

Adel Hatami-Marbini [a], Aliasghar Arabmaldar [b,*], Amin Hosseinian-Far [b]

[a] Huddersfield Business School, University of Huddersfield, Huddersfield, United Kingdom
[b] Hertfordshire Business School, University of Hertfordshire, Hatfield, United Kingdom

ARTICLE INFO

Keywords:
Data envelopment analysis (DEA)
Robust optimisation
Efficiency evaluation
Uncertainty modelling
Correlated and order statistics uncertainty sets

ABSTRACT

In engineering and management, multiplier Data Envelopment Analysis (DEA) models provide decision-makers with a foundational and interpretable framework by defining efficiency as the ratio of weighted outputs to weighted inputs. However, the pervasive presence of data uncertainty across applications is often overlooked, compromising the robustness of the results. This paper builds on the widely adopted budgeted uncertainty model in robust optimisation to develop two novel robust multiplier-based DEA models: one that incorporates correlated budgeted uncertainty to account for dependencies among uncertain variables, and another that employs order statistics to capture distribution-based risk scenarios. The proposed models are examined through both theoretical analysis and computational experiments using real and simulated datasets. The results demonstrate that the correlated model strikes an effective balance between robustness and computational efficiency, whereas the order statistics model enhances reliability by explicitly capturing distributional characteristics of the uncertain data. Together, these models are complementary, with the correlated budgeted model suited to large-scale or routine evaluations, and the order statistics model preferable when resilience under extreme uncertainty is the priority. This study provides decision-makers with practical guidance for selecting the most suitable robust DEA framework in environments where data uncertainty is a critical concern.

1. Introduction

Data Envelopment Analysis (DEA) is a well-established non-parametric method for evaluating the relative efficiency of a group of homogeneous decision-making units (DMUs) that use multiple inputs to produce multiple outputs (Mergoni et al., 2025). Foundational DEA models including CCR (Charnes et al., 1978) and BCC (Banker et al., 1984) have been extensively applied across various sectors such as healthcare, education, manufacturing, and banking (Mergoni et al., 2025). These models enable performance benchmarking without requiring a predefined functional form (Cook & Seiford, 2009). Within the DEA framework, efficiency can be evaluated using envelopment or multiplier models, where the former aims to construct a best-practice frontier and the latter derives optimal input-output weights (Cooper et al., 2007). A fundamental limitation of classical DEA lies in the assumption of deterministic input and output data. In real-world problems, however, performance measurements are frequently affected by data uncertainty arising from measurement errors, missing values, and temporal fluctuations (Lu, 2015; Ehrgott et al., 2018; Kao & Liu, 2022). These data imperfections can considerably compromise the reliability and robustness of efficiency results and rankings obtained from DEA models, particularly in dynamic or data-constrained environments (Hatami-Marbini & Arabmaldar, 2021). Resultantly, addressing data uncertainty has lately emerged as a pivotal theme within the DEA literature (Zhu, 2003; Olesen & Petersen, 2016).

Several methodological families have been developed to handle uncertainty in DEA. Stochastic and chance-constrained frameworks incorporate randomness through probabilistic constraints (Olesen & Petersen, 1995, 2016), whereas bootstrap DEA improves the statistical reliability of efficiency estimates through resampling techniques (Simar & Wilson, 1998; Pham et al., 2023). Fuzzy DEA models capture qualitative or vague information using fuzzy sets theory (Hatami-Marbini et al., 2017; Aparicio et al., 2023). Interval (imprecise) DEA represents inputs and outputs as bounded ranges and typically measures optimistic (best-case) and pessimistic (worst-case) efficiency scores at the interval extremes (Despotis & Smirlis, 2002; Kao, 2006). As Shokouhi et al. (2010) discussed, interval DEA can be interpreted as a special case of

* Corresponding author.
E-mail addresses: a.hatamimarbini@hud.ac.uk (A. Hatami-Marbini), a.arabmaldar@herts.ac.uk (A. Arabmaldar), a.hosseinian-far@herts.ac.uk (A. Hosseinian-Far).

https://doi.org/10.1016/j.ejor.2026.03.013
Received 23 June 2025; Accepted 9 March 2026
Available online 10 March 2026

robust DEA because it evaluates efficiency only at the extreme points of the uncertainty set; consequently, it yields informative bounds but does not characterise behaviour for intermediate realisations inside the interval. This limitation motivates the robustness-path perspective adopted in this study. Robust DEA has recently attracted considerable attention as an alternative approach, as it protects efficiency scores against worst-case deviations in input and output data within a prescribed uncertainty set while avoiding the need for full distributional assumptions (Arabmaldar et al., 2023; Shokouhi et al., 2010).

We prioritise robust DEA in this study for three practical reasons. First, robust DEA handles data uncertainty without requiring full distributional assumptions, resampling, or analyst-specified membership functions (Bertsimas and Sim, 2004). Second, it produces a single worst-case efficiency score that also aids comparison and ranking (Hatami-Marbini and Arabmaldar, 2021). Third, robust DEA is flexible, as the uncertainty set and its adjustable parameters permit decision-makers to trade off robustness and conservatism in a transparent manner. It should be noted that interval DEA remains complementary to robust DEA and can be viewed as a special case of budget-type uncertainty (Shokouhi et al., 2010). The two approaches provide different practical and managerial purposes—explicit [LB, UB] interval efficiency based on pessimistic and optimistic viewpoints versus a single efficiency score based on the uncertainty budget—and can be applied together depending on the application. In this context, the uncertainty budget Γ provides a practical lever for decision-makers, allowing them to adjust the level of conservatism and effectively span the range between optimistic and pessimistic scenarios within a single production possibility set (PPS).

Robust optimisation, as a methodological backbone for our approach, aims to accept a suboptimal solution for the nominal values of the data in order to ensure that the solution remains feasible and near optimal when the data changes under a set of plausible data realisations (Ben-Tal & Nemirovski, 2000; Bertsimas & Sim, 2004). At the core of this methodology is the definition of an uncertainty set, which specifies the range of data variations considered during optimisation. The *budgeted uncertainty* set, introduced by Bertsimas and Sim (2004), is particularly popular due to its tractability and flexibility in improving robustness, i.e., the trade-off between the solution feasibility and near optimality, through an adjustable conservatism parameter. Unlike earlier models based on interval sets (Soyster, 1973) or ellipsoidal sets (Ben-Tal & Nemirovski, 2000), the budgeted uncertainty model allows decision-makers to explicitly control the degree of conservatism through a budget parameter. This provides a structured trade-off, with robustness quantified through the price of robustness (PoR).

The budgeted approach has since become a cornerstone of robust optimisation, and various refinements have been proposed to improve its practicality. Poss (2013, 2014) developed a robust combinatorial optimisation model incorporating a *variable budgeted uncertainty* set, where the level of conservatism adapts based on decision variables, which is particularly effective for sparse solutions. Ke et al. (2013) introduced a proportion-based robust optimisation approach, allowing a specified proportion of uncertain coefficients in each solution to vary. Both approaches extend the original formulation while preserving its probabilistic guarantees. It should be noted that, although the original Bertsimas and Sim (2004) approach was based on independent uncertainty, it has since been extended to account for correlated variations, a common feature in real-world applications such as portfolio optimisation in finance and demand response modelling in the electricity industry (Gregory et al., 2011; Ferreira et al., 2012).

Recent advances propose even more refined formulations, including data-driven uncertainty sets (Bertsimas et al., 2018), order statistics sets (Zhang & Gupta, 2023), and methods that filter out ineffective worst-case regions (Dehghani Filabadi & Mahmoudzadeh, 2022). These developments aim to reduce conservatism while maintaining robustness, yet the budgeted approach remains a cornerstone due to its simplicity, tractability, and adaptability, particularly when designing new variants tailored to specific application needs. For instance, Zhang and Gupta (2023) presented a new uncertainty set for robust linear optimisation based on order statistics and quantiles of independent random variables, aiming to unify several existing sets such as interval, budget, and demand uncertainty under a single, flexible framework. Dehghani Filabadi and Mahmoudzadeh (2022) introduced a refined two-stage robust optimisation method that improves upon the traditional budgeted approach by identifying and excluding ineffective regions. Their model achieves less conservative and more efficient solutions through a focus on effective worst-case scenarios.

Robust DEA adapts robust optimisation principles to enhance the reliability of efficiency assessments under uncertainty. Specifically, existing robust DEA models are often built on the budgeted uncertainty set to protect efficiency scores against worst-case deviations in inputs and outputs. Early models addressed output uncertainty only (Sadjadi & Omrani, 2008), later expanding to include both input and output uncertainty under constant and variable returns to scale technologies (Shokouhi et al., 2010; Arabmaldar et al., 2017; Toloo & Mensah, 2019; A. Hatami-Marbini et al., 2022). More recent contributions have proposed non-radial, cost efficiency, cross-efficiency, directional distance, and fractional robust models with adjustable levels of robustness (Salahi et al., 2021; Hatami-Marbini et al., 2022a, 2022b; Toloo et al., 2022; Wu et al., 2024; Arabmaldar et al., 2023). Applications particularly in the banking sector have reinforced the value of budgeted and related uncertainty sets (Arabmaldar et al., 2024).

Despite these studies, many existing robust DEA models rely on simplifying assumptions, such as independence among uncertain variables or symmetric deviations in data, that may be unrealistic in practice. Inputs and outputs often exhibit interdependencies, skewness and heavy tails; ignoring these features can lead to underestimating or overestimating the true impact of uncertainty on efficiency.

It is important to emphasise that the motivation for robust DEA is not to ensure feasibility—which is already guaranteed in classical DEA models such as CCR and BCC—but to enhance the robustness and reliability of efficiency scores under data uncertainty. We adopt the CCR model as the baseline formulation in this paper because it is widely used, transparent, and provides a tractable setting for developing and testing new uncertainty sets without confounding effects from varying returns to scale assumptions. Building on this baseline, we introduce two novel robust DEA models that reshape the uncertainty region and ensure efficiency results remain valid across admissible perturbations. The first model incorporates a *correlated budgeted uncertainty* set, allowing for interdependencies among input and output variations, particularly relevant in sectors such as banking or supply chains, where variables often move together. The second model is developed based on an *order statistics uncertainty* set, using ranked data to identify and protect against worst-case scenarios without relying on full probabilistic distributions.

Although inspired by the growing literature on robust DEA, our approach differs fundamentally from recent models, particularly Arabmaldar et al. (2024), who developed a non-linear robust DEA model based on a variable budgeted uncertainty set in which the degree of protection is endogenously adjusted. In contrast, our models retain the standard multiplier DEA framework form under a linear optimisation structure and a single PPS, while introducing robustness by reshaping the uncertainty region through the two uncertainty sets described above. In addition to its theoretical developments, this study also makes empirical contributions. The proposed models are evaluated against a conventional robust DEA model using synthetic data for comparative analysis and an empirical application involving 168 European banks. Evaluation criteria include efficiency scores, ranking stability, computational complexity, and the PoR. Results highlight the practical implications of choosing different robustness schemes under uncertainty.

The remainder of the paper is organised as follows: Section 2 reviews the mathematical background of classical and robust DEA models. Section 3 presents the proposed DEA models incorporating correlated budgeted and order statistics uncertainty sets. Section 4 provides a comparative analysis using synthetic data. Section 5 validates the models empirically and discusses managerial implications. Section 6 concludes with key findings and directions for future research.

2. Theoretical preliminaries

DEA is a non-parametric technique widely used to measure the relative efficiencies of DMUs that utilise multiple inputs to produce multiple outputs. Consider n DMUs indexed as $\text{DMU}_j;\ j \in J = \{1, \ldots, n\}$ where each DMU consumes m semi-positive input $\boldsymbol{x}_j = (\cdots, x_{ij}, \cdots)^T;\ i \in I = \{1, \ldots, m\}$ to produce s semi-positive outputs $\boldsymbol{y}_j = (\cdots, y_{rj}, \cdots)^T;\ r \in R = \{1, \ldots, s\}$. The following fractional programming formulation proposed by Charnes et al. (1978) assesses the efficiency of DMU_o relative to the entire set of DMUs:

$$\begin{aligned} &\max \frac{\boldsymbol{u}\boldsymbol{y}_o}{\boldsymbol{v}\boldsymbol{x}_o} \\ &\text{s.t.} \\ &\frac{\boldsymbol{u}\boldsymbol{y}_j}{\boldsymbol{v}\boldsymbol{x}_j} \le 1,\ \forall j, \\ &\boldsymbol{u}, \boldsymbol{v} \ge 0, \end{aligned} \tag{1}$$

where $\boldsymbol{v} = (v_1, \cdots, v_m) \in \mathbb{R}^m$ and $\boldsymbol{u} = (u_1, \cdots, u_s) \in \mathbb{R}^s$ are the weights (decision variables) associated with the i^{th} input and r^{th} output, respectively. A well-established approach for transforming this nonlinear fractional programming model into a linear form, as outlined by Charnes and Cooper (1962). However, the presence of inherent uncertainty in input and output data must be explicitly considered, as it can significantly impact the accuracy and reliability of efficiency assessments (Zhu, 2003). Ignoring such uncertainty may lead to misleading conclusions about the performance of DMUs, particularly in environments subject to variability, data imprecision, or estimation errors (Park, 2010). To address this, the robust counterpart of model (1) can be formulated using a worst-case scenario framework, which captures uncertainty in both the objective function and the constraints. Let $\mathscr{U}_j \subseteq \mathbb{R}^{m+s}$ denote an uncertainty set for DMU_j representing all admissible realisations $(\tilde{\boldsymbol{x}}_j, \tilde{\boldsymbol{y}}_j)$. The robust model can then be expressed as follows:

$$\begin{aligned} &\max_{(\boldsymbol{u},\boldsymbol{v})} \min_{(\tilde{\boldsymbol{x}}_o,\tilde{\boldsymbol{y}}_o) \in \mathscr{U}_o} \left\{ \frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_o}{\boldsymbol{v}\tilde{\boldsymbol{x}}_o} \right\} \\ &\text{s.t.} \\ &\min_{(\tilde{\boldsymbol{x}}_o,\tilde{\boldsymbol{y}}_o) \in \mathscr{U}_o} \left\{ \frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_o}{\boldsymbol{v}\tilde{\boldsymbol{x}}_o} \right\} \le 1, \\ &\max_{(\tilde{\boldsymbol{x}}_j,\tilde{\boldsymbol{y}}_j) \in \mathscr{U}_j} \left\{ \frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_j}{\boldsymbol{v}\tilde{\boldsymbol{x}}_j} \right\} \le 1,\ \forall j \ne o, \\ &\boldsymbol{v}, \boldsymbol{u} \ge 0. \end{aligned} \tag{2}$$

Here $\mathscr{U}_j$ defines the bounded region of possible input–output observations for DMU_j. In the robust model (2), the n constraints of model (1) are reformulated as a single constraint that controls the worst-case performance of the evaluated DMU (DMU_o) and $n-1$ constraints that control the best-case performance of the peer DMUs (i.e., $\text{DMU}_{j \ne o}$).

The rationale for employing the max–min objective and the min/max operators in the constraints of model (2) is outlined below. Because efficiency under uncertainty is not a single scalar but a range of possible ratios $\frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_j}{\boldsymbol{v}\tilde{\boldsymbol{x}}_j}$ for $(\tilde{\boldsymbol{x}}_j, \tilde{\boldsymbol{y}}_j) \in \mathscr{U}_j$, robust optimisation seeks weights $(\boldsymbol{u}, \boldsymbol{v})$ that guarantee acceptable performance across all admissible realisations. The objective $\max_{(\boldsymbol{u},\boldsymbol{v})} \min_{(\tilde{\boldsymbol{x}}_o,\tilde{\boldsymbol{y}}_o) \in \mathscr{U}_o} \left\{ \frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_o}{\boldsymbol{v}\tilde{\boldsymbol{x}}_o} \right\}$ therefore chooses weights that maximise DMU_o's worst-case efficiency — i.e. the highest guaranteed ratio that holds for every realisation inside $\mathscr{U}_o$. The first constraint enforces that this worst-case ratio does not exceed one for DMU_o (the DEA frontier), thereby preserving the DEA feasibility condition for DMU_o under all realisations in $\mathscr{U}_j$. Conversely, the peer constraints (for $\text{DMU}_{j \ne o}$) require that no competitor can, under any favourable realisation within its uncertainty set, achieve a ratio exceeding 1; this is expressed by $\max_{(\tilde{\boldsymbol{x}}_j,\tilde{\boldsymbol{y}}_j) \in \mathscr{U}_j} \left\{ \frac{\boldsymbol{u}\tilde{\boldsymbol{y}}_j}{\boldsymbol{v}\tilde{\boldsymbol{x}}_j} \right\} \le 1$. Together, these requirements, consistent with the relevant literature (see e.g., Arabmaldar et al., 2024; Park, 2007; Toloo et al., 2022), ensure that the obtained efficiency is both conservative (robust against adverse realisations of DMU_o) and fair (precluding artificially high peer performance under advantageous realisations).[1]

To represent the uncertainty defined by the set $\mathscr{U}_j$, the true values of the uncertain inputs and outputs are modelled as $\tilde{\boldsymbol{x}}_j = \boldsymbol{x}_j + \boldsymbol{\eta}_j^x \hat{\boldsymbol{x}}_j$ and $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j + \boldsymbol{\eta}_j^y \hat{\boldsymbol{y}}_j$ for all j. Here, $\hat{\boldsymbol{x}}_j = e_i \boldsymbol{x}_j$ and $\hat{\boldsymbol{y}}_j = e_r \boldsymbol{y}_j$ denote the deviations from the nominal values $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$, where e_i and e_r represent the percentage perturbations that define the level of uncertainty in $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$, respectively. A DMU_k, characterised by the input-output pair $(\boldsymbol{x}_k, \boldsymbol{y}_k)$, is referred to as *uncertain* if there exists at least one $i \in I_k$ or $r \in R_k$ such as $x_{ij}^k > 0,\ \forall j$ or $y_{rj}^k > 0,\ \forall j$. I_k and R_k denote the sets of uncertain inputs and outputs, respectively. When only a subset of the input and output data is uncertain, with the remainder known precisely, the notations I and R are used, where $I_j \subseteq I$ and $R_j \subseteq R$. The uncertainty vectors $\boldsymbol{\eta}_j^x = (\eta_{1j}^x, \ldots, \eta_{mj}^x)$, $\boldsymbol{\eta}_j^y = (\eta_{1j}^y, \ldots, \eta_{sj}^y)$ consist of independently distributed random variables bounded within the interval [-1, 1]. Consequently, the uncertain input and output values $\tilde{\boldsymbol{x}}_j$ and $\tilde{\boldsymbol{y}}_j$ lie within the symmetric intervals $[\boldsymbol{x}_j - \hat{\boldsymbol{x}}_j, \boldsymbol{x}_j + \hat{\boldsymbol{x}}_j]$ and $[\boldsymbol{y}_j - \hat{\boldsymbol{y}}_j, \boldsymbol{y}_j + \hat{\boldsymbol{y}}_j]$, respectively. We assume that $\Gamma_j^x \in [0, |I_j|]$, and $\Gamma_j^y \in [0, |R_j|]$ are the budget of uncertainty parameters, representing the maximum number of uncertain input and output variables, respectively, that are allowed to vary. In addition, we impose the constraints $\mathbf{1}_{|I_j|} |\boldsymbol{\eta}_j^x| \le \Gamma_j^x$ and $\mathbf{1}_{|R_j|} |\boldsymbol{\eta}_j^y| \le \Gamma_j^x$, where $\mathbf{1}_{|I_j|}$ and $\mathbf{1}_{|R_j|}$ are vectors of ones of appropriate dimensions. Under this setup, the budgeted uncertainty set is defined as:

$$\mathscr{U}(\Gamma_j^x + \Gamma_j^y) = \left\{ (\tilde{\boldsymbol{x}}_j, \tilde{\boldsymbol{y}}_j) \in \mathbb{R}^{|I_j|+|R_j|} \left| \begin{array}{l} \tilde{\boldsymbol{x}}_j = \boldsymbol{x}_j + \boldsymbol{\eta}_j^x \hat{\boldsymbol{x}}_j,\ |\boldsymbol{\eta}_j^x| \le 1, \mathbf{1}_{|I_j|} |\boldsymbol{\eta}_j^x| \le \Gamma_j^x, \forall j \\ \tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j + \boldsymbol{\eta}_j^y \hat{\boldsymbol{y}}_j,\ |\boldsymbol{\eta}_j^y| \le 1, \mathbf{1}_{|R_j|} |\boldsymbol{\eta}_j^y| \le \Gamma_j^y, \forall j \end{array} \right. \right\} \tag{3}$$

The budget parameter Γ_j determines the extent of allowable uncertainty within the model's constraints. Following the budgeted uncertainty framework proposed by Bertsimas and Sim (2004), the following robust counterpart DEA formulation of model (2) is presented, which accommodates varying levels of uncertainty:

[1] It is important to highlight the distinction between *bounded data* and *robust protection*. Boundedness (the existence of $\mathscr{U}_j$) merely defines the region of possible data values. Classical bounded-data approaches —such as optimistic/pessimistic interval DEA (Despotis & Smirlis, 2002; Kao, 2006)—evaluate performance at boundary points only; by contrast, our robust formulation enforces validity simultaneously for all realisations in the uncertainty set. Thus, bounded data provide the domain for uncertainty, while the min/max operators implement worst- and best-case protection inside that domain.

$$\begin{aligned}
&\max\ \boldsymbol{u}\boldsymbol{y}_o + \beta_0\left(\boldsymbol{u}^*, \mathbf{0}_m, \Gamma_o^y\right)\\
&\text{s.t.}\\
&\boldsymbol{v}\boldsymbol{x}_o + \beta_0\left(\mathbf{0}_s, \boldsymbol{v}^*, \Gamma_o^x\right) \le 1,\\
&\boldsymbol{u}\boldsymbol{y}_o - \boldsymbol{v}\boldsymbol{x}_o - \beta_0\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_0^x + \Gamma_o^y\right) \le 0,\\
&\boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j + \beta_j\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_j^x + \Gamma_j^y\right) \le 0,\ \forall j \ne o,\\
&\boldsymbol{u}, \boldsymbol{v} \ge 0,
\end{aligned} \tag{4}$$

where $\beta_j\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_j^x + \Gamma_j^y\right)$ represents a protection function that quantifies the impact of uncertain data on each constraint $j \ne o$. The terms $\beta_0\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_0^x + \Gamma_o^y\right)$, $\beta\left(\boldsymbol{u}^*, \mathbf{0}_m, \Gamma_o^y\right)$ and $\beta\left(\mathbf{0}_s, \boldsymbol{v}^*, \Gamma_o^x\right)$ in model (4) address uncertainty in the inputs and outputs for DMU$_o$. The protection function is formulated as:

$$\beta_j\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_j^x + \Gamma_j^y\right) = \max_{\mathscr{S}_j}\{\xi\} \tag{5}$$

where
$\xi = \sum_{r \in S_j} u_r \widehat{y}_{rj} + \left(\Gamma_j^y - \lfloor \Gamma_j^y \rfloor\right) u_t \widehat{y}_{rt_j} + \sum_{i \in \overline{S}_j} v_i \widehat{x}_{ij} + \left(\Gamma_j^x - \lfloor \Gamma_j^x \rfloor\right) v_t \widehat{x}_{it_j}$.
Here, $S_j \subseteq R_j$ and $\overline{S}_j \subseteq I_j$ are subsets of the output and input indices, respectively, representing the worst-case deviations permitted within the uncertainty budgets. Specifically, the sets S_j and $\overline{S}_j$ contain the $\lfloor \Gamma_j^y \rfloor$ and $\lfloor \Gamma_j^x \rfloor$ elements associated with the largest potential deviations in outputs and inputs, respectively. The elements $t_j \in R_j \backslash S_j$ and $\bar{t}_j \in I_j \backslash \overline{S}_j$, which correspond to the fractional parts of the uncertainty budgets Γ_j^y and Γ_j^x, represent the most impactful uncertain variables beyond those included in S_j and $\overline{S}_j$. Their effects are weighted by the fractional components $\Gamma_j^y - \lfloor \Gamma_j^y \rfloor$ or $\Gamma_j^x - \lfloor \Gamma_j^x \rfloor$, respectively. The index set $\mathscr{S}_j$ thus enumerates all valid combinations of these subsets and fractional elements, defined as: $\mathscr{S}_j = \left\{ \begin{array}{l} S_j \cup \{t_j\} \mid S_j \subseteq R_j,\ |S_j| = \lfloor \Gamma_j^y \rfloor,\ t_j \in R_j \backslash S_j \\ \overline{S}_j \cup \{\bar{t}_j\} \mid \overline{S}_j \subseteq I_j,\ |\overline{S}_j| = \lfloor \Gamma_j^x \rfloor,\ \bar{t}_j \in I_j \backslash \overline{S}_j \end{array} \right\}$.

The protection function $\beta_j\left(\boldsymbol{u}^*, \boldsymbol{v}^*, \Gamma_j^x + \Gamma_j^y\right)$ therefore collects the worst-case perturbation that can affect constraint j when up to $\lfloor \Gamma_j^y \rfloor$ outputs and $\lfloor \Gamma_j^x \rfloor$ inputs take their full deviations, and at most one addition term is scaled by the fractional part of the budget. Intuitively, the function selects the components that maximise the weighted perturbation relative to the multipliers $(\boldsymbol{u}, \boldsymbol{v})$ and sums their impacts, thereby protecting the left-hand side of the constraint against those worst-case scenarios.

This formulation aligns with the classical budgeted uncertainty model of Bertsimas and Sim (2004), in which uncertainty is modelled independently across inputs and outputs. The budget parameters $\Gamma_j^x \in [0, |I_j|]$, and $\Gamma_j^y \in [0, |R_j|]$ allow decision-makers to control the level of conservatism. When all inputs and outputs are fully subject to uncertainty, i.e., $\Gamma_j^x = |I_j|$ and $\Gamma_j^y = |R_j|$, the model represents the worst-case scenario. Conversely, setting $\Gamma_j^x = \Gamma_j^y = 0$ yields the nominal (uncertainty-free case). By adjusting $\Gamma_j^x \in [0, |I_j|]$, and $\Gamma_j^y \in [0, |R_j|]$, decision-makers can effectively balance robustness against conservatism in performance evaluation.

Following the approach developed by Bertsimas and Sim (2004), the robust counterpart model (4), which has been extended and successfully applied in the DEA literature (see Arabmaldar et al., 2017; Toloo et al., 2022), is formulated as follows:

$$\begin{aligned}
&\theta_R = \max\ \boldsymbol{u}\boldsymbol{y}_o - p_o^y \Gamma_o^y - \mathbf{1}_{|R_o|}^T \boldsymbol{w}_o^y\\
&\text{s.t.}\\
&\boldsymbol{v}\boldsymbol{x}_o + p_o^x \Gamma_o^x + \mathbf{1}_{|I_o|}^T \boldsymbol{w}_o^x \le 1,\\
&\boldsymbol{u}\boldsymbol{y}_o - \boldsymbol{v}\boldsymbol{x}_o - p_o^x \Gamma_o^x - p_o^y \Gamma_o^y - \mathbf{1}_{|R_o|}^T \boldsymbol{w}_o^y - \mathbf{1}_{|I_o|}^T \boldsymbol{w}_o^x \le 0,\\
&\boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j + p_j^x \Gamma_j^x + p_j^y \Gamma_j^y + \mathbf{1}_{|R_j|}^T \boldsymbol{w}_j^y + \mathbf{1}_{|I_j|}^T \boldsymbol{w}_j^x \le 0,\ \forall j \ne o,\\
&p_j^y + \boldsymbol{w}_j^y \ge \boldsymbol{u}\widehat{\boldsymbol{y}}_j,\ \forall j,\\
&p_j^x + \boldsymbol{w}_j^x \ge \boldsymbol{v}\widehat{\boldsymbol{x}}_j,\ \forall j,\\
&\boldsymbol{w}_j^y, \boldsymbol{w}_j^x, p_j^y, p_j^x,\ \boldsymbol{v},\ \boldsymbol{u} \ge 0,\ \forall j.
\end{aligned} \tag{6}$$

The purpose of model (6) is to measure the robust efficiency score (θ_R) of DMU$_o$ by incorporating decision [auxiliary protection] variables and the budget parameters that protect both the objective function and constraints from the effect of data uncertainty. Specifically, the budget parameters Γ_o^x and Γ_o^y define the level of the uncertainty budget, while the variables $\left(p_j^x, p_j^y\right)$ and $\left(\boldsymbol{w}_j^y, \boldsymbol{w}_j^x\right)$, with $\boldsymbol{w}_j^x = \left(w_{1j}^x, \dots, w_{mj}^x\right)$ and $\boldsymbol{w}_j^y = \left(w_{1j}^y, \dots, w_{sj}^y\right)$, capture the sensitivity of the inputs and outputs data to variations in uncertainty levels. Besides, $p_j^x \Gamma_j^x + \mathbf{1}_m^T \boldsymbol{w}_j^x$ and $p_j^y \Gamma_j^y + \mathbf{1}_s^T \boldsymbol{w}_j^y$ (for $j = 1, .., n$) represent the worst-case deviations of uncertain inputs and outputs from their nominal values under the given uncertainty budgets.

Since model (6) is the linear robust counterpart obtained under the budgeted uncertainty framework of Bertsimas and Sim (2004) (as used in the DEA literature; see Arabmaldar et al., 2017; Toloo et al., 2022), its constraints are affine functions of the decision variable vector $\boldsymbol{z} = \left(\boldsymbol{u}, \boldsymbol{v}, p_j^x, p_j^y, \boldsymbol{w}_j^x, \boldsymbol{w}_j^y\right)$ for all j. Because the budget parameters Γ are fixed exogenous parameters, the additional protection terms introduced by the Bertsimas–Sim reformulation (e.g. $p_j^x \Gamma_j^x + \mathbf{1}^\top \boldsymbol{w}_j^x$) are linear in the auxiliary variables p and $\boldsymbol{w}$. Therefore, the feasible set for $\boldsymbol{z}$ can be written in the form $A\boldsymbol{z} \le b$, $\boldsymbol{z} \ge 0$ i.e. as a finite intersection of closed half-spaces and hence a polyhedron. In finite-dimensional spaces polyhedra are closed and convex, so the feasible region of model (6) is a convex polyhedron. The corresponding robust PPS is the projection of this polyhedron onto the input–output coordinates; projections of polyhedra remain polyhedral (and therefore closed and convex), so the standard convexity properties of the CCR technology are preserved under the budgeted uncertainty set (see e.g. Podinovski et al., 2016; Podinovski & Bouzdine-Chameeva, 2019; Mehdiloo et al., 2026). This establishes that the robust CCR technology obtained through the Bertsimas–Sim budgeted model remains convex and closed.

While the Bertsimas & Sim-based robust DEA approach suggests a powerful mechanism for incorporating adjustable uncertainty budgets into efficiency analysis, its application to real-world datasets, particularly those with heterogeneous uncertainty profiles, may result in overly conservative efficiency estimates or increased computational complexity. To address these challenges and obtain more adaptive robustness, the next section introduces two novel robust multiplier DEA models that extend and refine the existing formulations in the related literature. These models are developed to enhance flexibility, improve ranking stability, and deliver more practical robustness for performance evaluation under uncertainty. In this study, we adopt the standard technology postulates commonly used to characterise DEA production sets under uncertainty (see Arabmaldar et al., 2023): inclusion of observations (Q1), free disposability (Q2), convexity (Q3), and ray unboundedness (Q4).[2] In CCR-type settings, Q3–Q4 jointly imply a conical

[2] Q1 (inclusion of observations) requires that all observed DMUs belong to the PPS; Q2 (free disposability) allows inputs to be increased and outputs to be decreased without violating feasibility; Q3 (convexity) permits convex combinations of feasible activities; and Q4 (ray unboundedness) ensures scalability of activities under constant returns to scale. These postulates are standard in CCR-type DEA technologies.

technology (constant returns to scale), and when the technology admits a polyhedral (half-space) representation, the resulting robust extension is a closed convex cone. Because model (6) yields a polyhedral feasible set in the decision variables, these postulates remain satisfied by the robust CCR technology, and hence the usual conical and convexity properties of the CCR technology carry over under the budgeted uncertainty specification (Kuosmanen & Podinovski, 2009; Podinovski & Bouzdine-Chameeva, 2019; Mehdiloo et al., 2026).

3. The proposed models

This section presents two novel robust multiplier DEA models to improve efficiency analysis under uncertainty. The first model, introduced in SubSection 3.1, employs a correlated budgeted uncertainty set, which captures dependencies among uncertain parameters to better reflect real-world conditions. The second model, presented in SubSection 3.2, utilises an order statistics uncertainty set to improve robustness by focusing on worst-case scenarios derived from ranked data. A theoretical analysis is provided to establish the key properties and advantages of both models. Finally, SubSection 3.3 presents a computational study that benchmarks the proposed models against the traditional robust DEA model based on a conventional budgeted uncertainty set.

3.1. Correlated budgeted uncertainty set

In the existing literature, data uncertainties have often been treated as independent. However, in many real-world situations, multiple parameters are simultaneously affected by a limited number of underlying sources of uncertainty, resulting in correlated data. To better capture this inherent structure, the present study adopts the redesigned budgeted uncertainty set introduced by Bertsimas and Sim (2004) to reflect areas where data points are concentrated, rather than sparsely populated regions. Specifically, due to correlations among the coefficients, random points tend to align along the diagonals.

Fig. 1 illustrates three types of uncertainty sets—interval, budgeted, and correlated budgeted—in a two-dimensional space of uncertain coefficients (η_1, η_2), where each axis corresponds to one uncertain parameter. The interval set (solid line) allows each parameter (η_1 and η_2) to vary independently within the range $[-1,1]$, forming a rectangular region. The budgeted set (dotted) restricts the total value of simultaneous deviations (here $\Gamma = 1$), so the admissible region shrinks toward the centre and, in two dimensions, takes the form of a diamond nested inside the square. The correlated budgeted set (dashed) accounts for dependencies among uncertain parameters (η_1 and η_2); its boundary curves toward the diagonal directions, representing typical patterns of co-movement as parameters are correlated—so the budget is spent on combinations (η_1,η_2) rather than on the axes individually. Furthermore, the extent of boundary curvature responds to the strength of correlation: higher correlation levels lead to greater curvature and stronger concentration of the uncertainty set around the diagonals; with zero

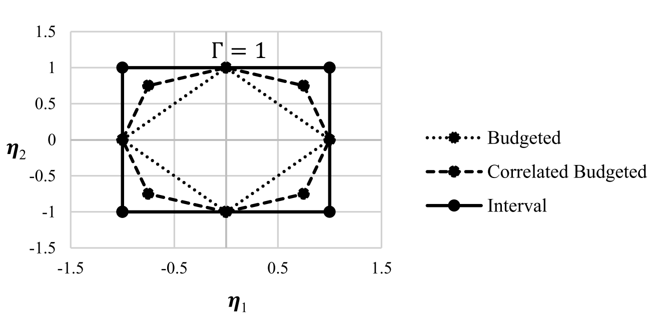


Fig. 1. Correlated versus standard interval and budgeted uncertainty sets.

correlation, it reverts to the classical budgeted polygon, and with very high correlation, most admissible points lie close to the diagonals. In summary, as observed in Fig. 1, the curvature visible in the correlated budgeted set is the visual signature of allocating the uncertainty budget to joint movements instead of independent ones, causing the set to align more closely with high-density regions of the data cloud (typically near the diagonals when variables co-move). These representations demonstrate different approaches to modelling uncertainty in robust optimisation frameworks and how the structure of the uncertainty set changes under different assumptions. The correlated budgeted set also improves coverage around dense data regions while avoiding overextension into void areas.

To formally account for the correlation structure among uncertain parameters, consider a particular constraint j, where I_j and R_j represent the sets of input and output coefficients within constraint j that are subject to uncertainty. Each uncertain element $\widetilde{\boldsymbol{x}}_j$ and $\widetilde{\boldsymbol{y}}_j$ for all $j \in J$ is modeled as: $\widetilde{x}_j = \boldsymbol{x}_j + \boldsymbol{\eta}_j^x \sum_{k^x \in K_j^x} g_{ik^x}$ $(\forall i, j)$ and $\widetilde{\boldsymbol{y}}_j = \boldsymbol{y}_j + \boldsymbol{\eta}_j^y \sum_{k^y \in K_j^y} g_{rk^y}$ $(\forall r, j)$, where $k_j = k_j^x \cup k_j^y$ with $k^x \in k_j^x$ and $k^y \in k_j^y$. The random variables $\boldsymbol{\eta}_{k^x}^x$, $\boldsymbol{\eta}_{k^y}^y$ are assumed to be independent and symmetrically distributed and bounded within the interval $[-1,1]$. The terms g_{ik^x} and g_{rk^y} are the coefficients that represent the influence of uncertainty sources k^x and k^y on the uncertain constraint j associated with DMU$_j$.

Within this framework, there are only $|k_j| = |k_j^x \cup k_j^y|$ distinct sources of uncertainty influencing the data associated with constraint j. These shared uncertainty sources simultaneously impact all inputs x_{ij}, $i \in I_j$ and all outputs y_{rj}, $r \in R_j$. For instance, if $|k_j^x| = 1$, then all input data in a given constraint j are affected by a single random variable. A practical illustration arises in the performance evaluation of bank branches, where both performance variables—with inputs such as personnel expenses, operational costs, and total assets, and outputs including number of processed transactions, operating income, or loan disbursements—are often influenced by macroeconomic conditions. In such settings, external factors—such as interest rate fluctuations or nationwide economic downturns—are common sources of uncertainty that simultaneously affect inputs and outputs of multiple bank branches and their associated performance metrics. For instance, consider a nationwide change in monetary policy—such as a central bank increasing interest rates. This policy shift could simultaneously affect the volume of customer deposits (an output), the cost of capital (an input), and even the demand for loans (another output) across all branches, albeit to varying degrees. Building on the logic of the protection function introduced in the protection function (5), we now propose the following robust formulation to incorporate this structured, correlated uncertainty into the DEA model:

$$\beta\left(\boldsymbol{u}, \boldsymbol{v}, \Gamma_j^x + \Gamma_j^y\right) = \max_{\tilde{\mathscr{S}}_j}\{\xi\} \tag{7}$$

where $\xi = \sum_{k \in S_j} \left|\sum_{r \in R_j} g_{rk^y} u_r\right| + \left(\Gamma_j^y - \lfloor \Gamma_j^y \rfloor\right)\left|\sum_{r \in R_j} g_{t_j r} u_r\right| + \sum_{i \in \overline{S}_j} \left|\sum_{i \in I_j} g_{ik^x} v_i\right| + \left(\Gamma_j^x - \lfloor \Gamma_j^x \rfloor\right)\left|\sum_{i \in I_j} g_{t_j i} v_i\right|$ and $\widetilde{\mathscr{S}}_j$ denotes all combinations of subsets of the input and output uncertainty index sets K_j^x and K_j^y, defined as: $\widetilde{\mathscr{S}}_j = \left\{\begin{matrix} S_j \cup \{t_j\} \,|\, S_j \subseteq K_j^x, |S_j| = \lfloor \Gamma_j^y \rfloor, t_j \in K_j^x \backslash S_j \\ \overline{S}_j \cup \{\bar{t}_j\} \,|\, \overline{S}_j \subseteq K_j^y, |\overline{S}_j| = \lfloor \Gamma_j^x \rfloor, \bar{t}_j \in K_j^y \backslash \overline{S}_j \end{matrix}\right\}$.

This construction reflects the joint budgeted uncertainty affecting both inputs and outputs, enabling the model to capture partial realisations through t_j and $\bar{t}_j$ in a computationally manageable way. The interpretation of the budget parameters Γ_j^x and Γ_j^y remains conceptually consistent with their definitions in Section 2. Specifically, $\Gamma_j^x(\Gamma_j^y)$ can be continuously adjusted within the interval $[0, k_j^x]([0, k_j^y])$, thereby controlling the degree of conservatism in the model. This adjustment ensures robustness by allowing $\lfloor \Gamma_j^y \rfloor(\lfloor \Gamma_j^x \rfloor)$ of the associated random variables $\boldsymbol{\eta}_j^x(\boldsymbol{\eta}_j^y)$ in the protection function (5) to vary fully within the range $[-1, 1]$, while at most one variable is permitted to deviate proportionally by the fractional part, i.e., $\Gamma_j^x - \lfloor \Gamma_j^x \rfloor \left(\Gamma_j^y - \lfloor \Gamma_j^y \rfloor \right)$. Unlike the protection function in (5), which perturbs each input and output independently, the correlated model represents deviations along data-driven factor directions, yielding a more realistic uncertainty representation for robust efficiency analysis. We estimate these factor directions using principal component analysis (PCA).[3] The PCA-derived impact coefficients g_{ik^x} and g_{rk^y} weight each input and output by their corresponding factor loading, indicating how strongly a retained factor can move that variable in the worst case. Here, a loading is the coefficient of a variable in a principal component (after standardisation). It measures the contribution of that variable to the component, where larger absolute values imply stronger co-movement, and the sign reflects the direction of this association. Consequently, the protection term aggregates worst-case co-movements among variables that load on the same factors, and the uncertainty budget Γ allocates protection across these factor directions. In practice, variables are standardised before PCA, and the impact coefficients (g_{ik^x} and g_{rk^y}) are derived from the absolute (and optionally normalised) loadings to ensure that protection reflects relative factor sensitivity rather than measurement scale.

As with the earlier model, the proposed formulation can be equivalently reformulated as the following robust model:

$$
\begin{aligned}
&\theta_R = \max\ \boldsymbol{u}\boldsymbol{y}_o - p_o^y\Gamma_o^y - 1_{|K_r^y|}\boldsymbol{w}_o^y \\
&s.t. \\
&\boldsymbol{v}\boldsymbol{x}_o + p_o^x\Gamma_o^x + 1_{|K_i^x|}\boldsymbol{w}_o^x \le 1, \\
&\boldsymbol{u}\boldsymbol{y}_o - \boldsymbol{v}\boldsymbol{x}_o - p_o^x\Gamma_o^x - p_o^y\Gamma_o^y - 1_{|K_o^y|}\boldsymbol{w}_o^y - 1_{|K_o^x|}\boldsymbol{w}_o^x \le 0, \\
&\boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j + p_j^x\Gamma_j^x + p_j^y\Gamma_j^y + 1_{|K_j^y|}\boldsymbol{w}_j^y + 1_{|K_j^x|}\boldsymbol{w}_j^x \le 0, && \forall j \neq o, \\
&p_j^y + \boldsymbol{w}_j^y \ge \sum_{r\in R_j} g_{rk^y}\boldsymbol{u}, && \forall j,\ \forall k^y \in K_j^y, \\
&p_j^x + \boldsymbol{w}_j^x \ge \sum_{i\in I_j} g_{ik^x}\boldsymbol{v}, && \forall j,\ \forall k^x \in K_j^x, \\
&\boldsymbol{w}_j^y, \boldsymbol{w}_j^x, p_j^y, p_j^x,\ \boldsymbol{v},\ \boldsymbol{u} \ge 0, && \forall i,\ r,\ j.
\end{aligned}
\tag{8}
$$

In this model, the uncertainty sets for inputs and outputs are formulated by incorporating the estimated impact factors g_{ik^x} and g_{rk^y}. Each impact factor is positive when the corresponding uncertain input k_j^x and/or uncertain output k_j^y significantly influences the efficiency of DMU$_j$; otherwise, g_{ik^x} and g_{rk^y} are set to zero. By aggregating these factors, we focus attention on the most influential sources of variability (uncertainty) for each DMU. Dependence is therefore handled directly on the multiplier side. Correlated perturbations are projected onto data-estimated impact directions and act on the weights (u, v) without altering the PPS or the convexity of the technology (see Lemma 1). The uncertainty budget Γ continues to control the level of protection within a single linear programming (LP) model, making the efficiency–price-of-robustness trade-off observable in practice through ranking stability. The construction also permits DMU-specific uncertainty sets while preserving a common PPS.

Lemma 1. *For fixed robustness budgets Γ_j^x and Γ_j^y, and fixed PCA-based coefficients $g_{ik}^x, g_{rk}^y \ge 0$, the correlated budgeted robust DEA model (8) is an LP problem in the variables $\left(\boldsymbol{u}, \boldsymbol{v}, p_j^x, p_j^y, \boldsymbol{w}_j^x, \boldsymbol{w}_j^y\right)$ for all j. Its feasible region is a non-empty, closed, convex polyhedron, and the associated robust PPS is closed and convex.*

[3] PCA is a statistical technique used to reduce the dimensionality of a dataset while retaining the most significant variance in the data. PCA transforms correlated variables into a new set of uncorrelated principal components, which helps identify the most influential patterns in the data. This technique is particularly useful for improving data interpretability and eliminating redundancy in complex datasets.

Proof. For fixed parameters Γ_j^x and Γ_j^y, and $g_{ik}^x, g_{rk}^y \ge 0$, all coefficients in model (8) are constant. The objective $\theta_R = \boldsymbol{u}\boldsymbol{y}_o - p_o^y\Gamma_o^y - \mathbf{1}_{|K_r^y|}\boldsymbol{w}_o^y$ is linear in $\left(\boldsymbol{u}, p_o^y, \boldsymbol{w}_o^y\right)$, and every constraint in model (8)—the normalisation, the robust inequalities for DMU$_o$ and for $j \neq o$, the link constraints $p_j^y + \boldsymbol{w}_j^y \ge \sum_{r\in R_j} g_{rk^y}\boldsymbol{u}$, and $p_j^x + \boldsymbol{w}_j^x \ge \sum_{i\in I_j} g_{ik^x}\boldsymbol{v}$, and the nonnegativity conditions—is affine in $\left(\boldsymbol{u}, \boldsymbol{v}, p_j^x, p_j^y, \boldsymbol{w}_j^x, \boldsymbol{w}_j^y\right)$ for all j. Hence, the feasible set is the intersection of finitely many closed half-spaces and can be written in standard LP form as $\{z : Az \le b, z \ge 0\}$, i.e., a closed convex polyhedron. Non-emptiness is immediate; e.g., the feasible solution $\boldsymbol{u} = \mathbf{0}_s$, $\boldsymbol{v} = \mathbf{0}_m$ and $p_j^x = p_j^y = 0$, $\boldsymbol{w}_j^x = \mathbf{0}_{|K_j^x|}$, $\boldsymbol{w}_j^y = \mathbf{0}_{|K_j^y|}$ for all j, satisfies the normalisation and all other inequalities. This demonstrates that the feasible set is non-empty; the normalisation constraint and maximisation of θ_R ensure that efficiency scores are non-trivial and meaningful in practice. The convexity of the robust technology follows from the supporting-half-space representation where the feasible multiplier vectors define linear inequalities of the form $\boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j \le 0$, and the intersection of these half-spaces is closed and convex. With ray unboundedness and free disposability (Q4–Q2), the robust PPS forms a closed convex cone. Furthermore, the robust PPS is obtained by projecting the feasible polyhedral set onto the input–output space. Since projections of polyhedra preserve polyhedrality, the robust PPS maintains polyhedrality and convexity. ∎

The uncertainty set is constructed analogously to the uncorrelated case, but the deviations are now defined using the aggregated impact factors. Specifically, the aggregated impact widths are $\widehat{\boldsymbol{x}}_j = \sum_{k^x \in K_j^x} g_{ik^x}$ and $\widehat{\boldsymbol{y}}_j = \sum_{k^y \in K_j^y} g_{rk^y}$. As a result, the uncertain inputs $\widetilde{\boldsymbol{x}}_j$ and outputs $\widetilde{\boldsymbol{y}}_j$ lie within symmetric intervals as $\left[\boldsymbol{x}_j - \sum_{k^x \in K_j^x} g_{ik^x}, \boldsymbol{x}_j + \sum_{k^x \in K_j^x} g_{ik^x}\right]$ and $\left[\boldsymbol{y}_j - \sum_{k^y \in K_j^y} g_{rk^y}, \boldsymbol{y}_j + \sum_{k^y \in K_j^y} g_{rk^y}\right]$, respectively. These intervals capture the range of potential data fluctuations, where the widths are directly influenced by the estimated significance of each uncertainty source. The uncorrelated model introduced in model (6) remains applicable under this structure, where the uncertainty bounds are tailored to reflect the values of the contributing factors k^x and k^y for each DMU$_j$.

As mentioned earlier, to estimate the impact factors g_{ik^x} and g_{rk^y} used in this study, we employ PCA (Jackson, 1991) to identify the primary sources of variability within the dataset. We perform correlation-matrix PCA on standardised inputs and on standardised outputs separately to keep interpretation clear and correlated subsets compact. Let $L^x \in \mathbb{R}^{h_x \times |I|}$ and $L^y \in \mathbb{R}^{h_y \times |R|}$ denote the retained loading matrices, with eigenvalues $\{\lambda_{k^x}\}_{k^x=1}^{h_x}$, $\{\lambda_{k^y}\}_{k^y=1}^{h_y}$ (variance explained). We translate component weights into impact coefficients through $g_{ik^x} = |L_{ik^x}|\sqrt{\lambda_{k^x}}$ and $g_{rk^y} = |L_{rk^y}|\sqrt{\lambda_{k^y}}$, where absolute values remove arbitrary sign. We apply light sparsification by setting g_{ik^x} (or g_{rk^y}) to zero whenever $|L^x|$ (or $|L^y|$) is below a small threshold τ. The uncertainty widths are then computed as $\widehat{\boldsymbol{x}}_j = \sum_{k^x=1}^{h_x} g_{ik^x}$ and $\widehat{\boldsymbol{y}}_j = \sum_{k^y=1}^{h_y} g_{rk^y}$, yielding symmetric intervals $\left[\boldsymbol{x}_j \pm \widehat{\boldsymbol{x}}_j\right]$ and $\left[\boldsymbol{y}_j \pm \widehat{\boldsymbol{y}}_j\right]$.

The per-variable input and output uncertainty contribution implied by the retained components is measured as the standard-deviation-like quantity $\sigma_i^x = \sqrt{\sum_{k^x=1}^{h_x} \left(g_{ik^x}\right)^2}$ and $\sigma_r^y = \sqrt{\sum_{k^y=1}^{h_y} \left(g_{rk^y}\right)^2}$. We then set the uncertainty half-widths as $\widehat{\boldsymbol{x}}_j = \kappa\sigma_i^x$ and $\widehat{\boldsymbol{y}}_j = \kappa\sigma_r^y$ (with κ a chosen scaling factor, e.g. $\kappa = 1$). Accordingly, for each DMU$_j$ the uncertain input and output variables lie in symmetric intervals; $\widehat{\boldsymbol{x}}_j \in \left[\boldsymbol{x}_j \pm \widehat{\boldsymbol{x}}_j\right]$ and $\widehat{\boldsymbol{y}}_j \in \left[\boldsymbol{y}_j \pm \widehat{\boldsymbol{y}}_j\right]$. The same coefficients appear in the correlated protection terms. Correlated shocks act along component directions, producing deviations of the form $\sum_k g_{ik^x}\xi_k$ for inputs and $\sum_k g_{rk^y}\zeta_k$ for outputs (with ξ_k and ζ_k the bounded component shocks). Thus, the uncertainty budget allocates protection across empirically estimated co-movements rather than treating each variable independently.

In the baseline, we retain the smallest h_x and h_y that explain at least

90% of total variance and choose τ to remove very small loadings while preserving interpretability; qualitative conclusions are robust to varying the variance cut-off within 80–95%, modest changes in τ, and an optional varimax rotation. A combined inputs-and-outputs PCA yields similar leading directions but is not used in the baseline to keep correlated subsets compact. This PCA-based construction provides insights into the contributions of various inputs and outputs to overall variability, which is essential for constructing representative uncertainty sets and improving discrimination in efficiency analysis (Adler & Yazhemsky, 2010; Premachandra, 2001; Hatami-Marbini, Hekmat & Agrell, 2022).

Based on the PCA results, the bounds for each input i and output r are determined using the standard deviations associated with the most influential principal components. This leads to the formation of uncertainty intervals for the i^{th} input and r^{th} output as $[\bar{x}_j - \sqrt{\sum g_{ik^x}}, \bar{x}_j + \sqrt{\sum g_{ik^x}}]$ and $[\bar{y}_j - \sqrt{\sum g_{rk^y}}, \bar{y}_j + \sqrt{\sum g_{rk^y}}]$, which reflect the underlying data structure and variability patterns. By systematically identifying key sources of variability through PCA, the robust optimisation framework ensures reliable efficiency evaluations under uncertainty, with the derived g_{ik^x} and g_{rk^y} as being adjustment parameters that refine the uncertainty sets and enhance the model robustness and precision.

3.2. Order statistics uncertainty

This subsection introduces a new robust DEA model that incorporates the order statistics uncertainty set developed by Zhang and Gupta (2023), with the aim of minimising the cost associated with uncertainty. For simplicity, we omit the input and output constraint indices and focus on a representative constraint throughout the discussion.

Let $\boldsymbol{\eta}_j^x = (\eta_{1j}^x, \dots, \eta_{mj}^x)$ and $\boldsymbol{\eta}_j^y = (\eta_{1j}^y, \dots, \eta_{sj}^y)$ represent continuous, independently distributed random variables over the interval $[0,1]$, each following an arbitrary continuous distribution with unknown cumulative distribution functions $\mathscr{F}_j^x$ and $\mathscr{F}_j^y$. Let $U_j^x = \mathscr{F}_j^x(\boldsymbol{\eta}_j^x)$ and $U_j^y = \mathscr{F}_j^y(\boldsymbol{\eta}_j^y)$ denote random variables for each j, where U_j^x and U_j^y are uniformly distributed on $[0,1]$.

The corresponding order statistics are represented by $U_{(1)}^x, \cdots, U_{(|I_j|)}^x$ and $U_{(1)}^y, \cdots, U_{(|R_j|)}^y$ obtained by sorting the realisations of U_j^x and U_j^y in ascending order. Unlike the original random variables, the k^{th} order statistic $U_{(k)}$ follows a Beta distribution with parameters $(k,\ |J^x| + 1 - k)$. Let $I_t(k,\ |J^x| + 1 - k)$ denote the cumulative distribution function (CDF) of this Beta distribution, and let Q_k^t represent the corresponding quantile function, defined as $Q_k^t = \inf\{\tau :\ I_\tau(k,\ |J^x| + 1 - k) = t\}$. We then define the following uncertainty sets for the input and output variables based on these order statistics (see Zhang & Gupta, 2023, p. 1026 for further details):

$$\mathscr{U}_j^{x^{OS}}(\boldsymbol{\varepsilon}) = \left\{\boldsymbol{\eta}_j^x \mid \mathscr{F}_j^x(\boldsymbol{\eta}_j^x) = U_j^x,\ \forall j \in J,\ \text{and}\ U_{(k^x)}^x \le Q_{k^x}^{(1-\varepsilon_{k^x})},\ \forall k^x \in I_j\right\},$$

$$\mathscr{U}_j^{y^{OS}}(\boldsymbol{\varepsilon}) = \left\{\boldsymbol{\eta}_j^y \mid \mathscr{F}_j^y(\boldsymbol{\eta}_j^y) = U_j^y,\ \forall j \in J,\ \text{and}\ U_{(k^y)}^y \le Q_{k^y}^{(1-\varepsilon_{k^y})},\ \forall k^y \in R_j\right\} \tag{9}$$

where $Q_{k^x}^{(1-\varepsilon_{k^x})}$ represents the upper limit for the (k^x)-th order statistics, satisfying $\text{Prob}\left(U_{(k^x)}^x \le Q_{k^x}^{(1-\varepsilon_{k^x})}\right) = 1 - \varepsilon_{k^x}$, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{|J|})$ and each $\varepsilon_j \in [0,1]$. A similar structure is used to define the output uncertainty set $\mathscr{U}_j^{y^{OS}}$, where the output values are constrained by their respective order statistics in a manner analogous to the inputs.

As highlighted by Zhang and Gupta (2023), the direct use of the order statistics uncertainty set defined in (9) leads to intractable robust optimisation problems. To address this challenge, they proposed an assignment-based reformulation strategy that enables a tractable representation of the protection functions associated with these uncertainty sets. In line with their methodology, this study adopts the assignment-based approach to derive a computationally feasible formulation for the order statistics-based uncertainty set, which can be used for both inputs and outputs in the development of a new robust DEA model.

The order statistics uncertainty set shares conceptual similarities with the interval and budgeted uncertainty sets, although it is driven by distinct statistical principles. Fig. 2 illustrates this relationship using a numerical example with $|j| = 10$. The x-axis represents the index k associated with the order of the statistics, while the y-axis shows the corresponding $z(k)$ values. Under interval uncertainty (solid line), all optimal Z_j values are equal to 1. In the case of the budgeted uncertainty set (dashed line), at most one Z_j can take a fractional value between 0 and 1, while the remaining values are either 0 or 1. In contrast, the order statistics uncertainty set (dotted line) allows the Z_j values to vary gradually and take fractional values based on quantiles that correspond to the specified uncertainty level, reflecting the flexible and data-driven nature of this set. Unlike the rigid structures of the interval and budgeted sets, the order statistics set results in significant variation among the Z_j values and allows for up to $|J|!$ distinct solutions, corresponding to all possible permutations of indices. Furthermore, the geometric structure of the order statistics uncertainty set can be flexibly modulated through the adjustment of up to $|J|^2$ parameters.

Building on Proposition 2.3 of Zhang and Gupta (2023), we derive a tractable formulation for the input and output protection functions, denoted as $\beta_j(\boldsymbol{u}^*, \boldsymbol{v}^*, \mathscr{U}_j^{OS}(\boldsymbol{\varepsilon}))$, which quantifies the worst-case deviation under the order statistics uncertainty set. Specifically, let the input protection function $q_{ik^x}^x$ represent the quantile of $\boldsymbol{\eta}_j^x$ associated with the quantile of order $Q_{k^x}^{(1-\varepsilon_{k^x})}$, defined as $q_{ik^x}^x = \inf\left\{\boldsymbol{v} : F_j^x(\boldsymbol{v}) \ge Q_{k^x}^{(1-\varepsilon_{k^x})}\right\}$ for all $i, k^x \in I_j$ and $j \in J$. An analogous definition applies to the output protection function given by $q_{rk^y}^y = \inf\left\{\boldsymbol{u} : F_j^y(\boldsymbol{u}) \ge Q_{k^y}^{(1-\varepsilon_{k^y})}\right\}$ for all r, $k^y \in R_j$ and $j \in J$.

The following proposition formalises the relationship between the protection function $\beta_j(\boldsymbol{u}^*, \boldsymbol{v}^*, \mathscr{U}_j^{OS}(\boldsymbol{\varepsilon}))$ under the order statistics uncertainty set and its computational tractability. It shows that, for a given pair $(\boldsymbol{v}, \boldsymbol{u})$, the optimal objective value of the protection function can be obtained by solving linear optimisation problems. This provides a computationally tractable approach to evaluate the protection function within the robust DEA framework. In essence, the order statistics protection is computed by assigning each component to a quantile weight and selecting the worst-case combination. The linear assignment problems in models (10) and (11) implement this mechanism directly, yielding the protection value used in the robust constraints.

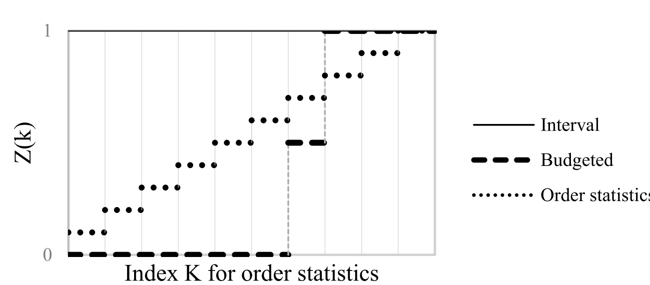


Fig. 2. Comparison of Z_j order statistics under different uncertainty sets.

Proposition 1. *Given a pair $(\boldsymbol{v},\boldsymbol{u})$, the optimal objective value of the protection function $\beta(\boldsymbol{u}^*,\boldsymbol{v}^*,\mathscr{U}_j^{OS}(\boldsymbol{\varepsilon}))$ associated with uncertain inputs and outputs, can be obtained by solving the following linear optimisation problems:*

$$\begin{aligned}
&\max_{\eta}\sum_{r\in R_j}\widehat{\boldsymbol{y}}_j|\boldsymbol{u}|.\left(\sum_{k^y\in R_j}q^y_{rk^y}z^y_{rk^y}\right)\\
&\text{s.t.}\\
&\sum_{k^y}z^y_{rk^y}=1, && \forall r\in R_j,\\
&\sum_{r}z^y_{rk_r}=1, && \forall k^y\in R_j,\\
&z^y_{rk^y}\geq 0, && \forall r,k^y\in R_j.
\end{aligned}\tag{10}$$

and

$$\begin{aligned}
&\max_{\eta}\sum_{i\in I_j}\widehat{\boldsymbol{x}}_j|\boldsymbol{v}|.\left(\sum_{k^x\in I_j}q^x_{ik^x}z^x_{ik^x}\right)\\
&\text{s.t.}\\
&\sum_{k^x}z^x_{ik^x}=1, && \forall i\in I_j,\\
&\sum_{i}z^x_{ik^x}=1, && \forall k^x\in I_j,\\
&z^x_{ik^x}\geq 0, && \forall i,k^x\in I_j.
\end{aligned}\tag{11}$$

Proof. The proof closely follows Zhang and Gupta (2023, Proposition 2.3) and is therefore omitted. ∎

The assignment-based problems formulated above yield the optimal values of the output-side and input-side protection functions, respectively. Specifically, solving model (10) provides the value of $\beta_j(\boldsymbol{u}^*,\boldsymbol{0}_m,\mathscr{U}^{OS}(\boldsymbol{q}^y))$ corresponding to output-side uncertainty, while solving model (11) yields $\beta_j(\boldsymbol{0}_s,\boldsymbol{v}^*,\mathscr{U}^{OS}(\boldsymbol{q}^x))$ capturing input-side uncertainty. When uncertainty is present in both inputs and outputs, the full protection function $\beta_j(\boldsymbol{u}^*,\boldsymbol{v}^*,\mathscr{U}^{OS}(\boldsymbol{q}))$ is computed as the sum of these two components, reflecting the additive nature of input- and output-side protection within the robust DEA framework. Henceforth, the resulting protection functions are denoted by $\beta_j(\boldsymbol{u}^*,\boldsymbol{v}^*,\mathscr{U}_i^{OS}(\boldsymbol{q}))$, where $q=(q^x,q^y)$ represents the quantile-based uncertainty across all inputs and outputs.

By incorporating the defined uncertainty set $\mathscr{U}^{OS}(\boldsymbol{\varepsilon})$ into the classical CCR model (2), we arrive at the following robust counterpart:

$$\begin{aligned}
&\theta_o^{OS}=\max\ \boldsymbol{u}\boldsymbol{y}_o+\beta_o(\boldsymbol{u}^*,\boldsymbol{0}_m,\mathscr{U}^{OS}(\boldsymbol{q}^y))\\
&\text{s.t.}\\
&\boldsymbol{v}\boldsymbol{x}_o+\beta_o(\boldsymbol{0}_s,\boldsymbol{v}^*,\mathscr{U}^{OS}(\boldsymbol{q}^x))\leq 1,\\
&\boldsymbol{u}\boldsymbol{y}_o-\boldsymbol{v}\boldsymbol{x}_o-\beta_o(\boldsymbol{u}^*,\boldsymbol{v}^*,\mathscr{U}^{OS}(\boldsymbol{q}))\leq 0,\\
&\boldsymbol{u}\boldsymbol{y}_j+u_0-\boldsymbol{v}\boldsymbol{x}_j+\beta_j(\boldsymbol{u}^*,\boldsymbol{v}^*,\mathscr{U}^{OS}(\boldsymbol{q}))\leq 0,\ \forall j\neq o,\\
&\boldsymbol{v},\ \boldsymbol{u}\geq 0,
\end{aligned}\tag{12}$$

where the protection terms $\beta_j(.)$ for each DMU is directly calculated from models (10) and (11), depending on whether the uncertainty pertains to inputs, outputs, or both. The robust formulation of the CCR model (2) under the order statistics uncertainty sets $\mathscr{U}_i^{OS}(\varepsilon)$ and $\mathscr{U}_r^{OS}(\varepsilon)$ is presented in the following theorem.

Theorem 1. *Model (12) is equivalent to the following LP formulation:*

$$\begin{aligned}
&\theta_o^{OS}=\max\ \boldsymbol{u}\boldsymbol{y}_o-\boldsymbol{1}^T_{|R_o|}(\boldsymbol{\varphi}^y_o+\boldsymbol{\psi}^y_o)\\
&\text{s.t.}\\
&\boldsymbol{v}\boldsymbol{x}_o+\boldsymbol{1}^T_{|I_o|}(\boldsymbol{\varphi}^x_o+\boldsymbol{\psi}^x_o)\leq 1,\\
&\boldsymbol{u}\boldsymbol{y}_o-\boldsymbol{v}\boldsymbol{x}_o-\boldsymbol{1}^T_{|I_o|}(\boldsymbol{\varphi}^x_o+\boldsymbol{\psi}^x_o)-\boldsymbol{1}^T_{|R_o|}(\boldsymbol{\varphi}^y_o+\boldsymbol{\psi}^y_o)\leq 0,\\
&\boldsymbol{u}\boldsymbol{y}_j-\boldsymbol{v}\boldsymbol{x}_j+\boldsymbol{1}^T_{|I_j|}(\boldsymbol{\varphi}^x_j+\boldsymbol{\psi}^x_j)+\boldsymbol{1}^T_{|R_j|}(\boldsymbol{\varphi}^y_j+\boldsymbol{\psi}^y_j)\leq 0, && \forall j\neq o,\\
&\boldsymbol{\varphi}^x_j+\boldsymbol{\psi}^x_j\geq\widehat{\boldsymbol{x}}_j\boldsymbol{v}q^x_{ik^xj}, && \forall i,k^x\in I_j,\ \forall j,\\
&\boldsymbol{\varphi}^y_j+\boldsymbol{\psi}^y_j\geq\widehat{\boldsymbol{y}}_j\boldsymbol{u}q^y_{rk^yj}, && \forall r,k^y\in R_j,\ \forall j,\\
&\boldsymbol{v},\ \boldsymbol{u}\geq 0,\\
&\boldsymbol{\varphi}^x_j,\boldsymbol{\psi}^x_j,\boldsymbol{\varphi}^y_j,\boldsymbol{\psi}^y_j,\ \text{free in sign} && .
\end{aligned}\tag{13}$$

Proof. The derivation relies on the protection functions defined in (5), Proposition 1, and the dual formulations of models (10) and (11), given respectively by:

$$\begin{aligned}
&\min\ \boldsymbol{1}^T_{|R_j|}(\boldsymbol{\varphi}^y_j+\boldsymbol{\psi}^y_j)\\
&\text{s.t.}\\
&\boldsymbol{\varphi}^y_j+\boldsymbol{\psi}^y_j\geq\widehat{\boldsymbol{y}}_j|\boldsymbol{u}|q^y_{rk^yj},\ \forall r,k^y\in R_j,\ \forall j.
\end{aligned}\tag{14}$$

and

$$\begin{aligned}
&\min\ \boldsymbol{1}^T_{|I_j|}(\boldsymbol{\varphi}^x_j+\boldsymbol{\psi}^x_j)\\
&\text{s.t.}\\
&\boldsymbol{\varphi}^x_j+\boldsymbol{\psi}^x_j\geq\widehat{\boldsymbol{x}}_j|\boldsymbol{v}|q^x_{ik^xj},\ \forall i,k^x\in I_j,\ \forall j.
\end{aligned}\tag{15}$$

By applying strong duality and substituting the dual problems (14) and (15) into the robust counterpart model (12), the robust DEA model (13) follows directly. ∎

The next theorem establishes that the proposed robust DEA model (13) remains consistent with the foundational principles of traditional DEA models.

Theorem 2. *(i) Model (13) is guaranteed to be feasible, and (ii) the optimal value satisfies $\theta_o^{OS*}\leq 1$.*

Proof. (i) Feasibility of model (13) can be verified by setting $\boldsymbol{v}=\boldsymbol{0}_m$, $\boldsymbol{u}=\boldsymbol{0}_s$, and $\boldsymbol{\varphi}^x_j=\boldsymbol{\psi}^x_j=\boldsymbol{0}_{|I_j|}$, $\boldsymbol{\varphi}^y_j=\boldsymbol{\psi}^y_j=\boldsymbol{0}_{|R_j|}$ for all $j\in J$. Under this setting, all constraints are satisfied trivially, thereby demonstrating that the model is feasible. (ii) At the optimal solution, the first constraint yields $\boldsymbol{v}^*\boldsymbol{x}_o+\boldsymbol{1}^T_{|I_o|}(\boldsymbol{\varphi}^{x^*}_o+\boldsymbol{\psi}^{x^*}_o)\leq 1$. From the second constraint, we also obtain $\boldsymbol{u}^*\boldsymbol{y}_o-\boldsymbol{1}^T_{|R_o|}(\boldsymbol{\varphi}^{y^*}_o+\boldsymbol{\psi}^{y^*}_o)\leq\boldsymbol{v}^*\boldsymbol{x}_o+\boldsymbol{1}^T_{|I_o|}(\boldsymbol{\varphi}^{x^*}_o+\boldsymbol{\psi}^{x^*}_o)\leq 1$. Thus, by the definition of the objective function, it directly follows that $\theta_o^{OS*}\leq 1$. ∎

The following theorem looks into the equivalence between the conventional robust DEA model (6), which employs a budgeted uncertainty set with budget Γ, and the proposed robust DEA model (13), which is constructed based on an order statistics uncertainty set, under an appropriate definition of the quantile values q_{rk^y} or q_{ik^x}.

Theorem 3. *The conventional robust DEA model (6) and the proposed model (13) are equivalent if the quantile values q_{ik^x} are defined as:*

$$q_{ik^x}=\begin{cases}0, & \text{if } k^x\in\left[1,|I_j|-\lfloor\boldsymbol{\Gamma}^x\rfloor-1\right], & \forall i\in I_j\\ \boldsymbol{\Gamma}^x-\lfloor\boldsymbol{\Gamma}^x\rfloor, & \text{if } k^x=|I_j|-\lfloor\boldsymbol{\Gamma}^x\rfloor, & \forall i\in I_j\\ 1, & \text{if } k^x\in\left[|I_j|-\lfloor\boldsymbol{\Gamma}^x\rfloor+1,|I_j|\right], & \forall i\in I_j.\end{cases}$$

Proof. Without loss of generality, we focus on the input-side protection function and assume a single constraint for both the budgeted and order statistic uncertainty sets. For simplicity, we suppress the subscript *i*.

To prove the equivalence, we compare the objective values of $\beta(\boldsymbol{v}^*,\boldsymbol{u}^*,\mathscr{U}^{BS}(\boldsymbol{\Gamma}^x))$ from model (4) and $\beta(\boldsymbol{v}^*,\boldsymbol{u}^*,\mathscr{U}^{OS}(\boldsymbol{\varepsilon}))$ from model (12).

Recall from earlier results that model (11) represents a linear relaxation of a maximum weight assignment problem, which yields an integer optimal solution. Therefore, for each $i \in I_j$, there exists a unique $k^x \in I_j$ such that $\eta^x_{ik^x} = 1$. In this case, the term $\hat{x}_{ij}|v_i|$ is paired with the corresponding quantile q_{ik^x}. Considering the definitions of $\beta(\boldsymbol{v}^*, \boldsymbol{u}^*, \mathscr{U}^{BS}(\boldsymbol{\Gamma}^x))$ from (4), and $\beta_j(\boldsymbol{v}^*, \boldsymbol{u}^*, \mathscr{U}^{OS}_j(\boldsymbol{q}^x))$ from Proposition 1, we define the following three cases:

(i) If $1 \le k^x \le |I_j| - \lfloor \Gamma^x \rfloor - 1$, then $q_{ik^x} = 0$.
(ii) If $k^x = |I_j| - \lfloor \Gamma^x \rfloor$, then $q_{ik^x} = \Gamma^x - \lfloor \Gamma^x \rfloor$, representing the fractional part of the budget.
(iii) If $|I_j| - \lfloor \Gamma^x \rfloor + 1 \le k^x \le |I_j|$, then $q_{ik^x} = 1$.

Accordingly, among all terms $\hat{x}_{1j}|v_1|,\ \hat{x}_{2j}|v_2|, \cdots,\ \hat{x}_{|I_j|j}|v_{|I_j|}|$, exactly $\lfloor \Gamma^x \rfloor$ are paired with a weight of 1, one with a weight of $\Gamma^x - \lfloor \Gamma^x \rfloor$, and the rest with 0. As a result, the total protection term computed under the order statistics uncertainty set matches exactly that under the budgeted uncertainty set. This proves that the optimal objective values for $\beta(\boldsymbol{v}^*, \boldsymbol{u}^*, \mathscr{U}^{BS}(\boldsymbol{\Gamma}^x))$ and $\beta(\boldsymbol{v}^*, \boldsymbol{u}^*, \mathscr{U}^{OS}(\boldsymbol{q}^x))$ are equivalent. An analogous argument holds for the output side. Therefore, the two models are equivalent under the stated quantile specification. ∎

We implement the order statistics set as a rank-weighted budget that emphasises tail behaviour. The corresponding multiplier counterpart remains linear because the worst-case, rank-weighted penalties for inputs and outputs are computed through assignment-type constraints, and consequently both convexity and the PPS are preserved. This construction therefore provides a DEA-specific mechanism to capture extreme realisations (tail risk) rather than modelling correlation structure.

Lemma 2. *For any fixed quantile parameters q^x and q^y, the feasible region of model (13) in the variables $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{\varphi}^x_j, \boldsymbol{\psi}^x_j, \boldsymbol{\varphi}^y_j, \boldsymbol{\psi}^y_j)$ is a non-empty, closed, and convex polyhedron. In particular, the robust technology associated with the order statistics model (13) is closed and convex, and the optimal value θ^{OS*}_o is attained.*

Proof. By Theorem 1, model (12) admits the LP reformulation given in model (13). For fixed quantile parameters q^x and q^y, all coefficients in (13) are constants. The objective function $\theta^{\text{OS}}_o = \boldsymbol{u}\boldsymbol{y}_o - \mathbf{1}^\top(\boldsymbol{\varphi}^y_o + \boldsymbol{\psi}^y_o)$ is linear in the decision variable vector $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{\varphi}^x, \boldsymbol{\psi}^x, \boldsymbol{\varphi}^y, \boldsymbol{\psi}^y)$, and every constraint of (13)— the normalisation, the robust inequalities (for DMU$_o$ and for $j \neq o$), the link constraints $\boldsymbol{\varphi}^x_j + \boldsymbol{\psi}^x_j \ge \hat{\boldsymbol{x}}_j \boldsymbol{v} q^x_{ik^x j}, \boldsymbol{\varphi}^y_j + \boldsymbol{\psi}^y_j \ge \hat{\boldsymbol{y}}_j \boldsymbol{u} q^y_{rk^y j}$, and the sign constraints— is affine. Hence, the feasible set is the intersection of finitely many closed half-spaces and can be written in the standard LP form as $\{z : Az \le b, z \ge 0\}$, i.e., a closed convex polyhedron. Theorem 2 guarantees that this polyhedral region is non-empty and that the objective function value is bounded above (by 1). Therefore, the linear maximisation problem (13) attains its optimum $\theta^{\text{OS}*}_o$ on this feasible set. Finally, the robust PPS is the projection of this polyhedral feasible set onto the input–output coordinates; projections of polyhedra preserve polyhedrality, closedness, and convexity. Consequently, the robust technology corresponding to model (13) is closed and convex. ∎

Building on robust DEA models developed under the budgeted uncertainty set, this section extends the framework in two key directions. First, we introduce a *correlated budgeted uncertainty set* in model (8) to address the limitation of model (6), which—despite suggesting a tractable and intuitive approach to uncertainty—assumes independence among uncertain parameters. Incorporating correlations allows model (8) to present a more realistic representation of uncertainty. Second, model (13) introduces an *order statistics uncertainty set*, which enables a more detailed characterisation of uncertainty by modelling the distributional behaviour of extreme values. Together, these extensions enhance both realism and robustness in performance evaluation under uncertainty.

More importantly, convexity is preserved across all three robust DEA models. For the budgeted model (6), the uncertainty set is polyhedral, and the robust counterpart is the Bertsimas & Sim–based LP problem, hence the feasible weight region (multiplier feasible set) is a convex polyhedron and the CCR technology remains a convex cone. For the correlated budgeted model (8), Lemma 1 shows that the PCA-based protection terms enter linearly and again yield a convex polyhedral feasible set. For the order statistics model (13), Theorem 1 and Proposition 1 establish an LP reformulation through an assignment-based approach, Lemma 2 then shows that the feasible region is a non-empty convex polyhedron, and that the resulting robust technology (PPS) forms a tightening of the original CCR cone rather than a non-convex deformation. In all cases, the robust constraints remain linear (polyhedral), hence the resulting robust technologies remain within the class of closed convex (polyhedral) DEA technologies. This aligns with the established convex-polyhedral structure of DEA frontiers (Podinovski, 2005; Podinovski et al., 2016; Podinovski & Bouzdine-Chameeva, 2019; Mehdiloo et al., 2026).

We emphasise that both models (8) and (13) are formulated as LP problems with respect to the decision variables u and v and associated auxiliary variables $(\boldsymbol{w}^x_j, \boldsymbol{w}^y_j, p^x_j, p^y_j)$ for model (8) and $(\boldsymbol{\varphi}^x_j, \boldsymbol{\psi}^x_j,\ \boldsymbol{\varphi}^y_j, \boldsymbol{\psi}^y_j)$ for model (13), provided that the parameters defining the uncertainty sets (e.g., robust parameters, Γ^x_j and Γ^y_j, impact factors, g_{ik^x} and g_{rk^y}, quantile values q_{ik^x} and q_{rk^y} and perturbation levels e_i and e_r) are specified exogenously and not treated as decision variables. This approach follows the robust optimisation literature (e.g., Bertsimas & Sim, 2004), which defines uncertainty sets prior to optimisation. Because uncertainty is encoded in the set specification rather than in the optimisation variables, the models remain linear and tractable, allowing them to be solved using standard LP solvers without introducing nonlinearities into the DEA framework.

While both models (8) and (13) build on the budgeted uncertainty framework, their focus and strengths differ. The correlated budgeted model (8) improves realism and computational efficiency by capturing interdependencies among uncertain parameters, making it well suited for large datasets and routine operational applications. The order statistics model (13) strengthens robustness by incorporating distributional information about extreme values, providing stability in efficiency rankings under severe uncertainty and supporting high-stakes decision-making. These complementary strengths broaden the applicability of robust DEA across diverse decision-making contexts. Together, these complementary features extend the applicability of robust DEA to a wider range of decision-making contexts. This distinction also provides a basis for the subsequent subsection, which examines the computational complexity and performance implications of these robust DEA models.

3.3. Computational complexity

Table 1 presents a comparative analysis of the computational complexity associated with the three robust DEA models (6), (8), and (13). The comparison is based on the number of variables, constraints, and the number of model instances solved. Although all three models are LP problems and share a similar solution frequency, they differ substantially in terms of problem size and structural complexity.

Model (6), based on a standard budgeted uncertainty set, exhibits the most compact structure, with a linear increase in variables and constraints proportional to the size of the input/output reference sets $|I_j|$ and $|R_j|$. Model (8), which incorporates correlations through subsets $|K^x_j|$ and $|K^y_j|$, maintains a similar structure, but may reduce dimensionality when correlated subsets are smaller. In contrast, model (13), built on order statistics uncertainty set, results in a significantly larger problem size. This increase arises from the quadratic growth in the number of constraints, proportional to $|I_j|^2$ and $|R_j|^2$, resulting from the pairwise

Table 1
Computational comparison of robust DEA models.

Criteria	Model (6)	Model (8)	Model (13)
Model type	LP	LP	LP
No. of variables	$m+s+n(1+\|I_j\|+\|R_j\|)$	$m+s+n(1+\|K_j^x\|+\|K_j^y\|)$	$2n(\|I_j\|+\|R_j\|)+m+s$
No. of constraints	$1+n(1+\|I_j\|+\|R_j\|)$	$1+n(1+\|K_j^x\|+\|K_j^y\|)$	$1+n(1+\|I_j\|^2+\|R_j\|^2)$
No. of models	n	n	n

indexing in the uncertainty formulation. While model (13) remains an LP problem, this increased structural complexity leads to a more computationally intensive formulation.

Despite all models being LPs, the scale and structure of model (13) may pose scalability challenges, particularly for large-scale datasets with numerous inputs and outputs. In practice, the choice of model should consider both the trade-off robustness and tractability. For scenarios with limited computational resources or large DMU populations, model (6) often suggests a practical and interpretable solution. Model (8) presents a middle ground, capturing interdependencies with moderate complexity. When extreme-value behaviour or tail-risk sensitivity is critical—such as in healthcare, finance, or disaster logistics— model (13) provides superior granularity and adaptability in handling uncertainty, albeit at a higher computational cost.

In model (13), the order statistics assignment blocks introduce $O(|I_j|^2)$ and $O(|R_j|^2)$ constraints per DMU. This enlarges and densifies the constraint matrix, raises memory requirements, and typically requires greater solver effort than models (6) and (8), particularly when the number of inputs and outputs per DMU is high. Nevertheless, all three models (6), (8) and (13) remain linear when uncertainty set parameters are fixed exogenously, with no mixed integer components introduced. The main trade-off lies between the extra detail provided by the order statistics construction, which is useful when tail-risk sensi-

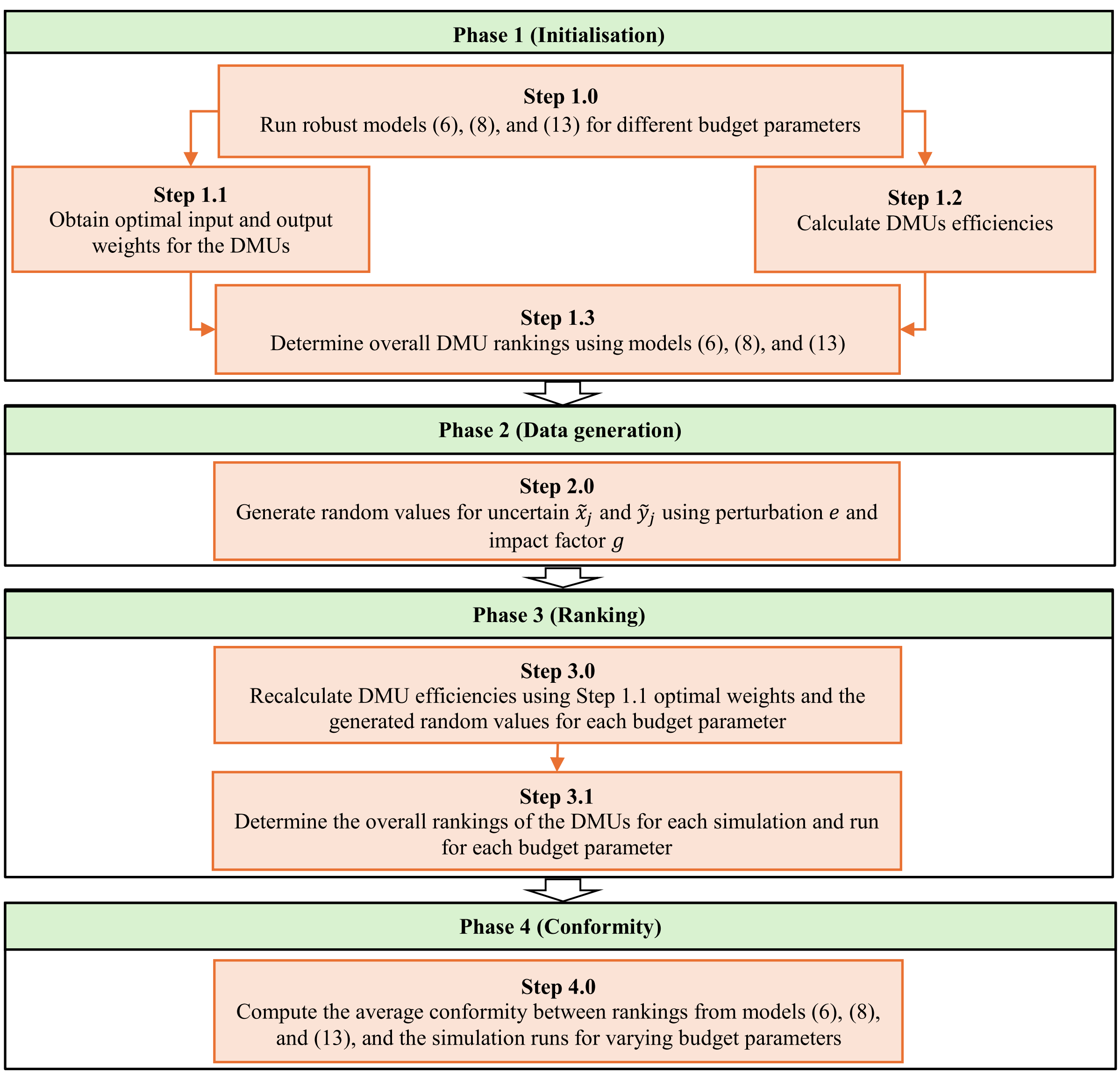


Fig. 3. The scheme of the proposed models and Monte Carlo simulation.

tivity is important, and scalability in high-dimensional instances. When computational efficiency is paramount, the more compact structures of models (6) and (8) may be preferable.

4. A comparative analysis

This section presents a comparative example that benchmarks the proposed robust DEA models against the conventional budgeted uncertainty approach. A Monte Carlo simulation—structured in four phases: (i) Initialisation, (ii) Data generation, (iii) Ranking, and (iv) Conformity (see Fig. 3)—is employed to assess the stability of DMU rankings under uncertainty. For each budget parameter (Γ for models (6) and (8) and q for model (13)), 5,000 random samples are generated within predefined input-output bounds. Efficiencies are then recalculated using the optimal weights derived from the corresponding robust DEA models. A binary scoring system is used to compute conformity percentages, providing insights into the stability of rankings and informing the selection of suitable budget parameter values for models (6), (8), and (13).

Interval DEA and robust DEA fulfil related but distinct decision purposes under bounded data. Interval DEA addresses data uncertainty through efficiency bounds $[h^L, h^U]$, which is valuable for screening and for identifying DMUs that remain efficient under all admissible realisations. However, when the practical objective is to form a single ordering under uncertainty, endpoint-based rankings can be fragile since many DMUs may share the same upper endpoint (often $h^U = 1$) and efficiency intervals may overlap, requiring additional post-processing to discriminate ranks. By contrast, the robust multiplier models proposed here provide a guaranteed efficiency level at each chosen protection parameter and can be interpreted as a robustness path $\theta(\Gamma)$ for models (6)–(8) or $\theta(q)$ for model (13), characterising how efficiency and rankings vary as protection increases.

The proposed models are developed within the robust optimisation literature, which highlights two practical advantages; they are distribution-free, guaranteeing feasibility for every realisation inside the uncertainty set, and adjustable and tractable, since parameters Γ or q control the degree of conservatism, and many robust counterparts admit linear or convex reformulations. In addition, the robust optimisation-based models have the ability to represent uncertainty in ways that align with how data are observed in practice, with interval sets reflecting measurement error or reporting ranges, budgeted sets capturing situations in which only a limited number of variables deviate simultaneously, and order statistics or quantile-based sets exploiting partial historical information when full distributions are unavailable. Accordingly, we interpret robust DEA scores jointly alongside two complementary diagnostics; the PoR, which quantifies the cost of conservatism, and Monte Carlo conformity, which measures how reproducible the resulting ranking is under repeated draws within the same bounds. This combined reporting avoids treating the worst-case score as a stand-alone ranking and instead supports the selection of protection levels using explicit evidence on ranking stability. Robustness in this study is therefore defined through the protection parameters (Γ or q), which provide a transparent and auditable mechanism for stressing the feasible weight region and guaranteeing efficiency performance over all admissible realisations within the uncertainty set.

Table 2 presents input and output data for seven academic departments (DMUs) within a university, originally introduced by Wong and Beasley (1990) for efficiency analysis. Each department is evaluated based on three inputs—academic staff (x_1), research funding (x_2), and administrative support (x_3)—and three outputs—undergraduate degrees awarded (y_1), postgraduate degrees awarded (y_2), and research publications (y_3). This dataset is used to compare models (6), (8), and (13), allowing for the assessment of their effectiveness in handling different data patterns and providing robust efficiency evaluations in an academic context.

As discussed earlier in Subsection 3.1, the impact factors g_{ki} and g_{kr} for the robust DEA model (8) are calculated using PCA on the standardised dataset. The first principal component (PC1) accounted for 89.4% of the total variance, making it a suitable basis for identifying the most influential variables. Squaring the PC1 loadings yields the values of g_{ik^x} and g_{rk^y}, which quantify the contribution of each input and output to overall variability. These factors are then used to define variable-specific uncertainty bounds in the robust DEA model (8), enabling the model to reflect the relative importance of each input and output more accurately under data uncertainty.

The three box-and-whisker plots illustrate the efficiency scores of seven DMUs under different robust DEA models, each based on a distinct uncertainty set: budgeted (Fig. 4(a)), correlated budgeted (Fig. 4(b)), and order statistics (Fig. 4(c)). Each figure displays the distribution, median, and variability of the efficiency scores, allowing a visual comparison of how the models respond to different uncertainty structures. The budgeted and correlated budgeted sets lead to relatively similar distributions; however, the correlated set captures dependencies among uncertain parameters, resulting in noticeable but moderate variations in the efficiency range. In contrast, the order statistics set in Fig. 4(c) exhibits a slightly broader spread in some DMUs (e.g., DMU2 and DMU4), indicating greater sensitivity to uncertainty. Overall, these figures illustrate how the choice of uncertainty set affects the robustness of DEA scores.

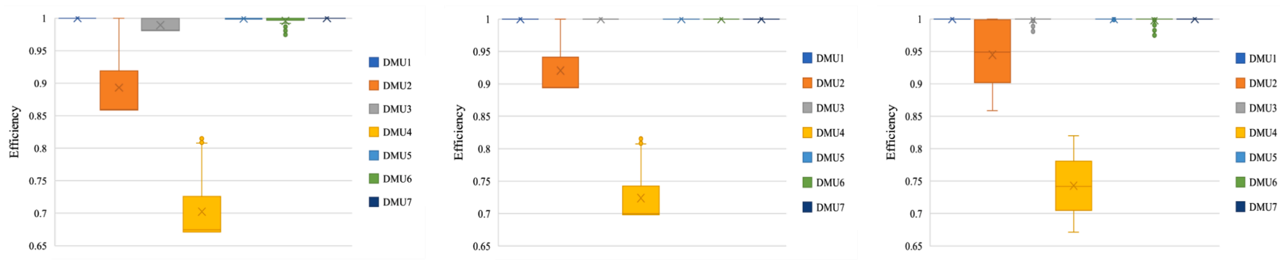
Fig. 5 illustrates the average conformity between the DMU rankings obtained from the three robust DEA models (6), (8), and (13) and the rankings derived from 5,000 Monte Carlo simulation runs, under varying levels of uncertainty. A total of 5,000 simulation runs is selected to ensure statistical stability of the results, as further increases yielded negligible changes in the estimated ranking distributions. This level of replication aligns with standard practice in simulation-based DEA studies under uncertainty (e.g., Torres-Jiménez et al., 2015; Tsionas, 2021).

In Fig. 5(a) (budgeted set), conformity starts at approximately 52% for $\Gamma = 0$ (the deterministic case). The highest conformity (91%) occurs within $\Gamma \in [2.27, 2.35]$. At the maximum uncertainty level ($\Gamma = 6$), conformity drops to about 29%, indicating a substantial reduction in ranking consistency. Fig. 5(b) (correlated budgeted set) shows conformity 61% at $\Gamma = 0$, peaks at 97% for $\Gamma \in [2.75, 5.92]$, and remains near 63% at $\Gamma = 6$. Fig. 5(c) (order statistics) presents conformity ranging from 46% at $q = 0$ to 43% at $q = 1$, with a maximum of 95% in $q \in [0.76, 0.87]$. These findings indicate that while increasing uncertainty can reduce ranking consistency, appropriately chosen uncertainty constructions and intermediate protection levels (especially under order statistics sets) can yield high empirical stability.

Table 3 presents the efficiency scores and conformity values of seven DMUs under three robust DEA models—budgeted, correlated budgeted, and order statistics—evaluated at three levels of uncertainty: deterministic ($\Gamma = q = 0$), maximum average conformity ($\Gamma = 2.5$ or $q = 0.75$), and maximum uncertainty level ($\Gamma = 6$ or $q = 1$). In the deterministic case, all models assign a full efficiency score of 1 to all DMUs except for DMU4, which is the only inefficient unit with a score of 0.8197, and the corresponding conformity values are 52% for model (6)

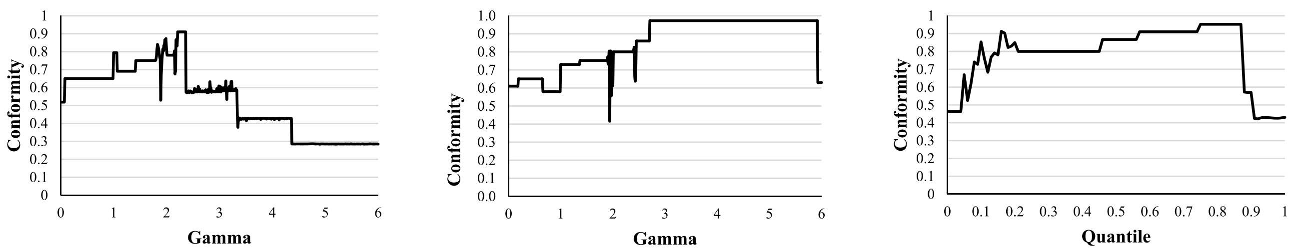
Table 2
Data for seven university departments.

	Inputs			Outputs		
	x_1	x_2	x_3	y_1	y_2	y_3
DMU1	12	400	20	60	35	17
DMU2	19	750	70	139	41	40
DMU3	42	1,500	70	225	68	75
DMU4	15	600	100	90	12	17
DMU5	45	2,000	250	253	145	130
DMU6	19	730	50	132	45	45
DMU7	41	2,350	600	305	159	97



(a) (b) (c)

Fig. 4. Box-and-whisker plots of efficiency scores for seven DMUs under three uncertainty sets (a) budgeted, (b) correlated, and (c) order statistics.



(a) (b) (c)

Fig. 5. Average conformity between DMU rankings from three uncertainty sets (a) budgeted, (b) correlated, and (c) order statistics and Monte Carlo simulation results.

Table 3
Efficiency and ranking conformity of seven DMUs under robust and interval DEA models.

	Budgeted	Correlated	Order statistics	Budgeted	Correlated	Order statistics	Budgeted	Correlated	Order statistics	Interval DEA	
	$\Gamma = 0$	$\Gamma = 0$	$q = 0$	$\Gamma = 2.26$	$\Gamma = 2.72$	$q = 0.75$	$\Gamma = 6$	$\Gamma = 6$	$q = 1$	h^L	h^U
DMU1	1	1	1	1	1	1	1	1	1	1	$1(E^{++})$
DMU2	1	1	1	0.8809	0.8972	0.8813	0.8586	0.8937	0.8586	0.8586	$1(E^{+})$
DMU3	1	1	1	1	1	1	0.9809	1	0.9809	0.9809	$1(E^{+})$
DMU4	0.8197	0.8197	0.8197	0.6960	0.7077	0.6887	0.6710	0.6985	0.6710	0.6710	$1(E^{+})$
DMU5	1	1	1	1	1	1	0.9987	1	0.9987	0.9987	$1(E^{+})$
DMU6	1	1	1	1	1	1	0.9747	1	0.9747	0.9747	$1(E^{+})$
DMU7	1	1	1	1	1	1	1	1	1	1	$1(E^{++})$
Conformity	52%	61%	46%	91%	97%	95%	29%	63%	43%	-	-

with the budgeted set, 61% for model (8) with the correlated budgeted set, and 46% for model (13) with the order statistics set. At the maximum [average] conformity level, slight efficiency reductions are observed for DMU2 and DMU4 across all models, while conformity improves significantly, reaching 91% (budgeted), 97% (correlated), and 95% (order statistics)—indicating strong alignment with simulated rankings. Under the maximum uncertainty level, efficiency scores drop further for DMUs 2, 3, 4, 5, and 6, particularly in models with the budgeted and order statistics sets, consistent with expectations from Theorem 3. Conformity values also drop sharply, most notably to 29% for the budgeted case, whereas the correlated (63%) and order statistics (43%) show comparatively greater robustness in maintaining performance and ranking consistency under uncertainty. These results, supported by the earlier box-and-whisker and conformity plots, indicate that although all models are affected by uncertainty, the correlated budgeted and order statistics models demonstrate greater resilience and consistency in both efficiency measurement and the reliability of DMU rankings under uncertain conditions.

To benchmark against interval DEA, we also compute endpoint efficiencies $[h^L, h^U]$ under symmetric $\pm 5\%$ bounds on all inputs and outputs and report them in the last two columns of Table 3 together with the Despotis and Smirlis (2002) class labels E^{++}, E^{+}, and E^{-}. Under interval DEA, the interval efficiency of each DMU is evaluated from pessimistic and optimistic perspectives obtained at the extreme points of bounded input and output data. The Despotis–Smirlis classification is defined as follows: $E^{(++)}$ if $h^L = 1$, E^{+} if $h^L < 1$ and $h^U = 1$, and E^{-} otherwise. While interval DEA is therefore useful for screening, in bounded-data applications it frequently yields many E^{+} DMUs (sharing $h^U = 1$) and overlapping efficiency intervals, so endpoint information alone offers limited discrimination and often requires additional post-processing to produce a unique ordering. By contrast, models (6), (8), and (13) yield a single guaranteed efficiency at each chosen protection level (Γ or q), providing a robustness path $\theta(\Gamma)$ or $\theta(q)$ that characterises how guaranteed efficiency and ranks change as protection increases. We report these robust scores jointly with the PoR and Monte Carlo conformity (rank-stability) so that protection levels can be selected on the basis of a direct trade-off between discrimination, conservatism, and empirical ranking stability. The numerical results in Table 3 illustrate these differences. Interval DEA assigns $h^U = 1$ to all DMUs and classifies five DMUs as E^{+}, so endpoint information alone provides poor discrimination among efficient candidates. On the contrary, at intermediate protection levels (e.g. $\Gamma = 2.72$ or $q = 0.75$) the robust models produce differentiated guaranteed scores while achieving much higher conformity with the simulated rankings (up to 97% for the correlated model). This indicates that these protection ranges yield reproducible (robust) orderings under the same $\pm 5\%$ bounds. Moreover, the interval lower bound h^L coincides with the budgeted robust score at $\Gamma = 6$, confirming that interval DEA corresponds to robust DEA and that the robust framework is strictly more flexible and informative. In a nutshell, interval DEA remains a valuable screening tool, but the proposed robust optimisation approach provides an operationally meaningful ranking mechanism with an adjustable protection parameter — allowing practitioners to choose protection levels that balance discrimination, conservatism (PoR), and empirical stability (conformity).

We compare the computational complexity of the three LP robust DEA models—(6), (8), and (13)—using their formulation size for 7 DMUs here. The models differ in problem size (the number of variables and constraints) and thus in runtime and memory requirements. Model (13) is the most computationally intensive, with 90 variables and 134 constraints per run. Model (6) involves 55 variables and 50 constraints, while model (8) is the most computationally efficient, requiring only 44 variables and 29 constraints. These results indicate that model (8) is the most computationally efficient, whereas model (13) may demand substantially more processing time and memory, especially for larger datasets. From a managerial perspective, these results inform model selection. Model (8) is the most computationally efficient, with the fewest variables and constraints, making it well-suited for large-scale applications or scenarios requiring real-time analysis. Model (6) strikes a balance between robustness and computational complexity, though it might yield lower ranking conformity and fitness under high uncertainty levels. Model (13), despite being the most computationally intensive, demonstrates superior stability in efficiency scores and retains relatively high fitness—especially under correlated or order statistics uncertainty sets. Thus, practitioners should select a model according to operational scale, resource constraints, and the required level of robustness; prefer model (8) for speed and scalability, and model (13) when reliability under uncertainty is the priority and computational cost is acceptable.

5. Application to European banks

To demonstrate the applicability of the proposed robust DEA models, we analyse a dataset of 168 banks from the UK, Germany, France, Spain, and Italy, based on financial data from the 2015 accounting year (Alfiero et al., 2019). These countries represent a mix of Europe's largest banking sectors and varying regulatory environments, providing a meaningful ground for cross-country efficiency analysis. These institutions vary in size and market reach, from large international banks such as BNP Paribas, Deutsche Bank, and HSBC to smaller banks operating primarily within national markets. Given the volatility of financial markets and the sensitivity of bank performance to regulatory shifts, macroeconomic shocks, and geopolitical uncertainties, robust DEA provides a more realistic and reliable assessment framework. Explicitly accounting for data variability and ambiguity enables more stable efficiency evaluations and reduces the risk of misleading conclusions driven by atypical observations or uncertain conditions. This has clear practical implications for regulators and decision-makers; it supports more resilient benchmarking, helps identify consistently efficient performers, and informs strategic planning in uncertain and volatile financial environments.

The dataset includes key financial indicators sourced from the Bankscope database,[4] including total assets, number of employees, personnel expenses, equity, loans, net interest income, interbank deposits (i.e., deposits received from other banks), operating income, and net fees and commissions. Monetary variables are reported in millions of euros, whereas the number of employees, including both banking professionals and support staff, is recorded in absolute values. To ensure model consistency and prevent numerical instability in the DEA computations, the data were scaled prior to analysis. The selection of these inputs and outputs follows the well-established *intermediary approach*, which views banks as entities that utilise labour and financial resources to generate loans and income (see Berger & Humphrey, 1997; Fethi & Pasiouras, 2010). Inputs include the number of employees, total assets, personnel expenses, and equity, representing resource usage. Outputs include loans, operating income, net interest revenue, interbank deposits, and net fees and commissions, which indicate the services and revenue generated through financial intermediation. We note that the role of deposits in bank production is treated differently across the DEA literature; some studies treat deposits as an input, others as an output, and some recent contributions model deposits as an intermediate product or in two-stage frameworks (Boďa & Piklová, 2021; Cook & Zhu, 2007; Holod & Lewis, 2011). In our specification, total assets capture the bank's overall resource base (which includes deposit-funded resources in part), while interbank deposits are included among outputs to reflect the bank's intermediation activity in the interbank market. This choice aligns with the intermediary perspective and also avoids specification issues that can arise when including both highly collinear aggregates (e.g., total assets and customer-deposit aggregates) as separate variables (see Fethi & Pasiouras, 2010; Boďa & Piklová, 2021). Fig. 6 summarises the approach adopted for the empirical application in this study.

The banks are ranked based on total asset size, enabling sub-regional comparisons across the five selected countries—France, Germany, Spain, Italy, and the UK. Descriptive statistics for all inputs and outputs of those countries are presented in Table 4. The inputs exhibit considerable variability, indicating a wide range in the scale of banking operations. For instance, the number of employees ranges from 217 to 193,863, and total asset size varies from €10,124 to over €1.99 million, showing the inclusion of both small and large banks. Other outputs such as net interest revenue, loans, and operating income, also show substantial dispersion. Loans range from €692 to €758,505, and net interest revenue spans from €209 to €263,121. The high standard deviations across most variables suggest significant heterogeneity in bank size and performance, which is crucial for benchmarking and efficiency analysis.

Fig. 7 compares the distribution of total assets (in million euros) across five countries. The UK shows both the highest median and the widest interquartile range, indicating greater variability in total asset values. France, Germany, Spain, and Italy exhibit lower medians and more compact distributions. All countries have several outliers, with France and the UK showing some of the highest individual total asset values. This suggests a more dispersed total asset distribution among banks in the UK and France compared to the other considered countries.

The 2015 cross-section is used here as a methodological demonstration. Our objective is to show how alternative robust uncertainty sets (budgeted, correlated budgeted, and order statistics) reshape DEA weights, scores and rankings when deterministic observations are perturbed in empirically plausible ways. To operationalise plausible perturbations, we (i) impose a baseline $\pm 5\%$ perturbation band (alternative magnitudes are reported below), (ii) derive PCA-based impact directions from the standardised inputs and outputs so that shocks follow observed co-movements (e.g., assets, loans and net interest income). These checks convert the 2015 snapshot into a set of meaningful stress- and noise-scenarios (measurement error, short-term volatility, minor revisions) and demonstrate that the correlated budgeted model delivers substantially greater rank-stability and lower PoR in this realistic testing environment.

We report the results from applying the proposed models to evaluate the relative efficiency of the banks. First, we specify how uncertainty in the selected bank measures is represented in the robust models. Following standard practice in the robust DEA literature (e.g., Toloo et al., 2022; Arabmaldar et al., 2024), we impose a $\pm 5\%$ perturbation band on all inputs and outputs. Accordingly, uncertain realisations lie within $\widetilde{x}_{ij} \in \left[x_{ij} \pm 0.05 x_{ij}\right]$ and $\widetilde{y}_{rj} \in \left[y_{rj} \pm 0.05 y_{rj}\right]$.

[4] https://bankfocus.bvdinfo.com/

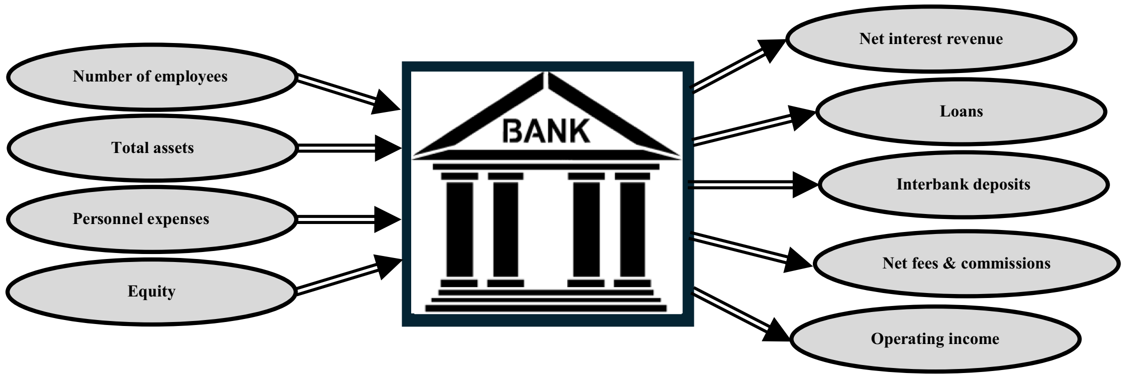


Fig. 6. Inputs and outputs variables.

Table 4
Descriptive statistics of inputs and outputs for the selected countries.

Variables	Mean	SD	Min	Max
Inputs				
Employees	16,206.05	34,661.55	217	193,863
Total assets	167,761.26	360,920.86	10,124	1994,193
Personnel expenses	9,765.66	19,389.02	226.30	100,077
Equity	1,224.43	2,681.69	14.20	16,061
Outputs				
Net interest revenue	23,876.95	47,554.08	208.60	263,121
Loans	73,593.21	141,153.25	691.80	758,505
Interbank Deposits	2,095.97	4,666.34	52.30	33,267
Net fees & commissions	1,501.33	3,311.83	20.60	19,805
Operating income	950.54	2,036.94	6.60	12,765

Using standardised inputs and outputs for 2015, a correlation-matrix PCA (Table 5) shows that PC1 explains 91.60% of the variance and PC2 adds 3.34%, yielding 94.93% cumulative variance. PC1 loads positively and nearly uniformly across variables (a common scale factor), whereas PC2 captures the principal contrast pattern. We therefore retain PC1–PC2 as the baseline and compute impact factors as $g = |L|\sqrt{\lambda}$ with light sparsification; these components produce the impact factors g_{ik^x} and g_{rk^y} that define the uncertainty bounds for the correlated budgeted uncertainty set in the robust DEA model (8).

Because the correlated budgeted set allocates protection along empirical co-movements, we verify that our conclusions do not hinge on a particular PCA construction. Re-estimating the impact factors from the correlation-matrix PCA on standardised inputs and outputs, we consider: (i) the baseline PC1–PC2 retention (at least 90% variance), (ii) an extended PC1–PC3 retention (at least 95% variance), (iii) PC1 only, (iv) an orthogonal varimax rotation of the retained loadings, and (v) light sparsification with thresholds $\tau \in \{0,\ 0.05,\ 0.10\}$. Across representative robustness budgets $\Gamma \in \{0,\ 4,\ 8\}$, summary statistics and rankings from model (8) are indistinguishable at the reporting precision (ΔMean ≤ 0.002; Δ (No. efficient) ≤ 1; Kendall's $\tau \geq 0.98$ relative to the baseline). These robustness checks confirm that model (8)'s rank stability and results under uncertainty are not an artefact of the PCA retention, rotation, or sparsification choices.

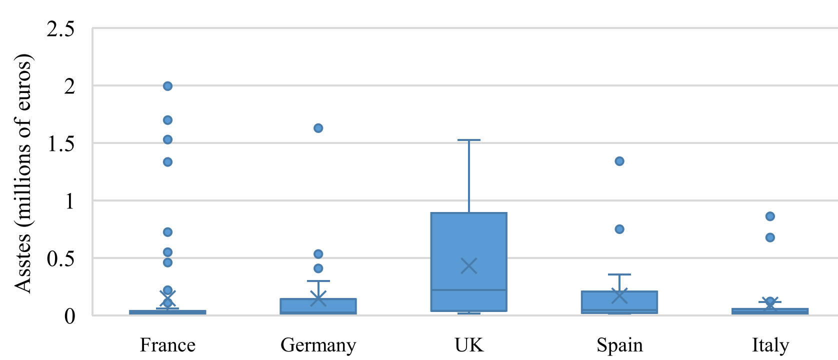


Fig. 7. Total asset distribution across banks in five European countries.

Table 5
Coefficients of principal components (PCs).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Employees (I)	0.3371	0.0428	-0.0310	0.6569	0.5321	-0.1535	-0.3354	0.1426	-0.1122
Assets (I)	0.3406	0.0311	-0.0667	-0.5622	0.1018	0.2628	-0.4015	0.4699	-0.3178
Equity (I)	0.3440	-0.1361	-0.1447	-0.1382	0.0857	0.1928	-0.1664	-0.8396	-0.2141
Personnel Expenses (I)	0.3433	-0.0714	0.2208	-0.1783	0.2623	0.1810	0.0190	-0.0044	0.8334
Interbank Deposits (O)	0.3030	0.8606	-0.2702	0.0634	-0.2630	0.0167	0.0816	-0.0467	0.1111
Loans (O)	0.3313	-0.3127	-0.4551	-0.1553	-0.2062	-0.7026	0.0077	0.0703	0.1419
Net Interest Revenue (O)	0.3341	-0.3321	-0.3139	0.3010	-0.1945	0.5104	0.4877	0.2159	-0.0781
Operating Income (O)	0.3359	0.1106	0.4586	-0.1847	0.2674	-0.2852	0.6094	-0.0048	-0.3260
Net Fees & Commissions (O)	0.3289	-0.1127	0.5796	0.2206	-0.6398	-0.0406	-0.2869	0.0079	-0.0347
% of variance	91.59	3.33	2.51	1.12	0.59	0.32	0.22	0.15	0.13
% of cumulative variance	91.59	94.93	97.44	98.57	99.16	99.48	99.71	99.86	100.00

Table 6
Descriptive statistics of robust efficiency scores.

			Mean	SD	Min	Max	No. Eff
Model (6)	Γ	0	0.8551	0.1190	0.5247	1	33
		0.8	0.8219	0.1204	0.5022	1	22
		1.6	0.7914	0.1217	0.4822	1	19
		2.4	0.7698	0.1238	0.4659	1	16
		3.2	0.7539	0.1247	0.4523	1	14
		4	0.7421	0.1253	0.4425	1	13
		4.8	0.7324	0.1264	0.4359	1	12
		5.6	0.7257	0.1279	0.4325	1	12
		6.4	0.7221	0.1293	0.4300	1	12
		7.2	0.7205	0.1299	0.4295	1	12
		8	0.7201	0.1301	0.4295	1	12
Model (8)	Γ	0	0.8551	0.1190	0.5247	1	33
		0.8	0.8475	0.1195	0.5190	1	29
		1.6	0.8413	0.1200	0.5148	1	28
		2.4	0.8367	0.1206	0.5113	1	27
		3.2	0.8338	0.1211	0.5082	1	25
		4	0.8315	0.1215	0.5058	1	24
		4.8	0.8299	0.1221	0.5045	1	24
		5.6	0.8288	0.1225	0.5041	1	23
		6.4	0.8283	0.1228	0.5041	1	23
		7.2	0.8282	0.1229	0.5041	1	23
		8	0.8281	0.1230	0.5041	1	22
Model (13)	q	0	0.8552	0.1189	0.5247	1	33
		0.1	0.8431	0.1209	0.5153	1	28
		0.2	0.8297	0.1227	0.5051	1	25
		0.3	0.8157	0.1238	0.4951	1	20
		0.4	0.8021	0.1253	0.4853	1	20
		0.5	0.7885	0.1266	0.4757	1	19
		0.6	0.7747	0.1271	0.4662	1	17
		0.7	0.7611	0.1285	0.4570	1	17
		0.8	0.7480	0.1293	0.4479	1	13
		0.9	0.7347	0.1298	0.4391	1	13
		1	0.7214	0.1300	0.4303	1	12

Table 6 presents the descriptive statistics of robust efficiency scores for models (6), (8), and (13) under increasing levels of uncertainty, represented by the robustness parameter Γ for models (6) and (8), and the quantile q for model (13). At $\Gamma = 0$ (or $q = 0$), all three models yield almost an identical average efficiency score of approximately 0.855, with a standard deviation of about 0.119 and a minimum score of 0.5247, identifying 33 DMUs as efficient. As uncertainty increases, all models show a gradual decline in both mean efficiency and the number of efficient units. For model (6), the mean efficiency score drops to 0.7201 at $\Gamma = 8$, with only 12 DMUs remaining efficient. In contrast, model (8) demonstrates greater resilience, maintaining a higher average efficiency of 0.8281 and identifying 22 efficient DMUs at $\Gamma = 8$. Model (13) also shows a steady decline in efficiency as q increases; that said, the mean score drops to 0.7214 with only 12 efficient DMUs at $q = 1$. These results carry important real-world implications. They suggest that small changes or uncertainty in data such as fluctuations in financial indicators due to market volatility or data reporting inconsistencies can lead to noticeably different efficiency assessments. In regulatory and managerial contexts, relying on models without robustness may lead to overestimating or underestimating certain banks' performance. Model (8), however, demonstrates strong robustness under uncertainty by consistently providing higher average efficiency scores and identifying a larger number of efficient DMUs. This robustness makes it especially well-suited for practical applications where data ambiguity is unavoidable, enabling more reliable benchmarking and more equitable evaluation of bank performance over time.

In summary, although all models are affected by increasing uncertainty, model (8) exhibits the most stable performance, and can serve as a valuable tool for decision-makers who require reliable efficiency insights in uncertain environments.

Fig. 8 presents box plots of robust efficiency scores across five countries under three different models: (a) model (6), (b) model (8), and (c) model (13). Across all three figures, efficiency distributions show some variation across countries, with median values generally exceeding 0.7. The results driven from model (8) show slightly higher and more stable efficiency scores, especially for Germany and the UK, evidenced by narrower interquartile ranges. In contrast, models (6) and (13) exhibit greater dispersion in efficiency scores for countries such as France and Spain, indicating higher sensitivity to uncertainty. Notably, German and UK banks consistently achieve higher and more stable efficiency scores across all models, signifying stronger operational resilience, whereas French and Spanish banks show greater variability and lower median efficiencies, particularly under model (6), implying a higher sensitivity to uncertainty. Overall, these illustrations align with the descriptive statistics in Table 7, confirming that model (8) maintains superior robustness and consistency across different settings.

To assess whether the rankings obtained from the three robust DEA models differ significantly, we employ non-parametric statistical tests, appropriate given the bounded and non-normal distribution of the rankings. Pairwise comparisons using the Mann–Whitney U test reveals statistically significant differences across all model pairs (see Table 7): model (6) vs. model (8) ($U = 9162$, $p < 0.0001$, $r = -0.303$), model (6) vs. model (13) ($U = 11966$, $p = 0.0159$, $r = -0.132$), and model (8) vs. model (13) ($U = 17108$, $p = 0.0008$, $r = 0.184$). These results indicate that rankings are notably influenced by the choice of model, with the largest difference observed between models (6) and (8), as reflected by a moderate effect size, suggesting a practically relevant difference in rankings.

To further validate these findings, the Kruskal–Wallis H test is conducted to test whether the efficiency scores differ significantly across the

Table 7
The results of the Mann–Whitney U test.

Comparison	U-statistic	p-value	Z-score	Effect size (r)
Model (6) vs. Model (8)	9,162	<0.0001	-5.559959	-0.303321
Model (6) vs. Model (13)	11,966	0.015939	-2.410439	-0.131500
Model (8) vs. Model (13)	17,108	0.000762	3.365179	0.183586

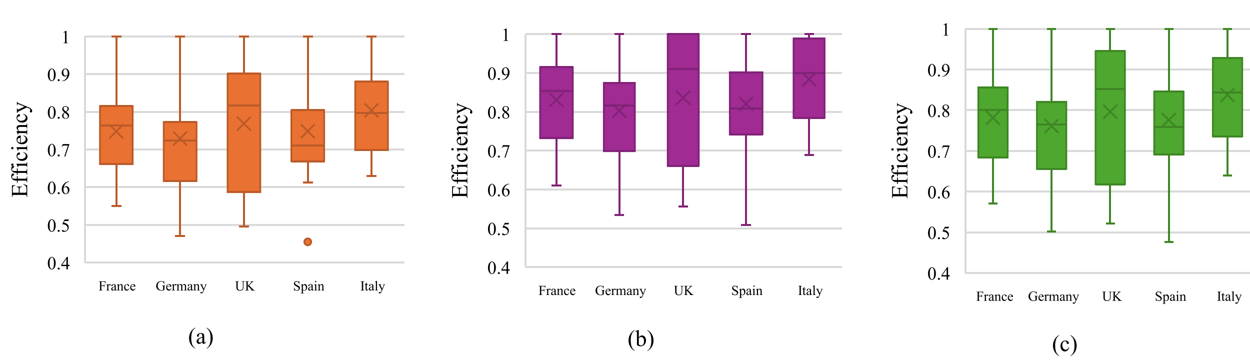


Fig. 8. Efficiency changes under (a) budgeted, (b) correlated, and (c) order statistics uncertainty sets.

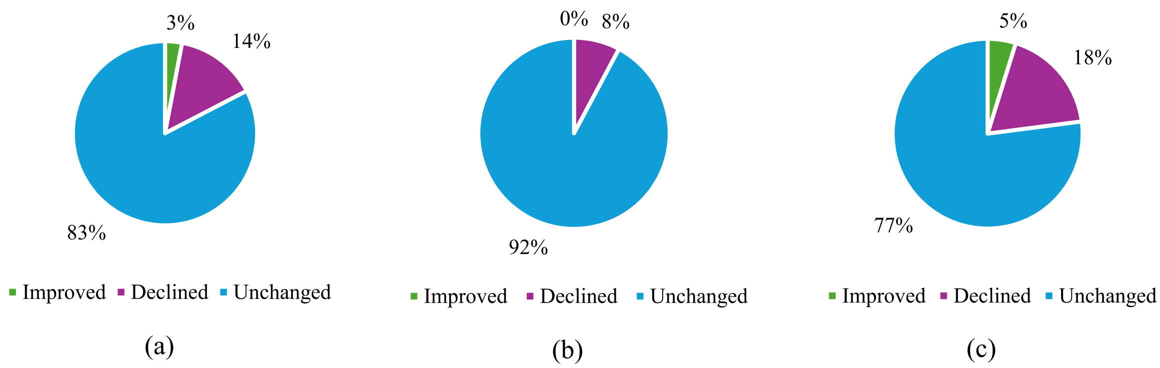


Fig. 9. Percentage changes in bank rankings under (a) budgeted, (b) correlated, and (c) order statistics uncertainty sets.

three DEA models. The test yields a statistically significant result ($H = 32.078$, $p < 0.0001$), indicating that at least one model produces rankings that differ from the others. These results highlight the importance of model choice in robust DEA and reinforce the role of conducting robustness checks when interpreting efficiency rankings.

Fig. 9 illustrates the percentage of changes in bank rankings under increasing uncertainty for three models: (a) model (6), (b) model (8), and (c) model (13). In model (6), 3% of DMUs improved, 14% declined, and 83% remained unchanged. Model (8) shows the highest stability, with 92% of banks retaining their original ranks, 8% declining, and none improving. Conversely, model (13) exhibits the greatest volatility, with 5% improving, 18% declining, and 77% remaining unchanged. These

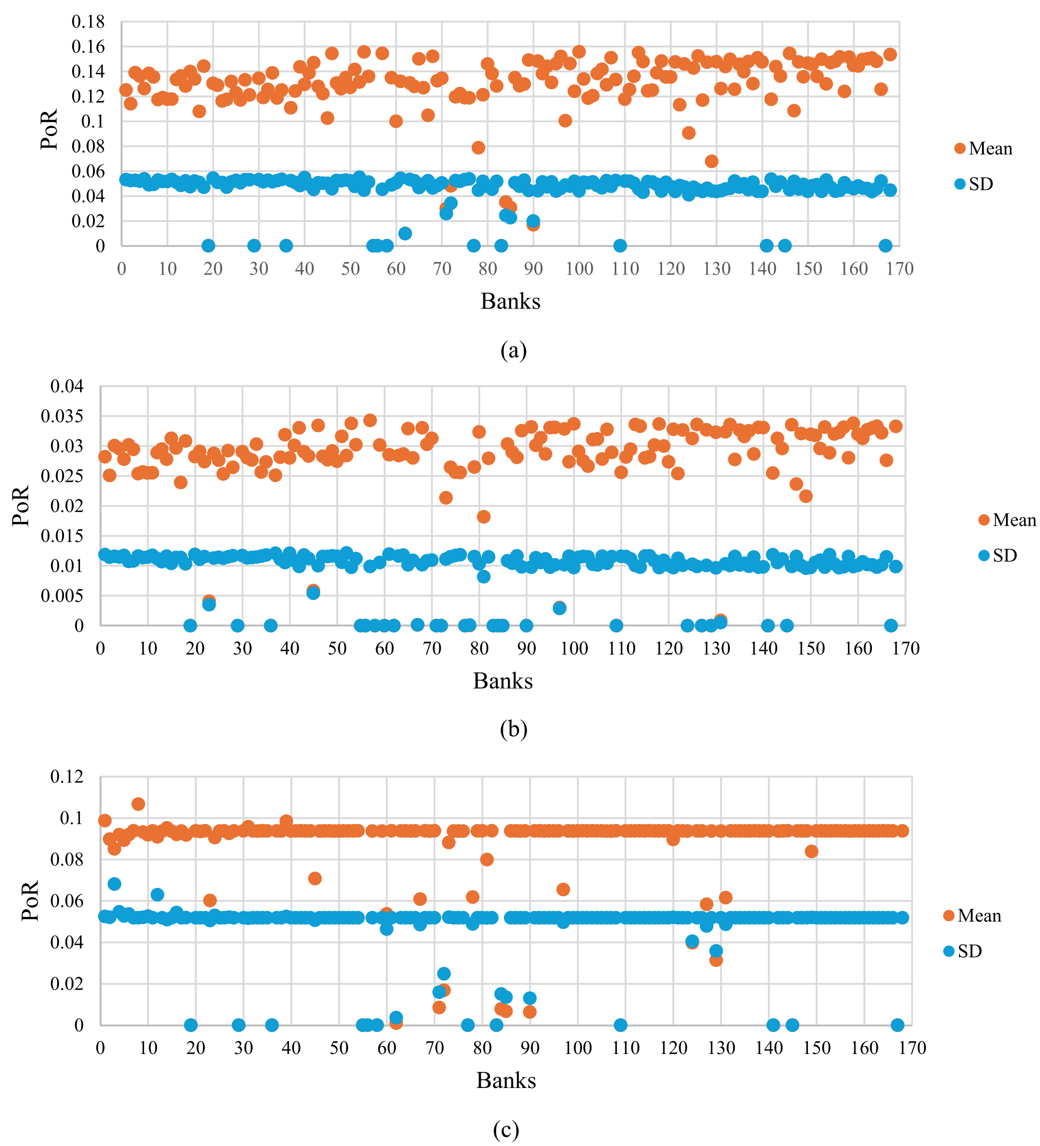


Fig. 10. Mean vs. SD of bank robustness under (a) budgeted, (b) correlated, and (c) order statistic uncertainty sets.

results suggest that model (8) provides the most consistent rankings under uncertainty, while model (13) is more sensitive to data variability.

Fig. 10 presents scatter plots of the mean and standard deviation of the PoR for each DMU under three different models: (a) model (6), (b) model (8), and (c) model (13).

In all cases, most banks have relatively low PoR values, indicating minor efficiency losses due to the inclusion of uncertainty. Model (8) shows the lowest PoR values, with both mean and standard deviation tightly clustered around zero, highlighting its robustness and minimal performance degradation. By contrast, models (6) and (13) exhibit wider distributions and higher peaks in both mean and standard deviation of the PoR, indicating greater sensitivity to uncertainty. Overall, the results highlight model (8) as the more stable and cost effective choice when balancing robustness with efficiency.

Fig. 11 presents box plots of the PoR for five countries under three robust DEA models (6), (8), and (13). The results reveal notable differences in the values and variability of PoR across models and countries. Model (8) consistently shows the lowest PoR values, with median values ranging between 0.005 and 0.015 across countries, and narrow interquartile ranges, indicating high stability. In contrast, model (6) exhibits higher PoR values, with median values approaching 0.12 and upper extremes near 0.18 in countries such as France and Germany.

Model (13) shows a moderate level of robustness, with median PoR values higher than those of model (8) but lower than those of model (6), ranging from 0.06 to 0.08 across most countries. French and German banks demonstrate the greatest performance degradation under model (6), indicating higher vulnerability to uncertainty. In contrast, UK and Italian banks show the least variability across all models, highlighting more stable and resilient performance. These findings confirm that model (8) provides the most cost efficient robustness, introducing minimal efficiency losses across different settings.

5.1. Sensitivity analysis

To assess the robustness of the proposed models to both data uncertainty and macroeconomic shocks, we consider two complementary scenarios: (i) varying levels of measurement noise and short-term volatility, and (ii) a sector-wide revenue stress. These analyses allow us to evaluate how the budgeted, correlated, and order statistics uncertainty formulations—represented by models (6), (8), and (13), respectively—respond to fluctuations in input–output data and to systematic contractions in banking revenues.

In the first scenario, we apply symmetric perturbation bands at ±1%, ±2%, ±3%, ±5%, and ±10% to all inputs and outputs and re-run models (6), (8) and (13). Table 8 reports four summary metrics at $\Gamma = 8$ for models (6) and (8), and $q = 1$ for model (13); mean efficiency, number of efficient banks, percentage of unchanged ranks (relative to the nominal ordering, $\Gamma = 0$, $q = 0$), and the median PoR.

The correlated budgeted model (8) is the least conservative and the most rank-stable, exhibiting higher mean scores, larger shares of unchanged ranks, and the lowest median PoR), while models (6) and (13) show larger score declines and more re-ordering as the perturbation (measurement noise) band widens. At the 5% baseline, model (8) preserves 92% of ranks (median PoR = 0.01), versus 83% (PoR = 0.12) for model (6) and 77% (PoR = 0.07) for model (13). Overall, the models remain informative for both small measurement noises (±1–3%) and larger short-term volatility (±5–10%), but model choice significantly affects ranking stability. Country-level patterns provide additional practical insight. Germany and the UK display tighter interquartile ranges and higher median efficiencies across all three models, suggesting more homogeneous performance and greater resilience to data noise; model (8) further compresses dispersion, indicating that it dampens distinctive volatility in these systems. France and Spain exhibit wider spreads and more frequent rank movements under models (6) and (13), consistent with greater sensitivity to measurement noise (which may reflect more heterogeneous bank size, revenue concentration, or data variability). Italy falls between these groups, with moderate shifts that are notably reduced by model (8).

In the second scenario, we extend the analysis to an economically interpretable sector stress event by imposing a stylised 10% reduction in *Loans* and *Net interest revenue* for all banks (a contraction consistent with a mild recessionary impulse), and re-evaluating efficiency. Table 9 summarises retention of the nominally efficient set,[5] the number of banks with rank changes, and the median absolute rank shift. Results are computed at $\Gamma = 8$ for models (6) and (8) and $q = 1$ for model (13); retention is measured relative to the nominal efficient set ($n = 33$ at $\Gamma = q = 0$).

Under the −10% revenue stress, model (8) retains 26 of the 33 originally efficient banks (79% retention) and triggers rank changes for only 22 banks with a median shift of 1 place. In contrast, model (6) retains 15 (45% retention) with 48 banks changing rank (median shift 5), while model (13) retains 17/33 (52% retention) with 40 rank changing rank (median shift 3). These results demonstrate that the correlated budgeted formulation (model (8)) not only preserves a larger portion of the efficient frontier but also ensures smoother rank transitions under systemic shocks, highlighting its structural robustness. Large, diversified institutions such as BNP Paribas, Crédit Agricole, Banco Santander, Barclays, and HSBC remain efficient under model (8), indicating their ability to absorb revenue contractions through diversified balance sheets and global operations. In contrast, mid-tier banks with narrower business models and concentrated revenue flows (e.g., Bankia, Banco Popular Español, Banco Sabadell, UBI Banca, Banco Popolare, and Banca Monte dei Paschi di Siena) experience greater efficiency losses and downward rank shifts, highlighting their vulnerability to revenue shocks.

Fig. 12 presents the country-level efficient-set retention under this −10% revenue stress. Germany and the UK exhibit the highest rank stability, with model (8) retaining a large majority of originally efficient large banks. This reflects more diversified banking architectures and stronger earnings buffers. France and Spain, by contrast, record substantial increases in median $|\Delta\text{rank}|$ under models (6) and (13), indicating that more fragmented or domestically concentrated systems are disproportionately affected by revenue contractions. Italy again occupies an intermediate position, where its system displays notable rank mobility, but model (8) substantially limits this volatility relative to the other two models.

Taken together, both sensitivity exercises present clear, actionable guidance. The correlated budgeted formulation (model (8)) suggests superior robustness, dampening measurement-noise effects, preserving the efficient frontier under co-moving revenue shocks, and producing significantly less rank volatility than alternative specifications, so it should be preferred for regulatory benchmarking and resilience assessments. Country diagnostics matter—diversified systems (Germany, UK) show tighter distributions and benefit less dramatically but still gain stability from model (8), whereas more fragmented or revenue-concentrated systems (France, Spain, and, to a lesser extent, Italy) are markedly more sensitive to noise and stress and therefore benefit most from correlated uncertainty sets. Practically, we recommend adopting model (8) for formal benchmarking, reporting sensitivity bands (e.g., ±1–10%) alongside point estimates, and supplying country-level diagnostics (interquartile range, median, efficient-set retention) to help stakeholders distinguish persistent structural inefficiency from transient shocks.

5.2. Managerial implications

The comparative analysis of models (6), (8), and (13) under various

[5] Efficient DMUs retained (count/%) refers to the number of banks remaining efficient after the stress imposition and the corresponding percentage relative to the 33 nominally efficient banks.

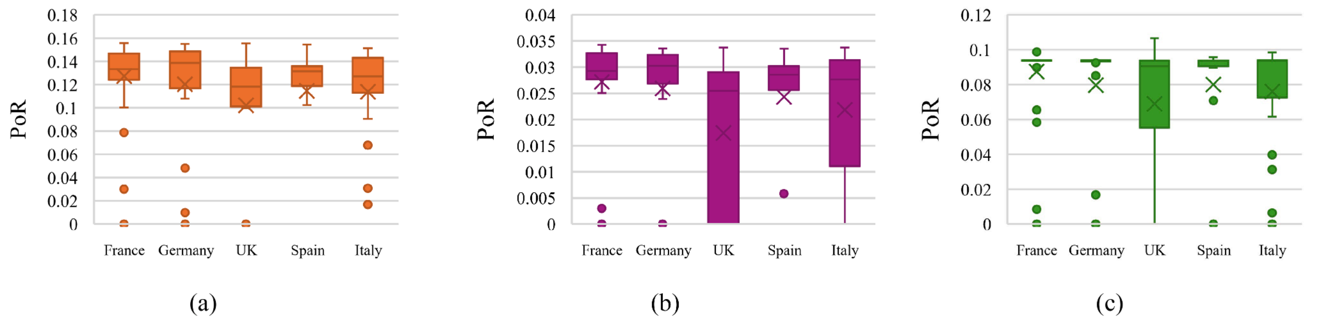


Fig. 11. Country robustness under (a) budgeted, (b) correlated, and (c) order statistic uncertainty sets.

Table 8
Sensitivity to different perturbation levels.

	Perturbation band (%)	Mean	No. efficient	unchanged ranks %	Median PoR
Model (6)	1	0.828	27	92	0.025
	2	0.791	24	88	0.055
	3	0.755	20	85	0.085
	5	0.720	12	83	0.120
	10	0.585	6	70	0.240
Model (8)	1	0.850	31	97	0.002
	2	0.840	29	95	0.004
	3	0.834	26	93	0.006
	5	0.828	22	92	0.010
	10	0.801	18	85	0.020
Model (13)	1	0.828	26	92	0.012
	2	0.802	23	86	0.030
	3	0.777	20	82	0.050
	5	0.721	12	77	0.070
	10	0.588	7	68	0.140

Table 9
Targeted sector stress (−10% on Loans and Net interest revenue for all banks).

Models	Efficient retained (n,%)	No. efficient lost	No. new efficient	No. rank change	Median $\|\Delta$ rank$\|$
Model (6)	15 (45%)	18	3	48	5.0
Model (8)	26 (79%)	7	2	22	1.0
Model (13)	17 (52%)	16	2	40	3.0

uncertainty settings provides valuable insights and practical guidance for bank managers, regulators, and performance analysts engaged in performance assessment, risk oversight, and strategic planning. Our findings highlight clear trade-offs between robustness to uncertainty, computational complexity, and the stability in bank rankings. These trade-offs should drive model selection according to the decision purpose, the institution's tolerance for conservatism, and available computational resources. Building on the sensitivity analysis in Sub-Section 5.1, managers can also monitor stability indicators such as median PoR and percentage of unchanged ranks to guide escalation decisions and stress-testing priorities.

For routine benchmarking and operational monitoring, model (8) is the recommended baseline. Empirically, it achieves the lowest median PoR (median between 0.005 and 0.015 across countries) and preserves ranking stability (92% of banks retain their positions as uncertainty increases). Its formulation with fewer constraints and variables translates into shorter solution times and straightforward integration into decision-support systems and performance dashboards. Managers should therefore use model (8) for regular reporting cycles (for example, monthly or quarterly KPIs), for early-warning systems, and as the primary metric to detect process inefficiencies and underperformance that require operational corrective actions. Model (8) also effectively limits ranking volatility across countries with diversified banking systems, such as Germany and the UK, while providing significant stability benefits in more fragmented or revenue-concentrated systems (France, Spain, Italy).

For stress testing, capital planning, regulatory submissions, and other high-stakes decisions where conservatism and worst-case assessment are priorities, model (13) is the preferred tool. The order statistics formulation addresses extreme-value behaviour and produces more conservative efficiency estimates that are stable under severe perturbations. Model (13) is appropriate for scenario-driven strategic decisions, informing capital buffers, contingency funding plans, and systemic-risk assessments. Given its greater computational demands, managers should run model (13) on a less frequent cadence (e.g., quarterly or ad hoc during stress cycles), allocate sufficient computational resources, or apply it to representative samples when full-sample runs are impractical.

Model (6) remains a valuable diagnostic instrument. Its sensitivity to the uncertainty budget parameter Γ makes it well suited to exploratory analysis and to mapping how efficiency scores evolve as uncertainty increases. Managers can use model (6) to identify vulnerability points and to prioritise where deeper investigation or remedial action is needed. For instance, a notable empirical finding is that model (6) shows substantial shifts in efficiency as Γ grows, with the number of efficient banks dropped from 33 to 12 and mean efficiency declined from 0.8551 to 0.7201 in our sample—signals that merit managerial attention in stress scenarios.

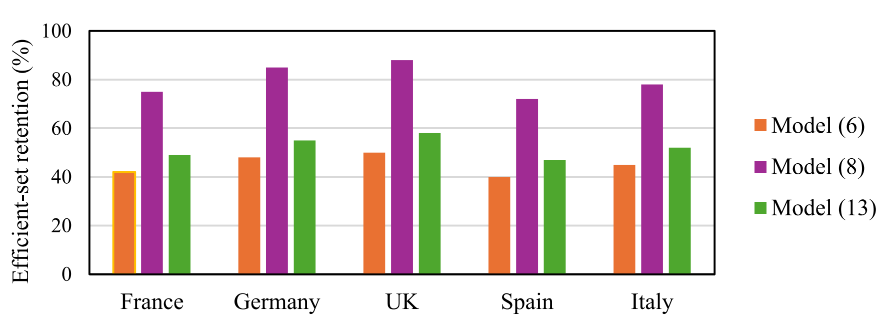


Fig. 12. Country-level retention of the efficient set under a −10% revenue stress.

To translate these findings into practice, we propose a tiered implementation protocol. First, adopt model (8) as the operational baseline for continuous monitoring and embed it within automated reporting. Second, implement routine checks on stability indicators derived from SubSection 5.1 (median PoR, percentage of unchanged rankings, and historical rank volatility) to detect early signs of performance instability. If these indicators breach predefined thresholds—for example, a sustained median PoR above approximately 0.01 or abrupt ranking changes beyond historical variability—escalate the case to targeted runs of models (6) and (13) to obtain conservative estimates and to investigate tail risks. Third, incorporate model (13) outcomes into formal stress-testing exercises and into regulatory or board-level reports where conservative, worst-case assessments are required. This structured approach maintains operational tractability while enabling risk-sensitive evaluation when conditions justify escalation.

Good governance and operational practices are essential for reliable deployment. Ensure standardised data pipelines, metadata tracking, and version control for inputs so results are reproducible and auditable. Document model choices and parameter settings (notably Γ and any correlated subsets) and record interpretation rules for stakeholders. Where computational constraints exist, consider sampling strategies, parallelised solver runs, or off-peak batch processing for model (13). When communicating with boards, senior management, or regulators, translate technical robustness into managerial metrics—report point estimates together with stability indicators and present comparative visuals that juxtapose routine (model (8)) and conservative (model (13)) assessments so decision-makers can see the range of plausible performance outcomes.

All in all, these results accentuate that understanding the nature, distribution, and impact of uncertainty is critical for effective model interpretation and decision-making. The observed differences in model behaviour across countries highlight that uncertainty can influence efficiency scores and rankings in varied ways. The two proposed models are complementary rather than competing. The correlated budgeted model (8) is recommended when computational efficiency and practicality are priorities, while the order statistics model (13) is preferable when robustness and stability under extreme uncertainty are the primary objectives. In practice, choosing a robust model also influences the type of managerial insights generated. Model (6), with its sensitivity to uncertainty, highlights potential stress points and supports risk-aware scenario planning. Model (8), with its computational efficiency and stable performance, is ideal for routine benchmarking and operational monitoring, enabling managers to track performance and address inefficiencies proactively. Model (13), providing conservative efficiency estimates, is most appropriate for strategic or regulatory contexts such as capital planning or systemic risk assessment, where robustness under extreme conditions is essential. These distinctions enable bank managers to align model selection with institutional priorities, achieving a balance between tractability and robustness that supports actionable, context-specific performance insights.

6. Conclusions

This study presents a comprehensive comparative evaluation of three robust DEA models—model (6) (budgeted), model (8) (correlated budgeted), and model (13) (order statistics)—across multiple dimensions, including efficiency scores, ranking stability, computational complexity, and PoR. Drawing on both simulation-based performance analysis and empirical evidence from real-world datasets, the findings reveal meaningful differences in each model's practical utility and performance under uncertainty. Among the three, the proposed robust DEA model incorporating a correlated budgeted uncertainty set (model (8)) emerges as the most balanced and practically effective. It consistently produces robust efficiency scores with minimal ranking variation, features the lowest PoR, and maintains high computational efficiency. These features make it particularly appropriate for large-scale evaluations and routine benchmarking exercises, where both robustness and operational feasibility are essential without imposing excessive computational burden. Conversely, the two robust DEA models based on budgeted and order statistics sets exhibit greater sensitivity to data uncertainty and higher computational complexity. Despite these challenges, both models provide valuable capabilities, particularly in risk-sensitive or highly scrutinised decision-making environments where precise discrimination among units or more conservative efficiency assessments are necessary. In particular, the robust DEA model using order statistics strikes a balance between robustness and computational complexity, making it a viable option in contexts requiring moderately conservative assessments. For practitioners and decision-makers, these findings accentuate the critical importance of aligning DEA model selection with organisational objectives, data characteristics, and risk tolerance. A carefully chosen robust DEA model enables more informed, equitable, and dependable performance evaluations, which in turn support improved strategic decision-making, fairer resource allocation, and enhanced operational outcomes (Cook & Seiford, 2009; Ehrgott et al., 2018; Georgiou et al., 2025).

This study is subject to several limitations that should be acknowledged. First, the order statistics formulation (model (13)) can become computationally intensive as the number of quantile terms and DMUs increases, leading to higher computation times and memory requirements, particularly in large-scale applications. Second, the correlated budgeted model (model (8)) relies on PCA-derived impact factors to estimate dependence; this assumes approximately linear co-movements and stationarity over the evaluation window and can be sensitive to variable scaling and sample size. Third, robustness parameters such as the uncertainty budget Γ and quantile levels are adjusted exogenously; different settings shift the efficiency–price-of-robustness trade-off and can affect rankings. Fourth, the empirical application to European banks is subject to data and modelling constraints—cross-country accounting heterogeneity, the use of interbank deposits as a proxy for deposit services, sample representativeness, and reliance on a static CCR baseline—which may limit direct generalisability to other applications or regulatory contexts.

These limitations suggest several promising directions for future research. Scalable solution methods such as decomposition, screening, and warm-start techniques could be developed for order statistics models (Zhang & Gupta, 2023), while time-varying dependence structures might be attained through dynamic PCA approaches (Atkins & Cummins, 2023). Further exploration of data-driven adjustment of robustness parameters would also enhance adaptability. Further extensions include applying the framework to dynamic and network DEA settings (Tone & Tsutsui, 2014; Fukuyama et al., 2023), integrating predictive analytics and machine learning techniques (Bertsimas & Kim, 2024), and validating the approach empirically across diverse industry sectors. In addition, extending the proposed robust DEA framework beyond the constant returns to scale assumption to other returns to scale assumptions represents a natural direction for future research, further broadening the applicability of the models. Such explorations would provide additional insights and potentially expand the practical applicability and robustness of these methods for real-world efficiency assessment challenges.

CRediT authorship contribution statement

Adel Hatami-Marbini: Writing – review & editing, Writing – original draft, Validation, Conceptualization. **Aliasghar Arabmaldar:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Amin Hosseinian-Far:** Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adler, N., & Yazhemsky, E. (2010). Improving discrimination in data envelopment analysis: PCA–DEA or variable reduction. *European Journal of Operational Research, 202*(1), 273–284.

Alfiero, S., Esposito, A., Mensah, E. K., & Toloo, M. (2019). A dataset of European banks in performance evaluation under uncertainty. *Data in Brief, 22*, 214–217.

Aparicio, J., Kapelko, M., & Ortiz, L. (2023). Enhancing the measurement of firm inefficiency accounting for corporate social responsibility: A dynamic data envelopment analysis fuzzy approach. *European Journal of Operational Research, 306*(2), 986–997.

Arabmaldar, A., Hatami-Marbini, A., Loske, D., Hammerschmidt, M., & Klumpp, M. (2024). Robust data envelopment analysis with variable budgeted uncertainty. *European Journal of Operational Research, 315*(2), 626–641.

Arabmaldar, A., Jablonsky, J., & Hosseinzadeh Saljooghi, F. (2017). A new robust DEA model and super-efficiency measure. *Optimization, 66*(5), 723–736.

Arabmaldar, A., Sahoo, B. K., & Ghiyasi, M. (2023). A generalized robust data envelopment analysis model based on directional distance function. *European Journal of Operational Research, 311*(2), 617–632.

Atkins, P. J., & Cummins, M. (2023). Improved scalability and risk factor proxying with a two-step principal component analysis for multi-curve modelling. *European Journal of Operational Research, 304*(3), 1331–1348.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.

Ben-Tal, A., & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming, 88*(3), 411–424.

Berger, A. N., & Humphrey, D. B. (1997). Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research, 98*(2), 175–212.

Bertsimas, D., Gupta, V., & Kallus, N. (2018). Data-driven robust optimization. *Mathematical Programming, 167*, 235–292.

Bertsimas, D., & Kim, C. W. (2024). A machine learning approach to two-stage adaptive robust optimization. *European Journal of Operational Research, 319*(1), 16–30.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research, 52*(1), 35–53.

Boďa, M., & Piklová, Z. (2021). Impact of an input-output specification on efficiency scores in data envelopment analysis: A banking case study. *RAIRO - Operations Research, 55*, S1551–S1583.

Charnes, A., & Cooper, W. W. (1962). Programming with linear fractional functionalsle. *Naval Research Logistics, 9*, 181–185.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research, 2*(6), 429–444.

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA) - Thirty years on. *European Journal of Operational Research, 192*(1), 1–17.

Cook, W. D., & Zhu, J. (2007). Classifying inputs and outputs in data envelopment analysis. *European Journal of Operational Research, 180*(2), 692–699.

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

Dehghani Filabadi, M., & Mahmoudzadeh, H. (2022). Effective budget of uncertainty for classes of robust optimization. *INFORMS Journal on Optimization, 4*(3), 249–277.

Despotis, D. K., & Smirlis, Y. G. (2002). Data envelopment analysis with imprecise data. *European Journal of Operational Research, 140*(1), 24–36.

Ehrgott, M., Holder, A., & Nohadani, O. (2018). Uncertain data envelopment analysis. *European Journal of Operational Research, 268*(1), 231–242.

Ferreira, R. S., Barroso, L. A., & Carvalho, M. M. (2012). Demand response models with correlated price data: A robust optimization approach. *Applied Energy, 96*, 133–149.

Fethi, M. D., & Pasiouras, F. (2010). Assessing bank efficiency and performance with operational research and artificial intelligence techniques: A survey. *European Journal of Operational Research, 204*(2), 189–198.

Fukuyama, H., Tsionas, M., & Tan, Y. (2023). Dynamic network data envelopment analysis with a sequential structure and behavioural-causal analysis: Application to the Chinese banking industry. *European Journal of Operational Research, 307*(3), 1360–1373.

Georgiou, A. C., Tsaples, G., & Thanassoulis, E. (2025). Planning methods using data envelopment analysis and markov systems. *European Journal of Operational Research, 326*(3), 569–584.

Gregory, C., Darby-Dowman, K., & Mitra, G. (2011). Robust optimization and portfolio selection: The cost of robustness. *European Journal of Operational Research, 212*(2), 417–428.

Hatami-Marbini, A., Agrell, P. J., Fukuyama, H., Gholami, K., & Khoshnevis, P. (2017). The role of multiplier bounds in fuzzy data envelopment analysis. *Annals of Operations Research, 250*(1), 249–276.

Hatami-Marbini, A., & Arabmaldar, A. (2021). Robustness of Farrell cost efficiency measurement under data perturbations: Evidence from a US manufacturing application. *European Journal of Operational Research, 295*(2), 604–620.

Hatami-Marbini, A., Arabmaldar, A., & Asu, J. O. (2022a). Robust productivity growth and efficiency measurement with undesirable outputs: Evidence from the oil industry. *OR Spectrum, 44*(4), 1213–1254.

Hatami-Marbini, A., Arabmaldar, A., Toloo, M., & Nehrani, A. M. (2022b). Robust non-radial data envelopment analysis models under data uncertainty. *Expert Systems with Applications, 207*.

Hatami-Marbini, A., Hekmat, S., & Agrell, P. J. (2022c). A strategy-based framework for supplier selection: A grey PCA-DEA approach. *Operational Research, 22*(1), 263–297.

Holod, D., & Lewis, H. F. (2011). Resolving the deposit dilemma: A new DEA bank efficiency model. *Journal of Banking & Finance, 35*(11), 2801–2810.

Jackson, J. E. (1991). *A use's guide to principal components*. Wiley.

Kao, C. (2006). Interval efficiency measures in data envelopment analysis with imprecise data. *European Journal of Operational Research, 174*(2), 1087–1099.

Kao, C., & Liu, S.-T. (2022). Stochastic efficiencies of network production systems with correlated stochastic data: The case of Taiwanese commercial banks. *Annals of Operations Research, 315*(2), 1151–1174.

Ke, L., Xu, Z., Feng, Z., Shang, K., & Qian, X. (2013). Proportion-based robust optimization and team orienteering problem with interval data. *European Journal of Operational Research, 226*(1), 19–31.

Kuosmanen, T., & Podinovski, V. (2009). Weak disposability in nonparametric production analysis: Reply to Färe and Grosskopf. *American Journal of Agricultural Economics, 91*(2), 539–545.

Lu, C. C. (2015). Robust data envelopment analysis approaches for evaluating algorithmic performance. *Computers & Industrial Engineering, 81*, 78–89.

Mehdiloo, M., Papaioannou, G., & Podinovski, V. V. (2026). Closed cone extensions of convex and nonconvex production technologies. *European Journal of Operational Research, 332*(2), 575–585.

Mergoni, A., Emrouznejad, A., & De Witte, K. (2025). Fifty years of data envelopment analysis. *European Journal of Operational Research, 326*(3), 389–412.

Olesen, O. B., & Petersen, N. C. (1995). Chance constrained efficiency evaluation. *Management Science, 41*(3), 442–457.

Olesen, O. B., & Petersen, N. C. (2016). Stochastic data envelopment analysis - A review. *European Journal of Operational Research, 251*(1), 2–21.

Park, K. S. (2007). Efficiency bounds and efficiency classifications in DEA with imprecise data. *Journal of the Operational Research Society, 58*(4), 533–540.

Park, K. S. (2010). Duality, efficiency computations and interpretations in imprecise DEA. *European Journal of Operational Research, 200*(1), 289–296.

Pham, M., Simar, L., & Zelenyuk, V. (2023). Statistical inference for aggregation of Malmquist productivity indices. *Operations Research, 72*(4), 1615–1629.

Podinovski, V. V. (2005). Selective convexity in DEA models. *European Journal of Operational Research, 161*(2), 552–563.

Podinovski, V. V., & Bouzdine-Chameeva, T. (2019). Cone extensions of polyhedral production technologies. *European Journal of Operational Research, 276*(2), 736–743.

Podinovski, V. V., Chambers, R. G., Atici, K. B., & Deineko, I. D. (2016). Marginal values and returns to scale for nonparametric production frontiers. *Operations Research, 64*(1), 236–250.

Poss, M. (2013). Robust combinatorial optimization with variable budgeted uncertainty. *4OR, 11*(1), 75–92.

Poss, M. (2014). Robust combinatorial optimization with variable cost uncertainty. *European Journal of Operational Research, 237*(3), 836–845.

Premachandra, I. M. (2001). A note on DEA vs principal component analysis: An improvement to Joe Zhu's approach. *European Journal of Operational Research, 132*(3), 553–560.

Sadjadi, S. J., & Omrani, H. (2008). Data envelopment analysis with uncertain data: An application for Iranian electricity distribution companies. *Energy Policy, 36*(11), 4247–4254.

Salahi, M., Toloo, M., & Torabi, N. (2021). A new robust optimization approach to common weights formulation in DEA. *Journal of the Operational Research Society, 72*(6), 1390–1402.

Shokouhi, A. H., Hatami-Marbini, A., Tavana, M., & Saati, S. (2010). A robust optimization approach for imprecise data envelopment analysis. *Computers and Industrial Engineering, 59*(3), 387–397.

Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science, 44*(1), 49–61.

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research, 21*(5), 1154–1157.

Toloo, M., & Mensah, E. K. (2019). Robust optimization with nonnegative decision variables: A DEA approach. *Computers and Industrial Engineering, 127*, 313–325.

Toloo, M., Mensah, E. K., & Salahi, M. (2022). Robust optimization and its duality in data envelopment analysis. *Omega, 108*, Article 102583.

Tone, K., & Tsutsui, M. (2014). Dynamic DEA with network structure: A slacks-based measure approach. *Omega, 42*(1), 124–131.

Torres-Jiménez, M., García-Alonso, C. R., Salvador-Carulla, L., & Fernández-Rodríguez, V. (2015). Evaluation of system efficiency using the Monte Carlo DEA: The case of small health areas. *European Journal of Operational Research, 242*(2), 525–535.

Tsionas, M. G. (2021). Optimal combinations of stochastic frontier and data envelopment analysis models. *European Journal of Operational Research, 294*(2), 790–800.

Wong, Y.-H. B., & Beasley, J. E. (1990). Restricting weight flexibility in data envelopment analysis. *Journal of the Operational Research Society, 41*(9), 829–835.

Wu, J., Shen, L., Zhang, G., Zhou, Z., & Zhu, Q. (2024). Efficiency evaluation with data uncertainty. *Annals of Operations Research, 339*, 1379–1403.

Zhang, P., & Gupta, D. (2023). Robust optimization with order statistic uncertainty set. *European Journal of Operational Research, 311*(3), 1023–1039.

Zhu, J. (2003). Imprecise data envelopment analysis (IDEA): A review and improvement with an application. *European Journal of Operational Research, 144*(3), 513–529.