WORKING PAPER SERIES

2020-EQM-02

Procedures for Ranking
Technical and Cost Efficient Units:
With a Focus on Nonconvexity

Kristiaan Kerstens
IÉSEG School of Management and LEM-CNRS (UMR 9221), Lille, France

Jafar Sadeghi
IÉSEG School of Management and Ivey Business School, Western University, London, Ontario, Canada

Mehdi Toloo
Department of Systems Engeneering, VSB-Technical University of Ostrava, Czech Republic

Ignace Van de Woestyne
KU Leuven, Research Unit MEES, Brussels, Belgium

Procedures for Ranking
Technical and Cost Efficient Units:
With a Focus on Nonconvexity

Kristiaan Kerstens[*], Jafar Sadeghi[†], Mehdi Toloo[‡], Ignace Van de Woestyne[§]

May 4, 2020

Abstract

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 on 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.

JEL CODES: D24

KEYWORDS: Data Envelopment Analysis; Free Disposal Hull; Technical Efficiency; Cost Efficiency; Super-efficiency.

[*]Corresponding author: IESEG School of Management, CNRS-LEM (UMR 9221), Université de Lille, 3 rue de la Digue, F-59000 Lille, France, Tel: +33 320545892 (switchboard), Fax: +33 320574855, k.kerstens@ieseg.fr

[†]IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France & Ivey Business School, Western University, London, Ontario, Canada j.sadeghi1987@gmail.com

[‡]Department of Systems Engineering, VSB-Technical University of Ostrava, Czech Republic mehdi.toloo@vsb.cz

[§]KU Leuven, Research Unit MEES, Warmoesberg 26, B-1000 Brussels, Belgium, ignace.vandewoestyne@kuleuven.be

1 Introduction

The seminal articles of Farrell (1957) and Charnes, Cooper, and 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 and Yang (2018) surveys the first 40 years of scholarly literature in DEA over the period 1978 till 2016 and lists about 10300 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, and 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 and Pollitt (2000)). Finally, an example of a managerial application is the Sherman and 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, and 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 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 and 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 and 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 and Estellita Lins (2002), Soltanifar and Shahghobadi (2014) and Adler and Volta (2019)).

In addition to ranking efficient observations, super-efficiency models have also been used to develop tests for influential observations or outliers (examples include Banker and Chang (2006), Dusansky and Wilson (1994), and Thanassoulis (1999), among 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 and 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, and 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 ax-

---

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

iom 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, Cooper, and Rhodes (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, and Tulkens (1984). Kerstens and 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, and Vanden Eeckaut (2004) define nonconvex cost functions corresponding to the specific returns to scale assumptions in Kerstens and Vanden Eeckaut (1999).

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, Kerstens, and Vanden Eeckaut (2004) proof 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. This fact that convexity affects the incidence and amounts of all sorts of inefficiency makes the issue relevant in super-efficiency model context. Apart from the early contribution of Van Puyenbroeck (1998) and the more recent one by Aldamak, Hatami-Marbini, and 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. 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. Both the technical and cost super-efficiency definitions are illustrated with a figure in Section 4.

---

[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, and Tsionas (2013) is the seminal article allowing to represent a mixture of both fully efficient and inefficient firms. Rho and Schmidt (2015) and Tran and Tsionas (2016) extend this basic idea in several directions. However, we are unaware of the development of a super-efficiency notion in this part of the frontier literature.

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 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, and 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) $\Gamma \equiv \mathrm{CRS} = \{\delta \mid \delta \geq 0\}$;

    (ii) $\Gamma \equiv \mathrm{VRS} = \{\delta \mid \delta = 1\}$.

(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 for two returns to scale assumptions: i.e., constant returns to scale (CRS) and 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:

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

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:

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

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 ($w \in \mathbb{R}_+^N$):

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

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

---

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

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, Kerstens, and Vanden Eeckaut (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 $(x_k, y_k) \in \mathbb{R}_+^N \times \mathbb{R}_+^M$ ($k \in \{1, \ldots, K\}$) 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:

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

where

(i) $\Gamma \equiv \mathrm{CRS} = \{\delta \mid \delta \geq 0\}$;

(ii) $\Gamma \equiv \mathrm{VRS} = \{\delta \mid \delta = 1\}$;

and

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

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

The activity vector z operates subject to a convexity ($\Lambda \equiv \mathrm{C}$) or a nonconvexity ($\Lambda \equiv \mathrm{NC}$) 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.

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 and 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 and 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, Grosskopf, and Lovell (1985)):

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

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, Kerstens, and Vanden Eeckaut (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, Grosskopf, and Lovell (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:

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

Färe, Grosskopf, and Lovell (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, and Zofio (2015) and Aparicio, Ortiz, and Pastor (2017).

3 Super-Efficiency: New Definitions and Decompositions

3.1 Technical Super-Efficiency

Andersen and 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 p-th observation, the resulting technology can mathematically be formulated as follows:

$$T_p^{\Lambda,\Gamma} = \left\{ (x,y) \mid x \geq \sum_{k=1}^{K} x_k \delta z_k, y \leq \sum_{k=1}^{K} y_k \delta z_k, z_p = 0, z \in \Lambda, \delta \in \Gamma \right\}. \tag{12}$$

Observe that the removal of the p-th observation is realized by forcing the p-th component of the activity vector to be zero. Consequently, this p-th 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:

$$DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma}) = \min_{\lambda}\{\lambda \mid \lambda \geq 0, (\lambda x_p, y_p) \in T_p^{\Lambda,\Gamma}\}, \tag{13}$$

with $T_p^{\Lambda,\Gamma}$ defined in (12).

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 and 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 of which we now provide a selective overview. Thrall (1996) and Zhu (1996) note that the

input-oriented super-efficiency model under CRS may well be infeasible. Dula and Hickman (1997) point out that infeasibility shows that the observation under evaluation is efficient, though not lending itself to ranking, and that it is useful in locating endpoint positions of the extreme efficient observations. Seiford and Zhu (1999) prove some necessary and sufficient conditions under which various super-efficiency models are infeasible and also investigate the infeasibility when determining returns to scale based on super-efficiency models.

Recently, researchers have been proposing a variety of methods in an attempt to solving these infeasibility problems of the super-efficiency models. Mehrabian, Alirezaee, and Jahanshahloo (1999) propose the MAJ model that has received some attention in the literature: this model solves the infeasibility issue of the basic Andersen and Petersen (1993) model as well as the sensitivity of this model to small variations in the data when the observations have smaller values of inputs or outputs. Jahanshahloo, Lotfi, Shoja, Tohidi, and Razavyan (2004) and Jahanshahloo, Sanei, Lotfi, and Shoja (2004) present alternative versions of the MAJ model that avoid infeasibilities under some conditions. Chen (2004) also studies super-efficiency under the condition that not all inputs (or outputs) are simultaneously changed in the same proportion: this author suggests using both input- and output-oriented super-efficiency models to fully characterize the super-efficiency.

Chen (2005) provides a necessary and sufficient condition for simultaneous infeasibility of input- and output-oriented super-efficiency models in the VRS case. It is claimed that both of these models are infeasible only in exceptional situations. However, Soleimani-Damaneh, Jahanshahloo, and Foroughi (2006) show that this latter contention is not correct by providing some counterexamples. Ray (2008) proposes an alternative procedure using the directional distance function to measure super-efficiency by solving the infeasibility problem which generally leads to a complete ranking of the observations. Johnson and McGinnis (2009) show that using a non-radial hyperbolic measurement orientation in a super-efficiency model leads to find feasible solutions for certain cases when the requirement for all data to be strictly positive is relaxed.

Tone (2001) develops the slacks-based measure (SBM) of super-efficiency. Du, Liang, and Zhu (2010) extend the SBM of super-efficiency method proposed by Tone (2001) to the additive (slacks-based) DEA model and point out that -unlike the radial super-efficiency models- additive super-efficiency models are always feasible. Fang, Lee, Hwang, and Chung (2013) show that the projection identified by the super SBM model may not be strongly Pareto efficient. Hence, they develop an alternative super-efficiency SBM model for which the identified projection is strongly Pareto efficient.

Cook, Liang, Zha, and Zhu (2009) propose a modified super-efficiency model in which a super-efficiency score for efficient DMUs can be obtained even when feasible solutions do not exist. When the VRS super-efficiency model is infeasible, their method finds a (virtual) referent DMU formed by the remaining DMUs and yields a score that characterizes the super-efficiency in inputs and outputs. Aldamak, Hatami-Marbini, and Zolfaghari (2016) modify FDH super-efficiency models related to the optimistic and pessimistic FDH in both input and output orientations without infeasibility problems.

There have been several further attempts for dealing with infeasibility problem: examples include, e.g., Lee, Chu, and Zhu (2011); Chen and Liang (2011); Chen, Deng, and Gingras (2011); Chen, Du, and Huo (2013), among others.

We finish this selective overview of contributions by referring the reader to the surveys of the super-efficiency literature in Angulo-Meza and Estellita Lins (2002), Soltanifar and Shahghobadi (2014) and Adler and Volta (2019) for further details. Some prudent conclusions from this literature appear. First, infeasibilities are hard to avoid in general. Second, infeasibilities can sometimes only be avoided by resorting to the use of more general measurement orientations that also modify the interpretation of the original model.

Unless infeasibility is investigated explicitly in what follows, we avoid infeasibility issues by assuming that all required components are well-defined.

3.2 Cost Super-Efficiency

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 (12) as follows:

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

$$CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = \frac{C^{SE}(y_p, w \mid T_p^{\Lambda,\Gamma})}{w x_p}, \tag{14}$$

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

$$C^{SE}(y_p, w \mid T_p^{\Lambda,\Gamma}) = \min_x \{wx \mid x \geq 0, (x, y_p) \in T_p^{\Lambda,\Gamma}\}. \tag{15}$$

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

By analogy to Section 2.3, we now decompose the cost super-efficiency (14) 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:

$$AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) = \frac{CE^{SE}(w_p, y_p \mid T_p^{\Lambda,\Gamma})}{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})}. \tag{16}$$

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:

$$SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) = \frac{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})}{DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})}. \tag{17}$$

Clearly, $0 < SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) \leq 1$. 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 (12).

Furthermore, based on (16) and (17), CRS technical super-efficiency can be decomposed

into VRS technical super-efficiency and scale super-efficiency as follows:

$$CE^{SE}(w_p, y_p \mid T_p^{\Lambda,CRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS}) \times AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$$

$$= DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \times SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) \times AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS}). \quad (18)$$

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 (18) lead to the same results.

4   Numerical Illustration

Now we clarify the above Definitions 3.1 and 3.2 by means of Figure 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.
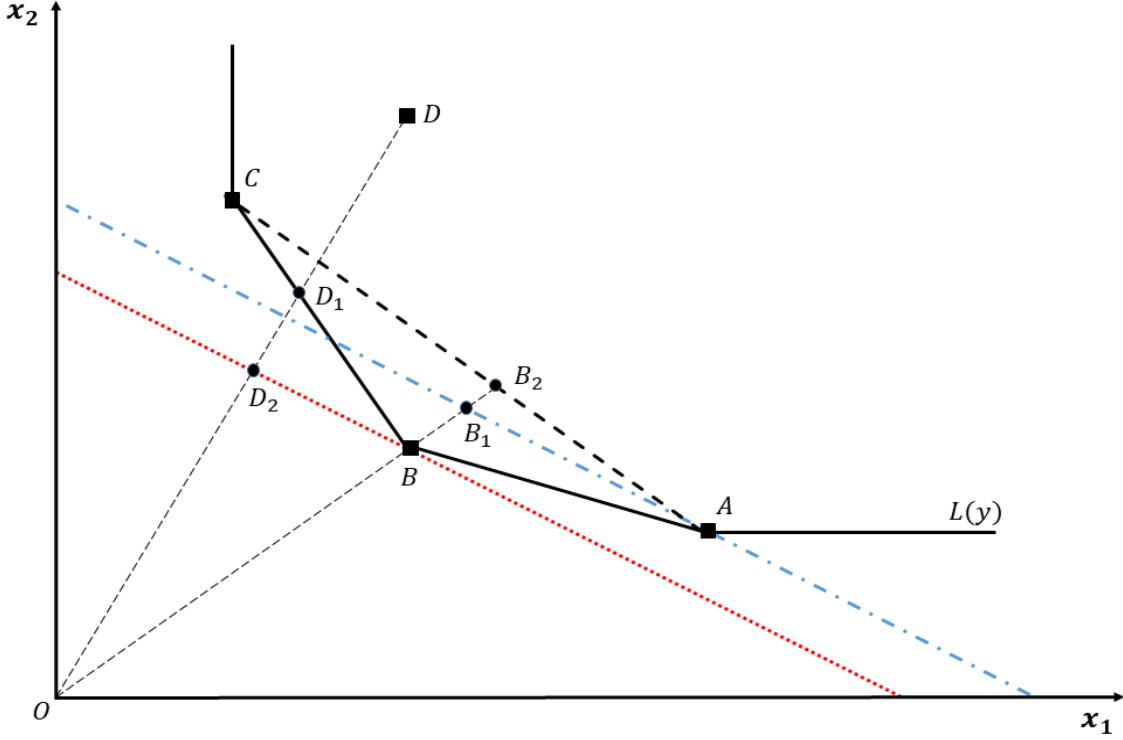
Figure 1: Isoquant with combinations of two inputs yielding a given output level y

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) $CE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \leq DF_i(x_p, y_p \mid T^{\Lambda,\Gamma}) \leq 1$.

(ii) *The number of cost efficient observations is smaller than or equal to the number of input-oriented technical efficient observations.*

(iii) $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,\Gamma}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,\Gamma})$.

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 $(x_p, y_p) \in T^{\Lambda,\Gamma}$: $0 < AE(x_p, y_p, w \mid T^{\Lambda,\Gamma}) \leq 1$.*

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

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) $DF_i(x_p, y_p \mid T^{C,\Gamma}) \leq DF_i(x_p, y_p \mid T^{NC,\Gamma})$;

(ii) $DF_i^{SE}(x_p, y_p \mid T_p^{C,\Gamma}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{NC,\Gamma})$;

(iii) $CE(x_p, y_p, w \mid T^{C,\Gamma}) \leq CE(x_p, y_p, w \mid T^{NC,\Gamma})$;

(iv) $CE^{SE}(x_p, y_p, w \mid T_p^{C,\Gamma}) \leq CE^{SE}(x_p, y_p, w \mid T_p^{NC,\Gamma})$.

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.

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.*

16

(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.*

Proof. See Appendix A. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

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. 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 problem of infeasibility is worse under nonconvexity because the volume of the production possibility set is smaller than or equal to the volume of the convex counterpart.

While Agrell, Bogetoft, and Tind (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. This existence result 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 Figure 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.
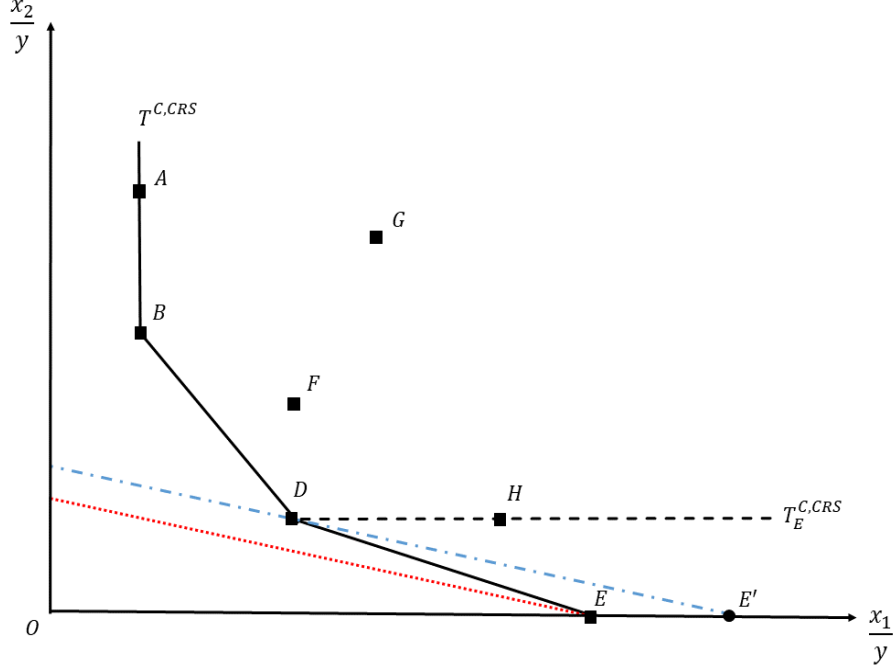
Figure 2: Isoquant with combinations of two inputs and single output under convexity and CRS assumptions

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.

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

18

feasible.

Figure 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.
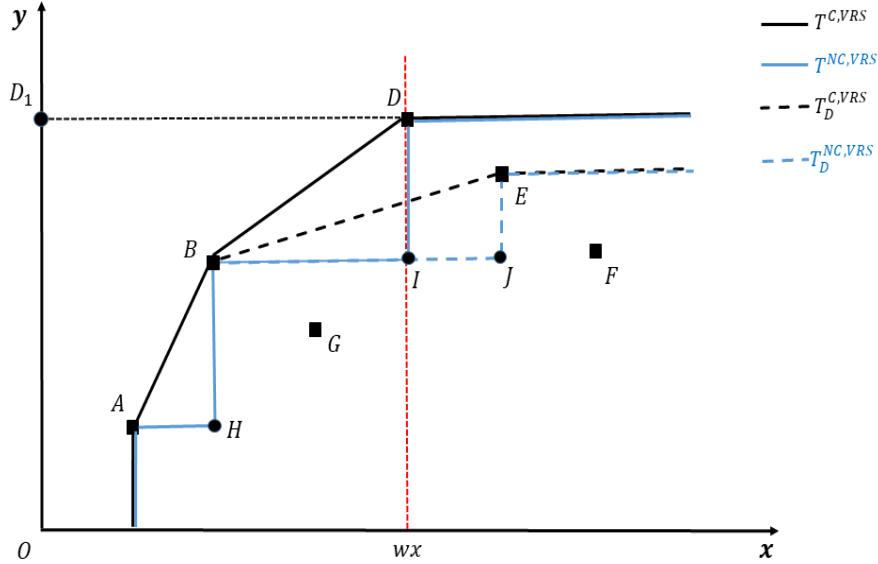

Figure 3: Convex and nonconvex technologies with single input and single output under VRS assumption

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 plotted as a dotted line. The corresponding technical and cost efficiency measures for observation D under both convex and nonconvex cases are $DF_i(x_D, y_D \mid T^{C,VRS}) = DF_i(x_D, y_D \mid T^{NC,VRS}) = \frac{D_1 D}{D_1 D} = 1$ and $CE(x_D, y_D, w \mid T^{C,VRS}) = CE(x_D, y_D, w \mid T^{NC,VRS}) = \frac{D_1 D}{D_1 D} = 1$, 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

19

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.

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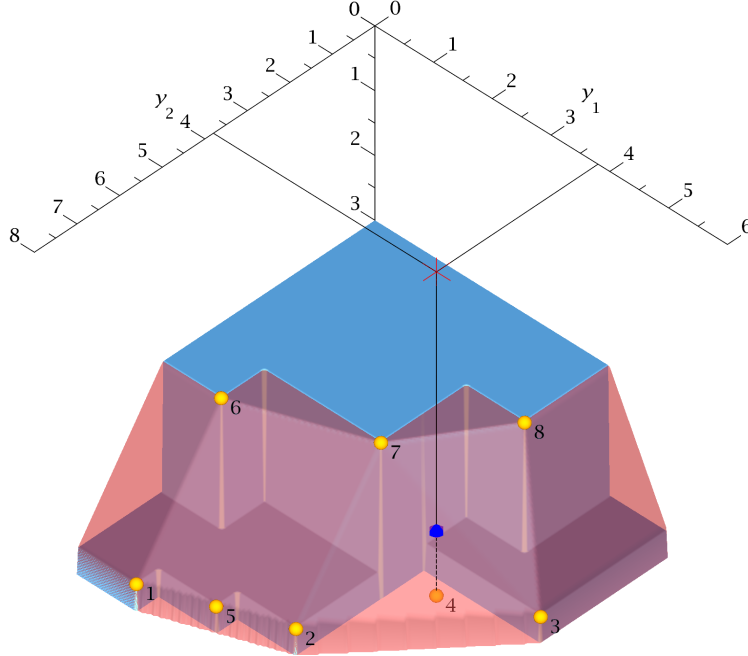
Figure 4: Convex and nonconvex technologies with single input and two outputs under VRS assumption

The visualization in Figure 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 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 and 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 supra), 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. It remains an open question whether this is possible.

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) $SCE(x_p, y_p \mid T^{\Lambda, \cdot\cdot}) \geq SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot\cdot})$

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 $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS}) \gtreqless CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

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 $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot\cdot}) \geq CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

(ii) If $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, VRS}) > 1$, then $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot\cdot}) \gtreqless CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

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 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.

We now turn to an empirical illustration of most of these 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*[4] for the sake of replicability. In particular, we opt for an unbalanced panel of three years of French fruit producers

---

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

collected by Ivaldi, Ladoux, Ossard, and 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, Ladoux, Ossard, and Simioni (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.

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 for 2 observations under convexity and 8 observations under nonconvexity in the VRS case, 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).

23

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

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.[5] 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.

We apply the decomposition of cost efficiency (11) and cost super-efficiency (18) 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 (18) yield the same results. Hence, to report the descriptive statistics for the components of the decomposition of cost super-efficiency (18), 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.

---

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

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

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.[6] 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.

---

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

To formally assess the above reported differences in distributions, we make use of a non-parametric test proposed initially by Li (1996) and further refined by Fan and Ullah (1999) and others. The most recent development is probably by Li, Maasoumi, and 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.[7]

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.

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(*).

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

---

[7]Note that dependency is a characteristic of nonparametric frontier estimators: i.e., technical and cost efficiency and super-efficiency levels depend on sample size, among others. Matlab code for this Li-test adopted here is developed by P.J. Kerstens based on Li, Maasoumi, and Racine (2009). This code is found at: https://github.com/kepiej/DEAUtils.

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 (13) 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 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.

Table 4: Spearman 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).

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.

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.

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 constant versus variable returns to scale.

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.

Among the avenues for future 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.

References

ADLER, N., AND N. VOLTA (2019): "Ranking Methods Within Data Envelopment Analysis," in *The Palgrave Handbook of Economic Performance Analysis*, ed. by T. ten Raa, and W. Greene, pp. 189–224. Palgrave Macmillan, Cham.

AFRIAT, S. (1972): "Efficiency Estimation of Production Functions," *International Economic Review*, 13(3), 568–598.

AGRELL, P., AND P. BOGETOFT (2017): "Regulatory Benchmarking: Models, Analyses and Applications," *Data Envelopment Analysis Journal*, 3(1–2), 49–91.

AGRELL, P., P. BOGETOFT, AND J. TIND (2002): "Incentive Plans for Productive Efficiency, Innovation and Learning," *International Journal of Production Economics*, 78(1), 1–11.

ALDAMAK, A., A. HATAMI-MARBINI, AND S. ZOLFAGHARI (2016): "Dual Frontiers Without Convexity," *Computers & Industrial Engineering*, 101, 466–478.

ANDERSEN, P., AND N. PETERSEN (1993): "A Procedure for Ranking Efficient Units in Data Envelopment Analysis," *Management Science*, 39(10), 1261–1264.

ANGULO-MEZA, L., AND M. ESTELLITA LINS (2002): "Review of Methods for Increasing Discrimination in Data Envelopment Analysis," *Annals of Operations Research*, 116(1–4), 225–242.

APARICIO, J., L. ORTIZ, AND J. PASTOR (2017): "Measuring and Decomposing Profit Inefficiency Through the Slacks-Based Measure," *European Journal of Operational Research*, 260(2), 650–654.

APARICIO, J., J. PASTOR, AND J. ZOFIO (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., AND H. CHANG (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.

——— (2000): "DEA and Activity Planning under Asymmetric Information," *Journal of Productivity Analysis*, 13(1), 7–48.

BRIEC, W., AND K. KERSTENS (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.

BRIEC, W., AND K. KERSTENS (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., K. KERSTENS, AND P. VANDEN EECKAUT (2004): "Non-convex Technologies and Cost Functions: Definitions, Duality and Nonparametric Tests of Convexity," *Journal of Economics*, 81(2), 155–192.

CHARNES, A., W. COOPER, AND E. RHODES (1978): "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, 2(6), 429–444.

CHEN, J.-X., M. DENG, AND S. GINGRAS (2011): "A Modified Super-Efficiency Measure Based on Simultaneous Input–Output Projection in Data Envelopment Analysis," *Computers & Operations Research*, 38(2), 496–504.

CHEN, Y. (2004): "Ranking Efficient Units in DEA," *Omega*, 32(3), 213–219.

——— (2005): "Measuring Super-Efficiency in DEA in the Presence of Infeasibility," *European Journal of Operational Research*, 161(2), 545–551.

CHEN, Y., J. DU, AND J. HUO (2013): "Super-Efficiency Based on a Modified Directional Distance Function," *Omega*, 41(3), 621–625.

CHEN, Y., AND L. LIANG (2011): "Super-Efficiency DEA in the Presence of Infeasibility: One Model Approach," *European Journal of Operational Research*, 213(1), 359–360.

COOK, W., L. LIANG, Y. ZHA, AND J. ZHU (2009): "A Modified Super-Efficiency DEA Model for Infeasibility," *Journal of the Operational Research Society*, 60(2), 276–281.

DEPRINS, D., L. SIMAR, AND H. TULKENS (1984): "Measuring Labor Efficiency in Post Offices," in *The Performance of Public Enterprises: Concepts and Measurements*, ed. by M. Marchand, P. Pestieau, and H. Tulkens, pp. 243–268. North Holland, Amsterdam.

DU, J., L. LIANG, AND J. ZHU (2010): "A Slacks-Based Measure of Super-Efficiency in Data Envelopment Analysis: A Comment," *European Journal of Operational Research*, 204(3), 694–697.

DULA, J., AND B. HICKMAN (1997): "Effects of Excluding the Column Being Scored from the DEA Envelopment LP Technology Matrix," *Journal of the Operational Research Society*, 48(10), 1001–1012.

DUSANSKY, R., AND P. WILSON (1994): "Technical Efficiency in the Decentralized Care of the Developmentally Disabled," *Review of Economics and Statistics*, 76(2), 340–345.

EMROUZNEJAD, A., AND G.-L. YANG (2018): "A Survey and Analysis of the First 40 Years of Scholarly Literature in DEA: 1978–2016," *Socio-Economic Planning Sciences*, 61, 4–8.

FAN, Y., AND A. ULLAH (1999): "On Goodness-of-fit Tests for Weakly Dependent Processes Using Kernel Method," *Journal of Nonparametric Statistics*, 11(1), 337–360.

FANG, H.-H., H.-S. LEE, S.-N. HWANG, AND C.-C. CHUNG (2013): "A Slacks-Based Measure of Super-Efficiency in Data Envelopment Analysis: An Alternative Approach," *Omega*, 41(4), 731–734.

FÄRE, R., S. GROSSKOPF, AND C. LOVELL (1985): *The Measurement of Efficiency of Production*. Kluwer, Boston.

———— (1994): *Production Frontiers*. Cambridge University Press, Cambridge.

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*. Springer, Berlin.

IVALDI, M., N. LADOUX, H. OSSARD, AND M. SIMIONI (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.

JAHANSHAHLOO, G., F. H. LOTFI, N. SHOJA, G. TOHIDI, AND S. RAZAVYAN (2004): "Ranking Using L1-norm in Data Envelopment Analysis," *Applied Mathematics and Computation*, 153(1), 215–224.

JAHANSHAHLOO, G., M. SANEI, F. LOTFI, AND N. SHOJA (2004): "Using the Gradient Line for Ranking DMUs in DEA," *Applied Mathematics and Computation*, 151(1), 209–219.

JAMASB, T., AND M. POLLITT (2000): "Benchmarking and Regulation: International Electricity Experience," *Utilities Policy*, 9(3), 107–130.

JOHNSON, A., AND L. MCGINNIS (2009): "The Hyperbolic-Oriented Efficiency Measure as a Remedy to Infeasibility of Super Efficiency Models," *Journal of the Operational Research Society*, 60(11), 1511–1517.

KERSTENS, K., AND I. VAN DE WOESTYNE (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., AND P. VANDEN EECKAUT (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.

KUMBHAKAR, S., C. PARMETER, AND E. TSIONAS (2013): "A Zero Inefficiency Stochastic Frontier Model," *Journal of Econometrics*, 172(1), 66–76.

LEE, H.-S., C.-W. CHU, AND J. ZHU (2011): "Super-Efficiency DEA in the Presence of Infeasibility," *European Journal of Operational Research*, 212(1), 141–147.

LI, Q. (1996): "Nonparametric Testing of Closeness between Two Unknown Distribution Functions," *Econometric Reviews*, 15(1), 261–274.

LI, Q., E. MAASOUMI, AND J. RACINE (2009): "A Nonparametric Test for Equality of Distributions with Mixed Categorical and Continuous Data," *Journal of Econometrics*, 148(2), 186–200.

LIU, J., W.-M. LU, L.Y.Y.AND LU, AND B. LIN (2013): "Data Envelopment Analysis 1978-2010: A Citation-Based Literature Survey," *Omega*, 41(1), 3–15.

MEHRABIAN, S., M. ALIREZAEE, AND G. JAHANSHAHLOO (1999): "A Complete Efficiency Ranking of Decision Making Units in Data Envelopment Analysis," *Computational Optimization and Applications*, 14(2), 261–266.

RAY, S. (2008): "The Directional Distance Function and Measurement of Super-Efficiency: An Application to Airlines Data," *Journal of the Operational Research Society*, 59(6), 788–797.

RHO, S., AND P. SCHMIDT (2015): "Are All Firms Inefficient?," *Journal of Productivity Analysis*, 43(3), 327–349.

SEIFORD, L. (1997): "A Bibliography for Data Envelopment Analysis (1978-1996)," *Annals of Operations Research*, 73, 393–438.

SEIFORD, L., AND J. ZHU (1999): "Infeasibility of Super-Efficiency Data Envelopment Analysis Models," *INFOR: Information Systems and Operational Research*, 37(2), 174–187.

SEITZ, W. (1970): "The Measurement of Efficiency Relative to a Frontier Production Function," *American Journal of Agricultural Economics*, 52(4), 505–511.

——— (1971): "Productive Efficiency in the Steam-Electric Generating Industry," *Journal of Political Economy*, 79(4), 878–886.

SHERMAN, H., AND G. LADINO (1995): "Managing Bank Productivity Using Data Envelopment Analysis (DEA)," *Interfaces*, 25(2), 60–73.

SOLEIMANI-DAMANEH, M., G. JAHANSHAHLOO, AND A. FOROUGHI (2006): "A Comment on "Measuring Super-Efficiency in DEA in the Presence of Infeasibility"," *European Journal of Operational Research*, 170(1), 323–325.

Soltanifar, M., and S. Shahghobadi (2014): "Survey on Rank Preservation and Rank Reversal in Data Envelopment Analysis," *Knowledge-Based Systems*, 60, 10–19.

Thanassoulis, E. (1999): "Setting Achievement Targets for School Children," *Education Economics*, 7(2), 101–119.

Thrall, R. (1996): "Duality, Classification and Slacks in DEA," *Annals of Operations Research*, 66(2), 109–138.

Tone, K. (2001): "A Slacks-Based Measure of Efficiency in Data Envelopment Analysis," *European Journal of Operational Research*, 130(3), 498–509.

Tran, K., and M. Tsionas (2016): "Zero-Inefficiency Stochastic Frontier Models with Varying Mixing Proportion: A Semiparametric Approach," *European Journal of Operational Research*, 249(3), 1113–1123.

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.

Appendices: Supplementary Material

A   Proofs of Propositions

Proof of Proposition 5.1:

(i) Suppose that λ^* is an optimal value of the model determining $DF_i(x_p, y_p \mid T^{\Lambda, \Gamma})$ (model (2)), therefore $(\lambda^* x_p, y_p) \in T^{\Lambda}$. Based on the technology (6), there is $z_k(k = 1, \cdots, K)$ such that:

$$\begin{aligned} &\lambda^* x_p \geq \sum_{k=1}^{K} x_k \delta z_k, \\ &y_p \leq \sum_{k=1}^{K} y_k \delta z_k, \\ &z \in \Lambda, \delta \in \Gamma. \end{aligned} \qquad (A.1)$$

If we let $x^* = \lambda^* x_p$, then (x^*, z_k) is a feasible solution of the model determining $CE(x_p, y_p, w \mid T^{\Lambda, \Gamma})$ (model (7)), thus we have: $C(y_p, w \mid T^{\Lambda, \Gamma}) \leq w x^* = \lambda^* w x_p$. Therefore, $\frac{C(y_p, w \mid T^{\Lambda, \Gamma})}{w x_p} \leq \lambda^*$. Hence, we have $CE(x_p, y_p, w \mid T^{\Lambda, \Gamma}) \leq DF_i(x_p, y_p \mid T^{\Lambda, \Gamma})$.

(ii) Based on the part (i), the proof is clear.

(iii) Suppose that λ^* is an optimal value of the model determining $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda})$ (model (13)), therefore $(\lambda^* x_p, y_p) \in T_p^{\Lambda}$. Based on the technology (12), there is $z_k(k = 1, \cdots, K), z_p = 0$ such that:

$$\begin{aligned} &\lambda^* x_p \geq \sum_{k=1}^{K} x_k \delta z_k, \\ &y_p \leq \sum_{k=1}^{K} y_k \delta z_k, \\ &z \in \Lambda, \delta \in \Gamma. \end{aligned} \qquad (A.2)$$

If we let $x^* = \lambda^* x_p$, then (x^*, z_k) is a feasible solution of the model determining $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, \Gamma})$ (model (15)), thus we have: $C(y_p, w \mid T^{\Lambda}) \leq w x^* = \lambda^* w x_p$. Therefore, $\frac{CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, \Gamma})}{w x_p} \leq \lambda^*$. Hence, we have $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, \Gamma}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, \Gamma})$.

Proof of Proposition 5.2:

(i) Suppose that (z_k^*, λ^*) is an optimal solution of the model determining $DF_i(x_p, y_p \mid T^{NC, \Gamma})$. Based on the technology (12), since we have $T^{NC, \Gamma} \subseteq T^{C, \Gamma}$, this optimal solution (z_k^*, λ^*) is a feasible solution of the model determining $DF_i(x_p, y_p \mid T^{C, \Gamma})$. Thus, we have

$DF_i(x_p, y_p \mid T^{C,\Gamma}) \leq \lambda^* = DF_i(x_p, y_p \mid T^{NC,\Gamma})$. In a similar way, the second part, i.e., $DF_i^{SE}(x_p, y_p \mid T^{C,\Gamma}) \leq DF_i^{SE}(x_p, y_p \mid T^{NC,\Gamma})$, can be proved.

(ii) Suppose that (z_k^*, x^*) is an optimal solution of the model determining $C(y_p, w \mid T^{NC,\Gamma})$. Based on the technology (12), since we have $T^{NC,\Gamma} \subseteq T^{C,\Gamma}$, this optimal solution (z_k^*, x^*) is a feasible solution of the model determining $C(y_p, w \mid T^{C,\Gamma})$. Thus, we have $C(y_p, w \mid T^{C,\Gamma}) \leq wx^* = C(y_p, w \mid T^{NC,\Gamma})$. Therefore, $CE(y_p, w \mid T^{C,\Gamma}) = \frac{C(y_p, w \mid T^{C,\Gamma})}{wx_p} \leq \frac{wx^*}{wx_p} = CE(y_p, w \mid T^{NC,\Gamma})$. In a similar way, the second part, i.e., $CE^{SE}(y_p, w \mid T^{C,\Gamma}) \leq CE^{SE}(y_p, w \mid T^{NC,\Gamma})$, can be proved.

Proof of Proposition 5.3:

(i) Assume that (x_q, y_q) is an observed unit such that $q \neq p$. Based on the constant returns to scale assumption, there is $\delta > 0$ such that $\delta(x_q, y_q) \in T^{\Lambda,CRS}$ and $\delta y_q \geq y_p$. Since, $(x_q, y_q) \in T_p^{\Lambda,CRS}$, therefore, $\delta(x_q, y_q) \in T_p^{\Lambda,CRS}$. Thus, based on technology $T_p^{\Lambda,CRS}$, there is $z_k (k = 1, \ldots, K), z_p = 0$ such that:

$$\delta x_q \geq \sum_{k=1}^{K} x_k z_k,$$

$$y_p \leq \delta y_q \leq \sum_{k=1}^{K} y_k z_k,$$

Hence, $z_k, k = 1, \ldots, K$ and $x = \delta x_q$ are a feasible solution of the CRS cost super-efficiency model $C^{SE}(y_p, w \mid T_p^{\Lambda,CRS})$.

(ii) Assume that (z_k^*, λ^*) is a feasible solution of model determining $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$. If we let $x^* = \lambda^* x_p$, then (z_k^*, x^*) is a feasible solution of model determining $C^{SE}(y_p, w \mid T_p^{\Lambda,VRS})$. Now, suppose that (z_k^*, x_i^*) is a feasible solution of model determining $C^{SE}(y_p, w \mid T_p^{\Lambda,VRS})$. We define $\lambda^* = \min\{\frac{x_i^*}{x_{ip}}, i = 1, \ldots, N\}$, therefore (z_k^*, λ^*) is a feasible solution of model determining $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS})$.

(iii) Based on the technology (12), since we have $T_p^{NC,\Gamma} \subseteq T^{C,\Gamma}$, therefore, if $T_p^{C,\Gamma} = \emptyset$, then $T_p^{NC,\Gamma} = \emptyset$.

(iv) Based on the technology (12), since we have $T_p^{NC,\Gamma} \subseteq T_p^{C,\Gamma}$, therefore, if $T_p^{C,\Gamma} = \emptyset$, then $T_p^{NC,\Gamma} = \emptyset$.

Proof of Proposition 5.4:

(i) If $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) < 1$, then based on Proposition 5.1, we have $CE(x_p, y_p, w \mid T^{\Lambda, CRS}) \leq DF_i(x_p, y_p \mid T^{\Lambda, CRS}) < 1$. Hence, $CE(x_p, y_p, w \mid T^{\Lambda, CRS}) = CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$ and $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS})$. Therefore, $AE(x_p, y_p, w \mid T^{\Lambda, CRS}) = AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

If $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = 1$, then based on Proposition 5.1, we have $CE(x_p, y_p, w \mid T^{\Lambda, CRS}) \leq 1$. Now, we have two items as follows:

(a) If $CE(x_p, y_p, w \mid T^{\Lambda, \Gamma}) < 1$, then $CE(x_p, y_p, w \mid T^{\Lambda, CRS}) = CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, \Gamma})$. Since $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS}) \geq 1$, we have $AE(x_p, y_p, w \mid T^{\Lambda, CRS}) \geq AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

(b) If $CE(x_p, y_p, w \mid T^{\Lambda, CRS}) = 1$, then since we have assumed that $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = 1$, thus we have $AE(x_p, y_p, w \mid T^{\Lambda, CRS}) = 1$. Also, based on Proposition 5.1, we have $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS})$, hence $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS}) \leq 1 = AE(x_p, y_p, w \mid T^{\Lambda, CRS})$.

(ii) Assume that $DF_i(x_p, y_p \mid T^{\Lambda, VRS}) < 1$. Hence $DF_i(x_p, y_p \mid T^{\Lambda, VRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, VRS})$. Since $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) \leq DF_i(x_p, y_p \mid T^{\Lambda, VRS}) < 1$, thus, we have $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS})$. Hence, $SCE(x_p, y_p \mid T^{\Lambda, \cdot \cdot}) = SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot \cdot})$.

Now, assume that $DF_i(x_p, y_p \mid T^{\Lambda, VRS}) = 1$. In this case, we have two items as follows:

(a) If $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) < 1$, then $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS})$. Since we have assumed that $DF_i(x_p, y_p \mid T^{\Lambda, VRS}) = 1$, hence $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, VRS}) \geq 1$. Therefore, we have $SCE(x_p, y_p \mid T^{\Lambda, \cdot \cdot}) \geq SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot \cdot})$.

(b) If $DF_i(x_p, y_p \mid T^{\Lambda, CRS}) = 1$, then since we have assumed that $DF_i(x_p, y_p \mid T^{\Lambda, VRS}) = 1$, thus we have $SCE(x_p, y_p \mid T^{\Lambda, \cdot \cdot}) = 1$. Also, we have $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS}) \leq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, VRS})$, hence $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda, \cdot \cdot}) \leq 1 = SCE(x_p, y_p \mid T^{\Lambda, \cdot \cdot})$.

Proof of Proposition 5.5:

(i) Based on equation (16), we have $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS}) \times DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, CRS}) = CE^{SE}(w_p, y_p \mid T_p^{\Lambda, CRS})$.

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) When $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda, VRS}) > 1$, the results of the empirical example show that $AE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$ can be equal, bigger or smaller than $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda, CRS})$.

A3

Proof of Proposition 5.6:

(i) Based on equation (17), we have $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) \times DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) = DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$.

If $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) \leq 1$, then $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) \geq DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,CRS})$.
Based on part (i) of Proposition 5.1, we have $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot}) \geq CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.

(ii) If $DF_i^{SE}(x_p, y_p \mid T_p^{\Lambda,VRS}) > 1$, then the results of the empirical example show that $SCE^{SE}(x_p, y_p \mid T_p^{\Lambda,\cdot\cdot})$ can be equal, bigger or smaller than $CE^{SE}(x_p, y_p, w \mid T_p^{\Lambda,CRS})$.