European Journal of Operational Research 300 (2022) 269–281

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

Procedures for ranking technical and cost efficient units: With a focus on nonconvexity[☆]

Kristiaan Kerstens[a,*], Jafar Sadeghi[b,c], Mehdi Toloo[d,e,f], Ignace Van de Woestyne[g]

[a] Univ. Lille, CNRS, IESEG School of Management, UMR 9221 - LEM - Lille Économie Management, Lille F-59000, France
[b] IESEG School of Management, 3 rue de la Digue, Lille F-59000, France
[c] Ivey Business School, Western University, London, Ontario, Canada
[d] Department of Business Transformation, Surrey Business School, University of Surrey, Guildford, GU2 7XH, UK
[e] Department of Systems Engineering, VSB-Technical University of Ostrava, Czech Republic
[f] Department of Operations Management & Business Statistics, College of Economics and Political Science, Sultan Qaboos University, Muscat, Oman
[g] KU Leuven, Research Unit MEES, Warmoesberg 26, Brussels B-1000, Belgium

A R T I C L E I N F O

Article history:
Received 6 July 2020
Accepted 11 October 2021
Available online 19 October 2021

Keywords:
Data envelopment analysis
Free disposal hull
Technical efficiency
Cost efficiency
Super-efficiency

A B S T R A C T

This contribution extends the literature on super-efficiency by focusing on ranking cost-efficient observations. To the best of our knowledge, the focus has always been on technical super-efficiency and this focus on ranking cost-efficient observations may well open up a new topic. Furthermore, since the convexity axiom has both an impact on technical and cost efficiency, we pay a particular attention to the effect of nonconvexity on both super-efficiency notions. Apart from a numerical example, we use a secondary data set guaranteeing replication to illustrate these efficiency and super-efficiency concepts. Two empirical conclusions emerge. First, the cost super-efficiency notion ranks differently from the technical super-efficiency concept. Second, both cost and technical super-efficiency notions rank differently under convex and nonconvex technologies.

1. Introduction

The seminal articles of Farrell (1957) and Charnes, Cooper, & Rhodes (1978) have contributed to making the nonparametric approach to production theory one of the grand success stories in the economics and operations research (OR) literatures in terms of both methodological developments and empirical applications. An early bibliographical survey article counts about 800 published articles and dissertations related to the so-called Data Envelopment Analysis (DEA) literature over the years 1978–1996 (see Seiford (1997)). One of the most recent bibliographic reviews of Emrouznejad & Yang (2018) surveys the first 40 years of scholarly literature in DEA over the period 1978 till 2016 and lists about 10,300 research articles. Empirical studies on efficiency and productivity using so-called frontier specifications are abundantly available and these frontier methodologies have become standard empirical tools that serve a variety of academic, regulatory and managerial purposes. Their widespread application in the academic literature analysing private and public sector performance-related issues can be glanced from, e.g., the Liu, Lu, & Lin (2013) survey of empirical frontier applications. But, also the implementation of incentive regulatory mechanisms (e.g. price cap regulation) using frontier-based performance benchmarks is, for instance, rather widespread in countries having liberalized their network industries (e.g., for the electricity industry, one may consult the survey by Jamasb & Pollitt (2000)). Finally, an example of a managerial application is the Sherman & Ladino (1995) study documenting how a US bank employs a basic frontier model to target sufficient savings in its branch network to fund its own expansion strategy.

While Farrell (1957) provided the first measurement scheme for the evaluation of technical and allocative efficiency resulting in cost efficiency in a frontier context, Färe, Grosskopf, & Lovell (1985, pp. 3-5) offer a more extended efficiency taxonomy by adding a scale efficiency component as well as a congestion component. In general, technical efficiency is solely based on information regarding physical inputs and outputs in the production process. The same holds true for congestion. By contrast, to evaluate cost efficiency one also needs information regarding input prices in addition to physical inputs and outputs. Scale efficiency can be measured in a technical way based on inputs and outputs alone, but very early on also a cost-based approach was suggested that also

☆ We acknowledge the most helpful comments of three constructive referees. The usual disclaimer applies.

* Corresponding author.
E-mail addresses: k.kerstens@ieseg.fr (K. Kerstens), j.sadeghi1987@gmail.com (J. Sadeghi), mehdi.toloo@vsb.cz (M. Toloo), ignace.vandewoestyne@kuleuven.be (I. Van de Woestyne).

https://doi.org/10.1016/j.ejor.2021.10.023
0377-2217/

requires information on input prices (see, e.g., Seitz (1970, 1971)). However, the variations on this basic measurement scheme are not central to our research questions.

The article of Andersen & Petersen (1993) is the first to ask a question with regard to the many observations that may obtain a similar relative technical efficiency status of unity. If all these observations are apparently equally technical efficient, is there any way to differentiate between these? Andersen & Petersen (1993, p. 1262) answer in the affirmative and state that "The basic idea is to compare the unit under evaluation with a linear combination of all other units in the sample, i.e., the DMU (Decision Making Unit) itself is excluded. It is conceivable that an efficient DMU may increase its input vector proportionally while preserving efficiency. The unit obtains in that case an efficiency score above one. The score reflects the radial distance from the DMU under evaluation to the production frontier estimated with that DMU excluded from the sample, i.e., the maximum proportional increase in inputs preserving efficiency." This basic idea has led to the so-called super-efficiency literature which investigates technical efficiency when the unit under evaluation is discarded from the technology.[1] This has led to explore issues of infeasibility under certain assumptions on technology and certain measurement orientations (see the surveys of Angulo-Meza & Estellita Lins (2002) and Adler & Volta (2019)).

In addition to ranking efficient observations, super-efficiency models have also been used to develop tests for influential observations or outliers (see, e.g., Banker & Chang (2006), among many others).

This same idea of super-efficiency has also been explored within the context of incentive-based regulation theory employing frontier methodologies (see Bogetoft (2000) and Agrell & Bogetoft (2017) for an early and a more recent survey). As far as we are aware of, Bogetoft (1994, p. 962) provides the first article within which a technical super-efficiency is defined to guarantee a proper incentive system. Agrell, Bogetoft, & Tind (2002, p. 6–7) similarly formulate a cost frontier from which the unit being evaluated is excluded from the technology definition. It seems that the existence of solutions for these technical and cost-based super-efficiency measures has not been given sufficient attention, especially in this incentive-based regulation theory employing frontier methodologies.

A first innovation of this contribution is to propose a super-efficiency model for the cost efficient observations. While cost efficiency is a sufficient condition for being technically efficient, technical efficiency is only a necessary but not a sufficient condition to being cost efficient. Thus, while there are fewer cost efficient than technically efficient observations, one could also wonder whether there is a way to discriminate between several of these cost efficient observations. This question seems never to have been treated in the literature.

A second innovation of this contribution is to focus on the impact of the convexity axiom on technical and cost super-efficiency models alike in the nonparametric, deterministic frontier tradition.[2] While it is tradition to impose convexity on the technology as a maintained axiom (see the seminal contributions of Farrell (1957) and Charnes et al. (1978)), Afriat (1972) is probably the first to mention a nonconvex Free Disposal Hull (FDH) model imposing only the assumptions of strong (free) disposal of inputs and outputs. This first nonconvex single output specification is generalized to the general multiple outputs case in the book chapter of Deprins, Simar, & Tulkens (1984). Kerstens & Vanden Eeckaut (1999) extend this basic FDH model by introducing specific returns to scale assumptions and by proposing a new goodness-of-fit method to characterize returns to scale for nonconvex technologies. Briec, Kerstens, & Vanden Eeckaut (2004) define nonconvex cost functions corresponding to the specific returns to scale assumptions in Kerstens & Vanden Eeckaut (1999).

Apart from these basic contributions, one can mention a variety of recent contributions to the further development of similar nonconvex production models. For instance, Tavakoli & Mostafaee (2019) are the first to extend network models to the class of nonconvex technologies as defined in Kerstens & Vanden Eeckaut (1999). Esteve, Aparicio, Rabasa, & Rodriguez-Sala (2020) combine regression trees from machine learning with the basic FDH model to estimate a new production frontier satisfying free disposability.

Most researchers are obviously aware of the fact that the convexity assumption affects the incidence and amount of technical inefficiency: under convexity less observations are technically efficient and the amount of technical inefficiency is higher. However, few people seem to realise that convexity also has consequences for the cost function: Briec et al. (2004) prove that convex cost functions are always smaller or equal to their convex counterparts for a given returns to scale assumption. There is only equality between convex and nonconvex cost functions in the case of a single output and constant returns to scale. In other words, the convexity assumption also affects the incidence and amount of cost inefficiency. Kerstens & Van de Woestyne (2021) empirically document the very substantial impact of convexity on cost function estimates in the literature (a total of 14 studies) as well as on two secondary data sets. In a similar vein, Kerstens, Sadeghi, & Van de Woestyne (2019) empirically document the substantial effect of convexity on technical efficiency and cost function based primal and cost-based capacity concepts.

This fact that convexity affects the incidence and amounts of all sorts of inefficiency makes the issue relevant in a super-efficiency model context. Apart from the early contribution of Van Puyenbroeck (1998) and the more recent one by Aldamak, Hatami-Marbini, & Zolfaghari (2016) focusing on technical super-efficiency under nonconvexity, we are unaware of any work focusing on the impact of convexity on technical and cost super-efficiency from a comparative perspective.

This contribution is structured as follows. Section 2 starts with the basic definitions of the technology, efficiency measures and the cost function. It also develops some basic efficiency decompositions. Then, to show the versatility of the approach, we discuss cost efficiency with incomplete information as introduced in Kuosmanen & Post (2001, 2003). Section 3 starts from the case of technical super-efficiency to introduce the new notion of cost super-efficiency and the ensuing super-efficiency decompositions. Thereafter, we also define cost super-efficiency for the incomplete information case. Both the technical and cost super-efficiency definitions are numerically illustrated with a figure in Section 4. Key results on technical and cost super-efficiency are developed in Section 5 and partially illustrated with numerical illustrations. An empirical illustration based on a secondary data set is developed in Section 6. We finish with some conclusions in the final Section 7.

2. Technology and cost function: basic definitions and decompositions

2.1. Technology and cost function: basic definitions

In this section, we define technology and some basic notation. Given N-dimensional input vectors $x \in \mathbb{R}^N_+$ and M-dimensional

[1] Banker & Chang (2006) mention an unpublished paper by Banker and Gifford from 1988 that delivered already the same idea.

[2] In the stochastic parametric and semi-parametric frontier models, all firms are inefficient since inefficiency is non-negative and the probability that inefficiency is exactly zero equals zero. Hence, there is no ranking problem of efficient observations. Kumbhakar, Parmeter, & Tsionas (2013) is the seminal article allowing to represent a mixture of both fully efficient and inefficient firms. This basic idea has been extended in several directions. However, we are unaware of the development of a super-efficiency notion in this part of the frontier literature.

output vectors $y \in \mathbb{R}^M_+$, the production possibility set or technology T can be defined as $T = \{(x, y) \mid x \text{ can produce at least } y\}$. It is common to impose some conditions on the input and output data defining the technology: (i) each producer uses nonnegative amounts of each input to produce nonnegative amounts of each output; (ii) there is an aggregate production of positive amounts of every output, and an aggregate use of positive amounts of every input; and (iii) each producer uses a positive amount of at least one input to produce a positive amount of at least one output (see, e.g., Färe, Grosskopf, & Lovell (1994, p. 44-45)). The input set $L(y) = \{x \mid (x, y) \in T\}$ associated with T holds all input vectors x capable of producing at least a given output vector y. In a similar way, the output set $P(x) = \{y \mid (x, y) \in T\}$ associated with T holds all output vectors y that can be produced from at most a given input vector x.

Throughout this contribution, technology T satisfies some combination of the following standard assumptions:

(T.1) Possibility of inaction and no free lunch, i.e., $(0, 0) \in T$ and if $(0, y) \in T$, then $y = 0$.
(T.2) T is a closed subset of $\mathbb{R}^N_+ \times \mathbb{R}^M_+$.
(T.3) Strong input and output disposal, i.e., if $(x, y) \in T$ and $(x', y') \in \mathbb{R}^N_+ \times \mathbb{R}^M_+$, then $(x', -y') \geq (x, -y) \Rightarrow (x', y') \in T$.
(T.4) $(x, y) \in T \Rightarrow \delta(x, y) \in T$ for $\delta \in \Gamma$, where:
  (i) $\Gamma \equiv \text{CRS} = \{\delta \mid \delta \geq 0\}$;
  (ii) $\Gamma \equiv \text{VRS} = \{\delta \mid \delta = 1\}$.
(T.5) T is convex.

Commenting these classical assumptions on the production technology, one can recall the following (see, e.g., Hackman (2008) for details). Inaction is possible and there exists no free lunch. Technology is a closed set. We assume free or strong disposability of both inputs and outputs in that inputs can be destroyed and outputs can be discarded at no opportunity costs. We allow independently for two returns to scale assumptions: either constant returns to scale (CRS), of variable returns to scale (VRS). Finally, the technology set is convex. Not all these axioms are maintained in the empirical analysis.[3] In particular, key assumptions distinguishing some of the technologies in the empirical analysis are CRS versus VRS, and convexity versus nonconvexity.

The input distance function completely characterizes the input set $L(y)$ and is defined as follows:

$$D_i(x, y \mid T) = \max_{\lambda}\{\lambda \mid \lambda \geq 0, (x/\lambda, y) \in T\} = \max_{\lambda}\{\lambda \mid \lambda \geq 0, x/\lambda \in L(y)\}. \tag{1}$$

Its main properties are: (i) $D_i(x, y \mid T) \geq 1$, with efficient production on the boundary (isoquant) of $L(y)$ represented by unity; (ii) it has a cost interpretation (see, e.g., Hackman (2008)). The inverse of this input distance function $DF_i(x, y \mid T) = [D_i(x, y \mid T)]^{-1}$ is known as the radial input efficiency measure. Therefore, this radial input efficiency measure is defined as:

$$DF_i(x, y \mid T) = \min_{\lambda}\{\lambda \mid \lambda \geq 0, (\lambda x, y) \in T\} = \min_{\lambda}\{\lambda \mid \lambda \geq 0, \lambda x \in L(y)\}. \tag{2}$$

Obviously, it is situated between zero and unity ($0 < DF_i(x, y) \leq 1$), with efficient production on the boundary (isoquant) of the input set $L(y)$ represented by unity.

Looking to a dual representation of technology, the cost function is defined as the minimum expenses required to produce a given output vector y for a given vector of semi-positive input prices ($w \in \mathbb{R}^N_+$):

$$C(y, w \mid T) = \min_{x}\{wx \mid (x, y) \in T\} = \min_{x}\{wx \mid x \in L(y)\}. \tag{3}$$

Duality relations link these primal and dual representations of technology. Duality allows a well-behaved technology to be reconstructed from the observations on cost minimizing producer behavior, and the reverse. The duality between input distance function (1) and cost function (3) can be stated as follows:

$$D_i(x, y \mid T) = \min_{w}\{wx \mid C(y, w \mid T) \geq 1\}, x \in L(y), \tag{4}$$

$$C(y, w \mid T) = \min_{x}\{wx \mid D_i(x, y \mid T) \geq 1\}, w > 0. \tag{5}$$

It is common to establish such duality relations under the hypothesis of a convex technology or a convex input set (e.g., Hackman, 2008, Ch. 7). Briec et al. (2004) are the first to establish a local duality result between nonconvex technologies subject to various scaling laws and their corresponding nonconvex cost functions.

2.2. Nonparametric frontier technology and cost function specification

For the methodological results and the empirical application, we assume a convex or nonconvex, nonparametric frontier technology under VRS and CRS assumptions. Let K input-output combinations $(x_k, y_k) \in \mathbb{R}^N_+ \times \mathbb{R}^M_+$ ($k \in \{1, \dots, K\}$) define the technology, then a unified algebraic representation of the corresponding convex and nonconvex nonparametric frontier technologies under the VRS and CRS assumptions is as follows:

$$T^{\Lambda,\Gamma} = \left\{(x, y) \mid x \geq \sum_{k=1}^{K} x_k \delta z_k, y \leq \sum_{k=1}^{K} y_k \delta z_k, z = (z_1, \dots, z_K) \in \Lambda, \delta \in \Gamma\right\}, \tag{6}$$

where

(i) $\Gamma \equiv \text{CRS} = \{\delta \mid \delta \geq 0\}$;
(ii) $\Gamma \equiv \text{VRS} = \{\delta \mid \delta = 1\}$;

and

(i) $\Lambda \equiv \text{C} = \left\{z \mid \sum_{k=1}^{K} z_k = 1 \text{ and } \forall k \in \{1, \dots, K\} : z_k \geq 0\right\}$;
(ii) $\Lambda \equiv \text{NC} = \left\{z \mid \sum_{k=1}^{K} z_k = 1 \text{ and } \forall k \in \{1, \dots, K\} : z_k \in \{0, 1\}\right\}$.

The activity vector z operates subject to a convexity ($\Lambda \equiv \text{C}$) or a nonconvexity ($\Lambda \equiv \text{NC}$) constraint. The activity vector z of real numbers summing to unity represents the convexity axiom, while this same sum constraint with each vector element being a binary represents nonconvexity. There is also a scaling parameter (δ) allowing for a particular scaling of all K observations determining the technology: this scaling parameter is free under CRS and fixed at unity under VRS. Briec et al. (2004) demonstrate that nonconvex CRS and VRS technologies satisfy a minimum extrapolation principle.

To compute the input efficiency measure (2) relative to convex technologies in (6) requires solving a nonlinear programming (NLP) problem for each evaluated observation. However, Briec & Kerstens (2006, Lemma 2.1) show how this NLP can be transformed to a linear programming (LP) problem by substituting $\bar{z}_k = \delta z_k$ in (6). For the nonconvex technologies, nonlinear binary mixed integer programs must be solved, but alternative solution strategies (including implicit enumeration) are available (see Kerstens & Van de Woestyne, 2014).

Assuming the nonparametric frontier technology $T^{\Lambda,\Gamma}$, the cost function (3) defined as the minimum expenditures needed to produce a given output vector $y_p \in \mathbb{R}^M_+$ for a given vector of semi-positive common input prices $w \in \mathbb{R}^N_+$ is obtained by solving the following programming problem (Färe et al., 1985):

$$C(y_p, w \mid T^{\Lambda,\Gamma}) = \min_{x}\{wx \mid (x, y_p) \in T^{\Lambda,\Gamma}\} = \min_{x}\{wx \mid x \in L(y_p)\}. \tag{7}$$

[3] Note that the convex VRS technology does not satisfy inaction.

Computing the cost function relative to convex technologies is straightforward when transforming the NLP to the corresponding LP problem as indicated above. Computing the cost function relative to nonconvex technologies can be done using implicit enumeration algorithms specified in Briec et al. (2004).

2.3. Overall or cost efficiency decomposition

Following Farrell (1957), the following basic overall or cost efficiency measure lends itself to a decomposition into technical and allocative efficiencies. A natural measure for cost efficiency is to take the ratio of minimum to actual observed cost:

$$CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = \frac{C(y_p, w \mid T^{\Lambda,\Gamma})}{wx_p}. \tag{8}$$

Similar to technical efficiency, cost efficiency is situated between zero and unity: i.e., $0 < CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \leq 1$.

Allocative efficiency is then defined as a ratio of cost efficiency over technical efficiency. Hence, the allocative efficiency of observation (x_p, y_p) is defined as a residual as follows:

$$AE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = \frac{CE(x_p, y_p, w \mid T^{\Lambda,\Gamma})}{DF_i(x_p, y_p \mid T^{\Lambda,\Gamma})}. \tag{9}$$

Note that the above components are defined for convex and nonconvex technologies under both CRS and VRS and it is easy to see that the allocative efficiency component always exceeds the cost efficiency measure: i.e., $AE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \geq CE(x_p, y_p, w \mid T^{\Lambda,\Gamma})$.

While Farrell (1957) already explores the notion of scale efficiency, Färe et al. (1985) define scale efficiency by a ratio of technical efficiency under CRS over technical efficiency under VRS. Hence, the scale efficiency of observation (x_p, y_p) is defined as follows:

$$SCE(x_p, y_p \mid T^{\Lambda,\cdot}) = \frac{DF_i(x_p, y_p \mid T^{\Lambda,CRS})}{DF_i(x_p, y_p \mid T^{\Lambda,VRS})}. \tag{10}$$

Clearly, scale efficiency is situated between zero and unity, i.e., $0 < SCE(x_p, y_p \mid T^{\Lambda,\cdot}) \leq 1$. The superscript Λ indicates that scale efficiency can be defined under both convex and nonconvex cases. Alternatively, Seitz (1971) defines a cost-based scale efficiency component as the ratio of cost functions under CRS and VRS.

Based on (9) and (10), CRS technical efficiency can be decomposed into VRS technical efficiency and scale efficiency:

$$\begin{aligned} CE(x_p, y_p, w \mid T^{\Lambda,CRS}) &= DF_i(x_p, y_p \mid T^{\Lambda,CRS}) \times AE(x_p, y_p, w \mid T^{\Lambda,CRS}) \\ &= DF_i(x_p, y_p \mid T^{\Lambda,VRS}) \times SCE(x_p, y_p \mid T^{\Lambda,\cdot}) \times AE(x_p, y_p, w \mid T^{\Lambda,CRS}). \end{aligned} \tag{11}$$

Färe et al. (1985) refine this decomposition by also distinguishing between pure technical efficiency and structural efficiency or congestion efficiency. Other decompositions are available in the literature: examples include Aparicio, Pastor, & Zofio (2015) and Aparicio, Ortiz, & Pastor (2017).

2.4. Overall or cost efficiency with incomplete prices

When price information is incomplete, Kuosmanen & Post (2001, 2003) derive upper and lower bounds for cost efficiency and allocative efficiency assuming incomplete price data in the form of a convex polyhedral cone. For this purpose, Kuosmanen & Post (2001) use an equivalent formulation of cost efficiency (8) which represents cost efficiency at normalised prices:

$$CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = \min_{x' \in L(y)} \{w'x' \mid w' = \alpha w, \alpha \in \mathbb{R}_+, w'x_p = 1\}. \tag{12}$$

In (12) input prices are normalized such that the normalized cost of DMU_p equals unity. To relax the information requirement on prices, Kuosmanen & Post (2001) present a convex polyhedral cone $W \subset \mathbb{R}^N_+$ as the price domain of inputs as follows:

$$W = \{w \in \mathbb{R}^N_+ \mid Aw \geq 0\}, \tag{13}$$

where A is a $L \times N$ matrix composed of L row vectors $A_1, \ldots, A_L$. Therefore, this cone W represents the input price domain by means of L linear inequalities.

To obtain an upper bound for cost efficiency, Kuosmanen & Post (2001) use the maximum value of cost efficiency over the price domain as follows:

$$\overline{CE}(x_p, y_p \mid T^{\Lambda,\Gamma}) = \max_{w \in W}\left(\min_{x' \in L(y_p)} \{wx' \mid wx_p = 1\}\right). \tag{14}$$

Kuosmanen & Post (2001, p. 52) prove that model (14) under input convexity, nonconvexity of outputs, and VRS can be equivalently obtained by solving the following model:

$$\begin{aligned} \overline{CE}(x_p, y_p) = \quad & \min_{\theta, z_k, \beta} \quad \theta \\ & s.t \quad \sum_{k \in \chi(y_p)} z_k x_k + \beta A \leq \theta x_p, \\ & \qquad \sum_{k \in \chi(y_p)} z_k = 1, \\ & \qquad \theta \geq 0, \beta \in \mathbb{R}^L_+, \\ & \qquad z_k \geq 0, \qquad k \in \chi(y_p), \end{aligned} \tag{15}$$

where $\chi(y_p) = \{k \mid y_k \geq y_p\}$ indicates all the observations that have a higher output level than y_p. Note that $p \in \chi(y_p)$, therefore $\chi(y_p) \neq \emptyset$.

To obtain a lower bound for cost efficiency, Kuosmanen & Post (2001, 2003) use the minimum value of cost efficiency over the price domain as follows:

$$\underline{CE}(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = \min_{w \in W}\left(\min_{x' \in L(y_p)} \{wx' \mid wx_p = 1\}\right). \tag{16}$$

To solve model (16), Kuosmanen & Post (2003, p. 456) reverse the order of the two minimization problems as follows:

$$\underline{CE}(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = \min_{x' \in L(y_p)}\left(\min_{w \in W}\{wx' \mid wx_p = 1\}\right). \tag{17}$$

Therefore, we first need to solve model $\min_{w \in W}\{wx_k \mid wx_p = 1\}$ for each $k \in \chi(y_p)$ and then we select the minimum of their optimal values as the lower bound of cost efficiency of DMU_p (see Section 4 in Kuosmanen & Post (2003, p. 456) for details).

Obviously, corresponding to these upper and lower bounds on overall or cost efficiency, it is straightforward to define a corresponding allocative efficiency component.

3. Super-efficiency: existing and new definitions and decompositions

3.1. Technical super-efficiency: existing definition

Andersen & Petersen (1993) develop a procedure to rank technically efficient units. Their method enables an efficient observation (x_p, y_p) to achieve an efficiency score greater than or equal to unity by removing this observation from those defining technology $T^{\Lambda,\Gamma}$. Referring with subscript p to the exclusion of this pth observation, the resulting technology can mathematically be formulated as follows:

$$T_p^{\Lambda,\Gamma} = \left\{(x, y) \mid x \geq \sum_{k=1}^{K} x_k \delta z_k, y \leq \sum_{k=1}^{K} y_k \delta z_k, z_p = 0, z \in \Lambda, \delta \in \Gamma\right\}. \tag{18}$$

Observe that the removal of the pth observation is realized by forcing the pth component of the activity vector to be zero. Consequently, this pth observation no longer influences the constraints.

The traditional technical efficiency measure (2) can now be transformed in a super-efficiency context as follows:

Definition 3.1. The input-oriented technical super-efficiency model is defined as:

$$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) = \min_{\lambda}\{\lambda \mid \lambda \geq 0, (\lambda x_p, y_p) \in T_p^{\Lambda,\Gamma}\}, \tag{19}$$

with $T_p^{\Lambda,\Gamma}$ defined in (18).

Note that $DF_i(x_p, y_p \mid T^{\Lambda,\Gamma}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$ for technically inefficient observations, while for technically efficient observations, $DF_i(x_p, y_p \mid T^{\Lambda,\Gamma}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$ (see (Andersen & Petersen, 1993, p. 1263)). Consequently, observation (x_p, y_p) is technically inefficient by the conventional model (2) if and only if the super-efficiency estimate $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) < 1$ also indicates inefficiency. This observation is technically efficient by the conventional model (2) if and only if $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) \geq 1$: it indicates efficiency when $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) = 1$ and it indicates super-efficiency when $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) > 1$.

Note that the super-efficiency model $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$ can lead to infeasibilities in practical applications. In fact, this observation has led to a rather substantial literature: the reader is referred to the surveys in Angulo-Meza & Estellita Lins (2002), Soltanifar & Shahghobadi (2014) and Adler & Volta (2019) for further details.

3.2. Cost super-efficiency: new definition

The previous section focused on the technical super-efficiency where input price information is not available. The super-efficiency concept seems never to have been applied in a cost function context. This section turns to the topic of cost efficiency to show how nonparametric frontiers can be used to identify a cost super-efficiency concept when information on input prices and costs are known exactly.

A natural measure for cost super-efficiency of observation (x_p, y_p) is to take the ratio of the minimum cost when this observation is removed from the set of observations, relative to the actual observed cost. Therefore, we define the cost super-efficiency of (x_p, y_p) by considering the production possibility set $T_p^{\Lambda,\Gamma}$ defined in (18) as follows:

Definition 3.2. The cost super-efficiency of observation (x_p, y_p) is defined by

$$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = \frac{C^{SE}(y_p, w \mid T_p^{\Lambda,\Gamma})}{wx_p}, \tag{20}$$

where $C^{SE}(y_p, w \mid T_p^{\Lambda,\Gamma})$ is obtained as follows:

$$C^{SE}(y_p, w \mid T_p^{\Lambda,\Gamma}) = \min_{x}\{wx \mid x \geq 0, (x, y_p) \in T_p^{\Lambda,\Gamma}\}. \tag{21}$$

Note that $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = CE(x_p, y_p, w \mid T^{\Lambda,\Gamma})$ for cost inefficient observations, while $CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \leq CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma})$ for cost efficient observations. Consequently, observation (x_p, y_p) is cost inefficient by definition (8) if and only if the cost super-efficiency estimate $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) < 1$, and it is cost efficient (i.e., $CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) = 1$) if and only if $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) \geq 1$.

3.3. Overall or cost super-efficiency decomposition: new definition

By analogy to Section 2.3, we now decompose the cost super-efficiency (20) into several efficiency components.

Allocative super-efficiency is defined as the ratio of cost super-efficiency over the technical super-efficiency. Hence, the allocative super-efficiency of observation (x_p, y_p) is defined as follows:

$$AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = \frac{CE^{SE}(w_p, y_p \mid T_p^{\Lambda,\Gamma})}{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})}. \tag{22}$$

Note again that the above components are defined for convex and nonconvex technologies alike under both CRS and VRS.

Based on the CRS and VRS technical super-efficiency scores, we can define a new scale super-efficiency by taking ratio of technical super-efficiency under CRS over technical super-efficiency under VRS. Hence, this scale super-efficiency of observation (x_p, y_p) is defined as follows:

$$SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot}) = \frac{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})}{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})}. \tag{23}$$

Clearly, $0 < SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot}) \leq 1$. Similar to Seitz (1971), one could equally define a cost-based scale super-efficiency component as the ratio of cost functions under CRS and VRS computed relative to technology (18).

Furthermore, based on (22) and (23), CRS technical super-efficiency can be decomposed into VRS technical super-efficiency and scale super-efficiency as follows:

$$\begin{aligned} CE^{SE}(w_p, y_p \mid T_p^{\Lambda,CRS}) &= DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS}) \times AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS}) \\ &= DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \times SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot}) \times AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS}). \end{aligned} \tag{24}$$

Note that if observation (x_p, y_p) is technically inefficient under VRS (i.e., $DF_i(x_p, y_p \mid T^{\Lambda,VRS}) < 1$), then $DF_i(x_p, y_p \mid T^{\Lambda,VRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$, $DF_i(x_p, y_p \mid T^{\Lambda,CRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$ and $CE(w_p, y_p \mid T^{\Lambda,CRS}) = CE^{SE}(w_p, y_p \mid T_p^{\Lambda,CRS})$. Hence, in this case, technical efficiency with technical super-efficiency, allocative efficiency with allocative super-efficiency and scale efficiency with scale super-efficiency are identical. Consequently, if the observation is technically inefficient under VRS, then its decomposition of overall or cost efficiency (11) and its decomposition of overall or cost super-efficiency (24) lead to the same results.

3.4. Overall or cost super-efficiency with incomplete prices: new definition

To calculate an upper bound of the cost super-efficiency under the maintained assumptions with respect to price domain W, we introduce the following model:

$$\begin{aligned} \overline{CE}^{SE}(x_p, y_p) = \quad & \min_{\theta, z_k, \beta} \quad \theta \\ & s.t \quad \sum_{k \in \chi(x_p)} z_k y_k + \beta A \leq \theta x_p, \\ & \qquad \sum_{k \in \chi(y_p)} z_k = 1, \\ & \qquad \theta \geq 0, \beta \in \mathbb{R}_+^L, \\ & \qquad z_p = 0, \\ & \qquad z_k \geq 0, \qquad k \in \chi(y_p), k \neq p, \end{aligned} \tag{25}$$

where $\chi(y_p)$ is defined as above. Observe that the removal of the pth observation is implemented by forcing the pth component of the activity vector to be equal to zero. Consequently, this pth observation no longer influences the constraints. If we have $\chi(y_p) = \{p\}$, then model (25) is infeasible and there is no upper bound for the cost super-efficiency of DMU$_p$.

To measure the lower bound of the cost super-efficiency under the maintained assumptions with respect to price domain W, we introduce the following model:

$$\underline{CE}^{SE}(x_p, y_p) = \min_{k \in \chi(y_p), k \neq p} \left(\min_{w \in W}\{wx_k \mid wx_p = 1\}\right). \tag{26}$$

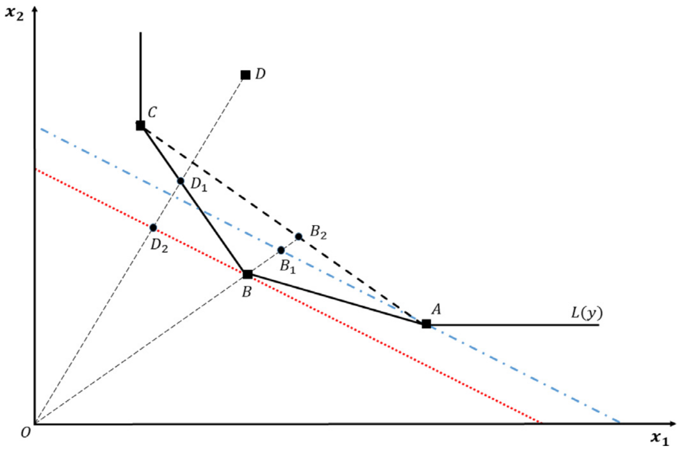

Fig. 1. Isoquant with combinations of two inputs yielding a given output level y.

Therefore, we first solve model $\min_{w\in W}\{wx_k \mid wx_p = 1\}$ for each $k \in \chi(y_p), k \neq p$ and then we select the minimum of their optimal values as the lower bound of the cost super-efficiency of DMU$_p$. If we have $\chi(y_p) = \{p\}$, then there is no lower bound for the cost super-efficiency of DMU$_p$.

4. Numerical illustration

Now we clarify the above Definitions 3.1 and 3.2 by means of Fig. 1. The isoquant indicating the combinations of two inputs x_1 and x_2 yielding a given output level y is represented by the polyline ABC and its horizontal and vertical extensions at A and C respectively. This isoquant is the boundary of the input set $L(y)$. We first focus on observation B to illustrate technical and cost efficiency as well as technical and cost super-efficiency measures.

The input prices of observation B are such that the cost-minimizing input vector is located at point B on the boundary of the input set $L(y)$: this cost function is plotted as a dotted line. The corresponding technical and cost efficiency measures for point B are $DF_i(x_B, y_B \mid T^{C,VRS}) = \frac{OB}{OB} = 1$ and $CE(x_B, y_B, w \mid T^{C,VRS}) = \frac{OB}{OB} = 1$, respectively. Therefore, since $DF_i(x_B, y_B \mid T^{C,VRS}) = CE(x_B, y_B, w \mid T^{C,VRS}) = 1$ this observation B is both technically and cost efficient.

By removing observation B from the observations defining the technology, the boundary of the corresponding input set at output level y is represented by the dashed line segment AC and its horizontal and vertical extensions at A and C respectively. In this case, the corresponding cost-minimizing input vector is located at point A: the corresponding cost function is drawn as a dash dotted line. In this situation, the technical and cost super-efficiency measures of observation B are $DF_i^{SE}(x_B, y_B \mid T_B^{C,VRS}) = \frac{OB_2}{OB} > 1$ and $CE^{SE}(x_B, y_B, w \mid T_B^{C,VRS}) = \frac{OB_1}{OB} > 1$, respectively. Since $OB_2 > OB_1$, $DF_i^{SE}(x_B, y_B \mid T^{C,VRS}) > CE^{SE}(x_B, y_B, w \mid T_B^{C,VRS}) > 1$ meaning that observation B is technically and cost super-efficient.

Now, we focus on observation D and assume it faces the same input prices as observation B. For this observation D, the corresponding technical and cost efficiency measures are $DF_i(x_D, y_D \mid T^{C,VRS}) = \frac{OD_1}{OD} < 1$ and $CE(x_D, y_D, w \mid T^{C,VRS}) = \frac{OD_2}{OD} < 1$, respectively. Since $OD_1 > OD_2$, $CE(x_D, y_D, w \mid T^{C,VRS}) < DF_i(x_D, y_D \mid T^{C,VRS}) < 1$ from which we conclude that observation D is both technically and cost inefficient. In this case, the boundary of the input set $L(y)$ for a given output level y, which is shown by the polyline ABC and its horizontal and vertical extensions at A and C respectively, remains unchanged after removing observation D from the observations defining technology. Thus, by removing this inefficient unit from technology, both technical and cost efficiency measures remain unchanged. Hence, $DF_i(x_D, y_D \mid T^{C,VRS}) = DF_i^{SE}(x_D, y_D \mid T_D^{C,VRS})$ and $CE(x_D, y_D, w \mid T^{C,VRS}) = CE^{SE}(x_D, y_D, w \mid T_D^{C,VRS})$.

5. Technical and cost super-efficiency: key results

A first Proposition 5.1 compares the technical and cost efficiency measures as well as the technical and cost super-efficiency measures:

Proposition 5.1. *For every observation* $(x_p, y_p) \in T^{\Lambda,\Gamma}$*:*

(i) $CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \le DF_i(x_p, y_p \mid T^{\Lambda,\Gamma}) \le 1$.
(ii) *The number of cost efficient observations is smaller than or equal to the number of input-oriented technical efficient observations.*
(iii) $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) \le DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$.

Proof. See Appendix A. □

Proposition 5.1 shows that the cost efficiency and cost super-efficiency measures are always smaller than or equal to the technical efficiency and technical super-efficiency measures, respectively, in both convex and nonconvex cases and under both CRS and VRS. Moreover, properties (i) and (iii) of Proposition 5.1 directly lead to the following corollary:

Corollary 5.1.

(i) *Allocative efficiency is situated between zero and unity, i.e., for every observation* $(x_p, y_p) \in T^{\Lambda,\Gamma}$*:* $0 < AE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \le 1$.
(ii) *Allocative super-efficiency is situated between zero and unity, i.e., for every observation* $(x_p, y_p) \in T^{\Lambda,\Gamma}$*:* $0 < AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) \le 1$.

The next Proposition 5.2 compares the technical efficiency and technical super-efficiency measures as well as the cost efficiency and cost super-efficiency measures in both convex and nonconvex cases and under both CRS and VRS:

Proposition 5.2. *For every observation* (x_p, y_p)*:*

(i) $DF_i(x_p, y_p \mid T^{C,\Gamma}) \le DF_i(x_p, y_p \mid T^{NC,\Gamma})$;
(ii) $DF_i^{SE}(x_p, y_p \mid T_p^{C,\Gamma}) \le DF_i^{SE}(x_p, y_p \mid T_p^{NC,\Gamma})$;
(iii) $CE(x_p, y_p, w \mid T^{C,\Gamma}) \le CE(x_p, y_p, w \mid T^{NC,\Gamma})$;
(iv) $CE^{SE}(x_p, y_p, w \mid T_p^{C,\Gamma}) \le CE^{SE}(x_p, y_p, w \mid T_p^{NC,\Gamma})$.

Proof. See Appendix A. □

When comparing convex and nonconvex results for all efficiency concepts (i.e., technical and cost efficiency as well as technical and cost super-efficiency measures), then the obtained results under the convexity assumption are always smaller than or equal to the ones obtained under the nonconvexity assumption. Note that Proposition 5.2(i) and (iii) is already proven in Briec et al. (2004, p. 178).

Zhu (1996) indicates that the input-oriented super-efficiency model under CRS is feasible unless certain patterns of zero data entries are present in the inputs. Therefore, if one assumes that all inputs data are strictly positive, then the input-oriented super-efficiency CRS model is always feasible. However, in Proposition 5.3(i), we prove that the cost super-efficiency model under CRS is always feasible even if there are some zero data entries in the inputs. Proposition 5.3 focuses on the potential for infeasibility for both technical and cost super-efficiency:

Proposition 5.3.

(i) *For every observation* $(x_p, y_p) \in T^{\Lambda,CRS}$*, the cost super-efficiency model* $C^{SE}(y_p, w \mid T_p^{\Lambda,CRS})$ *is always feasible.*

(ii) *Model* $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$ *is feasible if and only if model* $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$ *is feasible.*
(iii) *If model* $DF_i^{SE}(x_p, y_p \mid T_p^{C,\Gamma})$ *is infeasible, then model* $DF_i^{SE}(x_p, y_p \mid T_p^{NC,\Gamma})$ *is infeasible.*
(iv) *If model* $CE^{SE}(x_p, y_p, w \mid T_p^{C,VRS})$ *is infeasible, then model* $CE^{SE}(x_p, y_p, w \mid T_p^{NC,VRS})$ *is infeasible.*
(v) *Model* $\overline{CE}^{SE}(x_p, y_p)$ *is feasible if and only if model* $\underline{CE}^{SE}(x_p, y_p)$ *is feasible. Moreover, the infeasibility of these two models* $\overline{CE}^{SE}(x_p, y_p)$ *and* $\underline{CE}^{SE}(x_p, y_p)$ *is independent from the price domain of inputs* W.

Proof. See Appendix A. □

Proposition 5.3(i) shows that the cost super-efficiency model under CRS is always feasible. Proposition 5.3(ii) shows that infeasibilities occur for the same observations in both corresponding models $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$ under both convex and nonconvex cases. Consequently, the number of infeasible observations by solving the corresponding models $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$ is identical.

Note that Proposition 5.3(ii) shows that the infeasibility of model $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$ under both convex and nonconvex cases are independent from the input prices.

According to Proposition 5.3(iii) and (iv), if an observation results in an infeasibility under convexity, then this same observation also leads to an infeasibility under nonconvexity for both technical and cost super-efficiency measures. Note that the inverses of Proposition 5.3(iii) and (vi) are not satisfied. Consequently, the number of infeasible observations in the convex case is smaller than or equal to the number of infeasible observations in the nonconvex case for both technical and cost super-efficiency measures. The potential incidence of infeasibility is worse under nonconvexity because the volume of the technology is smaller than or equal to the volume of its convex counterpart.

Proposition 5.3 (v) shows that infeasibilities occur for the same observations in both upper and lower bounds models $\overline{CE}^{SE}(x_p, y_p)$ and $\underline{CE}^{SE}(x_p, y_p)$. Moreover, infeasibility of these two upper and lower bounds models are independent from the price domain of inputs W.

While Agrell et al. (2002, p. 6–7) define the cost frontier for CRS and VRS assumptions alike, it is clear that a super-efficiency cost frontier version may not exist under VRS, but it always exists under CRS. Unfortunately, this lack of feasibility under VRS is much more important from a policy viewpoint than this existence result under CRS, since few economists and regulators are advocating to impose CRS when determining a frontier-based cost norm. This existence result under CRS seems new to the incentive-based regulation theory employing frontier methodologies.

To illustrate the difference in the infeasibility behaviour between technical super-efficiency and cost super-efficiency in Proposition 5.3(i), consider the isoquant of a CRS technology in Fig. 2. This isoquant of technology $T^{C,CRS}$ is determined by 6 observations (A, B, D, E, F, G, H) denoted by square dots in the space of two inputs for a given single output. It is represented by the polyline $ABDE$ and its vertical and horizontal extensions at A and E, respectively.

We focus on observation E with a second input equal to zero to illustrate technical and cost efficiency as well as technical and cost super-efficiency measures. The input prices of observation E are such that the cost-minimizing input vector is located at E on the frontier of $T^{C,CRS}$: this cost function is plotted as a dotted line. The corresponding technical and cost efficiency measures for observation E are $DF_i(x_E, y_E \mid T^{C,CRS}) = CE(x_E, y_E, w \mid T^{C,CRS}) = \frac{OE}{OE} = 1$, respectively. Consequently, observation E is technically and cost efficient.

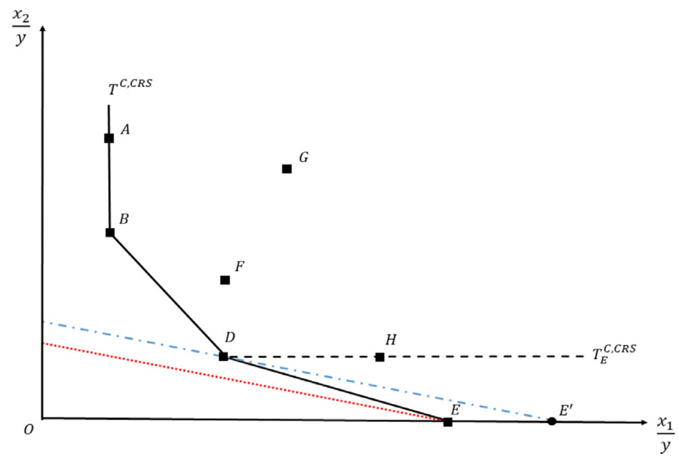

Fig. 2. Isoquant with combinations of two inputs and single output under convexity and CRS assumptions.

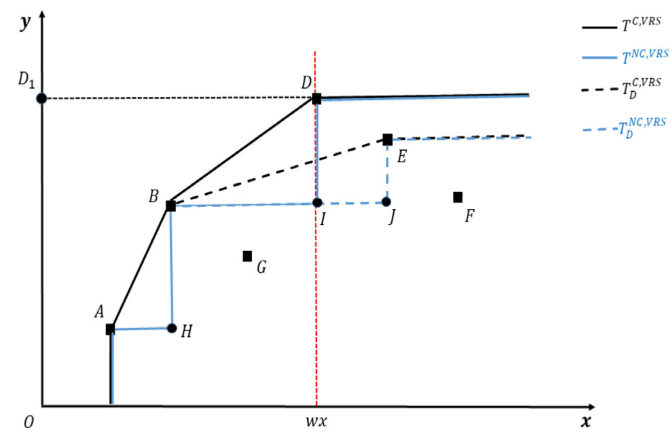

Fig. 3. Convex and nonconvex technologies with single input and single output under VRS assumption.

By removing observation E from the observations defining technology, the boundary of the altered technology $T_E^{C,CRS}$ is represented by the polyline $ABDH$ and its vertical and horizontal extensions at A and H, respectively. Clearly, there is no referent observation for E when scaling down or up both inputs in a radial way: to be precise, the ray from the origin through point E (which coincides with the horizontal axis) has no intersection with the altered technology $T_E^{C,CRS}$. Consequently, the input-oriented technical super-efficiency model under CRS is infeasible for observation E.

However, cost minimization is still possible when observation E is removed and results in the cost-minimizing input vector located at D on the frontier of technology $T^{C,CRS}$: this cost function is drawn in a dash dotted line. Hence, the corresponding cost super-efficiency measure of observation E is simply $CE^{SE}(x_E, y_E, w \mid T_E^{C,CRS}) = \frac{OE'}{OE} > 1$ and is perfectly feasible.

Fig. 3 illustrates Proposition 5.3(ii), namely how infeasibility for super-efficiency under VRS may occur for the case of a single input and a single output under both convex and nonconvex cases. The six observations (A, B, D, E, F, G) determining technologies $T^{C,VRS}$ and $T^{NC,VRS}$ are denoted by square dots. The boundary of technology $T^{C,VRS}$ is represented by the polyline ABD and its vertical and horizontal extensions at A and D, respectively. The boundary of technology $T^{NC,VRS}$ is represented by the polyline $AHBID$ and its vertical and horizontal extensions at A and D, respectively.

We focus on observation D to illustrate technical and cost efficiency as well as technical and cost super-efficiency measures. The cost function minimizing expenditures at the output level of D is

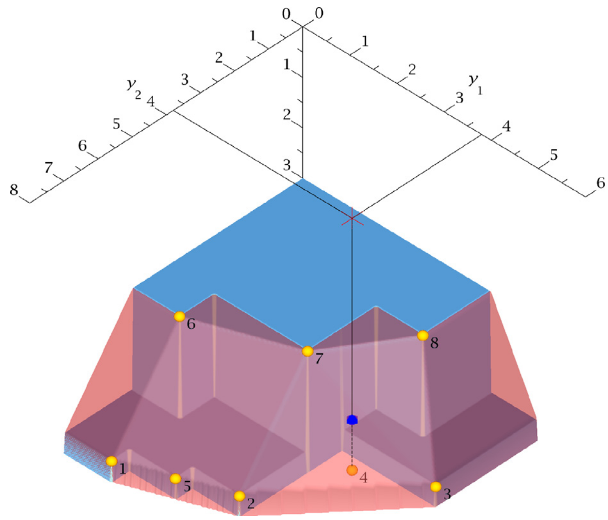

Fig. 4. Convex and nonconvex technologies with single input and two outputs under VRS assumption.

plotted as a dotted line. The corresponding technical and cost efficiency measures for observation D under both convex and nonconvex cases are $DF_i(x_D, y_D \mid T^{C,VRS}) = DF_i(x_D, y_D \mid T^{NC,VRS}) = \frac{D_1D}{D_1D} = 1$ and $CE(x_D, y_D, w \mid T^{C,VRS}) = CE(x_D, y_D, w \mid T^{NC,VRS}) = \frac{D_1D}{D_1D} = 1$, respectively. Therefore, we have $CE(x_D, y_D, w \mid T^{\Lambda,VRS}) = DF_i(x_D, y_D \mid T^{\Lambda,VRS}) = 1$ for $\Lambda = C$ and $\Lambda = NC$. Hence, observation D is technically and cost efficient under both convexity and nonconvexity.

Removing observation D alters technology in the convex case to the new technology $T_D^{C,VRS}$ represented by the polyline ABE and its vertical and horizontal extensions at A and E, respectively. Similarly, removal of observation D changes technology in the nonconvex case to the new technology $T_D^{NC,VRS}$ represented by the polyline $AHBJE$ and its vertical and horizontal extensions at A and E, respectively. Clearly, there are no observations in technologies $T_D^{C,VRS}$ and $T_D^{NC,VRS}$ with at least current output level of observation D. Therefore, both technical and cost super-efficiency measures under VRS and under convexity or nonconvexity are infeasible.

Fig. 4 illustrates Proposition 5.3(iii), namely if an observation results in an infeasibility under convexity, then this same observation also leads to an infeasibility under nonconvexity for the technical super-efficiency measure. Also, it shows that the inverse of Proposition 5.3(iii) is not satisfied: i.e., if an observation results in an infeasibility under nonconvexity, then this same observation may lead to a feasibility under convexity for the technical super-efficiency measure.

The visualization in Fig. 4 depicts the case of one input (vertical direction pointing downwards) and two outputs (upper horizontal plane). We focus on observation 4 indicated by a small sphere which is an inefficient unit under the convex case and an efficient unit under the nonconvex case. The visible production frontiers are those with observation 4 excluded: the nonconvex case is in a solid color, and the convex case is in a slightly transparent color. Observation 4 is scaled upward to the position indicated by the cube when optimizing in the input direction and assuming convexity. But, in the case of nonconvexity, there is no scaling up or down that reaches the boundary of technology with observation 4 excluded, whence the infeasibility.

The problem of infeasibilities can lead to several reactions. One position that has been defended in the literature on the infeasibility of the Luenberger productivity indicator is to accept the fact that distance functions need not always be defined and to simply report all infeasibilities whenever these occur (see, e.g., Briec & Kerstens (2009)). The same argument could be applied to the cost function. This is the position taken in this contribution. Another position is to remedy the infeasibilities in distance function and/or cost function whenever this is possible. While some undeniable progress has been made to reduce or eliminate the infeasibility problem for distance functions (see the surveys cited at the end of subsection 3.1), we are unaware of any solution to the infeasibility problem for the cost function. Moreover, depending on whether or not one attaches some importance to the static efficiency decomposition one may wish to have solutions that are structurally similar across different super-efficiency concepts. We think that it is desirable to find a solution to the infeasibility problem that works in a similar way for technical and cost super-efficiency concepts alike. It remains an open question whether this is possible.[4]

Proposition 5.4 compares the allocative and scale efficiency measures with the allocative and scale super-efficiency measures:

Proposition 5.4. *For every observation* (x_p, y_p)*:*

(i) $AE(x_p, y_p, w \mid T^{\Lambda,CRS}) \geq AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$
(ii) $SCE(x_p, y_p \mid T^{\Lambda,.,.}) \geq SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,.,.})$

Proof. See Appendix A. □

Proposition 5.4 shows that the allocative and scale efficiency measures are greater than or equal to the allocative and scale super-efficiency measures, respectively, under both convex and nonconvex cases.

Proposition 5.5 compares the allocative super-efficiency measure with the cost super-efficiency measure:

Proposition 5.5. *For every observation* (x_p, y_p)*:*

(i) *If* $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \leq 1$, *then* $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS}) \geq CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.
(ii) *If* $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) > 1$, *then* $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS}) \substack{\geq \\ = \\ <} CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.

Proof. See Appendix A. □

Note that the allocative efficiency measure is always higher than the cost efficiency measure, and Proposition 5.5 shows that we have the same result for allocative super-efficiency and cost super-efficiency measures when $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \leq 1$. Also, if DMU_p is a super-efficient unit, i.e., $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) > 1$, then we can not compare the allocative super-efficiency and the cost super-efficiency measures.

Finally, Proposition 5.6 compares the scale super-efficiency measure with the cost super-efficiency measure:

Proposition 5.6. *For every* DMU_p *we have:*

(i) *If* $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \leq 1$, *then* $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,.,.}) \geq CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.
(ii) *If* $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) > 1$, *then* $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,.,.}) \substack{\geq \\ = \\ <} CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.

Proof. See Appendix A. □

Recall that scale efficiency is always greater than or equal to cost efficiency. Proposition 5.6 proves that the same result holds true for scale super-efficiency and cost super-efficiency measures

[4] The infeasibility of the super-efficiency models and of the Luenberger productivity indicator is structurally similar: both consist in one (or more) observation(s) being situated outside a technology without there being a guarantee that the distance from the outside point to this technology can be achieved. One conjecture is that a solution to the infeasibility of both models could be structurally similar.

when $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \leq 1$. Also, if an observation is technically super-efficient (i.e., $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) > 1$), then there is no relation whatsoever between the scale super-efficiency and the cost super-efficiency measures.

It is important to note that the results obtained in this contribution are not available in both the Van Puyenbroeck (1998) and Aldamak et al. (2016) articles. Furthermore, our results normally carry over when one would only consider partial convexity rather than nonconvexity. We now turn to an empirical illustration of most of the obtained results.

6. Empirical illustration

6.1. Description of the sample

Our empirical illustration of the super-efficiency notions draws upon data that are publicly available in the data repository of the *Journal of Applied Econometrics*[5] for the sake of replicability. In particular, we opt for an unbalanced panel of three years of French fruit producers collected by Ivaldi, Ladoux, Ossard, & Simioni (1996) based on annual accounting data collected in a survey. Two criteria determine the selection of farms: (i) the production of apples is positive, and (ii) the acreage of the orchard is at least five acres. This short panel spans the three successive years from 1984 to 1986. As a description of the technology, one can say that three aggregate inputs deliver two aggregate outputs. The three inputs are: (i) capital (including land), (ii) labor, and (iii) materials. The two outputs are (i) the production of apples, and (ii) an aggregate of alternative products. Descriptive statistics for the 405 observations in total and details on the definitions of all variables are available in Appendix 2 in Ivaldi et al. (1996). A striking feature of this sample is the large heterogeneity in terms of size among the different inputs as well as the outputs. Due to the short length of the panel (just three years), we believe that the use of an intertemporal frontier approach ignoring technical change is justified.

Table B.1 presents basic descriptive statistics for the inputs, the outputs, and the input prices: see the Appendix B. One observes basically a lot of heterogeneity and a rather wide range for all inputs and outputs. The range for some of the input prices is smaller.

6.2. Empirical results

Table 1 reports the descriptive statistics of cost efficiency ($CE(x_p, y_p, w \mid T^{\Lambda,\Gamma})$, cost super-efficiency ($CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}$), input-oriented technical efficiency measure ($DF_i(x_p, y_p \mid T^{\Lambda,\Gamma})$) and input-oriented technical super-efficiency measure ($DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$) for DMUs using convex and nonconvex technologies under VRS and CRS, respectively. We report the average, the standard deviation, and the minima and maxima (depending on the context) for all these measures. This explains the four horizontal parts of this Table 1: the first two parts report VRS results, while the last two parts list CRS results, and each of these two major parts reports results under convexity and nonconvexity, respectively.

Note that in the VRS case for 2 and 8 observations under convexity and nonconvexity respectively, the corresponding models $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$ are infeasible. This illustrates that the inverses of Proposition 5.3(iii) and (iv) are not satisfied. As a result, the prevalence of infeasibility is higher in the nonconvex case. Therefore, we do not include these infeasible observations in the corresponding descriptive statistics computations under VRS. By contrast, no observations are infeasible in the CRS case: this is a clear illustration of Proposition 5.3(i).

Furthermore, to compute the descriptive statistics for the super-efficiency results (i.e., $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma})$ and $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$), we only consider the observations that are cost efficient (i.e., $CE(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = 1$) and input-oriented technically efficient (i.e., $DF_i(x_p, y_p \mid T_p^{\Lambda,\Gamma}) = 1$), respectively. Since the inefficient observations are in common between standard efficiency and super-efficiency concepts, we focus on the distribution of the super-efficiency results that are unity or higher. Therefore, the number of efficient units (#Efficient units), the number of infeasible units (#Infeasible units), as well as the number of observations that we include in the corresponding descriptive statistics computations (#Considered units) is shown in the three last lines of each of the four horizontal part of Table 1. Obviously, for the super-efficiency results the sum of #Infeasible units and #Considered units equals the #Efficient units for the standard efficiency results. Note also that the notion of #Efficient units is not applicable for the super-efficiency results.

Analyzing the results in Table 1, one can draw the following conclusions. First, as expected based on Proposition 5.1(i), the average of $DF_i(x_p, y_p \mid T^{\Lambda,VRS})$ is larger than the average of $CE(x_p, y_p, w \mid T^{\Lambda,VRS})$ under both the convex and nonconvex cases and under both VRS and CRS. Second, Proposition 5.1(iii) cannot be verified at the level of our descriptive statistics: while we indeed have $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$ under both convex and nonconvex cases, under the CRS case this relation does not hold under convexity and nonconvexity. This is due to the fact that the number of observations that we include in the corresponding descriptive statistics is different (see last line with #Considered units). Obviously, the relation in Proposition 5.1(iii) does hold at the level of the common individual observations.[6] Third, the minimum of $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma})$ and $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$ under both the convex and nonconvex cases and under both VRS and CRS is higher than unity. Therefore, both $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma})$ and $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$ separately yield a unique rank for all feasible observations.

Fourth, the obtained VRS results show that 22 units are located on the technically efficient frontier under convexity, among which 7 units are cost efficient, while 185 units are located on the technically efficient frontier under nonconvexity, among which 62 units are cost efficient. Fifth, the results reveal that in the CRS case, 9 units are located on the technically efficient frontier under convexity, whereby 2 units are also cost efficient, while 52 units are located on the technically efficient frontier under nonconvexity, whereby 8 units are cost efficient. Both these conclusions underscore Propositions 5.1(ii) and 5.2: the technical and cost efficiency scores are higher in the nonconvex case.

Fifth, the percentage differences between average nonconvex and convex costs relative to nonconvex costs can be computed by taking the ratio of standard cost efficiency ratios (see the first column).[7] Under VRS and CRS this amounts to cost differences of 31.00% and 30.95%, respectively.[8] This is in the mid range of the cost difference amounts reported in Kerstens & Van de Woestyne (2021).

We now turn to some basic empirical results of the lower and upper bounds on cost efficiency (i.e., $\underline{CE}(x_p, y_p, w \mid T^{\Lambda,\Gamma})$ and

[5] Web site: http://qed.econ.queensu.ca/jae/

[6] These results at the level of the individual observations are available upon request.

[7] In particular, taking a ratio of cost efficiency ratios based on cost frontier estimates under convexity and nonconvexity nets out the observed cost and reveals the difference in cost estimates under convexity and nonconvexity.

[8] Note that $31.00\% = (1 - 0.434/0.629)$ and $30.95\% = (1 - 0.261/0.378)$. Note also that under CRS the cost efficiency ratios under convexity and nonconvexity of 0.261 and 0.378 have earlier been reported in the last column of Table 5 in Kerstens et al. (2019).

Table 1
Descriptive statistics for technical and cost efficiency and super-efficiency under VRS and CRS and in both convex and nonconvex cases.

VRS, Convex	$CE(x_p, y_p, w \mid T^{C,VRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{C,VRS})$	$DF_i(x_p, y_p \mid T^{C,VRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{C,VRS})$
Average	0.434	1.153	0.604	1.251
Stand. Dev.	0.190	0.144	0.193	0.169
Min	0.104	1.011	0.187	1.033
Max	1.000	1.374	1.000	1.661
#Efficient units	7	-	22	-
#Infeasible units	0	2	0	2
#Considered units	405	5	405	20
VRS, Nonconvex	$CE(x_p, y_p, w \mid T^{NC,VRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{NC,VRS})$	$DF_i(x_p, y_p \mid T^{NC,VRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{NC,VRS})$
Average	0.629	1.386	0.847	1.456
Stand. Dev.	0.248	0.401	0.190	0.488
Min	0.134	1.020	0.359	1.002
Max	1.000	3.037	1.000	3.513
#Efficient units	62	-	185	-
#Infeasible units	0	8	0	8
#Considered units	405	54	405	177
CRS, Convex	$CE(x_p, y_p, w \mid T^{C,CRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{C,CRS})$	$DF_i(x_p, y_p \mid T^{C,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{C,CRS})$
Average	0.261	1.380	0.375	1.301
Stand. Dev.	0.161	0.108	0.219	0.246
Min	0.036	1.303	0.048	1.080
Max	1.000	1.456	1.000	1.794
#Efficient units	2	-	9	-
#Infeasible units	0	0	0	0
#Considered units	405	2	405	9
CRS, Nonconvex	$CE(x_p, y_p, w \mid T^{NC,CRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{NC,CRS})$	$DF_i(x_p, y_p \mid T^{NC,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{NC,CRS})$
Average	0.378	1.333	0.582	1.284
Stand. Dev.	0.219	0.226	0.280	0.252
Min	0.039	1.051	0.049	1.008
Max	1.000	1.724	1.000	2.042
#Efficient units	8	-	52	-
#Infeasible units	0	0	0	0
#Considered units	405	8	405	52

$\overline{CE}(x_p, y_p \mid T^{\Lambda,\Gamma}))$ and cost super-efficiency (i.e., $\underline{CE}^{SE}(x_p, y_p)$ and $\overline{CE}^{SE}(x_p, y_p)$). Due to space limitations, these empirical results are presented in Appendix D.

We apply the decomposition of cost efficiency (11) and cost super-efficiency (24) to this same empirical example. The convex and nonconvex results for each of these two decompositions are exhibited in Table 2. This explains the four horizontal parts of this table. As mentioned in subsection 3.3, if observation (x_p, y_p) is inefficient under VRS (i.e., $DF_i(x_p, y_p \mid T^{\Lambda,VRS}) < 1$), then its decomposition of overall or cost efficiency (11) and its decomposition of overall or cost super-efficiency (24) yield the same results. Hence, to report the descriptive statistics for the components of the decomposition of cost super-efficiency (24), we only consider the observations that are input-oriented technically efficient. The number of observations that we include in the corresponding descriptive statistics computations (#Considered units) is shown in the last line of every part of Table 2. Note that this #Considered units for the decomposition of cost super-efficiency corresponds to the #Considered units for the VRS technical super-efficiency results in Table 1.

The following pertinent conclusions emerge. First, the average of $CE(x_p, y_p, w \mid T^{NC,CRS})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{C,CRS})$ are smaller than the average of their decomposition components under both convex and nonconvex technologies. Second, Proposition 5.4 cannot be verified at the level of our descriptive statistics: while the allocative efficiency measure is greater than or equal to the allocative super-efficiency measure under both convex and nonconvex cases, it is not the case that the scale efficiency measure is greater than or equal to the scale super-efficiency measure under both convex and nonconvex cases. Again, this is because the number of observations included in the corresponding descriptive statistics is different (see last line with #Considered units). But, of course this relation in Proposition 5.4 holds true at the level of the common individual observations.[9] Third, the results show that the average of allocative super-efficiency is smaller than the average of scale super-efficiency, under both convex and nonconvex cases. Fourth, the average of scale efficiency under nonconvexity is higher than under convexity, while the average of scale super-efficiency under convexity is higher than under nonconvexity. Fifth, the average of allocative efficiency under convexity is higher than under nonconvexity while the average on allocative super-efficiency under convexity is smaller than under nonconvexity.

To formally assess the above reported differences in distributions, we make use of a nonparametric test proposed initially by Li (1996) and further refined by Fan & Ullah (1999) and others. The most recent development is probably by Li, Maasoumi, & Racine (2009). This nonparametric test focuses on the differences between entire distributions instead of looking at, for instance, first moments (e.g., the Wilcoxon signed-ranks test). It tests for the eventual statistical significance of differences between two kernel-based estimates of density functions f and g of a random variable x. The null hypothesis states the equality of both density functions almost everywhere: $H_0 : f(x) = g(x)$ for all x. By contrast, the alternative hypothesis negates this equality of both density functions: $H_1 : f(x) \neq g(x)$ for some x. This test is valid for both dependent and independent variables.[10]

[9] These results at the level of the individual observations are available upon request.

[10] Note that dependency is intrinsic to nonparametric frontier estimators: i.e., technical and cost efficiency and super-efficiency levels depend on sample size, among others. Matlab code developed by P.J. Kerstens based on Li et al. (2009) is found at: `https://github.com/kepiej/DEAUtils`.

Table 2
Descriptive statistics for overall efficiency and overall super-efficiency decomposition.

Decomposition of cost efficiency				
Convex	$CE(x_p, y_p, w \mid T^{C,CRS})$	$DF_i(x_p, y_p \mid T^{C,VRS})$	$SCE(x_p, y_p \mid T^{C,\cdot})$	$AE(x_p, y_p, w \mid T^{C,CRS})$
Average	0.261	0.604	0.612	0.722
Stand. Dev.	0.161	0.193	0.259	0.177
Min	0.036	0.187	0.073	0.242
Max	1.000	1.000	1.000	1.000
#Considered units	405	405	405	405
Nonconvex	$CE(x_p, y_p, w \mid T^{NC,CRS})$	$DF_i(x_p, y_p \mid T^{NC,VRS})$	$SCE(x_p, y_p \mid T^{NC,\cdot})$	$AE(x_p, y_p, w \mid T^{NC,CRS})$
Average	0.378	0.847	0.664	0.659
Stand. Dev.	0.219	0.190	0.238	0.180
Min	0.039	0.359	0.079	0.253
Max	1.000	1.000	1.000	1.000
#Considered units	405	405	405	405
Decomposition of cost super-efficiency				
Convex	$CE^{SE}(x_p, y_p, w \mid T_p^{C,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{C,VRS})$	$SCE^{SE}(x_p, y_p \mid T_p^{C,\cdot})$	$AE^{SE}(x_p, y_p, w \mid T_p^{C,CRS})$
Average	0.441	1.251	0.644	0.558
Stand. Dev.	0.314	0.169	0.289	0.228
Min	0.052	1.033	0.066	0.247
Max	1.303	1.661	0.999	0.892
#Considered units	20	20	20	20
Nonconvex	$CE^{SE}(x_p, y_p, w \mid T_p^{NC,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{NC,VRS})$	$SCE^{SE}(x_p, y_p \mid T_p^{NC,\cdot})$	$AE^{SE}(x_p, y_p, w \mid T_p^{NC,CRS})$
Average	0.526	1.456	0.614	0.610
Stand. Dev.	0.253	0.488	0.165	0.188
Min	0.081	1.002	0.111	0.186
Max	1.724	3.513	0.982	0.978
#Considered units	177	177	177	177

Table 3
Li-test among all technical and cost efficiency as well as technical and cost super-efficiency notions under VRS and CRS.

VRS	$DF_i(x_p, y_p \mid T^{\Lambda,VRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$	$CE(x_p, y_p, w \mid T^{\Lambda,VRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$
$DF_i(x_p, y_p \mid T^{\Lambda,VRS})$	**167.945*****	*91.6179****	*108.0296****	*25.2885****
$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$	123.9854***	**52.6396*****	*19.8137****	*29.9605****
$CE(x_p, y_p, w \mid T^{\Lambda,VRS})$	275.9821***	-5.3169***	**49.9251*****	*22.6708****
$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$	167.7283***	30.5489***	32.55***	**21.3846*****
CRS	$DF_i(x_p, y_p \mid T^{\Lambda,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$	$CE(x_p, y_p, w \mid T^{\Lambda,CRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$
$DF_i(x_p, y_p \mid T^{\Lambda,CRS})$	**154.7688*****	*58.1697****	*56.8176****	*12.8614****
$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$	282.095***	**21.8972*****	*-2.6741****	*22.8463****
$CE(x_p, y_p, w \mid T^{\Lambda,CRS})$	138.9707***	-5.3709***	**20.434*****	*22.2282****
$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$	154.7681***	10.5353***	10.5306***	**12.7865*****

Li-test: critical values at 1% level= 2.33(***); 5% level= 1.64(**); 10%level= 1.28(*).

Table 3 reports the Li-test results among all technical and cost efficiency as well as technical and cost super-efficiency notions under VRS and CRS, respectively. Both VRS and CRS parts of this Table are structured as follows. First, components on the diagonal (in bold) depict the Li-test statistic between the convex and nonconvex cases. Second, the components under the diagonal show the Li-test statistic between convex efficiency measures, and the components above the diagonal show the Li-test statistic between nonconvex efficiency notions.

The following three conclusions emerge from studying Table 3. First, for the convex efficiency and super-efficiency notions (below the diagonal) all efficiency concepts, under both VRS and CRS cases, follow two by two significantly different distributions. Second, for the nonconvex efficiency and super-efficiency notions (above the diagonal) all efficiency concepts, under both VRS and CRS cases, follow two by two significantly different distributions. Third, all efficiency and super-efficiency notions, under both VRS and CRS cases, follow different distributions under convexity compared to nonconvexity (on the diagonal).

Table 4 reports the Spearman rank correlation coefficients for technical and cost efficiency as well as for technical and cost super-efficiency notions under VRS and CRS, respectively. In both VRS and CRS parts, the components on the diagonal show the rank correlations between convex and nonconvex cases. The components under the diagonal show the rank correlations between convex efficiency notions, and the components above the diagonal show the rank correlations between nonconvex efficiency notions. Note that the numbers in the parentheses indicate the number of observations in the computations of the Spearman rank correlation coefficients between the corresponding components.

To obtain the rank correlations between convex and nonconvex cases on the diagonals, we consider all feasible observations. To be precise, by solving the super-efficiency model (19) under VRS, we obtain 403 feasible observations under convexity, but only 397 of these are feasible under nonconvexity. Hence, we consider only these 397 units in common to obtain the rank correlations between convex and nonconvex cases. Note that in the CRS case all 405 units are feasible and are included in the computations.

To obtain the rank correlations between convex (under the diagonal) and nonconvex (above the diagonal) efficiency notions under both VRS and CRS respectively, we only consider the input-oriented technically efficient observations. Note that the numbers

Table 4
Spearman rank correlation among all technical and cost efficiency as well as technical and cost super-efficiency notions under VRS and CRS.

VRS	$DF_i(x_p, y_p \mid T^{\Lambda,VRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$	$CE(x_p, y_p, w \mid T^{\Lambda,VRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$
$DF_i(x_p, y_p \mid T^{\Lambda,VRS})$	**0.730****(397)	-	-	-
$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$	-	**0.760****(397)	*0.457***(177)	*0.496***(177)
$CE(x_p, y_p, w \mid T^{\Lambda,VRS})$	-	-0.192(20)	**0.812****(397)	*0.986***(177)
$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,VRS})$	-	-0.152(20)	0.992**(20)	**0.815****(397)
CRS	$DF_i(x_p, y_p \mid T^{\Lambda,CRS})$	$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$	$CE(x_p, y_p, w \mid T^{\Lambda,CRS})$	$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$
$DF_i(x_p, y_p \mid T^{\Lambda,CRS})$	**0.933****(405)	-	-	-
$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$	-	**0.935****(405)	*0.293**(52)	*0.302**(52)
$CE(x_p, y_p, w \mid T^{\Lambda,CRS})$	-	0.577(9)	**0.957****(405)	*0.998***(52)
$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$	-	0.583(9)	0.996**(9)	**0.957****(405)

*Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

20 and 177 for the VRS case and the numbers 9 and 52 for the CRS case coincide with the #Considered reported in Table 1. Furthermore, since for all input-oriented technically efficient observations we obviously have $DF_i(x_p, y_p \mid T^{\Lambda,\Gamma}) = 1$, we cannot compute Spearman rank correlations between $DF_i(x_p, y_p \mid T^{\Lambda,\Gamma})$ and the other components under both VRS and CRS.

The following conclusions emerge from studying Table 4. First, for the convex results (under the diagonal) in both VRS and CRS, one can observe that $CE(x_p, y_p, w \mid T^{C,\Gamma})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{C,\Gamma})$ have the highest rank correlation among other efficiency notions. Actually, we compare 20 and 9 units under VRS and CRS, respectively such that only 5 and 2 of these units are cost efficient (see #Considered reported in Table 1). Hence, we obtain high rank correlations only for these observations. Similar results are obtained for the rank correlations between $CE(x_p, y_p, w \mid T^{C,\Gamma})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{C,\Gamma})$ for the nonconvex results (above the diagonal) in both VRS and CRS. Second, comparing convex and nonconvex results on the diagonal under the VRS case, the rank correlations among technical standard and super-efficiency notions is lower than the rank correlations among cost-based standard and super-efficiency notions. Furthermore, the rank correlations among super-efficiency notions are higher than the rank correlations among standard efficiency notions. Third, comparing convex and nonconvex results on the diagonal under the CRS case, the rank correlations are remarkably high overall among efficiency notions, and these are highest for $CE(x_p, y_p, w \mid T^{\Lambda,CRS})$ and $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$ compared to $DF_i(x_p, y_p \mid T^{\Lambda,CRS})$ and $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$. Fourth, the rank correlations between technical super-efficiency and cost super-efficiency are not significant under the convex case (under the diagonal) for both VRS and CRS. By contrast, the rank correlations between these notions under the nonconvex case (above the diagonal) under both VRS and CRS is significant at the 0.01 and 0.05 levels, respectively.[11]

We can end with the following overall conclusions. First, the new notion of cost super-efficiency is clearly distinct in terms of ranking from the existing technical super-efficiency concept. This is especially true in the convex case, while in the nonconvex setting the correlations are at least somewhat similar under VRS and CRS in terms of statistical significance. Second, the technical super-efficiency concept is more clearly distinct in terms of ranking than the new cost-based super-efficiency concept when comparing convex and nonconvex cases. Furthermore, the distinction in terms of ranking for these super-efficiency concepts between convex and nonconvex cases is largest in the VRS case and smallest in the CRS case.

[11] Following the suggestion of a referee, we also report the Kruskal-Wallis rank test to determine differences in rank among groups of efficiency components. For reasons of space, these statistics are reported in Appendix C.

7. Conclusions

While the technical super-efficiency concept has been around for about two and a half decades, the cost super-efficiency concept seems entirely new. Apart from the -to our knowledge- two articles treating technical super-efficiency under nonconvexity, we believe we are the first to offer a truly comparative perspective on technical and cost super-efficiency conditioned on this important convexity axiom. In particular, this contribution explores technical and cost super-efficiency from a methodological and empirical perspective conditioning on traditional convex and less common nonconvex technologies and on CRS versus VRS.

After defining all super-efficiency concepts and numerically illustrating some basic issues, we derive a series of theoretical results regarding traditional efficiency concepts as well as super-efficiency concepts. We equally develop some results regarding traditional efficiency and super-efficiency concepts conditional on the axiom of convexity. Finally, we chart the potential infeasibilities governing these super-efficiency concepts. Most of our theoretical results are new to the frontier literature and even broaden our knowledge about the basic properties of the cost function within the context of incentive-based regulation theory employing frontier methodologies. An empirical section based on secondary data serves to illustrate some of these theoretical results. One main lesson is that the cost super-efficiency notion yields different rankings from the technical super-efficiency concept. Another lesson is that both cost and technical super-efficiency notions rank differently under convex and nonconvex technologies. The latter lesson underscores the differences between convex and nonconvex costs already listed in Kerstens & Van de Woestyne (2021) and the substantial effect of convexity on technical efficiency and cost function based capacity concepts documented in Kerstens et al. (2019).

While the consideration of cost super-efficiency does not change anything to the infeasibility problem for the corresponding technical super-efficiency problem, it is clear that nonconvexity aggravates the infeasibility problem for both the cost super-efficiency and the technical super-efficiency problems. Thus, instead of resolving the infeasibility problem that haunts the super-efficiency literature since its inception, we have deepened this problem. Perhaps, sometimes problems need to be aggravated before these can be successfully resolved.

Among the avenues for future methodological research one can list the following topics. One extension for future research is to analyse the super-efficiency notion in the context of either the revenue function or the profit function. Another avenue is to search for solutions for the infeasibility problem that can be equally applied to the efficiency measures as well as the cost function.

Finally, it may be worthwhile to apply these new models in the context of different technical and economic efficiency measures: examples include the directional distance function (Chambers, Chung, & Färe, 1998), but also the weighted additive model

(Cooper, Pastor, Aparicio, & Borras, 2011) or the slacks-based measure (Aparicio et al., 2017).

Acknowledgments

The research of Mehdi Toloo was supported by the Czech Science Foundation (GAČR 19-13946S).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2021.10.023.

References

Adler, N., & Volta, N. (2019). Ranking methods within data envelopment analysis. In T. ten Raa, & W. Greene (Eds.), *The Palgrave handbook of economic performance analysis* (pp. 189–224). Cham: Palgrave Macmillan.
Afriat, S. (1972). Efficiency estimation of production functions. *International Economic Review, 13*(3), 568–598.
Agrell, P., & Bogetoft, P. (2017). Regulatory benchmarking: Models, analyses and applications. *Data Envelopment Analysis Journal, 3*(1–2), 49–91.
Agrell, P., Bogetoft, P., & Tind, J. (2002). Incentive plans for productive efficiency, innovation and learning. *International Journal of Production Economics, 78*(1), 1–11.
Aldamak, A., Hatami-Marbini, A., & Zolfaghari, S. (2016). Dual frontiers without convexity. *Computers & Industrial Engineering, 101*, 466–478.
Andersen, P., & Petersen, N. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science, 39*(10), 1261–1264.
Angulo-Meza, L., & Estellita Lins, M. (2002). Review of methods for increasing discrimination in data envelopment analysis. *Annals of Operations Research, 116*(1–4), 225–242.
Aparicio, J., Ortiz, L., & Pastor, J. (2017). Measuring and decomposing profit inefficiency through the slacks-based measure. *European Journal of Operational Research, 260*(2), 650–654.
Aparicio, J., Pastor, J., & Zofio, J. (2015). How to properly decompose economic efficiency using technical and allocative criteria with non-homothetic DEA technologies. *European Journal of Operational Research, 240*(3), 882–891.
Banker, R., & Chang, H. (2006). The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research, 175*(2), 1311–1320.
Bogetoft, P. (1994). Incentive efficient production frontiers: An agency perspective on DEA. *Management Science, 40*(8), 959–968.
Bogetoft, P. (2000). DEA and activity planning under asymmetric information. *Journal of Productivity Analysis, 13*(1), 7–48.
Briec, W., & Kerstens, K. (2006). Input, output and graph technical efficiency measures on non-convex FDH models with various scaling laws: An integrated approach based upon implicit enumeration algorithms. *TOP, 14*(1), 135–166. https://doi.org/10.1007/BF02579006.
Briec, W., & Kerstens, K. (2009). Infeasibilities and directional distance functions with application to the determinateness of the Luenberger productivity indicator. *Journal of Optimization Theory and Applications, 141*(1), 55–73.
Briec, W., Kerstens, K., & Vanden Eeckaut, P. (2004). Non-convex technologies and cost functions: Definitions, duality and nonparametric tests of convexity. *Journal of Economics, 81*(2), 155–192.
Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions, and Nerlovian efficiency. *Journal of Optimization Theory and Applications, 98*(2), 351–364.
Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2*(6), 429–444.
Cooper, W., Pastor, J. T., Aparicio, J., & Borras, F. (2011). Decomposing profit inefficiency in DEA through the weighted additive model. *European Journal of Operational Research, 212*(2), 411–416.
Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor efficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: Concepts and measurements* (pp. 243–268). Amsterdam: North Holland.
Emrouznejad, A., & Yang, G.-l. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences, 61*, 4–8.
Esteve, M., Aparicio, J., Rabasa, A., & Rodriguez-Sala, J. J. (2020). Efficiency analysis trees: A new methodology for estimating production frontiers through decision trees. *Expert Systems with Applications, 162*, 113783.
Fan, Y., & Ullah, A. (1999). On goodness-of-fit tests for weakly dependent processes using kernel method. *Journal of Nonparametric Statistics, 11*(1), 337–360.
Färe, R., Grosskopf, S., & Lovell, C. (1985). *The measurement of efficiency of production*. Boston: Kluwer.
Färe, R., Grosskopf, S., & Lovell, C. (1994). *Production frontiers*. Cambridge: Cambridge University Press.
Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A: General, 120*(3), 253–281.
Hackman, S. (2008). *Production economics: integrating the microeconomic and engineering perspectives*. Berlin: Springer.
Ivaldi, M., Ladoux, N., Ossard, H., & Simioni, M. (1996). Comparing Fourier and translog specifications of multiproduct technology: Evidence from an incomplete panel of French farmers. *Journal of Applied Econometrics, 11*(6), 649–668.
Jamasb, T., & Pollitt, M. (2000). Benchmarking and regulation: International electricity experience. *Utilities Policy, 9*(3), 107–130.
Kerstens, K., Sadeghi, J., & Van de Woestyne, I. (2019). Convex and nonconvex input-oriented technical and economic capacity measures: An empirical comparison. *European Journal of Operational Research, 276*(2), 699–709.
Kerstens, K., & Vanden Eeckaut, P. (1999). Estimating returns to scale using nonparametric deterministic technologies: A new method based on goodness-of-fit. *European Journal of Operational Research, 113*(1), 206–214.
Kerstens, K., & Van de Woestyne, I. (2014). Solution methods for nonconvex free disposal hull models: A review and some critical comments. *Asia-Pacific Journal of Operational Research, 31*(1).
Kerstens, K., & Van de Woestyne, I. (2021). Cost functions are nonconvex in the outputs when the technology is nonconvex: Convexification is not harmless. *Annals of Operations Research, 305*(1-2), 81–106.
Kumbhakar, S., Parmeter, C., & Tsionas, E. (2013). A zero inefficiency stochastic frontier model. *Journal of Econometrics, 172*(1), 66–76.
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.
Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews, 15*(1), 261–274.
Li, Q., Maasoumi, E., & Racine, J. (2009). A nonparametric test for equality of distributions with mixed categorical and continuous data. *Journal of Econometrics, 148*(2), 186–200.
Liu, J., Lu, W.-M., Lu, L. Y. Y., & Lin, B. (2013). Data envelopment analysis 1978–2010: A citation-based literature survey. *Omega, 41*(1), 3–15.
Seiford, L. (1997). A bibliography for data envelopment analysis (1978–1996). *Annals of Operations Research, 73*, 393–438.
Seitz, W. (1970). The measurement of efficiency relative to a frontier production function. *American Journal of Agricultural Economics, 52*(4), 505–511.
Seitz, W. (1971). Productive efficiency in the steam-electric generating industry. *Journal of Political Economy, 79*(4), 878–886.
Sherman, H., & Ladino, G. (1995). Managing bank productivity using data envelopment analysis (DEA). *Interfaces, 25*(2), 60–73.
Soltanifar, M., & Shahghobadi, S. (2014). Survey on rank preservation and rank reversal in data envelopment analysis. *Knowledge-Based Systems, 60*, 10–19.
Tavakoli, I., & Mostafaee, A. (2019). Free disposal hull efficiency scores of units with network structures. *European Journal of Operational Research, 277*(3), 1027–1036.
Van Puyenbroeck, T. (1998). Some remarks on modified FDH. *Journal of Productivity Analysis, 9*(1), 81–94.
Zhu, J. (1996). Robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research, 90*(3), 451–460.