Robustness of Farrell cost efficiency measurement under data perturbations: Evidence from a US manufacturing application

Adel Hatami-Marbini[a1], Aliasghar Arabmaldar[b]

[a] Department of Management and Entrepreneurship, Leicester Castle Business School, De Montfort University, Hugh Aston Building, The Gateway, Leicester LE1 9BH, UK
[b] Department of Mathematics, Faculty of Sciences, Golestan University, Gorgan, Iran

Abstract

Measuring economic and cost efficiency receives ever-increasing attention of the executives and managers of small-and medium-sized enterprises (SMEs) to minimise the total production costs. The conventional Farrell cost efficiency (CE) as a key determinant requires the precise information on inputs, outputs and input prices , while in praxis uncertainty is inherent and inevitable in data and its negligence conceivably results in a dire approximation for CE measures. This paper is concerned with Farrell CE in situations of both *endogenous* and *exogenous* uncertainty. The source of uncertainty allows us to define two different scenarios; (*i*) in situations of endogenous uncertainty in input and output data where the uncertainty is affected by the decision maker, and *(ii)* in situations of uncertain prices for inputs where the uncertainty is exogenously given. In the first scenario, the theory of robust optimisation is adopted to develop the robust data envelopment analysis (DEA) models with the aim of grappling uncertainties in input and output data when measuring technical and cost efficiencies. The second scenario aims to accommodate uncertainties on price information by developing a pair of robust DEA models based upon robust optimisation estimating the upper and lower bounds for CE measures. This unprecedented study helps us to provide a generalised framework for economic efficiency with uncertainties in which conventional properties of Farrell measures are fulfilled. In addition to comparing the developed approach in this paper with other existing approaches through a simple numerical example, the usefulness and applicability of the suggested framework are minutely studied in an empirical application in the context of allocation problems.

Keywords: Data envelopment analysis; Cost efficiency; Uncertainty; Robust optimisation.

1. Introduction

Efficiency and productivity assessment lie at the heart of management's strategies and activities. Modern efficiency measurement and benchmarking approaches are mostly based on the economic theory foundation found in pioneering works of Farrell (1957), Debreu (1951) and Koopmans (1951). Farrell (1957) and later Färe, Grosskopf, and Lovell (1985) showed that a *cost* (*economic*) *efficiency* measure (CE) can be calculated by a combination of *technical efficiency* (TE) and *allocative efficiency* (AE), viz., CE=(TE) ×

[1] Corresponding author.
E-mail addresses: adel.hatamimarbini@dmu.ac.uk, https://orcid.org/0000-0002-7956-687X (A. Hatami-Marbini); n.arab@pgs.usb.ac.ir, https://orcid.org/0000-0001-6412-9514 (A. Arabmaldar)

(AE). TE measures the ability of a firm to attain the maximum outputs from a certain set of inputs, while AE measures the ability of a firm to use the inputs in optimal proportions when prices and costs are available. Putting TE and AE all together, CE aims to assess the ability of a firm to produce current outputs at minimal cost. Charnes, Cooper, and Rhodes (1978) initially developed Farrell (1957)'s idea to propose data envelopment analysis (DEA), so-called *CCR model*, which is a mathematical programming technique and aims to assess [relative] TE measures of a set of decision-making units (DMUs) that use multiple inputs to produce multiple outputs given the assumption of constant returns to scale (CRS). Afterwards, Banker, Charnes, and Cooper (1984) extended DEA and introduced the *BCC model* which allows variable returns to scale (VRS). Neglecting price and cost information helps TE studies gain a significant level of popularity and interest in the related literature (see e.g., Cook and Seiford, 2009; Emrouznejad and Yang, 2018). The conventional CE models are not applicable unless information on input prices is exactly known while prices may fluctuate in the short term (Cooper et al. 1996). This data requirement leads to emphatically reduced consideration of CE in real-world applications.

Uncertainty in data is inherently inevitable since the data values are often affected by unknown and unpredictable events. Uncertain data can be divided into *exogenous* and *endogenous* types (Lappas and Gounaris, 2018). Exogenous uncertainty is not often affected by the decision maker while uncertainty is endogenous in a plethora of application settings and subject to the decision maker's decisions, strategies and judgments (see e.g., Peeta et al, 2010; Edirisinghe and Zhang, 2010; Hassanzadeh et al., 2014). For instance, in a monopolistic market, decisions depend on the information that the decision maker has about the [endogenous] uncertainty, whereas the decision maker often faces exogenous uncertainty about decisions in a perfectly comparative market. Decision making under uncertainty can be traced back to the mid-1950s by seminal work of Charnes and Cooper (1959) and Dantzig (1955) who, respectively, developed stochastic programming and optimisation under probabilistic constraints. These classes of optimisation models are built on the assumption that the probability distributions of the random variables are perfectly known. However, in the plethora of real-world problems the decision-maker is not sure whether this assumption is valid due to the lack of ample evidence. Consequently, the literature has frequently advocated the necessity of alternative approaches to cope with uncertainty in rapidly evolving situations where stochastic models cannot be successfully applied. Mathematically speaking, uncertain optimisation problems often face two challenges in practice; *(i)* the solution might not be feasible, and *(ii)* the solution, if feasible, might result in total cost (revenue) that is far greater (smaller) than the truly optimal strategy (Bertsimas and Thiele, 2006). *Robust optimisation* (RO) was initially proposed in the early 1970s by Soyster (1973) for linear programming problems when probabilistic models are intractable. RO aims to find a solution in the presence of uncertain data that would remain feasible for all possible values of the data parameters while attaining the best possible worst-case performance. Numerous real-life problems encompass both *here-and-now* and *wait-and-see* decisions. As such, here-and-now decision variables are those that have to be identified right away, and there is no time to hold on for further or full information on uncertain parameters, whereas wait-and-see variables are associated with those situations where we can wait to get more information on uncertain parameters prior to making decisions. In situations of here-and-now and wait-and-see decisions, static RO and adjustable RO can be applied, respectively, to deal with data uncertainty (Yanıkoğlu et al., 2019).

Since tractability and insightfulness of the optimal solution are of imperative importance in the presence of uncertain data, it is hugely worthwhile to consider exogenous and endogenous uncertainties in CE models by way of two scenarios in this paper. The first scenario is concerned with a short-term focus when prices are deterministic at each DMU and some or all of inputs and/or outputs are assumed to be endogenously uncertain. In such situation, we adopt the approach of static RO to develop robust counterparts of the deterministic Farrell TE and CE models to tackle severe uncertainty in the input-output setting as well as to make here-and-now decisions. This approach features the tractability prominently in order to optimise this worst-case criterion efficiently in a way that attains a high level of robustness (protection against uncertainty) and high-quality objective values (close to the objective of the nominal problem). The second scenario focuses on situations where input and output data are perfectly known but some or all of input prices are exogenously uncertain. In situations of price uncertainty, a pair of static robust CE models are suggested in this paper to obtain bounded intervals for CE measures in which the lower and upper bounds are used to classify and further discriminate the DMUs in terms of the variability of their CE measures. To decrease over-conservatism in some existing RO approaches that can be criticised by practitioners, Bertsimas and Sim (2004)'s approach is adapted in this paper to wholly control the conservatism level for each constraint through a parameter that can be adjusted by the decision-maker. Throughout this paper, it is showed the developed models provide an effective framework for studying technical, cost, and allocative efficiencies under uncertainty along with being a practical tool to determine best practices and construct achievable targets based upon observations and technology characteristics.

The paper unfolds as follows. Section 2 reviews the related literature and Section 3 provides the methodological DEA background for measuring TE and CE measures. Section 4 develops the robust counterpart of TE and CE measures in two different situations; endogenous uncertainty in input and/or output data and exogenous uncertainty in input prices. In the latter situation, a pair of mathematical programming models is presented to estimate the upper and lower bounds of CE and AE measures. Section 5 compares the proposed models with three existent studies in the literature. Section 6 contains an illustrative case study of the optimal allocation of US manufacturing to accentuate the applicability of the models developed in this paper. Section 7 concludes the current study and gives some directions for future research.

2. Literature review

DEA as a non-parametric technique plays a pivotal role in the literature of performance evaluation, particularly in situations where multiple precise inputs are consumed to produce multiple precise outputs. A part of the existing DEA literature has investigated CE where precise information of input prices is available (see e.g., Ray et al., 2008; Sahoo, Mehdiloozad, and Tone, 2014; Thanassoulis, Shiraz, and Maniadakis, 2015; Ray, 2016; Simar and Wilson, 2020). In the CE literature, only a few studies have considered exogenous uncertainty in input prices as reviewed concisely here. To the best of our knowledge, the original thrust can be found in Schaffnit et al. (1997) in which the pricing information is partially known as the bounds. Alternatively, Kuosmanen and Post (2001, 2003) developed DEA models to measure the upper and lower bounds of CE measures in which incomplete information on prices is represented by the polyhedral convex cone. Afterwards, some studies in situations of price bounds have been conducted to obtain a bounded interval of CE measures from the pessimistic and optimistic standpoints (see e.g., Camanho and Dyson, 2005; Toloo and Ertay, 2014). Camanho and Dyson (2005) employed the standard weight restriction techniques by way of input cone assurance regions to obtain the upper and lower bounds for CE measures from optimistic and pessimistic views, respectively. When assessing the lower bound of CE measures, the authors flagged up that the proposed model is often skewed in the presence of multiple outputs. In line with those foregoing studies, a few authors mounted the level of complexity by considering uncertainty in input and output data, and proposed DEA models to attain the lower and upper bounds of CE measures with precise and imprecise input prices (Mostafaee and Saljooghi 2010; Khanjani Shiraz et al., 2018). Mostafaee and Saljooghi (2010) first explored a situation where the input and output data are known to vary within bounded intervals, and their proposed method aimed at estimating the upper and lower bounds for CE measures. They then generalised the idea to a case where all the inputs, outputs and input prices lie within intervals. When the input prices are characterised by a convex set, they also showed that the upper and lower bounds of the CE can arrive at extreme points of the convex set even though Fang and Li (2012) provided some counterexamples to show that this is not invariably valid. Fang and Li (2013a, 2013b) developed an alternative model to estimate the lower bound of CE measures with the aim of addressing the shortcomings of Camanho and Dyson (2005) and Mostafaee and Saljooghi (2010). A detailed review of Camanho and Dyson (2005), Mostafaee and Saljooghi (2010), and Fang and Li (2013b) will be presented in Section 5. It might be worth noting here that assessing units based upon the bounded intervals of TE and CE measures would be questionable since much information and discriminatory power are lost (Shokouhi et al., 2010).

Making use of fuzzy logic can be viewed in Bagherzadeh Valami (2009) to gauge fuzzy CE measures in which input prices are characterised by triangular fuzzy numbers. The classical CE and revenue efficiency models with undesirable outputs were focused on by Puri and Yadav (2016) and further extended to fully fuzzy environments where the inputs, outputs, and input prices are presented by triangular fuzzy numbers. The developed fully fuzzy CE and revenue efficiency models were transformed into a family of conventional crisp models using a defuzzification approach in order to assess CE and revenue efficiency measures. Likewise, the developed CE model of Pourmahmoud and Bafekr Sharak (2018) was for fully fuzzy environments to deal with imprecise and vague data and input prices, simultaneously, using the α-level approach. Needless to say, the fuzzy approach often faces some controversial issues such as the deprivation of a clear method for defining the membership functions of fuzzy inputs and outputs, and some theoretical and computational burdens (Hatami-Marbini et al., 2011). Table 1 presents a summary of the aforesaid key DEA studies on imprecise the CE models in terms of their basic characteristics.

Table 1. Summary of existing CE studies under uncertainty

Study	Model’s form	Uncertainty		CE measures	Model’s type
		Input prices	Inputs and Outputs		
Schaffnit et al. (1997)	Multiplier	Bounded interval	Precise	Upper bound	LP*
Kuosmanen and Post (2001)	Envelopment	Convex polyhedral cone	Precise	Bounded interval	LP
Camanho and Dyson (2005)	Multiplier	Bounded interval	Precise	Bounded interval	LP
Bagherzadeh Valami (2009)	Envelopment	Fuzzy numbers	Precise	Fuzzy CE	FLP¢
Mostafaee and Saljooghi (2010)	Multiplier/ Envelopment	Bounded interval	Bounded interval	Bounded interval	LP
Fang and Li (2013a, 2013b)	Multiplier	Bounded interval	Precise	Lower bound	BLPŋ
Toloo and Ertay (2014)	Multiplier	Bounded interval	Precise	Bounded interval	MILP↑
Puri and Yadav (2016)	Envelopment	Fuzzy numbers	Fuzzy numbers	Precise CE	FFLP
Pourmahmoud and Bafekr Sharak (2018)	Envelopment	Fuzzy numbers	Fuzzy numbers	Fuzzy CE	FFLP
Khanjani Shiraz et al. (2018)	Envelopment	Precise	Stochastic	Stochastic	QPP†

*LP: Linear programming, ¢FLP: Fuzzy LP, ŋBLP: Bi-level LP, ↑MILP: Mixed integer LP, †QPP: Quadratic programming problem.

Robust optimisation (RO) was begun in operational research by Soyster (1973) in situations of uncertain data. The aim of RO is to look for a robust solution that remains feasible for all various values of the uncertain parameters as well as being the best possible worst-case performance. Soyster (1973) considered the worst-case value of each uncertain parameter in the linear programming (LP) model to find a solution that is immune against data uncertainty. However, Soyster (1973)’s model has been practically and correctly criticised due to being too conservative. To tackle the problem of over-conservatism, several significant attempts were made by Ben-Tal and Nemirovski (1998, 1999, 2000) and El-Ghaoui and Lebret (1997) and El-Ghaoui, Oustry, and Lebret (1998) to shape the underlying principles for modern RO models that can be applied to real-life problems by controlling the conservatism level with ellipsoidal uncertainty sets[2]. The considerable weakness of resulting RO models subject to ellipsoidal uncertainty sets is that the derived robust counterpart is computationally expensive and complex in comparison to the nominal counterpart due to the added nonlinearity. For instance, the robust counterpart of LP problems is formulated as second-order cone programming problems (SOCPs). The outstanding thrust of introducing a family of polyhedral uncertainty sets[3] by Bertsimas and Sim (2004) reduces over-conservatism along with fully controlling the level of conservatism for each constraint through a parameter that can be fine-tuned by the decision-maker. The aim of Bertsimas and Sim (2004)'s work is to immunise a solution against uncertain parameters to give the chance to take their worst-case values. Above all, the robust counterpart of an LP model in Bertsimas and Sim (2004) remains linear, providing the opportunity to take advantage of computational efficiency. In particular, Gabrel and Murat (2010) studied the robustness and duality of LP problems with uncertain coefficients in the right-hand side in which the duality was first used to transfer uncertainty from the right-hand side to the objective function and they went on with the idea of Bertsimas and Sim (2004) and generalised the best-case scenario for uncertain parameters. In light of the advantages made by the approach of Bertsimas and Sim (2004), many researchers, scholars and practitioners have widely used it in different areas such as portfolio selection (see e.g., Gregory, Darby-Dowman, and Mitra, 2011), production planning (see e.g., Lai, Xu, and Shang, 2019), logistics and transportation (see e.g., Marla, Vaze, and Barnhart, 2018), and performance assessment. Let us now draw attention to the relevant RO studies that are available in the DEA literature.

A survey has been conducted by Peykani et al. (2020) to provide an overarching look at DEA models through the use of RO. As far as we are aware, the RO approaches of Bertsimas and Sim (2004) and Ben-Tal and Nemirovski (2000) have been initially adapted to the traditional CCR model by Sadjadi and Omrani (2008). It should be noted that their developed models can be only employed for situations where output data is uncertain. Afterwards, there have been several robust DEA studies based upon Bertsimas and Sim (2004) as reported in Table 2. In effect, the concept of Bertsimas and Sim (2004) has been successfully applied to a range of existing DEA models. The robust CCR model of Sadjadi and Omrani (2008) was featured by bootstrapping in Sadjadi and Omrani (2010) to overcome the perturbation effect and sampling error. Shokouhi et al. (2010, 2014) focused on interval CCR models developed by Despotis and Smirlis (2002) and Wang et al. (2005) in which uncertain input and output data are partially known and lie within the upper and lower bounds. The authors first used Bertsimas and Sim (2004)'s approach to develop their non-linear robust DEA model and then applied Monte-Carlo simulation to obtain the conformity level of

---

[2] Mulvey et al. (1995) developed a scenario-based method which was also called robust optimisation. This is a different approach that was built on the stochastic programming problem aiming to optimise a weighted combination of the stochastic objective function and a penalty function, which penalises violations of the constraints.

[3] Note that this type of uncertainty sets has been dubbed "polyhedral sets", "budgeted uncertainty sets" or "Bertsimas and Sim uncertainty sets" in the literature (Ben-Tal et al., 2009).

the rankings of DMUs. Sadjadi et al. (2011a) laid emphasis on Estellita et al. (2004) to develop an interactive robust DEA model based upon Bertsimas and Sim (2004) that can attend to uncertainty on both inputs and outputs. Their study aims to identify correct targets for inputs and outputs when the decision makers' preferences matter. In the presence of uncertain input and output data, Omrani (2013) leveraged Bertsimas and Sim (2004)'s approach to introduce the robust counterpart of a common weight DEA model of Zohrehbandian et al. (2010) to find the common set of weights (CSWs) for all DMUs. Similar to the study conducted by Sadjadi and Omrani (2008), Lu (2015) presented robust BCC models based on Ben-Tal and Nemirovski (2000) and Bertsimas and Sim (2004) to deal with uncertain desirable and undesirable outputs. Moreover, they considered their models to assess a group of parameter settings of a genetic algorithm and a simulated annealing heuristic. The emphasis of Aghayi and Maleki (2016) was on the directional distance function model developed by Zanella et al. (2015) to investigate the interval and robust DEA models by capturing the ideas of Despotis and Smirlis (2002) and Bertsimas and Sim (2004), respectively. One limitation of Sadjadi and Omrani (2008) is that outputs were assumed to be uncertain only. To address the obstacle, Arabmaldar et al. (2017) regarded an effective way to develop a robust DEA model that combats uncertainty on both input and output data. Furthermore, Arabmaldar et al. (2017) established a robust super-efficiency DEA model using the approach of Bertsimas and Sim (2004). The VRS version of Arabmaldar et al. (2017)'s model can be found in Toloo and Mensah (2019). Of late, Salahi et al. (2020) first obtained the robust CCR solutions based on the approach of Bertsimas and Sim (2004) and then controlled the weight flexibility by finding the robust CSWs. The robust cross-efficiency DEA approach proposed by Tavana et al. (2020) is based upon Bertsimas and Sim (2004), aiming at circumventing degenerate optimal weights and uncertain data in the presence of undesirable outputs. Some researchers such as Sadjadi et al. (2011b), Lu et al. (2020) and Mensah (2020) have also tried to develop robust DEA models based upon the ellipsoidal uncertainty set, which is a different class of uncertainty sets proposed by Ben-Tal and Nemirovski (1999, 2000). Sadjadi et al. (2011b) proposed a robust super-efficiency DEA model in which data uncertainty is assumed to be within the ellipsoidal uncertainty set. The ellipsoidal uncertainty set was also presumed by Lu et al. (2020) to develop a second-order cone based robust DEA model with translation invariance property. Given the ellipsoidal and interval-based ellipsoidal uncertainty sets, Mensah (2020) first proposed a robust CCR and robust additive models and then explored the relationship between the proposed models under the two uncertainty sets. While here our focus is mostly limited to robust DEA models based upon Bertsimas and Sim (2004), a superb and comprehensive review of robust DEA models can be found in Mensah (2019). Table 2 shows some key differences among the aforesaid approaches and current research in terms of some criteria. Notably, the computational efficiency of the developed models can be observed in the last column of Table 2.

Table 2. Summary of existing robust DEA models based on BS* and BN†

Study	Model's form	Basic DEA model	Uncertain variables	Robust approach	Model's type
Sadjadi & Omrani (2008)	Multiplier	CCR (Charnes et al., 1978)	Outputs	BS[*] & BN[†]	NLP[♯] & LP
Shokouhi et al. (2010)	Multiplier	Interval CCR (Despotis and Smirlis, 2002)	Inputs & outputs	BS	NLP
Sadjadi & Omrani (2010)	Multiplier	CCR (Charnes et al., 1978)	Outputs	BS	LP
Sadjadi et al. (2011a)	Envelopment	MORO[ƛ] (Estellita et al., 2004)	Inputs & outputs	BS	MOLP[$]
Sadjadi et al. (2011b)	Envelopment	Super-efficiency (Andersen and Petersen, 1993)	Inputs & outputs	BN	SOCP[¢]
Omrani (2013)	Multiplier	CWDEA[ª] (Zohrehbandian et al., 2010)	Inputs & outputs	BS	LP
Shokouhi et al. (2014)	Multiplier	Interval CCR (Wang et al., 2005)	Inputs & outputs	BS	LP
Lu (2015)	Multiplier	BCC (Banker et al., 1984)	Outputs	BN & BS	SOCP & LP
Aghayi & Maleki (2016)	N/A	DDF[ʧ] (Zanella et al., 2015)	Inputs & outputs	BS	NLP
Arabmaldar et al. (2017)	Multiplier & envelopment	CCR (Charnes et al., 1978) & super-efficiency (Andersen and Petersen, 1993)	Inputs & outputs	BS	LP
Toloo & Mensah (2019)	Multiplier	BCC (Banker et al., 1984)	Inputs & outputs	BS	LP
Lu. et al. (2020)	Envelopment	MORO (Estellita et al., 2004)	Inputs & outputs	BN	SOCP
Salahi et al. (2020)	Multiplier & envelopment	CWDEA (Zohrehbandian et al., 2010)	Inputs & outputs	BS	LP
Mensah (2020)	Multiplier	CCR (Charnes et al., 1978) & additive model (Charnes et al., 1985)	Inputs & outputs	BN	SOCP
Tavana et al. (2020)	Multiplier	Interval CCR (Despotis and Smirlis, 2002)	Inputs & outputs	BS	NLP
The proposed approach	Envelopment	CCR (Charnes et al., 1978) & CE Farrell (1957)	Inputs & outputs Input prices	BS & GM[🠅]	LP & CQP[ŋ]

[†]BN: Ben-Tal and Nemirovski (1999, 2000), [*]BS: Bertsimas & Sim (2004), [♯]NLP: Non-linear programming, [ƛ]MORO: Multi-objective ratio optimisation, , [$]MOLP: Multi-objective LP, [¢]SOCP: Second-order cone programming problem, [ª]CWDEA: Common weight DEA, [ʧ]DDF: Directional distance function, [🠅]GM: Gabrel and Murat (2010), [ŋ]CQP: Concave quadratic programming.

The foregoing literature demonstrates the lack of ample attention to CE analysis in situations of uncertainty, particularly when the inputs and outputs are uncertain, and their probability distribution functions are unknown. Hence, the current paper shows that the RO approach is emphatically appropriate for tackling uncertainty in the input and output data and input prices while measuring TE and CE measures of DMUs.

3. TE and CE models

This section is designed to review the basic DEA models which are commonly used for measuring TE and CE measures of DMUs.

3.1. TE measure

Assume that there are *n* DMUs to be assessed with *m* inputs and *s* outputs. For $\text{DMU}_j; j = 1, \dots, n$, $\boldsymbol{y}_j = (y_{1j}, \dots, y_{sj})^T \in \mathbb{R}_+^s$ are produced by using $\boldsymbol{x}_j = (x_{1j}, \dots, x_{mj})^T \in \mathbb{R}_+^m$, where superscript T represents the transpose of a vector, and x_{ij} and y_{rj} $(i = 1, \dots, m; r = 1, \dots, s; j = 1, \dots, n)$ represent, respectively, the value of input i and output r of DMU_j. Let $X = [x_1, x_2, \dots, x_n]$ and $Y = [y_1, y_2, \dots, y_n]$ denote $m \times n$ and $s \times n$ matrices of inputs and outputs, respectively. The production possibility set (PPS), T_c, contains all input and output combinations. The boundary of this set is called the *production frontier* and the assumptions as per technology T_c articulate that: *(i)* there is no free lunch, *(ii)* the set of feasible input-output combinations encompasses all the points on its boundary (closedness), *(iii)* observed inputs and outputs are freely (or strongly) disposable, and *(iv)* PPS is convex. In addition, technology T_c is assumed to satisfy returns to scale. The frontier is said to exhibit constant returns to scale (CRS) if $(\boldsymbol{x}, \boldsymbol{y}) \in T_c$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_j) \in [0,1] \Rightarrow (\boldsymbol{\lambda x}, \boldsymbol{\lambda y}) \in T_c$. In view of the *minimum extrapolation principle*, the CRS technology can be mathematically defined as $T_c^{CRS} = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}_+^{m+s} | \boldsymbol{x} \geq \boldsymbol{\lambda} X, \boldsymbol{y} \leq \boldsymbol{\lambda} Y, \boldsymbol{\lambda} \geq \boldsymbol{0}\}$.

Following Farrell's idea of measuring efficiency as a proportional improvement opportunity, the input-oriented TE of a given DMU_o is calculated using the following LP problem (Charnes et al., 1978):

$$\theta_o^{TE} = \min_{\theta_o, \lambda} \left\{ \theta_o : \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \boldsymbol{x}_j + (\lambda_o - \theta_o) \boldsymbol{x}_o \leq 0, \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \boldsymbol{y}_j + (\lambda_o - 1) \boldsymbol{y}_o \geq 0, \lambda_j \geq 0, \forall j \right\}. \tag{1}$$

Notice that $(\lambda_o = 1, \lambda_j = 0; j \neq o, \theta_o = 1)$ as a feasible solution of model (1) implies that the model is feasible and bounded $(\theta_o^* \in (0,1])$.

Theorem 1. If $(\theta_o^*, \lambda_o^*)$ is the optimal solution of model (1), then $\theta_o^* \geq \lambda_o^*$ and consequently $\lambda_o^* \leq 1$.

Proof. Consider the first set of constraints in model (1). Obviously, it holds when $(\lambda_o^* - \theta_o^*) \leq 0$, viz., $\theta_o^* \geq \lambda_o^*$. Since $\theta_o^* \leq 1$, we have $\lambda_o^* \leq 1$ and this completes proof.□

3.2. CE measure

TE measures are often appropriate for situations where unit price and unit cost information are either unavailable or of questionable values due to high variability in prices and costs or measurement errors. However, if reliable and perfect price information is available, DEA is perfectly capable of identifying cost and allocative efficiencies. Farrell (1957) initially introduced the concept of *overall (cost) efficiency (CE)* and showed that CE can be decomposed into TE and *allocative efficiency (AE)*, viz. CE= *TE* × *AE*.

In the above-defined decomposition, the CE measure is the minimum cost for the current outputs concerning input prices, which can be calculated from the following model (Färe et al., 1985):

$$\omega_o = \min_{\boldsymbol{\lambda}, \boldsymbol{x}^0} \left\{ \boldsymbol{c}_o \boldsymbol{x}^0 : \sum_{j=1}^{n} \lambda_j \boldsymbol{x}_j \leq \boldsymbol{x}^0, \sum_{j=1}^{n} \lambda_j \boldsymbol{y}_j \geq \boldsymbol{y}_o, \lambda_j \geq o, \boldsymbol{x}^0 \geq 0, \forall j \right\}, \tag{2}$$

where $\boldsymbol{c}_o = (c_{1o}, c_{2o}, \dots, c_{mo}) \in \mathbb{R}_+^m$ is a given vector of the input prices for DMU_o under assessment and $\boldsymbol{x}^0 = (x_1^0, x_2^0, \dots, x_m^0)^T$ is a variable vector that, at optimality, yields the minimum input costs for DMU_o for producing the current outputs. Given the optimal costs ω_o^* obtained from the CE model (2), the CE measure of DMU_o is then calculated as the ratio of minimum cost with current prices, i.e., $\omega_o^{CE} = \frac{\omega_o^*}{\boldsymbol{c}_o \boldsymbol{x}_o}$. It should be noted that in the CE model (2), the feasible solution ($\lambda_o = 1, \lambda_j = 0; j \neq o, \boldsymbol{x}^0 = \boldsymbol{x}_o$) implies the feasibility and boundedness of the CE measure ($\omega_o^{CE} \in (0,1]$).

4. Developed robust efficiency assessment models

The traditional DEA models for obtaining TE and CE measures are assumed to use deterministic input and output data as well as complete and perfect input prices for all DMUs. However, volatility, uncertainty, and perturbation in data are inevitable over real-world problems. For instance, in an open electricity market, cost and demand fluctuations are increasingly translated into price fluctuations. Therefore, the treatment of the effects of volatility and uncertainty is indispensable to achieve more accurate and reliable results. Robust optimisation (RO) is a powerful and established tool that has been widely used in the engineering field to handle volatility and uncertainty in data, particularly when the data distributions are unknown. One of the most popular RO approaches in the literature was proposed by Bertsimas and Sim (2004) whereby a family of polyhedral uncertainty sets is defined in order to preserve the linearity of the linear optimisation problem in question, and Bertsimas and Sim (2004)'s approach thus takes advantages of tractability in large-scale settings. This section aims to deal with two distinct situations independently that can be viewed in the problems of performance analysis. The first situation occurs when some or all of input and/or output data are endogenously imprecise. For such situation, the RO approach is adopted to develop the new robust TE and CE models. For example, the factor of *the number of employees* as a commonly used variable in the CE analysis of universities is often assumed to be fixed and exact. However, the number of employees embraces unescapable perturbations since, in reality, employees repeatedly quit their jobs for a wide variety of reasons, and recruiting and training a new employee are time-consuming. The second situation occurs when input and output data are deterministic while the prices of inputs are exogenously uncertain. To tackle this situation, we develop a pair of new CE models to obtain the lower and upper bounds of the CE measure based upon the concept of RO. The unit cost of equity and debt capital as an example for this situation are often uncertain subject to market volatility, economic factors, risk premium, and so on.

4.1. Data uncertainty

Uncertainty is often inherent and inevitable in real-world problems on the one hand, and a small perturbation in data can lead to misleading optimal solutions in LP problems, unreliable efficiency measures or rankings in DEA models on the other hand (see e.g., Ben-Tal and Nemirovski, 2000 and Simar and Wilson, 1998). Therefore, considering robust counterparts of DEA models are essential to immunise the quality of results against data uncertainty. This subsection aims to develop robust counterparts of the deterministic TE model (1) and the deterministic CE model (2) where some or all of input and/or output data are merely assumed to be uncertain.

4.1.1. Robust TE measure

Let us here present the mathematical details of the robust TE model under endogenous uncertainty that can be derived from model (1). Assume that the true values of the uncertain input and output data are expressed as $\tilde{\boldsymbol{x}}_j = \boldsymbol{x}_j + \boldsymbol{\xi}_j^x \hat{\boldsymbol{x}}_j$ and $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j + \boldsymbol{\xi}_j^y \hat{\boldsymbol{y}}_j$ where $\boldsymbol{\xi}_j^x = \left(\xi_{1j}^x, \xi_{2j}^x, \dots, \xi_{mj}^x\right) \in \mathbb{R}^m$, and $\boldsymbol{\xi}_j^y = \left(\xi_{1j}^y, \xi_{2j}^y, \dots, \xi_{sj}^y\right) \in \mathbb{R}^s$ are the independent random variables which take values in the interval $[-1, 1]$, and $\hat{\boldsymbol{x}}_j = e_x \boldsymbol{x}_j$ and $\hat{\boldsymbol{y}}_j = e_y \boldsymbol{y}_j$ are referred to as the *maximum deviations*. Note that e_x and e_y, respectively, are the percentage of perturbation that determines the deviation amount of the uncertain input and output data from their true values. The total deviations in relation to the uncertain inputs and outputs are calculated as $\boldsymbol{\xi}_j^x = \left(\tilde{\boldsymbol{x}}_j - \boldsymbol{x}_j\right)/\hat{\boldsymbol{x}}_j$ and $\boldsymbol{\xi}_j^y = \left(\tilde{\boldsymbol{y}}_j - \boldsymbol{y}_j\right)/\hat{\boldsymbol{y}}_j$ that can take values at most the level of uncertainty parameters $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$, respectively, i.e., $\sum_{j=1}^n |\boldsymbol{\xi}_j^x| \leq \boldsymbol{\Gamma}^x$ and $\sum_{j=1}^n \left|\boldsymbol{\xi}_j^y\right| \leq \boldsymbol{\Gamma}^y$. Note that $\boldsymbol{\Gamma}^x = (\Gamma_1^x, \dots, \Gamma_m^x) \in \mathbb{R}_+^m$ and $\boldsymbol{\Gamma}^y = \left(\Gamma_1^y, \dots, \Gamma_s^y\right) \in \mathbb{R}_+^s$ are also called the vectors of *budget of uncertainty parameters* that take values within $[0, n]$. In the light of the aforesaid setting in the presence of uncertainty, model (1) for $\mathrm{DMU_o}$ is transformed as follows:

$$\theta_o^{RTE} = \min_{\theta_o, \boldsymbol{\lambda}} \{\theta_o : \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \tilde{\boldsymbol{x}}_j + (\lambda_o - \theta_o)\tilde{\boldsymbol{x}}_o \leq 0, \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \tilde{\boldsymbol{y}}_j + (\lambda_o - 1)\tilde{\boldsymbol{y}}_o \geq 0, \lambda_j \geq 0, \forall j\,\}. \tag{3}$$

Bertsimas and Sim (2004)'s approach is adopted to protect the constraints against the worst-case scenario, and, resultantly, the robust counterpart of model (1) is developed as follows:

$$\theta_o^{RTE} = \min_{\theta_o, \boldsymbol{\lambda}} \left\{\theta_o : \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{x}_j + (\lambda_o - \theta_o)\boldsymbol{x}_o + \beta_x \leq 0, \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{y}_j + (\lambda_o - 1)\boldsymbol{y}_o + \beta_y \geq 0, \lambda_j \geq 0, \forall j \right\}, \tag{4}$$

where $\beta_x = \max_{g^x} \left\{\sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{\xi}_j^x \hat{\boldsymbol{x}}_j + (\lambda_o - \theta_o)\boldsymbol{\xi}_o^x \hat{\boldsymbol{x}}_o\right\}$ and $\beta_y = \min_{g^y} \left\{\sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{\xi}_j^y \hat{\boldsymbol{y}}_j + (\lambda_o - 1)\boldsymbol{\xi}_o^y \hat{\boldsymbol{y}}_o\right\}$ are the *protection functions* associated with the uncertain input and output data, respectively, and $g^x = \left\{\boldsymbol{\xi}_j^x \middle| \sum_{j=1}^n |\boldsymbol{\xi}_j^x| \leq \boldsymbol{\Gamma}^x\right\}$ and $g^y = \left\{\boldsymbol{\xi}_j^y \middle| \sum_{j=1}^n \left|\boldsymbol{\xi}_j^y\right| \leq \boldsymbol{\Gamma}^y\right\}$. Let us first focus on the protection function of the input constraint, β_x, in model (4). Given the optimal solution vector $\boldsymbol{\lambda}^*$ and θ_o^*, the protection function β_x^* can be formulated as the following LP problem:

$$\beta_x^* = \max_{\xi_j^x}\left\{\sum_{\substack{j=1\\ j\neq o}}^{n} |\lambda_j^*| \boldsymbol{\xi}_j^x \widehat{\boldsymbol{x}}_j + |(\lambda_o^* - \theta_o^*)| \boldsymbol{\xi}_o^x \widehat{\boldsymbol{x}}_o : \sum_{j=1}^{n} \boldsymbol{\xi}_j^x \leq \boldsymbol{\Gamma}^x,\ 0 \leq \boldsymbol{\xi}_j^x \leq 1, \forall j \right\}. \tag{5}$$

By applying the strong duality theory, model (5) is transformed into a linear minimisation problem. In this regard, let $\boldsymbol{p}^x = (p_1^x, \dots, p_m^x)$ and $\boldsymbol{q}_j^x = (q_{1j}^x, \dots, q_{mj}^x)$, which are the corresponding vectors of dual variables for the first and second constraints of model (5), respectively. Given that $(\boldsymbol{\lambda}^*, \theta_o^*)$ is the optimal solution of model (5), $\alpha_j^x = |\lambda_j^*|, j \neq o$ and $\alpha_o^x = |(\lambda_o^* - \theta_o^*)|$ at optimality. The dual of model (5) is given below:

$$\beta_x^{D*} = \min_{\boldsymbol{p}^x, \boldsymbol{q}_j^x}\{\varphi^x : \boldsymbol{p}^x + \boldsymbol{q}_j^x \geq \widehat{\boldsymbol{x}}_j \alpha_j^x;\ \forall j \neq o, \boldsymbol{p}^x + \boldsymbol{q}_o^x \geq \widehat{\boldsymbol{x}}_o \alpha_o^x, -\alpha_j^x \leq \lambda_j^* \leq \alpha_j^x, -\alpha_o^x \leq \lambda_o^* - \theta_o^* \leq \alpha_o^x, \boldsymbol{p}^x \geq 0, \boldsymbol{q}_j^x \geq 0, \forall j\}, \tag{6}$$

where $\varphi^x = \boldsymbol{\Gamma}^x \boldsymbol{p}^x + \sum_{\substack{j=1\\ j\neq o}}^{n} \boldsymbol{q}_j^x + \boldsymbol{q}_o^x$. Since model (5) for all $\boldsymbol{\Gamma}^x$ is feasible and bounded, therefore its dual problem, i.e. model (6) is evidently feasible and bounded, and their objective function values are identical. From Theorem 1, $\theta_o^* \geq \lambda_o^*$ and consequently model (6) can be simplified as follows:

$$\beta_x^{D*} = \min_{\boldsymbol{p}^x, \boldsymbol{q}_j^x, \theta_o, \boldsymbol{\lambda}}\{\varphi^x : \boldsymbol{p}^x + \boldsymbol{q}_j^x \geq \widehat{\boldsymbol{x}}_j \lambda_j^*; \forall j \neq o, \boldsymbol{p}^x + \boldsymbol{q}_o^x \geq \widehat{\boldsymbol{x}}_o(\theta_o^* - \lambda_o^*), \boldsymbol{p}^x \geq 0, \boldsymbol{q}_j^x \geq 0,\ \forall j\}. \tag{7}$$

Let us return to the protection function β_y in model (4). Similarly, we have:

$$\beta_y^* = -\max_{\xi_j^y}\left\{\sum_{\substack{j=1\\ j\neq o}}^{n} |\lambda_j^*| \boldsymbol{\xi}_j^y \widehat{\boldsymbol{y}}_j + |\lambda_o^* - 1| \boldsymbol{\xi}_o^y \widehat{\boldsymbol{y}}_o : \sum_{j=1}^{n} \boldsymbol{\xi}_j^y \leq \boldsymbol{\Gamma}^y, 0 \leq \boldsymbol{\xi}_j^y \leq 1,\ \forall j \right\}. \tag{8}$$

Using the strong duality theory and Theorem 1, we have:

$$\beta_y^{D*} = \min_{\boldsymbol{p}^x, \boldsymbol{q}_j^x, \boldsymbol{\lambda}}\{\varphi^y : \boldsymbol{p}^y + \boldsymbol{q}_j^y \geq \widehat{\boldsymbol{y}}_j \lambda_j^*;\ \forall j \neq o, \boldsymbol{p}^y + \boldsymbol{q}_o^y \geq \widehat{\boldsymbol{y}}_o(1 - \lambda_o^*), \boldsymbol{p}^y \geq 0, \boldsymbol{q}_j^y, \geq 0, \forall j\}, \tag{9}$$

where $\lambda_o^* \leq 1$ and $\varphi^y = \boldsymbol{\Gamma}^y \boldsymbol{p}^y + \sum_{\substack{j=1\\ j\neq o}}^{n} \boldsymbol{q}_j^y + \boldsymbol{q}_o^y$. Substituting models (7) and (9) contemporaneously into model (4) yields the robust TE (θ_o^{RTE}) model as follows:

$$\theta_o^{RTE} = \min_{\theta_o, \boldsymbol{\lambda}, \boldsymbol{p}^x, \boldsymbol{p}^y, \boldsymbol{q}_j^x, \boldsymbol{q}_j^y}\left\{\theta_o : \sum_{\substack{j=1\\ j\neq o}}^{n} \lambda_j \boldsymbol{x}_j + (\lambda_o - \theta_o)\boldsymbol{x}_o + \varphi^x \leq 0, \sum_{\substack{j=1\\ j\neq o}}^{n} \lambda_j \boldsymbol{y}_j + (\lambda_o - 1)\boldsymbol{y}_o - \varphi^y \geq 0, \boldsymbol{p}^x + \boldsymbol{q}_j^x \geq \widehat{\boldsymbol{x}}_j \lambda_j, \boldsymbol{p}^y + \boldsymbol{q}_j^y \geq \widehat{\boldsymbol{y}}_j \lambda_j; \forall j \neq o, \boldsymbol{p}^x + \boldsymbol{q}_o^x \geq \widehat{\boldsymbol{x}}_o(\theta_o - \lambda_o), \boldsymbol{p}^y + \boldsymbol{q}_o^y \geq \widehat{\boldsymbol{y}}_o(1 - \lambda_o), \lambda_j, \boldsymbol{p}^x, \boldsymbol{p}^y, \boldsymbol{q}_j^x, \boldsymbol{q}_j^y \geq 0, \forall j \right\}, \tag{10}$$

where non-negative φ^{x} and φ^{y} aim to immunise the model against the violation of ith and rth constraints, respectively.

Theorem 2. *(i)* Model (10) is always feasible and *(ii)* the optimal objective function value of model (10) varies within (0, 1].

Proof. *(i)* Assume that $\lambda_o = 1$, $\lambda_j = 0\ (\forall j, j \neq o)$, $\theta_o = 1$, and $\boldsymbol{\Gamma}^y, \boldsymbol{\Gamma}^x, \boldsymbol{p}^x, \boldsymbol{q}_j^x, \boldsymbol{p}^y, \boldsymbol{q}_j^y = 0, (\forall j)$. This feasible solution completes the proof of the first part of the theorem.

(ii) Let us first prove that $\theta_o^* \leq 1$. Considering the feasible solution presented in (*i*) and model (10) as a minimisation model, the objective function value (θ_o^{RTE*})is invariably smaller than or equal to 1. On the other hand, it is shown that $\theta_o^* > 0$ by contradiction. In doing so, assume that $\theta_o^* = 0$ and the input constraints are changed to $\sum_{j=1, j\neq o}^{n} \lambda_j^* \boldsymbol{x}_j + (\lambda_o^* - 0)\boldsymbol{x}_o + \varphi^{x*} \leq 0 \Rightarrow \sum_{j=1}^{n} \lambda_j^* \boldsymbol{x}_j + \varphi^{x*} \leq 0 \Rightarrow \sum_{j=1}^{n} \lambda_j^* \boldsymbol{x}_j + \underbrace{\boldsymbol{\Gamma}^x \boldsymbol{p}^{x*} + \sum_{\substack{j=1 \\ j\neq o}}^{n} \boldsymbol{q}_j^{x*} + \boldsymbol{q}_o^{x*}}_{\varphi^{x*} \geq 0} \leq 0$. Since the nonzero (i.e., semipositive) assumption for the data and φ^{x*}, we have $\lambda_j^* = \boldsymbol{p}^{x*} = \boldsymbol{q}_{j\neq o}^{x*} = \boldsymbol{q}_o^{x*} = 0$. This is a contradiction because the constraint $\sum_{j=1}^{n} \lambda_j \boldsymbol{y}_j - \underbrace{\varphi^{y}}_{\geq 0} \geq \boldsymbol{y}_o$ forces λ_j to be nonzero. Putting these all together, we have $0 < \theta_o^{RTE*} \leq 1$.□

Theorem 3. The optimal objective function value of model (10) is greater than that of model (1).

Proof. Consider the following equations *(a)* and *(b)* that correspond to the input constraints in models (1) and (10), respectively:

$$\sum_{\substack{j=1 \\ j\neq o}}^{n} \lambda_j^* \boldsymbol{x}_j + \lambda_o^* \boldsymbol{x}_o - \theta_o^* \boldsymbol{x}_o \leq 0, \forall i \Rightarrow \theta_o^* \geq \frac{\sum_{j=1}^{n} \lambda_j^* \boldsymbol{x}_j}{\boldsymbol{x}_o}, \forall i, \qquad (a)$$

$$\sum_{\substack{j=1 \\ j\neq o}}^{n} \lambda_j^* \boldsymbol{x}_j + (\lambda_o^* - \theta_o^*)\boldsymbol{x}_o + \boldsymbol{\Gamma}^x \boldsymbol{p}^{x*} + \sum_{\substack{j=1 \\ j\neq o}}^{n} \boldsymbol{q}_j^{x*} + \boldsymbol{q}_o^{x*} \leq 0, \forall i \Rightarrow \theta_o^* \geq \frac{\sum_{j=1}^{n} \lambda_j^* \boldsymbol{x}_j + \boldsymbol{\Gamma}^x \boldsymbol{p}^{x*} + \sum_{\substack{j=1 \\ j\neq o}}^{n} \boldsymbol{q}_j^{x*} + \boldsymbol{q}_o^{x*}}{\boldsymbol{x}_o}, \forall i, \qquad (b)$$

where θ in *(a)* and *(b)* is a decision variable. By comparing the above ratios on the right-hand sides of inequalities, it is clear that the objective function value of model (10) is greater than that of model (1), i.e., by virtue of the non-negativity of $\boldsymbol{\Gamma}^x \boldsymbol{p}^{x*} + \sum_{\substack{j=1 \\ j\neq o}}^{n} \boldsymbol{q}_j^{x*} + \boldsymbol{q}_o^{x*}$.□

It should be noted that we find some similarities between our proposed models and robust TE models presented in Mensah (2019) and Salahi et al. (2020). However, the non-decreasing relationship between the deterministic TE and robust TE models was not explored in both studies. The importance of this issue in robust DEA models under the ellipsoidal uncertainty set was studied by Ehrgott et al. (2018) where data is uncertain and the related efficiency measures increase monotonically. It was accordingly discussed that the decision maker can catch the maximum possible efficiency of a DMU over all permissible uncertainties, and the minimal amount of uncertainty is required to achieve this efficiency. In this respect, Theorem 2 is in line with Definition 2 of Ehrgott et al. (2018, p. 233) showing that the maximal robust TE measure remains 1 under the polyhedral uncertainty set, and conforming to Proposition 1 of Ehrgott et al. (2018, p. 233), Theorem 3 demonstrates that the robust TE measure is non-decreasing when uncertainties increase. Despite the theoretical aspect, the applicability of these essential properties will be illustrated in an empirical application in Section 6.

4.1.2. Robust CE measure

There is no doubt that analysing CE with uncertain input and output data is of interest to researchers and practitioners in a sea of situations. Let us use the same notations in Subsection 4.1.1 to reformulate the CE model (2) in a way that gets to grips with uncertainty in input and output data:

$$\widetilde{\omega}_o = \min_{\boldsymbol{\lambda},\boldsymbol{x}^0} \left\{ \boldsymbol{c}_o \boldsymbol{x}^0 : \sum_{j=1}^{n} \lambda_j \widetilde{\boldsymbol{x}}_j \leq \boldsymbol{x}^0, \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \widetilde{\boldsymbol{y}}_j + (\lambda_o - 1)\widetilde{\boldsymbol{y}}_o \geq 0, \lambda_j \geq 0\ \boldsymbol{x}^0 \geq 0, \forall j \right\}. \tag{11}$$

Notably, the input prices in model (11) are assumed to be precise. Since the output constraints of model (11) are the same as those in model (3), the robust counterpart of model (11) hinges on the input constraints only. We hence define the robust counterpart of input constraints as $\sum_{j=1}^{n} \lambda_j \boldsymbol{x}_j + \bar{\beta}_x \leq 0$ where $\bar{\beta}_x = \max_{g^x} \left\{ \sum_{j=1}^{n} \lambda_j \boldsymbol{\xi}_j^x \widehat{\boldsymbol{x}}_j \right\}$ and $g^x = \left\{ \boldsymbol{\xi}_j^x \middle| \sum_{j=1}^{n} |\boldsymbol{\xi}_j^x| \leq \boldsymbol{\Gamma}^x \right\}$. In what follows, the robust counterpart of the CE model (2) can be expressed as follows:

$$\begin{aligned} \widetilde{\omega}_o = \min_{\boldsymbol{\lambda},\boldsymbol{x}^0,\boldsymbol{p}^x,\boldsymbol{q}_j^x,\boldsymbol{p}^y,\boldsymbol{q}_j^y} \Bigg\{ & \boldsymbol{c}_o \boldsymbol{x}^0 : \sum_{j=1}^{n} \lambda_j \boldsymbol{x}_j + \bar{\varphi}^x \leq \boldsymbol{x}^0, \sum_{\substack{j=1 \\ j \neq o}}^{n} \lambda_j \boldsymbol{y}_j + (\lambda_o - 1)\boldsymbol{y}_o - \varphi^y \geq 0, \boldsymbol{p}^x + \boldsymbol{q}_j^x \geq \\ & \widehat{\boldsymbol{x}}_j \lambda_j; \forall j,\ \boldsymbol{p}^y + \boldsymbol{q}_j^y \geq \widehat{\boldsymbol{y}}_j \lambda_j; \forall j \neq o, \boldsymbol{p}^y + \boldsymbol{q}_o^y \geq \widehat{\boldsymbol{y}}_o (1 - \lambda_o), \lambda_j, \boldsymbol{x}^0, \boldsymbol{p}^x, \boldsymbol{q}_j^x, \boldsymbol{p}^y, \boldsymbol{q}_j^y \geq 0, \forall j \Bigg\}, \end{aligned} \tag{12}$$

where $\bar{\varphi}^x = \boldsymbol{\Gamma}^x \boldsymbol{p}^x + \sum_{j=1}^{n} \boldsymbol{q}_j^x$, and $\varphi^y = \boldsymbol{\Gamma}^y \boldsymbol{p}^y + \sum_{\substack{j=1 \\ j \neq o}}^{n} \boldsymbol{q}_j^y + \boldsymbol{q}_o^y$. Notice that φ^y in model (12) is the same as that presented in model (9). Considering account of the optimal costs $\widetilde{\omega}_o^*$ calculated from model (12), the robust CE measure of the DMU under analysis is $\omega_o^{RCE} = \frac{\widetilde{\omega}_o^*}{\boldsymbol{C}_o \boldsymbol{x}_o}$.

The uncertainty levels of inputs and outputs can be altered and controlled by $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ in model (12) to mitigate over-conservatism and protect the ω_o^{RCE} measure against data fluctuations.

Theorem 4. (*i*) Model (12) is always feasible, (*ii*) $\omega_o^{RCE} \in (0,1]$, and (*iii*) the optimal objective function value of model (12) is greater than or equal to that of model (2), i.e., $\widetilde{\omega}_o^* \geq \omega_o^*$.

Proof. (*i*) It is easy to verify that $\lambda_o = 1, \lambda_j = 0\ (j \neq o), \boldsymbol{x}^0 = \boldsymbol{x}_o,\ \boldsymbol{\Gamma}^y, \boldsymbol{\Gamma}^x, \boldsymbol{p}^x, \boldsymbol{q}_j^x, \boldsymbol{p}^y, \boldsymbol{q}_j^y = 0\ (\forall j)$ is a feasible solution for model (12).

(*ii*) The feasible solution in (*i*) along with model (12) as a minimisation problem imply that $\omega_o^{RCE} \leq 1$. On the other hand, let us show that $\omega_o^{RCE} > 0$ by contradiction. In doing so, assume that $\omega_o^{RCE} = 0$. It means that $\widetilde{\omega}_o = \boldsymbol{c}_o\boldsymbol{x}^0 = 0$, and then we have $\boldsymbol{x}^0 = 0$ due to the semipositive assumption for $\mathbf{c}_o$. Given the first constraint of model (12) and non-negative variables $\boldsymbol{\Gamma}^x, \boldsymbol{p}^x, \boldsymbol{q}_j^x$ and λ_j, we have $\lambda_j = 0, \forall j$ (see the proof of Theorem 2). So, $\boldsymbol{y}_o \leq 0$ when considering the second constraint set of model (12) which yields a contradiction to $\boldsymbol{y}_0 \geq 0, \boldsymbol{y}_o \neq 0$. Thus, $\widetilde{\omega}_o > 0$ and $\omega_o^{RCE} > 0$.

(*iii*) Let $S(\boldsymbol{\Gamma}^x, \boldsymbol{\Gamma}^y)$ be a feasible region of model (12) for the given $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$. Clearly, $\boldsymbol{\Gamma}^x = \boldsymbol{\Gamma}^y = \mathbf{0}$, i.e., $S(\mathbf{0}, \mathbf{0})$[4] means that there is no data perturbation which can represent the feasible region of the CE model (2). Consequently, the feasible regions of model (2) and model (12) are identical and we have $\widetilde{\omega}_o^* = \omega_o^*$. In addition, if at least one of the parameters $\boldsymbol{\Gamma}^x$ or $\boldsymbol{\Gamma}^y$ takes positive values, i.e., $\exists$ $(\boldsymbol{\Gamma}^x > 0 \; or \; \boldsymbol{\Gamma}^y > 0)$, then $S(\mathbf{0}, \mathbf{0}) \subset S(\boldsymbol{\Gamma}^x, \boldsymbol{\Gamma}^y)$, and resultantly the optimal objective value of minimisation model (12) is greater than or equal to that of the CE model (2), $\widetilde{\omega}_o^* \geq \omega_o^*$.□

To sum up, the developed robust TE model (10) and the robust CE model (12) enable us to calculate TE and CE measures for varying levels of conservatism parameters $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ whereby all or some of input and/or output data are uncertain. Accordingly, the Farrell decomposition can be adapted to decompose the estimated robust CE into robust TE and robust AE (RAE), viz. $\omega_o^{RCE} = \theta_o^{RTE} \times$ RAE. This decomposition is based on the polyhedral uncertainty set and various conservatism levels to provide insights about the performance of DMUs.

In general, the decision maker is in need to choose the level of conservatism associated with uncertain parameters of each constraint. Put differently, the level of conservatism helps to identify the risk preference of the decision maker. Since the levels of conservatism for input and output data, $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$, vary within $[0, n]$, the decision maker can take one of the following decisions by changing the levels of conservatism:

(i) If $\boldsymbol{\Gamma}^x = \boldsymbol{\Gamma}^y = 0$, there is no protection against uncertainty and the model is transformed into the nominal problem.

(ii) If $\boldsymbol{\Gamma}^x = \boldsymbol{\Gamma}^y = n$, all the uncertain inputs and outputs are fully protected against uncertainty, leading to an excessively conservative solution.

(iii) If $\boldsymbol{\Gamma}^x \in (0, n)$ and $\boldsymbol{\Gamma}^y \in (0, n)$, the decision maker is able to control the levels of conservatism to reduce over-conservatism and fine-tune the analysis.

Although an experienced decision maker can choose an appropriate level of conservatism in terms of risk preference, there are two ways to choose the apt level of conservatism parameters $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ whose aims are to control the budget of uncertainty within input and output constraints of models (10) and (12) as well as helping the lay decision maker to select the most appropriate level of conservatism (see e.g. Shokouhi et al., 2010; Salahi et al., 2020). First, $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ are assumed to be the decision variables, and

[4] $\mathbf{0}$ is a zero vector.

$\boldsymbol{\Gamma}^x + \boldsymbol{\Gamma}^y = \boldsymbol{\Gamma}$ is included in the models where $\boldsymbol{\Gamma}$ is a parameter. Note that $\boldsymbol{\Gamma}$ takes a value within $[0, 2n]$ since there are at most n uncertain parameters in each input and output constraint set, i.e., $\boldsymbol{\Gamma}^x \in [0, n]$ and $\boldsymbol{\Gamma}^y \in [0, n]$. This approach helps a decision maker look into the effect of uncertainty on efficiency measures and rankings by varying the parameter $\boldsymbol{\Gamma}$. However, transforming the LP models (10) and (12) into the non-linear forms in the presence of $\boldsymbol{\Gamma}^x + \boldsymbol{\Gamma}^y = \boldsymbol{\Gamma}$ is the downside of the foregoing approach. The second way that is deployed in the present study is to consider the effect of perturbations of input and output data on results by allocating the various values within $[0, n]$ to the parameters $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$. Importantly, this approach takes advantage of holding the linearity form of the robust models (10) and (12) as well as providing the decision maker with the ability to control the conservativeness in terms of his/her risk preference. In our empirical application, we will focus on the latter approach by first assuming that $\boldsymbol{\Gamma}^x = \boldsymbol{\Gamma}^y = \boldsymbol{\Gamma}$ for the sake of simplicity, and then estimating the lower bound of conservatism level using $\boldsymbol{\Gamma} \geq \min\{1 + \Phi^{-1}(1-e)\sqrt{n}, n\}$ proposed by Bertsimas and Sim (2004) where Φ represents the cumulative distribution function of the standard normal distribution and n is the number of uncertain parameters in each input and output constraint.

4.2. Input price uncertainty

In some cases, input and output data are deterministic while the prices of inputs are exogenously uncertain. This sub-section aims to present a method to cope with such uncertainty in some or all of input prices. Obviously, the TE model that excludes any uncertainty is model (1). However, calculating CE measures might be a challenge in the presence of uncertain prices. We hence propose a pair of robust CE models to estimate interval CE measures in which the lower and upper bounds of these intervals are employed to classify and further discriminate the DMUs in terms of the variability of their CE measures. In effect, we adapt Bertsimas and Sim (2004)'s and Gabrel and Murat (2010)'s approach to obtain the upper and lower bounds of the interval CE measure, respectively.

Consider the objective function of the nominal CE model (2). Let I_o represent a set of uncertain input prices, denoted by $\tilde{\boldsymbol{c}}_o = (\tilde{c}_{1o}, \dots, \tilde{c}_{mo})$, that take values from a symmetric distribution $[\boldsymbol{c}_o - \hat{\boldsymbol{c}}_o,\ \boldsymbol{c}_o + \hat{\boldsymbol{c}}_o]$ whereby $\boldsymbol{c}_o$ is the nominal value of uncertain input prices. Moreover, we define m independent random variables $\boldsymbol{\xi}_o^c = (\xi_{1o}^c, \dots, \xi_{mo}^c)^T \in \mathbb{R}^m$ whereby $\boldsymbol{\xi}_o^c = (\tilde{\boldsymbol{c}}_o - \boldsymbol{c}_o)/\hat{\boldsymbol{c}}_o$ is a variable vector with the (unknown) symmetric distribution that takes values within $[-1,1]$. In other words, the nominal values of uncertain input prices can be expressed as $\tilde{c}_o = \boldsymbol{c}_o + \hat{\boldsymbol{c}}_o \boldsymbol{\xi}_o^c$ where the maximum deviations are defined as $\hat{\boldsymbol{c}}_o = e_o \boldsymbol{c}_o$. Note that $e_o > 0$ is a given uncertainty level which determines to what extent input prices are diverted from their nominal values. In addition, $\Gamma_o^c (\leq m)$ is referred to as the budget of uncertainty parameter associated with input prices whose value varies within a range from 0 to $|I_o|$. Let $|\mathbf{1}_m \boldsymbol{\xi}_o^c| \leq \Gamma_o^c$, where $\mathbf{1}_m = (1, \dots, 1)$ is an m-dimensional vector space and the parameter Γ_o^c plays a key part in adjusting the robustness against the level of conservatism in the robust solution (Bertsimas and Sim, 2004). In light of the above setting, the polyhedral uncertainty sets can be formulated as follows:

$$\mathcal{U}_{\Gamma_o^c} = \left\{ \widetilde{\boldsymbol{C}}_o \middle| \, \tilde{\boldsymbol{c}}_o = \boldsymbol{c}_o + \hat{\boldsymbol{c}}_o \boldsymbol{\xi}_o^c, |\mathbf{1}_m \xi_o^c| \leq \Gamma_o^c, \boldsymbol{\xi}_o^c \in [-1,1] \right\}. \tag{13}$$

Based on the approach of Bertsimas and Sim (2004), the *robust solution* of the robust counterpart is a feasible solution that is immunised against the highest level of uncertainty (worst-case scenario[5]). Accordingly, the robust solution can be described as the optimal value of a solution that the budget of uncertainty parameter Γ_o^c has the largest deviation from its nominal value. For instance, if there are five input prices $\boldsymbol{c}_o = (c_{1o}, c_{2o}, \dots, c_{5o})$ to form the CE model (2) and $\Gamma_o^c = 3.8$, then the 3 input prices of the $\boldsymbol{c}_o$ vector take their worst-case values as $c_{io} + \hat{c}_{io}, i = 1, \dots, 5$, and one input price of the $\boldsymbol{c}_o$ vector attains the value of $c_{io} + 0.8\, \hat{c}_{io}, i = 1, \dots, 5$. The worst-case scenario assesses the robustness against price uncertainty through adopting the approach of Bertsimas and Sim (2004) to lead to the robust counterpart of the CE model (2) as follows:

$$\omega_o^U = \min_{\boldsymbol{x}^0, \boldsymbol{\lambda}, \boldsymbol{\xi}_o^c} \left\{ \boldsymbol{C}_o \boldsymbol{x}^0 + \max_{\substack{|\mathbf{1}_m \boldsymbol{\xi}_o^c| \leq \Gamma_o^c \\ -1 \leq \boldsymbol{\xi}_o^c \leq 1}} \widehat{\boldsymbol{C}}_o \boldsymbol{\xi}_o^c \boldsymbol{x}^0 : constraints\ of\ model\ (2) \right\}, \tag{14}$$

where $\widehat{\boldsymbol{C}}_o$ is a diagonal matrix with diagonal elements $(\hat{c}_{1o}, \hat{c}_{2o} \dots, \hat{c}_{mo})$. The objective of model (14) is to calculate the *upper bound* of the CE measure. Since the inner optimisation problem of model (14) is a maximisation problem and $\boldsymbol{x}^0 \geq 0$, the objective function value of model (14) cannot decrease when $\boldsymbol{\xi}_o^c$ is non-negative. Therefore, model (14) can be rewritten as follows:

$$\omega_o^U = \min_{\boldsymbol{x}^0, \boldsymbol{\lambda}, \boldsymbol{\xi}_o^c} \left\{ \boldsymbol{C}_o \boldsymbol{x}^0 + \max_{\substack{|\mathbf{1}_m \boldsymbol{\xi}_o^c| \leq \Gamma_o^c \\ 0 \leq \boldsymbol{\xi}_o^c \leq 1}} \widehat{\boldsymbol{C}}_o \boldsymbol{\xi}_o^c \boldsymbol{x}^0 : \ constraints\ of\ model\ (2) \right\}. \tag{15}$$

Notably, model (15) is a non-linear programming (NLP) problem and the following theorem presents the equivalent linearised model that can be solved by commercial off-the-shelf optimisation software.

Theorem 5. The NLP model (15) is equivalent to the following LP model:

$$\omega_o^U = \min_{\boldsymbol{x}^0, \boldsymbol{\lambda}, p_o, q_o} \left\{ \begin{matrix} \boldsymbol{C}_o \boldsymbol{x}^0 + \Gamma_o^c p_o + \mathbf{1}_m \boldsymbol{q}_o : p_o + \boldsymbol{q}_o \geq \hat{\boldsymbol{c}}_o \boldsymbol{x}^0, p_o \geq 0, \boldsymbol{q}_o \geq \mathbf{0}_m \\ constraints\ of\ model\ (2) \end{matrix} \right\}. \tag{16}$$

Proof. Let $\boldsymbol{x}^{0*}$ be the optimal solution of model (14). In addition, assume that p_o and $\boldsymbol{q}_o = (q_{1o}, \dots, q_{mo})$ are the dual variables of the constraints $|\mathbf{1}_m \boldsymbol{\xi}_o^c| \leq \Gamma_o^c$ and $0 \leq \boldsymbol{\xi}_o^c \leq 1$, respectively. By providing the dual of the inner maximisation problem of model (14), we have:

[5] The worst-case scenario occurs when the Γ_o^c parameter takes its maximum value.

$$\omega_o^U = \min_{\boldsymbol{x}^0,\boldsymbol{\lambda},p_o,q_o} \left\{ \boldsymbol{C}_o\boldsymbol{x}^0 + \min_{\substack{p_o+q_o \geq \hat{c}_o\boldsymbol{x}^0,\\ p_o\geq 0,\boldsymbol{q}_o\geq \boldsymbol{0}_m}} (\Gamma_o^c p_o + \boldsymbol{1}_m\boldsymbol{q}_o) : \ constraints\ of\ model\ (2) \right\}. \qquad (17)$$

The inner problem of model (15) is feasible and bounded for all $\Gamma_o^c \in [0, |I_o|]$. Thus, by the strong duality theory, the dual inner minimisation of model (17) is also feasible and bounded. In such case, model (17) is equivalent to the LP problem (16).□

The bottom line is that the optimal solution of the CE model (16) is robust against price uncertainty even if prices of inputs maximally deviate from their nominal values and reach their worst values.

Theorem 6. The optimal objective function value of model (16) is an upper bound of the CE model (2).

Proof. Let us consider three cases**.**

(i) If $\Gamma_o^c = 0$ in model (16), then there is no uncertainty in input prices, that is, $\hat{\boldsymbol{c}}_o = \boldsymbol{0}_m$ and resultantly, $p_o = 0, \boldsymbol{q}_o = \boldsymbol{0}_m$. Thus, model (16) is transformed into model (2), i.e., $\omega_o^U = \omega_o$.

(ii) If the level of budget uncertainty in model (16) reaches out to its maximum value, i.e., $\Gamma_o^c = m$, the worst-case scenario occurs and $p_o + \boldsymbol{q}_o = \hat{\boldsymbol{c}}_o\boldsymbol{x}^0$ (Bertsimas and Thiele, 2006). Then, by summing up over all elements i, we have $\sum_{i=1}^m p_o + \sum_{i=1}^m q_i = \sum_{i=1}^m \hat{c}_{io}x^0 = \widehat{\boldsymbol{C}}_o\boldsymbol{x}^o$. Since $\sum_{i=1}^m p_o = mp_o$ and $\Gamma_o^c = m$, we have $\Gamma_o^c p_o + \boldsymbol{1}_m\boldsymbol{q}_o = \widehat{\boldsymbol{C}}_o\boldsymbol{x}^o$ that can be substituted into the objective function of model (16) in order to get $(\boldsymbol{C}_o + \widehat{\boldsymbol{C}}_o)\boldsymbol{x}^o$, i.e., $\omega_o \leq \omega_o^U(m)$.

(iii) It can be trivially verified that an increase in the level of budget uncertainty, Γ_o^c, all else held constant, will not decrease the objective value of model (16), that is, $\omega_o^U(\Gamma_o^c) \leq \omega_o^U(\Gamma_o^c + k)$ where k is a positive constant. If $0 < \Gamma_o^c < m$, then we $\omega_o \leq \omega_o^U(\Gamma_o^c) \leq \omega_o^U(m)$.

Putting it all together, the optimal objective function value of model (16) is the upper bound of the CE model (2). □

Model (16), first and foremost, is an LP problem for all $\Gamma_o^c \in [0, m]$. The upper bound of the CE measure of DMU_o is defined as the ratio of optimal costs ω_o^{U*} to current prices, viz. $\omega_o^{\overline{RCE}} = \frac{\omega_\text{o}^{U*}}{\boldsymbol{C}_o\boldsymbol{x}_o}$.

Let us now emphasise on a lower bound of the CE measure, which is the adaptation of Gabrel and Murat (2010)'s approach as a best-case scenario. The corresponding robust counterpart of the CE model (2) concerning the best-case scenario can be formulated as follows:

$$\omega_o^L = \min_{\boldsymbol{x}^0,\boldsymbol{\lambda},\boldsymbol{\xi}_o^c} \left\{ \boldsymbol{C}_o\boldsymbol{x}^0 - \max_{\substack{\boldsymbol{1}_m\boldsymbol{\xi}_o^c \leq \Gamma_o^c\\ 0\leq \boldsymbol{\xi}_o^c \leq 1}} \widehat{\boldsymbol{C}}_o\boldsymbol{\xi}_o^c\boldsymbol{x}^0 : constraints\ of\ model\ (2) \right\}. \qquad (18)$$

Similar to what we have carried out in Theorem 5, the following model is equivalent to model (18):

$$\omega_o^L = \min_{\boldsymbol{x}^0,\boldsymbol{\lambda},\boldsymbol{\xi}_o^c} \{\boldsymbol{C}_o\boldsymbol{x}^0 - \widehat{\boldsymbol{C}}_o\boldsymbol{\xi}_o^c\boldsymbol{x}^0 \colon \boldsymbol{1}_m\boldsymbol{\xi}_o^c \leq \Gamma_o^c, 0 \leq \boldsymbol{\xi}_o^c \leq 1, constraints\ of\ model\ (2)\}. \qquad (19)$$

The above model has the non-convex quadratic objective function by virtue of a negative semi-definite matrix $\widehat{\boldsymbol{C}}_o$[6] while its constraints set is convex. Therefore, model (19) is a concave quadratic programming problem for $\Gamma_o^c \in (0, m)$, but it can be transformed into an LP model when $\Gamma_o^c = 0$ (minimum conservatism level) and $\Gamma_o^c = m$ (maximum conservatism level). Some of more recent complex and sophisticated approaches for solving non-linear optimisation problems allow us to identify a global optimal or near-global optimal solution (see e.g., Pintér, 2007; Zamani, 2019). In this respect, BARON solver is used to solve the model (19) through general algebraic modelling systems (GAMS). Moreover, we verify the results by the LGO solver suite, which is frequently served for the analysis and global solution of concave quadratic models. The following theorem inspects the relationship between models (19) and (2).

Theorem 7. The optimal objective function value of model (19) is a lower bound of the CE model (2).

Proof. Similar to Theorem 6, let us consider the three following cases:

(*i*) If the level of budget uncertainty takes the minimum value in model (19), i.e., $\Gamma_o^c = 0$, then no uncertainty exists and model (19) is equal to the nominal CE model (2), therefore $\omega_o^L = \omega_o$.

(*ii*) If the level of budget uncertainty takes its maximum value, i.e., $\Gamma_o^c = m$. Since $\boldsymbol{x}^0 > 0$, the optimal values of the decision variable vector $\boldsymbol{\xi}_o^c$ is equal to 1. So, the objective function of model (19) is transformed into $(\boldsymbol{C}_o\boldsymbol{x}^0 - \widehat{\boldsymbol{C}}_o\boldsymbol{x}^0) = (\boldsymbol{C}_o - \widehat{\boldsymbol{C}}_o)\boldsymbol{x}^0$, which is the lower bound of CE model (2), i.e., $\omega_o^L(m) \le \omega_o$.

(*iii*) It can be easily showed that an increase in the level of budget uncertainty, Γ_o^c, all else held constant, will not increase the objective value of model (19), that is, $\omega_o^L(\Gamma_o^c + k) \le \omega_o^L(\Gamma_o^c)$ where k is a positive constant. If $0 < \Gamma_o^c < m$, then we $\omega_o^L(m) \le \omega_o^L(\Gamma_o^c) \le \omega_o$.

Considering all the above cases complete the proof.□

The lower bound of the CE measure of DMU_o is calculated as $\omega_o^{\underline{RCE}} = \frac{\omega_o^{L*}}{\boldsymbol{C}_o\boldsymbol{x}_o}$. Therefore, the resulting interval CE measures with uncertain prices, $\left[\omega_o^{\underline{RCE}}, \omega_o^{\overline{RCE}}\right]$, are referred to as "RoCE measures" for a given level of conservatism Γ_o^c as well as being decomposed into $\left[\theta_o^{TE} \times \underline{RAE}, \theta_o^{TE} \times \overline{RAE}\right]$ in which RoAE $=$ $\left[\underline{RAE}, \overline{RAE}\right]$ and θ_o^{TE} are the interval robust AE and TE (model (1)), respectively.

To sum up, in this section, we first argue how the robust TE, CE, and AE measures can be identified based upon RO for situations where (*i*) input and output data are subject to uncertainty, and (*ii*) only input prices are under uncertainty. It is, in turn, showed that the proposed robust TE and robust CE measures render the upper bound for their deterministic models. Resultantly, a DMU is said to be *robust cost efficient* if and only if it is both technically and allocatively efficient. In other words, a DMU is said to be *robust*

[6] A symmetric matrix C is positive semi-definite if $x^T C x \ge 0$ for every x.

allocative efficient relative to other DMUs if and only if ω_o^{RCE}=$\theta_o^{RTE} = 1$. In the second part of this section, we present a bounded interval for CE measures based upon RO for situations where some or all of input prices are uncertain only. Following Despotis and Smirlis (2002)'s approach, DMUs with the interval robust CE measures $\left[\omega_o^{\underline{RCE}}, \omega_o^{\overline{RCE}}\right]$ can be classified into three groups ($\Lambda^{++}, \Lambda^{+}$ and Λ^{-}); (*i*) a DMU is classified as Λ^{++}, so-called *strongly robust cost efficient*, if the lower and upper bounds of the RoCE measure are equal to unity, i.e., $\omega_o^{\underline{RCE}} = \omega_o^{\overline{RCE}} = 1$, (*ii*) a DMU is classified as Λ^{+}, so-called *robust cost efficient*, if the lower bound of the RoCE measure is smaller than unity and its upper bound of the RoCE measure is equal to unity, and (*iii*) a DMU is classified as Λ^{-}, so-called *robust cost inefficient*, if the upper bound of the RoCE measure is smaller than unity.

5. Comparison with related methods

It might be interesting to compare the developed models in this paper with Camanho and Dyson (2005)'s, Mostafaee and Saljooghi (2010)'s and Fang and Li (2013b)'s methods. In what follows, we first provide an overview of these three existent models in order to outline the discrepancies and similarities and then provide a small numerical example to illustrate the differences between the methods.

Camanho and Dyson (2005) (hereafter called CD) proposed a method to estimate the lower and upper bounds for the CE measure from the pessimistic and optimistic viewpoints in case of price uncertainty, that is, the minimal and maximal bounds estimated for the input prices of each DMU are available only. The CD method assesses CE in the light of the least favourable price scenario (pessimistic point of view) and the most favourable price scenario (optimistic point of view) where the range of input prices are considered. Let the price input i at $\mathrm{DMU_o}$ lie within $[\boldsymbol{C}_o^L, \boldsymbol{C}_o^U]$, and $W(\boldsymbol{v}) = \left\{v \middle| \frac{v_t C_{io}^L}{C_{to}^u} \le v_i \le \frac{v_t C_{io}^u}{C_{to}^L}, i < t, i, t = 1, \dots, m\right\}$ is weight restrictions by way of input cone assurance regions showing that the relative value of the input weights must be altered within the range of $\left[\frac{C_{io}^L}{C_{to}^u}, \frac{C_{io}^u}{C_{to}^L}\right]$. In view of Schaffnit et al. (1997)'s idea on CE, CD suggests the following pair of models to obtain the interval, $[\Psi_o^L, \Psi_o^U]$, for the CE measure:

$$\Psi_o^L = \min_{\boldsymbol{u},\boldsymbol{v}}\{\boldsymbol{u}\boldsymbol{y}_o : \boldsymbol{v}\boldsymbol{x}_o = 1, \boldsymbol{u}\boldsymbol{y}_P - \boldsymbol{v}\boldsymbol{x}_P = 0, \boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j \le \boldsymbol{0}_{\boldsymbol{n}}; \forall j, \boldsymbol{v} \in W(\boldsymbol{v}), \boldsymbol{u} \ge \boldsymbol{1}_{\boldsymbol{s}}\varepsilon\}, \tag{20}$$

$$\Psi_o^U = \max_{\boldsymbol{u},\boldsymbol{v}}\{\boldsymbol{u}\boldsymbol{y}_o : \boldsymbol{v}\boldsymbol{x}_o = 1, \boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j \le \boldsymbol{0}_{\boldsymbol{n}}; \forall j, \boldsymbol{v} \in W(\boldsymbol{v}), \boldsymbol{u} \ge \boldsymbol{1}_{\boldsymbol{s}}\varepsilon\}, \tag{21}$$

where the index P in model (20) represents the peer DMU, and the equality constraint $\boldsymbol{u}\boldsymbol{y}_P - \boldsymbol{v}\boldsymbol{x}_P = 0$ compels that $\mathrm{DMU_p}$ remains efficient when assessing the efficiency of $\mathrm{DMU_o}$. Notice that in addition to the computation complexity, CD suffers from the disadvantage of considering multiple outputs. In such circumstances, model (20) may give rise to very low values for the lower bound of CE measures which cannot be interpreted straightforwardly.

Mostafaee and Saljooghi (2010) (hereafter called MS) suggested a pair of DEA-based models to calculate the upper and lower bounds of CE measures in two distinct uncertainty situations; (*i*) uncertain data with deterministic input prices, and (*ii*) uncertain data and input prices. Assume that input and output data lie within bounded intervals, i.e., $[\boldsymbol{x}_j^L, \boldsymbol{x}_j^U]$ and $[\boldsymbol{y}_j^L, \boldsymbol{y}_j^U]$ where $\boldsymbol{x}_j^L > 0$ and $\boldsymbol{y}_j^L > 0$. The bounded intervals of CE measures, $[CE_o^L, CE_o^U]$, in situations of interval input and output data can be attained by the following pair of LP models:

$$CE_o^L = \max_{\boldsymbol{u}} \left\{ \boldsymbol{u}\boldsymbol{y}_o^L : \boldsymbol{u}\boldsymbol{y}_j^U \leq \frac{\boldsymbol{C}_o \boldsymbol{x}_j^L}{\boldsymbol{C}_o \boldsymbol{x}_o^U};\ \forall j \neq o, \boldsymbol{u}\boldsymbol{y}_o^L \leq 1, \boldsymbol{u} \geq \boldsymbol{0}_m \right\}, \tag{22}$$

$$CE_o^U = \max_{\boldsymbol{u}} \left\{ \boldsymbol{u}\boldsymbol{y}_o^U : \boldsymbol{u}\boldsymbol{y}_j^U \leq \frac{\boldsymbol{C}_o \boldsymbol{x}_j^U}{\boldsymbol{C}_o \boldsymbol{x}_o^L};\ \forall j \neq o, \boldsymbol{u}\boldsymbol{y}_o^U \leq 1, \boldsymbol{u} \geq \boldsymbol{0}_m \right\}. \tag{23}$$

The thrust used in formulating the above models is to consider the most favourable situation for $\mathrm{DMU_o}$ and the least favourable condition for the other DMUs. MS also attends to the second situation where there are also input prices by way of intervals, $[\boldsymbol{C}_o^L, \boldsymbol{C}_o^U]$ through the evaluation of CE. In this circumstance, MS formulates the following pair of CE models to obtain the bounded intervals of CE measures:

$$CE_o^L = \min_{v \in \boldsymbol{V_m}} \left\{ \min_{\boldsymbol{\lambda}, \boldsymbol{x}^o} \frac{\boldsymbol{C}_o^v \boldsymbol{x}^o}{\boldsymbol{C}_o^v \boldsymbol{x}_o^U} : \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{x}_j^L + \lambda_o \boldsymbol{x}_o^U = \boldsymbol{x}^o, \sum_{\substack{j=1 \\ j \neq o}}^n \lambda_j \boldsymbol{y}_j^U + \lambda_o \boldsymbol{y}_o^L \geq \boldsymbol{y}_o^L, \lambda_j \geq 0, \forall j \right\}, \tag{24}$$

$$CE_o^U = \max_{\boldsymbol{u}, \boldsymbol{\eta}} \left\{ \begin{array}{l} \boldsymbol{u}\boldsymbol{y}_o^U : \widehat{\boldsymbol{C}}_{\boldsymbol{o}} \boldsymbol{x}_o^L = 1, \boldsymbol{u}\boldsymbol{y}_j^L - \widehat{\boldsymbol{C}}_{\boldsymbol{o}} \boldsymbol{x}_{\boldsymbol{j}}^{\boldsymbol{U}} \leq \boldsymbol{0}_{\boldsymbol{n}};\ \forall j \neq o, \boldsymbol{u}\boldsymbol{y}_o^U \leq 1, \\ \eta \boldsymbol{C}_o^L \leq \widehat{\boldsymbol{C}}_{\boldsymbol{o}} \leq \eta \boldsymbol{C}_o^U, \boldsymbol{u} \geq \boldsymbol{0}_{\boldsymbol{s}}, \boldsymbol{\eta} \geq \boldsymbol{0}_{\boldsymbol{m}} \end{array} \right\}, \tag{25}$$

where V_m in $\{\boldsymbol{C}_o^v : v \in V_m\}$ of model (24) is a set of $\pm$1-vectors that is used to identify the set of all extreme points of the bounded interval set $\{\boldsymbol{C}_o : \boldsymbol{C}_o^L \leq \boldsymbol{C}_o \leq \boldsymbol{C}_o^U\}$, and, in model (25), $\widehat{\boldsymbol{C}}_{\boldsymbol{o}} = \eta[\boldsymbol{C}_o^L, \boldsymbol{C}_o^U]$ and $\eta = 1/\boldsymbol{x}_o^L[\boldsymbol{C}_o^L, \boldsymbol{C}_o^U]$. Although the MS method infers that CE_o^U can be calculated in a way similar to model (24) when the focus is on extreme points of the convex set, Fang and Li (2012, p.589) provided counterexamples to show that it is not always true.

Fang and Li (2013b) (hereafter called FL) only looked into the lower bound of CE measure from the pessimistic viewpoints in which the input prices take the form of bounded intervals. With an alternative view, the FL model is the following bi-level LP model based on the model of CD:

$$I_o^* = \min_{v \in W(\boldsymbol{v})} \max_{\boldsymbol{u}, \boldsymbol{v}} \{ \boldsymbol{u}\boldsymbol{y}_o : \boldsymbol{v}\boldsymbol{x}_o = 1, \boldsymbol{u}\boldsymbol{y}_j - \boldsymbol{v}\boldsymbol{x}_j \leq \boldsymbol{0}_{\boldsymbol{n}}; \forall j, \boldsymbol{u} \geq \boldsymbol{1}_{\boldsymbol{s}}\varepsilon, \boldsymbol{v} \geq \boldsymbol{1}_{\boldsymbol{m}}\varepsilon \}. \tag{26}$$

The main contribution of FL is extending a modified vertex-enumerating method to solve the above problem.

Let us here take a small numerical example from Fang and Li (2013b) to compare the results of these studies. Table 3 shows the data set on eight DMUs, each with two inputs and two outputs. It is presumed that the input and output data are precise, and the input prices associated with x_{1j} and x_{2j} for all DMUs are altered within the bounded interval [3, 4], viz. $3 \le c_1 \le 4$ and $3 \le c_2 \le 4$. Given the pre-determined input prices (the columns headed C_1 and C_2), the Farrell's CE measures are calculated by the CE model (2) as reported in the last column of Table 3.

Table 3. Data set and Farrell's CE measures

DMUs	x_1	x_2	y_1	y_2	$[c_1^L, c_1^U]$	$[c_2^L, c_2^U]$	Farrell's CE		
							c_1	c_2	ω_o^{CE}
A	2	7	1	3	[3, 4]	[3, 4]	3	4	0.794 (7)
B	3	5	1	4	[3, 4]	[3, 4]	4	3	1.000 (1)
C	5	3	1	7	[3, 4]	[3, 4]	4	3	1.000 (1)
D	7	2	1	2	[3, 4]	[3, 4]	3.8	3.2	0.830 (6)
E	3	7	1	5	[3, 4]	[3, 4]	3	4	0.730 (8)
F	5	5	1	8	[3, 4]	[3, 4]	3.8	3.2	0.934 (4)
G	9	2	1	9	[3, 4]	[3, 4]	4	3	0.888 (5)
H	10	2.5	1	10	[3, 4]	[3, 4]	3	4	0.964 (3)

The CD approach assesses the bounded intervals of CE measures from the pessimistic and optimistic viewpoints via models (20) and (21) as reported, respectively, in the 2nd and 3rd columns of Table 4. It is not difficult to view that the results occur at one of the extreme points of the set of input prices while in reality considering a scenario where the maximal and minimal price bounds are available for all DMUs cannot be often viable. In situations of uncertainty in input prices, it is are essential to deem models (24) and (25) of MS to calculate the bounded of CE measures. It should be noticed that $\boldsymbol{y}_j^L = \boldsymbol{y}_j^U$ and $\boldsymbol{x}_j^L = \boldsymbol{x}_j^U$ in models (24) and (25). The results of MS are showed in the 4th and 5th columns of Table 4. In addition to unescapable limitations of CD in setting the input prices, the upper bound of the CE for MS might not be invariably obtained at extreme points for input prices (Fang and Li, 2012). FL lays emphasis on the lower bound of CE measure from the pessimistic viewpoints and the results of their modified vertex-enumerating method is recorded in the 6th column of Table 4. Let us in turn focus on the RoCE measures $\omega_{\underline{o}}^{\underline{RCE}}$ and $\omega_o^{\overline{RCE}}$ developed in the current study in the presence of price uncertainty. For the sake of comparison, let $\boldsymbol{c}_o = (3.5, 3.5)$ as the nominal value of uncertain input prices rather than considering bounded intervals for input prices. It is also presumed that the perturbation level e_o of the input prices is set to 0.142 with the aim of resulting in $[\boldsymbol{c}_o - \hat{\boldsymbol{c}}_o,\ \boldsymbol{c}_o + \hat{\boldsymbol{c}}_o] = [3, 4]$. In such settings, the proposed lower and upper RoCE measures $\omega_{\underline{o}}^{\underline{RCE}}$ and $\omega_o^{\overline{RCE}}$ calculate the upper and lower bounds of CE measures over the bounded interval [3, 4]. Since there are two uncertain parameters pertinent to input prices, Γ_o is assumed to be 2 to ensure the full robustness against the level of conservatism in the robust solution. The last two columns of Table 4 present the bounded interval of our developed RoCE measures in view of the aforesaid settings. It should be noted that the ranks of DMUs, either the lower or upper bounds, associated with different CE methods appear in parentheses in Table 4. Given the upper bounds of CE measures, DMU_A and DMU_C are always cost efficient

(Λ^{+}) for all methods, which is the same as the Farrell CE. Notably, regardless of the lower and upper bounds of CE measures, DMU_C has the best performance, and DMU_E has the worst performance apart from the lower bound of CD. Having looked at DMU_B, the score calculated from model (20) seems to be underestimated against other methods, which is in line with the concern of multiple outputs raised by Camanho and Dyson (2005, p. 441).

Table 4. The results of various CE models

DMUs	CD		MS		FL	RoCE	
	LB*	UB**	LB	UB	LB	$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$
	Model (20)	Model (21)	Model (24)	Model (25)	Model (26)		
A	0.340 (7)	0.931 (6)	0.667 (6)	0.931 (6)	0.794 (6)	0.763 (6)	0.931 (6)
B	0.532 (5)	1.000 (1)	0.750 (1)	1.000 (1)	0.931 (2)	0.858 (1)	1.000 (1)
C	0.931 (1)	1.000 (1)	0.750 (1)	1.000 (1)	1.000 (1)	0.858 (1)	1.000 (1)
D	0.244 (8)	0.931 (6)	0.667 (6)	0.931 (6)	0.794 (6)	0.763 (6)	0.931 (6)
E	0.521 (6)	0.838 (8)	0.600 (8)	0.800 (8)	0.730 (8)	0.686 (8)	0.840 (8)
F	0.771 (2)	0.947 (5)	0.686 (4)	0.947 (5)	0.882 (4)	0.784 (4)	0.960 (5)
G	0.643 (3)	0.991 (3)	0.701 (3)	0.991 (3)	0.888 (3)	0.802 (3)	0.981 (3)
H	0.568 (4)	0.964 (4)	0.686 (4)	0.964 (4)	0.872 (5)	0.784 (4)	0.964 (4)

*LB: Lower bound; **UB: Upper bound

Considering the bounded interval of CE measures, the average fluctuation rate of all DMUs are 38%, 25%, and 16% for CD, MS and RoCE measures, respectively. To compare the differences and similarities between results, the average pairwise difference among four different methods in terms of the lower and upper bounds of CE measures is obtained, and, in turn, reported in Table 5. In the CD study, the identical comparison is made between the CD bounds and the CE Farrell measure while using the precise prices for inputs in the uncertain environment does not make sense and is incomprehensible.

Table 5. The average pairwise difference of results for various CE models

	Methods	CD	MS	FL	$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$
LB	CD	×	18.61%	8.89%	23.68%	×
	MS	18.61%	×	17.30%	9.89%	×
	FL	8.89%	17.30%	×	7.41%	×
	$\omega_o^{\underline{RCE}}$	23.68%	9.89%	7.41%	×	×
UB	CD	×	0.47%	×	×	0.31%
	MS	0.47%	×	×	×	0.79%
	$\omega_o^{\overline{RCE}}$	0.31%	0.79%	×	×	×

We test the consistency (correlation) and goodness of fit among these four CE methods by various non-parametric statistical tests (Siegel and Castellan, 1988). Brockett and Golany (1996) discussed that the non-parametric test is a more suitable statistical tool in the DEA problems where there is a small number of DMUs. The first non-parametric test is the *Wilcoxon–Mann–Whitney* test that is used to compare differences between two independent groups. The results of the *Wilcoxon–Mann–Whitney* test at the 5%

significance level are presented in Table 6. Having focused on the upper bounds of CE measures, the difference between each pair of methods is not statistically significant while for the lower bounds of CE measures all pairwise comparisons between methods are statistically significant apart from that between CD and MS, and FL and RoCE measures.

Table 6. Mann-Whitney test

	LB						UB		
	CD/MS	CD/FL	CD/RoCE	MS/FL	MS/RoCE	FL/RoCE	CD/MS	CD/RoCE	MS/RoCE
z	-1.52617	-2.52422	-2.15769	-3.10727	-2.90738	1.94863	-0.35454	-0.35388	-0.06541
p-value	0.12697	0.01159*	0.03095*	0.00189*	0.00364*	0.05134	0.72293	0.72343	0.94784

** Indicates significant at the 5%*

Since there are four methods (groups), the Kruskal–Wallis rank test which is an extension of the Mann–Whitney rank test is a more appropriate non-parametric test in a sense that the former test investigates if there are statistically significant differences between two or more groups of an independent variable (Kruskal and Wallis, 1952). In view of the lower bounds of CE measures, the results of the Kruskal–Wallis test at the 5% significance level (95% level of confidence) is H=16.7777 with 3 degrees of freedom and a p-value of 0.00079 indicating that there are statistically significant differences between CD, MS, FL and RoCE measures. On the other hand, comparing the upper bounds of CE measures of CD, MS, and RoCE using the Kruskal–Wallis test results in H=0.005 with 2 degrees of freedom and a p-value of 0.9975. We can conclude that there is no evidence statistically to show that the CD, MS and RoCE measures are different in terms of the upper bounds of CE measures.

It might be of interest to verify the goodness of fit between each pair of methods via the Spearman's rank correlation test. In this respect, the ranks calculated from the four CE methods are considered to compare the correlation between each other via Spearman's coefficient correlation. We hence calculate Spearman coefficients in terms of the lower and upper bounds of CE measures. Given the lower bound measures, there is a high degree of correlation between CD and FL (r=0.7545, p<0.05), MS and FL (r=0.98788, p<0.05), MS and RoCE (r=1, p<0.05), and FL and RoCE (r=0.98788, p<0.05), but the correlations between CD and MS (r=0.69102, p>0.05) and CD and RoCE (r=0.69102, p>0.05) are not significant. Unsurprisingly, the degree of correlation between pair of methods (CD, MS and RoCE) is perfectly positive, i.e., r=1, p<0.05, when focusing on the upper bound measures.

The final comparison criterion is concerned with the computational complexity of CD, MS, FL and RoCE measures from different aspects in view of price uncertainty (See Table 7). All the methods can be solved by any LP algorithms to assess the bounded interval CE measures. The number of variables and constraints is reliant on the number of inputs (m), outputs (s) and DMUs (n). Given the number of models being solved, models (20) and (24) associated with CD and MS are not computationally economical, and the FL method is highly dependent on obtaining all vertices of the feasible region. Overall, RoCE measures developed in this paper methods take advantage of computational practicality against other existing when the focus is on the maximal conservatism levels. Over and above, let us illustrate to what extent the computational burden of these models increases by considering their model's size of the above numerical example (see the numbers in parentheses in Table 7).

Table 7. Comparing models from the computational perspective with price uncertainty.

Criteria	LB models				UB models		
	CD	MS	FL	RoCE	CD	MS	RoCE
	Model (20)	Model (24)	Model (26)	$\omega_o^{\underline{RCE}}$	Model (21)	Model (25)	$\omega_o^{\overline{RCE}}$
Model type	LP	LP	LP*	LP#	LP	LP	LP
No. variables	$m+s$ (4)	$n+1$ (9)	$m+s$ (4)	$2m+n$ (12)	$m+s$ (4)	$m+s+1$ (5)	$2m+n+1$ (13)
No. constraints	$2\times C_2^m+n+2$ (12)	$m+s$ (9)	$2m+n+1$ (6)	$2m+s+1$ (7)	$2\times C_2^m+n+1$ (11)	$2m+n+1$ (13)	$2m+s$ (6)
No. models	n^2 (64)	$n2^m$ (32)	N/A@	n (8)	n (8)	n (8)	n (8)

* BLP: Bi-level LP; # In case of full conservatism levels; @ N/A: Not-applicable uncountable set

Note that the above models can be explored in situations of price uncertainty, even though a given situation might occur where there are precise input prices and uncertain inputs and/or outputs. The latter situation is studied in this paper by model (12) as well as models (22) and (23) in the MS method to measure robust CE and bounded interval CE, respectively. Table 8 compares these models from the computational perspective. It is obvious that MS has a smaller number of variables and constraints than that of our model (12), but n times more is necessary to solve the LP models in MS so as to obtain the lower and upper bounds of CE measures.

Table 8. Comparing models from the computational perspective without price uncertainty.

Criteria	MS		Our model
	LB	UB	Robust CE
	Model (22)	Model (23)	Model (12)
Model type	LP	LP	LP
No. variables	s	s	$(n+1)(m+s)+n+m$
No. constraints	n	n	$(n+1)(m+s)$
No. models	n	n	n

6. An empirical application: US manufacturing

The location selection problem is affected by various production cost when input prices are different across locations. Although the spatial economics and supply chain management literature focus on differences in transportation cost, less attention has been drawn to production cost. This section studies a production site(s) selection problem for a US manufacturer over several adjacent states within the US when data uncertainty matters. The aim here is to seek the cost-minimising location of the production site(s) in one or some states via the developed models in this study. The data set has been collated by Ray, Chen, and Mukherjee (2008) based on the US Economic Census in 2002 (United State Census Bureau, 2002). It is assumed that there is homogenous technology across the states. The available candidate locations consist of 48 states (DMUs) in which the inputs are *production labour* (PL), *non-production labour* (NL), *energy* (E), *materials* (M), and *capital* (K) and the output is the *Gross value of production* (GP) (see Ray et al. (2008) for further detail on these measures). The production site selection model can be illustrated in Figure 1.

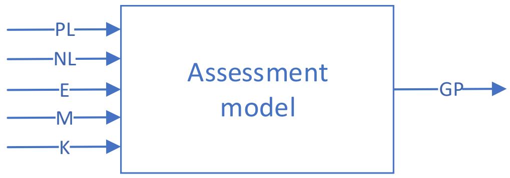

Figure 1. Assessment model with inputs and output

The definitions of input and output factors are listed in Table 9. The production and non-production labour prices, denoted by C_{PL} and C_{NL} are the hourly wage and emolument per labour, respectively. The energy price, C_E, is extracted from the industrial sector total energy price (US Energy Information Administration, 2008). The capital price, C_K, is computed by the sum of rent, depreciation and imputed interest expenses per dollar of the gross value of capital. The material price, C_M, is assumed to be equal to unity for all states. Table 10 shows the descriptive statistics for the input and output data as well as the input prices[7].

Table 9. Definitions of input and output factors

Factors	**Definition**
Production labour (PL)	The total number of hours worked by production labours per year (1000 hrs)
Non-production labour (NL)	The total number of non-production employees per year
Energy (E)	The total deflated expenditure of fuels and electricity per year ($1M)
Materials (M)	The total expenses on materials, parts, and components used for production ($1M)
Capital (K)	The average value of gross fixed assets such as land, buildings and equipment ($1M)
Gross value of production (GP)	The actual value of final goods and marketable by-products available for sale ($1M)

Table 10. Descriptive statistics for output, inputs, and input prices for 48 states in the US

Statistic	Output	Inputs					Input prices				
	GP	*PL*	*NL*	*E*	*M*	*K*	C_{PL}	C_{NL}	C_E	C_M	C_K
Min	4.045	21.53	3.92	6.34	1.883	2.134	13.16	41.818	3.81	1	0.11
Mean	11.506	60.134	11.954	36.731	5.449	5.314	16.278	52.903	7.096	1	0.136
Median	10.598	55.635	12.465	28.725	5.139	4.818	16.238	53.251	6.710	1	0.135
Sd.	4.758	21.646	2.926	32.644	2.985	2.652	1.541	4.486	2.080	0	0.015
Max	25.469	104.99	16.53	206.65	15.402	16.322	20.764	61.662	13.61	1	0.17

6.1. Data uncertainty

Uncertainty is unavoidable and can be observed in many ways. Let us first consider uncertainties endogenously in the inputs and outputs data. Such uncertainties may emanate from observational errors and forecasted values for the missing data that are commonly found within census data. It should be noted that input prices are assumed to be precisely measurable. To solve the robust TE model (10) and robust CE

[7] Details of data are available in Ray et al. (2008, p. 216).

model (12), for each location site (DMU_j) with $\hat{\boldsymbol{x}}_j = e_x \boldsymbol{x}_j$ and $\hat{\boldsymbol{y}}_j = e_y \boldsymbol{y}_j$, the percentage of perturbations e_x and e_y of the nominal data are set to 0.01 and 0.05 ($e_x = e_y = e = 0.01$ and 0.05). A snag that should be deliberated here is how to opt for an appropriate robustness level for $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ to protect the input and output constraints against perturbations. As discussed in Subsection 4.1.2, one way is to allocate the various values within $[0, n]$ to the parameters $\boldsymbol{\Gamma}^x$ and $\boldsymbol{\Gamma}^y$ in order to remain the linearity form of the robust models (10) and (12). For ease of exposition and without loss of generality, let us assume that $\boldsymbol{\Gamma}^x = \boldsymbol{\Gamma}^y = \boldsymbol{\Gamma}$, and in view of Bertsimas and Sim (2004), $\boldsymbol{\Gamma} \geq \min\left\{1 + \Phi^{-1}(1-e)\sqrt{n}, n\right\}$ where Φ represents the cumulative distribution function of the standard normal distribution and n is the number of uncertain parameters in each input and output constraint. In this application, the number of uncertainties is 48 depending on input and output constraints. Therefore, the suitable robustness levels, $\boldsymbol{\Gamma}^x(\boldsymbol{\Gamma}^y)$, are 17.11 and 12.40 for $e = 0.01$ and 0.05, respectively. These levels show that $17.11/48 \approx 35\%$ and $12.40/48 \approx 25\%$ of data uncertainty is deemed to produce robust solutions. Importantly, the decision maker must be mindful of hidden over-conservatism when there is the polyhedral uncertainty set with a budget of uncertainty constraint (Thiele, 2010; Liu et al., 2016). Hence, in this application, based upon Bertsimas and Sim (2004)'s approach we increase $\boldsymbol{\Gamma}$ by 1 from 0 to 17 and 12 for $e = 0.01$ and 0.05. In view of the foregoing settings, models (10) and (12) are solved to obtain robust TE, CE, and AE measures. Table 11 reports the descriptive statistics of θ_o^{RTE}, ω_o^{RCE}, and robust AE for different levels of $\boldsymbol{\Gamma}$s where the nominal data are fluctuated by 1% and 5%. The deterministic (Farrell) θ_o^{TE}, ω_o^{CE}, and AE efficiencies can be obtained when there are no uncertainties (i.e., all $\boldsymbol{\Gamma}$s are 0). Given results in Table 11, it can be inferred that the measures of θ_o^{RTE}, ω_o^{RCE}, and robust AE remain unchanged after the value of $\boldsymbol{\Gamma}$ is increased up to 5, 2, and 4 for $e = 0.01$, and 6, 2, and 3 for $e = 0.05$, respectively.

Table 11. Descriptive statistics for the robust CE decomposition under input and output uncertainties

		$e = 0.01$					$e = 0.05$				
	Γ	Min	Median	Mean	Max	No. eff	Min	Median	Mean	Max	No. eff
θ_o^{RTE}	0	0.795	0.940	0.939	1	11	0.795	0.940	0.939	1	11
	1	0.812	0.959	0.952	1	16	0.879	1	0.990	1	33
	2	0.818	0.972	0.959	1	17	0.914	1	0.997	1	45
	3	0.823	0.977	0.962	1	18	0.930	1	0.998	1	45
	4	0.826	0.978	0.963	1	18	0.943	1	0.999	1	47
	5	0.828	0.979	0.964	1	18	0.953	1	0.999	1	47
	6	0.828	0.979	0.964	1	18	0.961	1	0.999	1	47
ω_o^{RCE}	0	0.447	0.604	0.625	1	1	0.447	0.604	0.625	1	1
	1	0.452	0.610	0.631	1	1	0.470	0.636	0.656	1	1
	2	0.456	0.617	0.637	1	1	0.494	0.668	0.687	1	2
	3	0.456	0.617	0.637	1	1	0.494	0.668	0.687	1	2
	4	0.456	0.617	0.637	1	1	0.494	0.668	0.687	1	2
RAE	0	0.472	0.656	0.666	1	1	0.472	0.656	0.665	1	1

1	0.477	0.649	0.663	1	1	0.478	0.644	0.662	1	1
2	0.481	0.649	0.664	1	1	0.494	0.668	0.689	1	2
3	0.481	0.647	0.662	1	1	0.494	0.668	0.688	1	2
4	0.481	0.646	0.661	1	1	0.494	0.668	0.688	1	2
5	0.481	0.646	0.660	1	1	0.494	0.668	0.688	1	2

Notice that making a detailed inspection of results shows that the robust TE (robust CE) measures remain unchanged when the robustness level reaches 6.6 (1.95) and 8.96 (2) for 1% and 5% perturbations of input and output data. In other words, 13% and 18% of the total amount of uncertainty are ample to guard the solutions against 1% and 5% perturbations, respectively. Indeed, these results are in line with discussions presented at the end of Subsection 4.1.1 for the robust TE model (10). Table 11 and the above-mentioned details reveal the minimal required uncertainty level and the maximal efficiency measures. The average robust TE, robust CE, and robust AE measures under different levels of robustness are illustrated in Figure 2. As can be seen in Figure 2, the number of technical efficient units increases as $\boldsymbol{\Gamma}$ mounts. This finding demonstrates that the greater level of uncertainty (as a penalty value) worsens the discrimination power. The same phenomenon can be viewed for CE and AE measures. Notably, in line with Theorem 3 and Theorem 4, the results indicate that θ_o^{RTE} and ω_o^{RCE} for $\boldsymbol{\Gamma > 0}$ are greater than those for $\boldsymbol{\Gamma = 0}$.

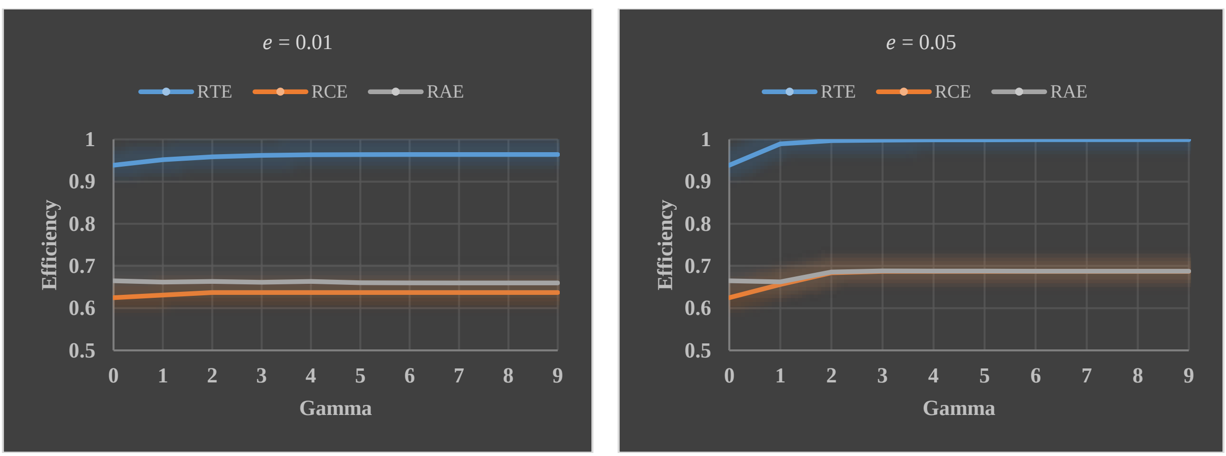

Figure 2. The comparison of robust TE, robust CE, and robust AE with various Gammas

Table 12. Optimal cost minimisation of AL for three selected cases

Γ (*e*)	Price	*PL*	*NL*	*E*	*M*	*K*	*Total cost*
0 (5%)	Optimal input	41.216	8.900	26.926	7.964	5.701	1201.550
	Actual input	85.72	12.46	82.7	6.648	7.289	2330.413
	Cost variation%	48	71	33	120*	78	52
1.95 (1%)	Optimal input	42.469	9.170	27.744	8.206	5.874	1238.066
	Actual input	85.72	12.46	82.7	6.648	7.289	2330.413
	Cost variation %	50	74	34	123	81	53
2 (5%)	Optimal input	47.832	10.328	31.248	9.242	6.616	1394.410
	Actual input	85.72	12.46	82.7	6.648	7.289	2330.413
	Cost variation %	56	83	38	139	91	60

*It shows a 20 percent rise in input M.

Let us centre on robust CE measures to provide the management team with more useful and practical information. Under 1% and 5% perturbations, Delaware (DE) is chosen as the cost-efficient location of the production site when $\Gamma = 0$. After further scrutinisation, DE is always robust cost efficient (Λ^{+}) for all the robustness levels when the perturbation amount is 1%. On the other hand, when 5% perturbation is imposed on input and output data, DE is still robust cost efficient (Λ^{+}) for $0 \leq \boldsymbol{\Gamma} \leq 1.41$, and DE and Louisiana (LA) for $\boldsymbol{\Gamma} \geq 1.42$. For instance, LA is the reference set for IN, ME, ND, TX when $e = 0.05$ and $\boldsymbol{\Gamma} = 1.5$. In addition, as an example of how the developed model can be worked, in view of $\Gamma = 0$, $\boldsymbol{\Gamma} = 1.95, e = 0.01$, and $\boldsymbol{\Gamma} = 2, e = 0.05$, the optimal and actual inputs of the site in AL as well as the percentage of cost variations are listed in Table 12. The last column of Table 12 shows to what extent the actual production cost can be reduced by varying its input mix to reach the minimum cost.

6.2. *Uncertainty in input prices*

The focus here is on RoCE measures $\omega_o^{\underline{RCE}}$ and $\omega_o^{\overline{RCE}}$ for 1%, 5%, 10%, and 15% perturbations whereby input prices are assumed to be exogenously uncertain. It should be noted that the larger value of perturbations indicates less available information on input prices. Likewise, Bertsimas and Sim (2004)'s approach, $\Gamma_o^c \geq \min\{1 + \Phi^{-1}(1 - e)\sqrt{m}, m\}$, is applied to pinpoint the most appropriate lower bound for the robustness level Γ_o^c for 1%, 5%, 10%, and 15% perturbations as reported in the first row of Table 13. The results of RoCE measures are presented in Table 13 to show the bounded interval of RoCE and RoAE, $\left[\omega_o^{\underline{RCE}}, \omega_o^{\overline{RCE}}\right]$ and $[\underline{RAE}, \overline{RAE}]$, for four cases. The bounded interval for RoAE is calculated by the ratio of RoCE to θ_o^{TE} where θ_o^{TE} is the Farrell CE model (1). Having looked at RoCE measures, it can be resulted that the fluctuation, viz. the difference between $\omega_o^{\underline{RCE}}$ and $\omega_o^{\overline{RCE}}$, is mounted by the increase in data perturbations. This is because the higher perturbations can give more flexibility to the models in order to find the larger range for the interval RoCE measures. Obviously, this finding can be viewed in RoAE measures since the TE is a constant value. Apropos of the upper bound of RoCE, the location site in DE is always robust cost efficient (Λ^{+}) and the number of the robust cost efficient units is various. In addition to DE, LA for $\Gamma_o^c = 4.67, e_o = 0.05$ and $\Gamma_o^c = 3.86, e_o = 0.1$ and DE, KY, LA, NM, SD, and WY for $\Gamma_o^c = 3.32, e_o = 0.15$ are robust cost efficient (Λ^{+}).

It is of interest to compare our results with the optimistic and pessimistic CD models which are literally in need of one output. The last two columns of Table 13 present the descriptive statistics of the results calculated from the CD models (20) and (21). The average fluctuation between the optimistic and pessimistic CD models is 9.3% while this proxy for RoCE measures are 2.5%, 12.2%, 24.4%, and 36.8% for 1%, 5%, 10%, and 15% perturbations, respectively. We also calculate the lower and upper bounds of AE measures of the CD method (see Table 13). In RoAE, the fluctuations are 2.6%, 13%, 26%, and 38.9%

on average where perturbations are presumed to be 1%, 5%, 10%, and 15%, respectively while this percentage for is 9.8% in the CD method.

Table 13. Descriptive statistics for RoCE and RoAE with uncertain input prices

Model	Statistics	$e_o = 0.01$ $\Gamma_o^c = 5$		$e_o = 0.05$ $\Gamma_o^c = 4.67$		$e_o = 0.1$ $\Gamma_o^c = 3.86$		$e_o = 0.15$ $\Gamma_o^c = 3.32$		CD	
		$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$	$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$	$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$	$\omega_o^{\underline{RCE}}$	$\omega_o^{\overline{RCE}}$	Model (20)	Model (21)
RoCE	Min	0.438	0.456	0.404	0.494	0.366	0.546	0.331	0.604	0.396	0.459
	Mean	0.612	0.637	0.565	0.687	0.511	0.755	0.462	0.829	0.567	0.659
	Median	0.592	0.616	0.547	0.667	0.494	0.738	0.447	0.817	0.563	0.647
	Sd.	0.109	0.112	0.100	0.112	0.091	0.113	0.082	0.114	0.100	0.112
	Max	0.980	1	0.905	1	0.818	1	0.740	1	0.958	1
	No. eff	0	1	0	2	0	2	0	6	0	2
	Statistics	$\underline{RAE}$	$\overline{RAE}$	$\underline{RAE}$	$\overline{RAE}$	$\underline{RAE}$	$\overline{RAE}$	$\underline{RAE}$	$\overline{RAE}$	LAE	UAE
RoAE	Min	0.462	0.481	0.427	0.521	0.386	0.576	0.349	0.637	0.396	0.501
	Mean	0.652	0.678	0.602	0.731	0.544	0.804	0.492	0.884	0.603	0.702
	Median	0.643	0.669	0.593	0.725	0.537	0.801	0.485	0.887	0.594	0.689
	Sd.	0.102	0.105	0.094	0.105	0.085	0.106	0.077	0.109	0.097	0.103
	Max	0.980	1	0.905	1	0.818	1	0.740	1	0.957	1
	No. eff	0	1	0	2	0	2	0	11	0	2

Contrary to the Farrell CE measure, we depict Figure 3 to illustrate the results in more detail. Figure 3(a) shows the average fluctuation between $\omega_o^{\overline{RCE}}$ and Farrell CE measures as well as $\omega_o^{\underline{RCE}}$ and Farrell CE measures for each location site. Likewise, Figure 3(b) is created based upon the bounded interval of RoAE and Farrell AE measures. As can be seen from Figure 3(a) and (b), the average fluctuations are on the rise when perturbations in input prices are mounted. Interestingly, it can be also observed that lower bounds CE (AE) measures are more robust in comparison with upper bounds for CE measures, particularly for the larger perturbations.

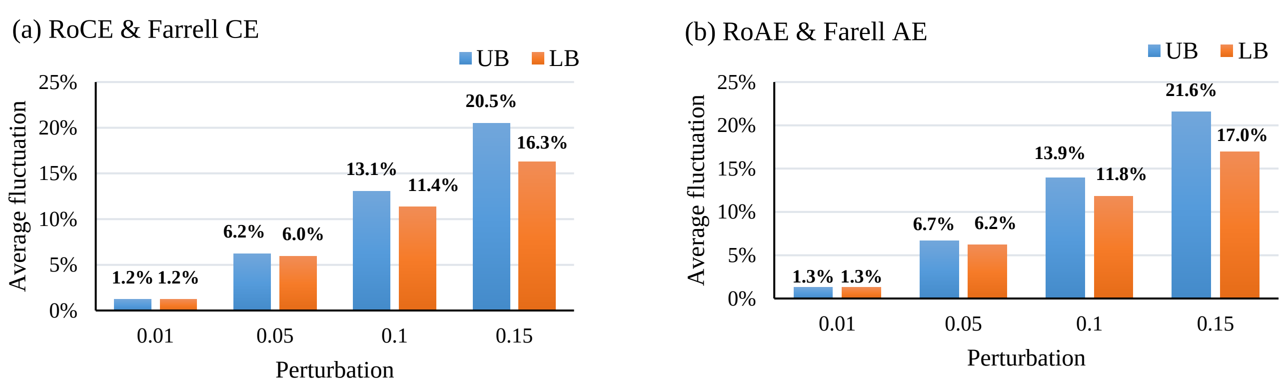

Figure 3. The average fluctuations for RoCE and RoAE

For instance, let us focus on the production site in New York (NY). Table 14 presents how this location could be altered based upon the interval optimal inputs in order to reduce the interval total production cost.

Obviously, the minimum case occurs when the target is the upper bounds. In case of $\Gamma = 0$, the Farrell CE results associated with NY are obtained. It should be noticed that Farrell CE measures invariably fall into the bounded intervals of CE and optimal input measures.

Table 14. The detailed analysis for the site in NY

Γ (e)	Price	PL	NL	E	M	K	*Total cost*
0	Optimal input	22.033	4.758	14.394	4.257	3.048	736.139
	Actual input	39.63	10.11	13.66	2.456	2.777	1311.743
	Cost variation%	55.60	47.06	105.37	173.33	109.76	56.12
5 (1%)	Optimal input	[21.812, 22.703]	[4.710, 4.902]	[14.250, 14.831	[4.214, 4.387]	[3.017, 3.14]	[728.778, 758. 481]
	Actual input	39.63	10.11	13.66	2.456	2.777	1311.743
	Cost variation%	[55.04,57.28]	[46.59,48.49]	[104.32,108.57]	[171.58,178.62]	[108.64, 113.07]	[55.56, 57.82]
4.67 (5%)	Optimal input	[19.195, 25.570]	[4.145, 5.521]	[12.540, 16.704]	[3.708, 4.941]	[2.655, 3.537]	[641.324, 854.263]
	Actual input	39.63	10.11	13.66	2.456	2.777	1311.743
	Cost variation%	[48.44, 64.52]	[41.00, 54.61]	[91.80, 122.28]	[150.98,201.18]	[95.61,127.37]	[48.89, 65.12]
3.86 (10%)	Optimal input	[16.745, 29.622]	[3.616, 6.396]	[10.939, 19.352]	[3.235, 5.724]	[2.316, 4.097]	[559.465, 989.649]
	Actual input	39.63	10.11	13.66	2.456	2.777	1311.743
	Cost variation%	[42.25, 74.75]	[35.77, 63.26]	[80.08,141.67]	[131.72, 233.06]	[83.40, 147.53]	[42.65, 75.45]
3.32 (15%)	Optimal input	[14.541, 34.281]	[3.140, 7.402]	[9.500, 22.395]	[2.809, 6.624]	[2.011, 4.742]	[485.852, 1145.298]
	Actual input	39.63	10.11	13.66	2.456	2.777	1311.743
	Cost variation%	[36.46,86.50]	[31.06, 73.21]	[69.65, 163.95]	[114.37, 269.71]	[72.42, 170.80]	[37.04, 87.31]

The implication of the proposed robust approach will be of particular interest to decision makers and system analysts to spot production inefficiencies. Though the performance assessment vividly plays a vital role in attaining the strategic objectives of organisations, this application tries to accentuate the importance of considering uncertainties in assessing TE and CE measures and assist managers in making better decisions in view of reliable and robust results. The managerial implication of this analysis is that the robust CE measures obtained from the proposed robust models are revolutionised in comparison with the conventional CE measures without price uncertainty. If the variability of the price is very high, being too conservative might lead to increasing fluctuations. This might not be what the decision maker is looking for. Thereby, this developed approach gives decision makers leeway to control the level of conservatism in terms of his/her risk preference. Another managerial insight derived from the results is that the developed robust models can provide planners and managers with useful and robust information in situations of high uncertainty. It may be needless to say that the proposed robust approach can be used in other applications that there are both endogenous and exogenous uncertainties.

7. Conclusions and future research directions

While original Farrell cost efficiency (CE) models are often subject to perfect and exact data, this study made an attempt to ease this condition by extending CE concepts in uncertain environments. We assumed that the uncertain inputs, outputs and input prices are represented by polyhedral uncertainty sets. First, the focus was on technical efficiency (TE) and CE models in situations of uncertainty in input and output data. We proposed the robust optimisation-based DEA models for estimating TE and CE measures where a problem involves endogenous and exogenous uncertainties. The resulting robust counterparts are linear and easily solvable using commercial off-the-shelf optimisation software. Second, we explored the CE assessment for a situation where uncertainty is observed only in input prices. In such cases, the development of our formulation included the worst- and best-case criteria based upon the concept of robust optimisation, which leads to the bounded intervals of CE estimates. The bottom line is that the resulting models constitutes not only the maximal and minimal CE bounds, but also enables managers to adjust the robustness level of uncertainty and analyse its effect on CE measures. The comparative study was conducted to examine the proposed models in comparison with the three existing approaches in the literature. Besides, a case study was presented in the context of the location analysis to demonstrate the applicability and usefulness of the models developed in this paper. The findings provide conclusive evidence that the robust CE measures are more resilient against uncertainty when more information on uncertain parameters are available.

For future research, it is worthwhile to explore other forms of uncertainty such as ellipsoidal and correlated polyhedral uncertainty sets for assessing CE in order to analyse in detail the effect of different uncertainty sets on CE measures. Moreover, the framework proposed in this paper could lend itself to a pile of situations where wait-and-see decisions are viably made. In such situations, our future work will be devoted to assessing CE measures using *adjustable RO* developed by Ben-Tal et al. (2004) because this dynamic RO approach is computationally tractable and leads to more flexible adjustable decisions in terms of variables and uncertainty sets. It is under no illusion that robust optimisation is incipient in the DEA literature and our approach can be also fully adapted to other variants, such as Russell measure models (Pastor et al., 1999), slacks-based measures (Tone, 2001), dynamic multilevel models (Färe and Grosskopf, 1996), FDH (Deprins et al., 1984), just to name a few, to open up new ways to tackle uncertainties.

References

Aghayi, N., & Maleki, B. (2016). Efficiency measurement of DMUs with undesirable outputs under uncertainty based on the directional distance function: Application on bank industry. *Energy*, *112*, 376–387.

Andersen, P., & Petersen, N. C. (1993). A procedure for ranking units in data envelopment analysis. *Management Science*, *39*, 1261–1264.

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

Bagherzadeh Valami, H. (2009). Cost efficiency with triangular fuzzy number input prices: An application of DEA. *Chaos, Solitons and Fractals*, *42*(3), 1631–1637.

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., El Ghaoui, L., & Nemirovski, A. (2009). *Robust Optimization*. Princeton University Press.

Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, *99*(2), 351–376.

Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of Operations Research*, *23*(4), 769–805.

Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, *25*(1), 1–13.

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

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

Bertsimas, D., & Thiele, A. (2006). Robust and data-driven optimization: Modern decision making under uncertainty. *Tutorials in Operations Research, INFORMS*, 95–122.

Brockett, P. L., & Golany, B. (1996). Using rank statistics for determining programmatic efficiency difference in DEA. *Management Science*, *42*(3), 466–472.

Camanho, A. S., & Dyson, R. G. (2005). Cost efficiency measurement with price uncertainty: A DEA application to bank branch assessments. *European Journal of Operational Research*, *161*(2), 432–446.

Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science*, *6*(1), 73–79.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, *30*(1–2), 91–107.

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.

Cooper, W.W., Z, Huang., & Li, S.X. (1996). Satisficing DEA models under chance constraints. *Annals of Operations Research*, *66*, 279–295.

Dantzig, G. B. (1955). Linear programming under uncertainty. *Management Science*, *1*(3), 197–206.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica*, *19*(3), 273–292.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor-efficiency in post offices. In *The performance of public enterprises: concepts and measurements* (pp. 243–268). Amsterdam: North-Holland.

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

Edirisinghe, N. C. P., & Zhang, X. (2010). An optimized DEA-based financial strength indicator of stock returns for U.S. markets. In *Lawrence, K.D. and Kleinman, G. (Ed.) Applications in Multicriteria Decision Making, Data Envelopment Analysis, and Finance (Applications of Management Science, Vol. 14)* (pp. 175–198). Emerald Group Publishing Limited.

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

El Ghaoui, L., & Lebret, H. (1997). Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, *18*(4), 1035–1064.

El Ghaoui, L., Oustry, F., & Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, *9*(1), 33–52.

Emrouznejad, A., & Yang, G. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA : 1978-2016. *Socio-Economic Planning Sciences*, *61*, 4–8.

Estellita Lins, M., Angulo-Meza, L., & Moreira da Silva, A. (2004). A multi-objective approach to determine alternative targets in data envelopment analysis. *Journal of the Operational Research Society*, *55*(10), 1090–1101.

Fang, L., & Li, H. (2012). A comment on "cost efficiency in data envelopment analysis with data uncertainty." *European Journal of Operational Research*, *220*(2), 588–590.

Fang, L., & Li, H. (2013a). Duality and efficiency computations in the cost efficiency model with price uncertainty. *Computers and Operations Research*, *40*(2), 594–602.

Fang, L., & Li, H. (2013b). Lower bound of cost efficiency measure in DEA with incomplete price information. *Journal of Productivity Analysis*, *40*, 219–226.

Färe, R., & Grosskopf, S. (1996). *Intertemporal production frontiers: With dynamic DEA* (1st ed.). Springer Netherlands.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiency of production*. Dordrecht: Springer Netherlands.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, *120*(3), 253.

Gabrel, V., & Murat, C. (2010). Robustness and duality in linear programming. *Journal of the Operational Research Society*, *61*(8), 1288–1296.

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.

Hassanzadeh, F., Modarres, M., Nemati, H. R., & Amoako-Gyampah, K. (2014). A robust R&D project portfolio optimization model for pharmaceutical contract research organizations. *International Journal of Production Economics, 158*, 18–27.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: Two decades in the making. *European Journal of Operational Research*, *214*(3), 457–472.

Khanjani Shiraz, R., Hatami-Marbini, A., Emrouznejad, A., & Fukuyama, H. (2018). Chance-constrained cost efficiency in data envelopment analysis model with random inputs and outputs. *Operational Research*, *20*, 1863–1898.

Koopmans, T. (1951). *Activity analysis of production and allocation : proceedings of a conference*. New York: Wiley.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621.

Kuosmanen, T., & Post, T. (2001). Measuring economic efficiency with incomplete price information: With an application to European commercial banks. *European Journal of Operational Research*, *134*(1), 43–58.

Kuosmanen, T., & Post, T. (2003). Measuring economic efficiency with incomplete price information. *European Journal of Operational Research*, *144*(2), 454–457.

Lai, C., Xu, L., & Shang, J. (2020). Optimal planning of technology roadmap under uncertainty. *Journal of the Operational Research Society*, *71*(4), 673–686.

Lappas, N. H., & Gounaris, C. E. (2018). Robust optimization for decision-making under endogenous uncertainty. *Computers and Chemical Engineering*, *111*, 252–266.

Liu, P., Yang, W., & Guo, T. (2016). A discussion on the conservatism of robust linear optimization problems. *Optimization*, *65*(8), 1641–1650.

Lu, Chao, Tao, J., An, Q., & Lai, X. (2020). A second-order cone programming based robust data envelopment analysis model for the new-energy vehicle industry. *Annals of Operations Research*, *292*, 321–339.

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

Marla, L., Vaze, V., & Barnhart, C. (2018). Robust optimization: Lessons learned from aircraft routing. *Computers & Operations Research*, *98*, 165–184.

Mensah, E. K. (2019). *Robust optimization in data envelopment analysis: Extended theory and applications*. University of Insubria, PhD Thesis.

Mensah, E. K. (2020). Robust data envelopment analysis via ellipsoidal uncertainty sets with application to the Italian banking industry. *Decisions in Economics and Finance*. https://doi.org/10.1007/s10203-020-00299-3.

Mostafaee, A., & Saljooghi, F. H. (2010). Cost efficiency measures in data envelopment analysis with data uncertainty. *European Journal of Operational Research*, *202*(2), 595–603.

Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations Research*, *43*(2), 264–281.

Omrani, H. (2013). Common weights data envelopment analysis with uncertain data: A robust optimization approach. *Computers & Industrial Engineering*, *66*(4), 1163–1170.

Pastor, J. T., Ruiz, J. L., & Sirvent, I. (1999). An enhanced DEA Russell graph efficiency measure. *European Journal of Operational Research*, *115*(3), 596–607.

Peeta, S., Sibel Salman, F., Gunnec, D., & Viswanath, K. (2010). Pre-disaster investment decisions for strengthening a highway network. *Computers and Operations Research*, *37*(10), 1708–1719.

Peykani, P., Mohammadi, E., Saen, R. F., Sadjadi, S. J., & Rostamy-Malkhalifeh, M. (2020). Data envelopment analysis and robust optimization: A review. *Expert Systems*. https://doi.org/10.1111/exsy.12534

Pintér, J. D. (2007). Nonlinear optimization with GAMS /LGO. *Journal of Global Optimization*, *38*(1), 79–101.

Pourmahmoud, J., & Bafekr Sharak, N. (2018). Measuring cost efficiency with new fuzzy DEA models. *International Journal of Fuzzy Systems*, *20*(1), 155–162.

Puri, J., & Yadav, S. P. (2016). A fully fuzzy DEA approach for cost and revenue efficiency measurements in the presence of undesirable outputs and its application to the banking sector in India. *International Journal of Fuzzy Systems*, *18*(2), 212–226.

Ray, S. (2016). Cost efficiency in an Indian bank branch network: A centralized resource allocation model. *Omega*, *65*, 69–81.

Ray, S. C., Chen, L., & Mukherjee, K. (2008). Input price variation across locations and a generalized measure of cost efficiency. *International Journal of Production Economics*, *116*(2), 208–218.

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.

Sadjadi, S. J., & Omrani, H. (2010). A bootstrapped robust data envelopment analysis model for efficiency estimating of telecommunication companies in Iran. *Telecommunications Policy*, *34*(4), 221–232.

Sadjadi, S. J., Omrani, H., Abdollahzadeh, S., Alinaghian, M., & Mohammadi, H. (2011b). A robust super-efficiency data envelopment analysis model for ranking of provincial gas companies in Iran. *Expert Systems with Applications*, *38*(9), 10875–10881.
Sadjadi, S. J., Omrani, H., Makui, A., & Shahanaghi, K. (2011a). An interactive robust data envelopment analysis model for determining alternative targets in Iranian electricity distribution companies. *Expert Systems with Applications*, *38*(8), 9830–9839.
Sahoo, B. K., Mehdiloozad, M., & Tone, K. (2014). Cost, revenue and profit efficiency measurement in DEA: A directional distance function approach. *European Journal of Operational Research*, *237*(3), 921–931.
Salahi, M., Toloo, M., & Torabi, N. (2020). A new robust optimization approach to common weights formulation in DEA A new robust optimization approach to common weights formulation. *Journal of the Operational Research Society*, 1–13. https://doi.org/10.1080/01605682.2020.1718016
Schaffnit, C., Rosen, D., & Paradi, J. C. (1997). Best practice analysis of bank branches: An application of DEA in a large Canadian bank. *European Journal of Operational Research*, *98*(2), 269–289.
Shokouhi, A. H., Hatami-Marbini, A., Tavana, M., & Saati, S. (2010). A robust optimization approach for imprecise data envelopment analysis. *Computers & Industrial Engineering*, *59*(3), 387–397.
Shokouhi, A. H., Shahriari, H., Agrell, P. J., & Hatami-Marbini, A. (2014). Consistent and robust ranking in imprecise data envelopment analysis under perturbations of random subsets of data. *OR Spectrum*, *36*(1), 133–160.
Siegel, S., & Castellan, N. J. J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, *44*(1), 49–61.
Simar, L., & Wilson, P. W. (2020). Technical, allocative and overall efficiency: Estimation and inference. *European Journal of Operational Research*, *282*(3), 1164–1176.
Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, *21*(5), 1154–1157.
Tavana, M., Toloo, M., Aghayi, N., & Arabmaldar, A. (2020). A robust cross-efficiency data envelopment analysis model with undesirable outputs. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2020.114117
Thanassoulis, E., Shiraz, R. K., & Maniadakis, N. (2015). A cost Malmquist productivity index capturing group performance. *European Journal of Operational Research*, *241*(3), 796–805.
Thiele, A. (2010). A note on issues of over-conservatism in robust optimization with cost uncertainty. *Optimization*, *59*(7), 1033–1040.
Toloo, M., & Ertay, T. (2014). The most cost efficient automotive vendor with price uncertainty: A new DEA approach. *Measurement*, *52*(1), 135–144.
Toloo, M., & Mensah, E. K. (2019). Robust optimization with nonnegative decision variables: A DEA approach. *Computers & Industrial Engineering*, *127*, 313–325.
Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, *130*(3), 498–509.
US Census Bureau. (2002). Economic Census, Manufacturing, Subject Series, General Summary.
Wang, Y.-M., Greatbanks, R., & Yang, J.-B. (2005). Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets and Systems*, *153*(3), 347–370.
Yanıkoğlu, İ., Gorissen, B. L., & den Hertog, D. (2019). A survey of adjustable robust optimization. *European Journal of Operational Research, 277*, 799–813.
Zamani, M. (2019). A new algorithm for concave quadratic programming. *Journal of Global Optimization*, *75*(3), 655–681.
Zanella, A., Camanho, A. S., & Dias, T. G. (2015). Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis. *European Journal of Operational Research*, *245*(2), 517–530.
Zohrehbandian, M., Makui, A., & Alinezhad, A. (2010). A compromise solution approach for finding common weights in DEA: an improvement to Kao and Hung's approach. *Journal of the Operational Research Society*, *61*(4), 604–610.