CONICAL FREE DISPOSAL HULL ESTIMATORS OF DIRECTIONAL DISTANCES AND LUENBERGER PRODUCTIVITY INDICES FOR GENERAL TECHNOLOGIES

Cinzia Daraio, Simone Di Leo, Léopold Simar

REPRINT | 2025 / 04

UCLouvain

LiDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication

ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Decision Support

Conical Free Disposal Hull estimators of directional distances and Luenberger productivity indices for general technologies

Cinzia Daraio [a,*], Simone Di Leo [a], Léopold Simar [b]

[a] Department of Computer, Control and Management Engineering A. Ruberti (DIAG), Sapienza University of Rome, via Ariosto 25, 00185, Rome, Italy
[b] Institut de Statistique, Biostatistique et Sciences Actuarielles (ISBA), LIDAM, Université Catholique de Louvain, Voie du Roman Pays 20, B 1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

JEL classification:
C1
C14
C13

Keywords:
Nonparametric production frontiers
Cone
DEA
FDH
Directional distances
Luenberger productivity indices

ABSTRACT

Directional distances are a popular tool in productivity and efficiency analysis due to their versatility in evaluating the distance of Decision Making Units (DMU) to the efficient frontier of the production set. The theoretical and statistical properties of these measures are well-established in various contexts. However, the measurement of directional distances to the cone spanned by the attainable set has not yet been explored. This cone is necessary to define the Luenberger indices for general technologies. This paper aims to fill this gap by presenting a method for defining and estimating directional distances to this cone, applicable to general technologies without imposing convexity. We also discuss the statistical properties of these measures, enabling us to measure distances to non-convex attainable sets under Constant Returns to Scale (CRS), as well as measure and estimate Luenberger productivity indices and their decompositions for general technologies. In addition, we provide a detailed description of how to make inferences on these indices. Finally, we offer simulated data and a practical example of inference on Luenberger productivity indices and their decompositions using a well-known real data set.

1. The framework and our contribution

1.1. Efficiency analysis and directional distance functions

In productivity and efficiency analysis, Decision-Making Units (DMUs) are benchmarked against a technical efficient frontier, characterized by the optimal combinations of the inputs (resources or factors of production) and the outputs (goods or services produced). Let Ψ denote the production set, i.e. the set of technically feasible combinations of the p inputs (x) and the q outputs (y). It can be defined as

$$\Psi = \{(x, y) \in \mathbb{R}_+^{p+q} \mid x \text{ can produce } y\}. \tag{1.1}$$

This set shares economic theory's usual characteristics (e.g. Shephard (1970)). A minimal set of assumptions is that Ψ is closed and freely disposable in the inputs and the outputs. All the inputs and outputs are freely (or strongly) disposable if $(x, y) \in \Psi \Rightarrow (\tilde{x}, \tilde{y}) \in \Psi$, $\forall \tilde{x} \geq x, \tilde{y} \leq y$. Sometimes, the convexity of Ψ is also assumed. In this paper, we will focus on more general technologies, i.e. without requiring the convexity of Ψ. See e.g. Kneip et al. (2023) and the references therein for a discussion on why the convexity of the production set may not be an appropriate assumption. The efficient boundary (frontier) of this set is the set of efficient combinations of inputs and outputs.

$$\Psi^\partial = \{(x, y) \in \Psi \mid (\gamma^{-1}x, \gamma y) \notin \Psi, \forall \gamma > 1\}. \tag{1.2}$$

Typically, the efficiency of a DMU (x, y) is measured by its distance from the boundary Ψ^∂. The Farrell–Debreu radial distances (input or output oriented) are the most popular (see Debreu (1951) and Farrell (1957)), but, in this paper, we will mainly work with Directional Distance Functions (DDFs, see Chambers et al. (1998)). The approach of DDFs is a general efficiency analysis method that extends traditional distance functions by allowing for flexible scaling of inputs and outputs simultaneously. Unlike traditional models that focus solely on radial expansions or contractions, DDFs provide a more general framework to measure inefficiency by setting a specific *direction* in which inputs are reduced and outputs are increased. DDFs quantify inefficiency by measuring how far a production unit can move in a given direction while still staying within the feasible production set. The main advantages of DDFs are (i) *flexibility*: DDFs allow for non-radial movements, making them suitable for settings where only specific inputs or outputs need

* Corresponding author.
E-mail addresses: daraio@diag.uniroma1.it (C. Daraio), dileo@diag.uniroma1.it (S. Di Leo), leopold.simar@uclouvain.be (L. Simar).

https://doi.org/10.1016/j.ejor.2024.12.025
Received 20 March 2024; Accepted 15 December 2024
Available online 21 December 2024
0377-2217/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

to be adjusted; (ii) *inclusion of undesirable outputs*: They allow to estimate the efficiency maximizing output while minimizing undesirable outputs.

Mathematically, DDFs are defined as

$$\delta(x, y) = \delta(x, y; d_x, d_y) := \sup\{\delta \,|\, (x - \delta d_x, y + \delta d_y) \in \Psi\}, \tag{1.3}$$

where $d_x \in \mathbb{R}^p_+$ and $d_y \in \mathbb{R}^q_+$. Here, the distance is measured along a path determined by a direction vector $d' = (-d'_x, d'_y)$ in an additive way. Clearly if $(x, y) \in \Psi$, $\delta(x, y) \geq 0$ and if (x, y) lies on the efficient frontier (1.2), $\delta(x, y) = 0$. When possible, and when there are no ambiguities, we will use the simpler notation $\delta(x, y)$ for $\delta(x, y; d_x, d_y)$, when it is understood that the directions have been well specified.

It is well-known that for $(x, y) \in \mathbb{R}^{p+q}_+$, the Farrell–Debreu radial measures are particular cases of the directional distances: for the input oriented case, the efficiency score $\theta(x, y) = 1 - \delta(x, y; x, 0_q) \leq 1$ represents the percentage of reduction of each input the unit (x, y) has to perform to reach the efficient frontier. Generally, 0_m denotes a vector of zeros of dimension m. Similarly, the output oriented measure of efficiency is $\lambda(x, y) = 1 + \delta(x, y; 0_p, y) \geq 1$. It represents the percentage of increase of each output to reach the efficient frontier.

There are many possible choices for the direction vector d, it can be specific for each DMU or we can choose, in a more "egalitarian" way, a common direction for all the DMUs. This shows the flexibility of the approach for the evaluation of efficiencies based on directional distance functions, see e.g. Daraio and Simar (2016) and Färe et al. (2008) for more discussions. Literature reviews on the applications of DDFs in different fields are available in Färe and Grosskopf (2000) and in Färe et al. (2015).

1.2. Limitation of convexity

In traditional models of production technology, the convexity of the production set Ψ has been a widely adopted assumption, as it simplifies the mathematical structure and estimation procedures. The reliance on convexity may lead to biases and inconsistencies in efficiency measurement because it does not accommodate non-convex technologies that may be present in many sectors. The convexity assumption may not be appropriate because real-world production sets often exhibit non-convexity, which violates the assumptions typically imposed by estimators like Data Envelopment Analysis (DEA). Kneip et al. (2023) mention that nonparametric envelopment estimators, such as DEA, typically rely on the convexity of the production set, meaning that all feasible combinations of inputs and outputs should form a convex set. Non-convexity can occur in situations where there are economies or diseconomies of scale, discrete inputs, or lumpy production technologies that do not scale smoothly. In fact, Kneip et al. (2023) and the references cited therein, describe cases where convexity has been rejected in empirical studies across various industries, including banks, hospitals, and local governments. This highlights that the convexity assumption often oversimplifies the complexity of real-world production technologies, leading to wrong efficiency and productivity measurements and inconsistent estimators.

1.3. Necessity of considering the cone spanned by Ψ

This paper introduces the notion of the cone spanned by the attainable set. This cone is defined as the union of all rays starting for the origin and passing through points of the set. So it is a pointed cone (i.e. it contains the origin). In the case of Constant Returns to Scale (CRS) of Ψ, the cone coincides with Ψ. Note also that if the attainable set is non-convex, the cone may be non-convex (as shown below in Fig. 1). So as shown below in Eq. (1.4), the cone spanned by Ψ, is obtained by radial expansion of its points.

Characterizing directional distances to the cone spanned by Ψ, and providing estimators, is useful for characterizing and estimating consistently directional efficiency scores in case of CRS, when the sets are possibly non-convex. But also, as described in the literature (see details below), the cone spanned by Ψ is needed to define Malmquist and Luenberger productivity indices, in terms of distances to its efficient boundary, even for cases where we do not have CRS.

Therefore, considering distances to the boundary of the cone spanned by Ψ for general technologies, is particularly useful in various setups where the production process exhibits non-convexities, independently of the assumption on the scale properties (CRS or non-CRS).

Formally, the cone spanned by Ψ is defined by

$$\mathcal{C}(\Psi) = \{(\tilde{x}, \tilde{y}) \mid \tilde{x} = ax, \tilde{y} = ay, \forall a \in \mathbb{R}_+ \text{ and } \forall (x, y) \in \Psi\}. \tag{1.4}$$

As a result of the non-convex nature of the attainable set Ψ, it is worth noting that the cone spanned $\mathcal{C}(\Psi)$ may also not be a convex set. In the particular case of Constant Returns to Scale (CRS), we have $\mathcal{C}(\Psi) = \Psi$, for both convex and non-convex cases, but in general $\Psi \subseteq \mathcal{C}(\Psi)$. Note that $\mathcal{C}(\Psi)$ is a cone pointed at zero.

Analogous to (1.2), we can define the frontier of the set $\mathcal{C}(\Psi)$ as

$$\mathcal{C}^\partial(\Psi) = \{(x, y) \in \mathcal{C}(\Psi) | (\gamma^{-1}x, \gamma y) \notin \mathcal{C}(\Psi), \forall \gamma > 1\}. \tag{1.5}$$

Fig. 1 illustrates a case for $q = 1$ and $p = 2$. In the given scenario, there is a single output and two inputs. The left panel of Fig. 1 highlights Variable Returns to Scale, as evidenced by an initial increase followed by a subsequent decrease. The isoquants in the input space, which are the boundaries of the sections, are defined by portions of circles. This is demonstrated by the contour plots on the floor of both panels in Fig. 1. It is to be noted that for this specific example, the resulting sections are non-convex. On the right panel of Fig. 1, the cone could also represent a non-convex production set with Constant Returns to Scale (CRS). Nevertheless, in this case, the cone is spanned by a general Ψ, which is analogous to the one exhibited in the left panel of Fig. 1.

Similarly, we can define the directional distance from a production plan (x, y) to the boundary of the cone spanned by the production set $\mathcal{C}(\Psi)$

$$\delta_{\mathcal{C}}(x, y) = \delta_{\mathcal{C}}(x, y; d_x, d_y) := \sup\{\delta \,|\, (x - \delta d_x, y + \delta d_y) \in \mathcal{C}(\Psi)\}. \tag{1.6}$$

This coincides to the directional efficiency measure for (x, y) under a CRS assumption, i.e. if $\mathcal{C}(\Psi) = \Psi$, $\delta_{\mathcal{C}}(x, y) = \delta(x, y)$. More generally, $\Psi \subseteq \mathcal{C}(\Psi)$ and $\delta_{\mathcal{C}}(x, y) \geq \delta(x, y)$ are needed for defining productivity indices in general cases (see below Section 4). Sometimes it will be useful, in the notation, to use the equivalence

$$\delta_{\mathcal{C}}(x, y) = \delta(x, y \mid \mathcal{C}(\Psi)), \tag{1.7}$$

which makes explicit the fact that we use a directional distance to the boundary of $\mathcal{C}(\Psi)$ and not of Ψ.

When data are available on the different units at several periods, researchers are interested in analyzing the change in productivity, and efficiency, but also to check if technical changes have occurred. All these issues can be analyzed through productivity indices among which the Malmquist Productivity Index (MPI) is very popular (see e.g. Caves et al. (1982), Färe et al. (1994) and Malmquist (1953). Surveys on Malmquist index theory and applications in different sectors are offered in Färe et al. (2008, 1998). The MPI is mainly defined in terms of radial efficiency measures. The statistical inference on the MPI, its aggregations and decompositions have been developed in recent years (see, e.g., Kneip et al. (2021, 2023) and Pham et al. (2023)) thanks in part to recent results on central limit theorems for efficiency estimators (see Kneip et al. (2015) and Simar and Zelenyuk (2018)) providing the possibility of assessing the statistical significance and confidence intervals of indices and their components. The Luenberger Productivity Indices (LPI) extend the MPI to account for directional distances (Chambers, 2002; Chambers et al., 1996). Boussemart et al. (2003) and Epure et al. (2011), using radial (proportional) directional

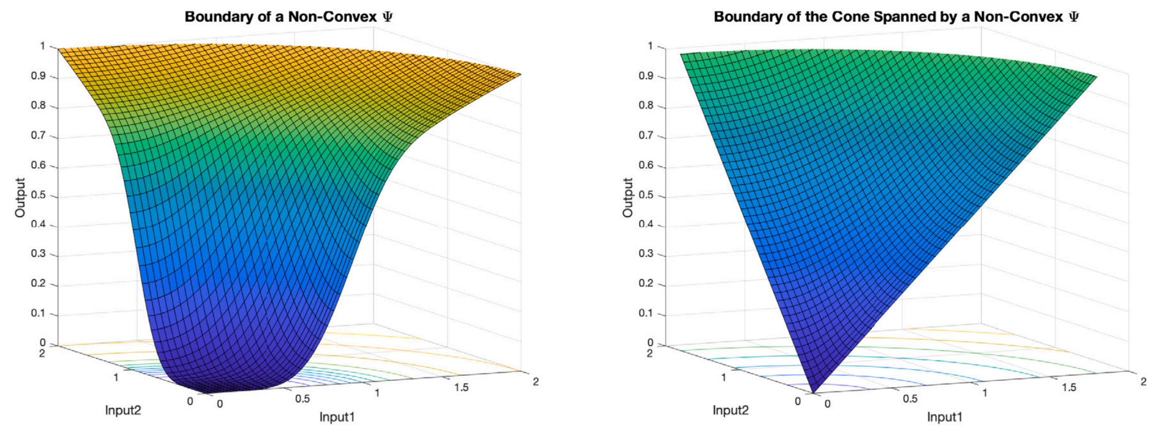

Fig. 1. Boundaries of a non-convex Ψ and of the cone spanned by a non-convex Ψ. The left panel shows a typical efficient boundary of a non-convex production set Ψ. The right panel displays the shape of $C(\Psi)$, the cone spanned by this type of set. The contour plots drawn on the floor are the isoquants in the input space.

distances notice that the LPI generalizes the MPI because it allows to investigate situations where simultaneous proportional reduction of inputs and increasing of outputs defines the efficient boundaries. Kevork et al. (2017) developed Shephard's type output directional efficiency measures to analyze European banks and more recently, Pastor et al. (2020) introduced a proportional directional distance function in the MPI and derived a new decomposition of the MPI.

In this paper, we generalize the previous literature, since, thanks to our $\delta_C(x, y; d_x, d_y)$ defined above, any directional distance vector d can be used to measure the distance of a DMU (x, y) to the boundary of $C(\Psi)$.

1.4. Contribution

Building on the conicity construction, this paper develops a comprehensive framework for defining and estimating directional distances to the cone spanned by the attainable set Ψ. Specifically, we propose a method to estimate directional distances under general technologies without assuming convexity. Our contributions are threefold:
(i) *Definition of Directional Distances to Cones*. We generalize previous studies by introducing directional distances of any kind to the boundary of the cone spanned by Ψ. This extends the work of Boussemart et al. (2003), Epure et al. (2011), Kevork et al. (2017) and Pastor et al. (2020) who assumed convexity and used radial or proportional directional distances;
(ii) *Estimation of Luenberger Productivity Indices for general technologies*, using these new directional distance estimators. We define and estimate Luenberger Productivity Indices (LPIs) for general technologies, allowing for decompositions to assess sources of productivity change. This is particularly important as our framework does not impose the CRS assumption;
(iii) *Statistical Inference*. We provide statistical tools for inference on the derived productivity indices, making it possible to test the significance of the directional distance measures and their decompositions under both convex and non-convex technologies. None of the existing studies on this topic has developed statistical inferences for directional distances of general technologies and the LPI and its decompositions.

The paper is organized as follows. Section 2 introduces the basic properties of $\delta_C(x, y)$, with the various versions one can obtain by selecting particular vectors of directions. Section 3 suggests nonparametric estimators and Section 4 shows how to define and estimate LPIs under general technologies. Appendix contains instructions for conducting inference. Section 5 demonstrates the various tools introduced in the paper with simulated samples. Practical inferences on LPIs and their components are illustrated using real data from the literature. Section 6 concludes the paper.

2. Computation and properties of $\delta_C(x, y)$

2.1. Basic properties

Before going to see how we can characterize $\delta_C(x, y)$, we reconsider the way to define $\delta(x, y)$ with its equivalent as described in Daraio et al. (2020). For simplicity, here we assume that all the elements of d_x and d_y are strictly positive. It would be easy, but at a high notational cost, to handle cases where some elements of d equal zero, as indicated by Daraio et al. (2020). We will see in Section 2.2 that the cases where $d_x = 0_p$ or the case where $d_y = 0_q$ are easy to handle.

We denote by (x_0, y_0) the point of interest. Directional distances are independent of the units of measurement as described in Färe et al. (2008) and are formally proven in Appendix in Simar and Vanhems (2012). As a consequence, we can rewrite (1.3) as

$$\delta(x_0, y_0; d_x, d_y) = \delta(x_0^\star, y_0^\star; i_p, i_q), \tag{2.1}$$

with $x_0^\star = x_0 \oslash d_x$, $y_0^\star = y_0 \oslash d_y$, where $\oslash$ is the Hadamard component-wise division of vectors, and i_m is the m-vector of ones. Similarly, below we will denote $x^\star = x \oslash d_x$ and $y^\star = y \oslash d_y$. As shown by Daraio et al. (2020), the last equation allows us to rewrite the directional distance as

$$\delta(x_0, y_0; d_x, d_y) = \sup_{(x,y)\in\Psi} \left\{ \min_{\substack{j=1,\dots,p \\ k=1,\dots,q}} [x_{0,j}^\star - x_j^\star, y_k^\star - y_{0,k}^\star] \right\}. \tag{2.2}$$

Daraio et al. (2020) introduce the Hadamard product to allow for more efficient computation of distances by focusing on component-wise changes. Hence, here again, the motivation for introducing the Hadamard product in the previous equations relates to simplifying the characterization of directional distances by scaling the inputs and outputs in a component-wise manner. In fact, it is shown that the directional distance in Eq. (1.3) is identical to the one computed through (2.2). The Hadamard notation is thus a mathematical trick to simplify the characterization of the directional distance and to facilitate its FDH estimation.

Now we can characterize more easily the distance $\delta_C(x_0, y_0)$. Fig. 2 illustrates the idea in the simplest case where $p = q = 1$. The point $A = (x_0, y_0)$ is projected in the direction d on the frontier of $C(\Psi)$ at $A_g^\partial = (x_0 - \delta_C(x_0, y_0)d_x, y_0 + \delta_C(x_0, y_0)d_y)$. Clearly, by properties of similar triangles, we have

$$\delta_C(x_0, y_0) = \sup_{a>0} \left\{ \frac{\delta(ax_0, ay_0)}{a} \mid (ax_0 - \delta(ax_0, ay_0)d_x, ay_0 + \delta(ax_0, ay_0)d_y) \in \Psi \right\}. \tag{2.3}$$

In Fig. 2, this supremum is achieved at $a = a_g$, defining the point $a_g A$ on the ray passing through $A = (x_0, y_0)$ and its projection P^∂ on the frontier

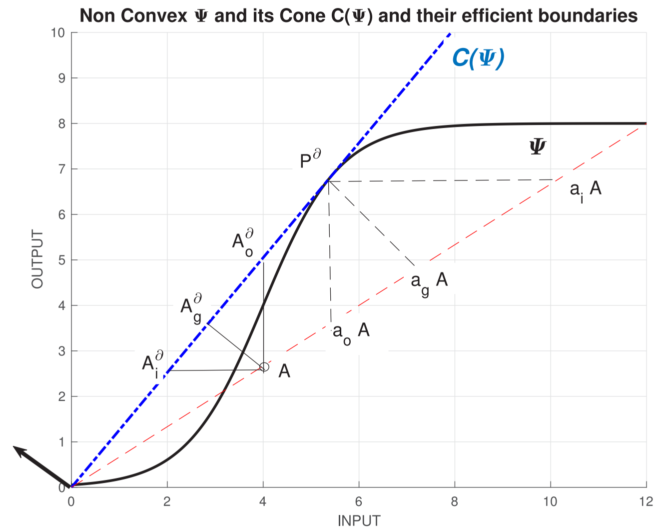

Fig. 2. Illustration of a production set and its boundary when $p = q = 1$. The point of interest is $A = (x_0, y_0)$, the solid black line is the efficient boundary of Ψ, the dash-dotted blue is the frontier of the cone $\mathcal{C}(\Psi)$ and the dashed red line is the ray passing through A. The arrow from the origin to the NW is the chosen direction $d = (-d_x, d_y)$.

in direction d. The latter has coordinates $(a_g x_0 - \delta(a_g x_0, a_g y_0)d_x, a_g y_0 + \delta(a_g x_0, a_g y_0)d_y)$. The distance between the points $a_g A$ and P^∂, given by $\|(a_g A)P^\partial\|$, is proportional to the distance $\|AA_g^\partial\|$ and the factor of proportionality is given by a_g.

In fact, by using the transformations above and (2.2), we have

$$\delta_C(x_0, y_0) = \sup_{a>0}\left\{\frac{1}{a} \sup_{(x,y)\in\Psi} \left\{ \min_{\substack{j=1,\dots,p \\ k=1,\dots,q}} [a x_{0,j}^\star - x_j^\star, y_k^\star - a y_{0,k}^\star]\right\}\right\}. \tag{2.4}$$

We will see below that in various particular cases, we have more explicit solutions to this problem and in particular how these provide, in practice, easy-to-compute estimators. In particular, it will be useful to notice that if $(x_0, y_0) \in \Psi$, we can find a value for a and a point $(x, y) \in \Psi$ such that the supremum is positive, so we should have $a x_0 \geq x$ and $a y_0 \leq y$ for characterizing the supremum. This provides lower and upper bounds for a_{opt}, the optimal a when $(x_0, y_0) \in \Psi$:

$$\min_{(x,y)\in\Psi}\left\{\max_{j=1,\dots,p} \frac{x_j}{x_{0,j}}\right\} \leq a_{opt} \leq \max_{(x,y)\in\Psi}\left\{\min_{k=1,\dots,q} \frac{y_k}{y_{0,k}}\right\}. \tag{2.5}$$

These correspond (for the case $p = q = 1$) to the two points where the ray aA intersects the frontier of Ψ.

2.2. Particular cases: input and output orientations

The basic property in (2.4) simplifies when we consider "oriented" directional distances, i.e. when either, $d_x = 0_p$ and only $d_y > 0_q$, or, $d_y = 0_q$ and $d_x > 0_p$. In the first case we have an output orientation, whereas in the second case, we have an input orientation. We first consider the input orientation, in this case, we have

$$\delta^{\text{inp}}(x_0, y_0) = \sup\{\delta \,|\, (x_0 - \delta d_x, y_0) \in \Psi\}, \tag{2.6}$$

and following Daraio et al. (2020), this can be written as

$$\delta^{\text{inp}}(x_0, y_0) = \sup_{(x,y)\in\Psi, y\geq y_0}\left\{\min_{j=1,\dots,p} [x_{0,j}^\star - x_j^\star]\right\}. \tag{2.7}$$

For defining the distance to the boundary of $C(\Psi)$, we still use Fig. 2 to illustrate the idea. The point $A = (x_0, y_0)$ is projected in the input direction d_x at $A_i^\partial = (x_0 - \delta_C(x_0, y_0)d_x, y_0)$. Again by properties of similar triangles, the distance $\|(a_i A)P^\partial\|$ is proportional to the distance $\|AA_i^\partial\|$ and the factor of proportionality is given by a_i which is the value of a that solves the optimization problem

$$\delta_C^{\text{inp}}(x_0, y_0) = \sup_{a>0}\left\{\frac{\delta(a x_0, a y_0)}{a} \,|\, (a x_0 - \delta(a x_0, a y_0)d_x, a y_0) \in \Psi\right\}. \tag{2.8}$$

which, by (2.7), is equivalent to

$$\delta_C^{\text{inp}}(x_0, y_0) = \sup_{a>0}\left\{\frac{1}{a} \sup_{(x,y)\in\Psi, y\geq a y_0}\left\{\min_{j=1,\dots,p} [a x_{0,j}^\star - x_j^\star]\right\}\right\},$$

$$= \sup_{a>0}\left\{\sup_{(x,y)\in\Psi, y\geq a y_0}\left\{\min_{j=1,\dots,p} [x_{0,j}^\star - a^{-1} x_j^\star]\right\}\right\} \tag{2.9}$$

Now define $\tilde{a}_0^{\text{out}} = \min_{k=1,\dots,q}[y_k / y_{0,k}]$ (note that $\tilde{a}_0^{\text{out}} = \tilde{a}_0^{\text{out}}(x, y, x_0, y_0)$). Clearly we have $\tilde{a}_0^{\text{out}} y_0 \leq y$ while $a y_0 \nleq y$ for any $a > \tilde{a}_0^{\text{out}}$. Now, $\min_{j=1,\dots,p}[x_{0,j}^\star - a^{-1} x_j^\star] \leq \min_{j=1,\dots,p}[x_{0,j}^\star - (\tilde{a}_0^{\text{out}})^{-1} x_j^\star]$ for any $a \leq \tilde{a}_0^{\text{out}}$, so the solution to (2.9) is given for this value of a. Then we have

$$\delta_C^{\text{inp}}(x_0, y_0) = \sup_{(x,y)\in\Psi}\left\{\frac{1}{\min_{k=1,\dots,q}[y_k / y_{0,k}]}\right.$$

$$\left.\left\{\min_{j=1,\dots,p}\left[\min_{k=1,\dots,q}[y_k / y_{0,k}] x_{0,j}^\star - x_j^\star\right]\right\}\right\}. \tag{2.10}$$

We will see in Section 3 that this expression provides an explicit formula for nonparametric estimators, where the unknown Ψ will be replaced by its Free Disposal Hull (FDH) estimator (see Deprins et al. (1984)).

For the output orientation, the argument is very similar. We have

$$\delta^{\text{out}}(x_0, y_0) = \sup\{\delta \,|\, (x_0, y_0 - \delta d_y) \in \Psi\}, \tag{2.11}$$

which can be written as (see Daraio et al. (2020))

$$\delta^{\text{out}}(x_0, y_0) = \sup_{(x,y)\in\Psi, x\leq x_0}\left\{\min_{k=1,\dots,q} [y_k^\star - y_{0,k}^\star]\right\}. \tag{2.12}$$

Here, in Fig. 2, we have to characterize the proportionality between $\|(a_o A)P^\partial\|$ and $\|AA_o^\partial\|$ and the factor of proportionality is given by a_o which is the value of a that solves the optimization problem

$$\delta_C^{\text{out}}(x_0, y_0) = \sup_{a>0}\left\{\frac{\delta(a x_0, a y_0)}{a} \,|\, (a x_0, a y_0 + \delta(a x_0, a y_0)d_y) \in \Psi\right\}. \tag{2.13}$$

which, by (2.12), is equivalent to

$$\delta_C^{\text{out}}(x_0, y_0) = \sup_{a>0}\left\{\frac{1}{a} \sup_{(x,y)\in\Psi, x\leq a x_0}\left\{\min_{k=1,\dots,q} [y_k^\star - a y_{0,k}^\star]\right\}\right\},$$

$$= \sup_{a>0}\left\{\sup_{(x,y)\in\Psi, x\leq a x_0}\left\{\min_{k=1,\dots,q} [a^{-1} y_k^\star - y_{0,k}^\star]\right\}\right\} \tag{2.14}$$

Now define $\tilde{a}_0^{\text{inp}} = \max_{j=1,\dots,p}[x_j / x_{0,j}]$. Clearly we have $\tilde{a}_0^{\text{inp}} x_0 \geq x$ while $a x_0 \ngeq x$ for any $a < \tilde{a}_0^{\text{inp}}$. Now, $\min_{k=1,\dots,q}[a^{-1} y_k^\star - y_{0,k}^\star] \leq \min_{k=1,\dots,q}[(\tilde{a}_0^{\text{inp}})^{-1} y_k^\star - y_{0,k}^\star]$ for any $a \geq \tilde{a}_0^{\text{out}}$, so the solution to (2.14) is given for this value of a. Then we have

$$\delta_C^{\text{out}}(x_0, y_0) = \sup_{(x,y)\in\Psi}\left\{\frac{1}{\max_{j=1,\dots,p}[x_j / x_{0,j}]}\right.$$

$$\left.\times\left\{\min_{k=1,\dots,q}\left[y_k^\star - \max_{j=1,\dots,p}[x_j / x_{0,j}] y_{0,k}^\star\right]\right\}\right\}. \tag{2.15}$$

Again, this expression will provide an explicit formula for nonparametric FDH estimators.

2.3. Radial orientations: proportional directional distances

The particular case of directional distances with radial orientations is often used in practice and all the above expressions for defining $\delta_C(x_0, y_0)$ greatly simplifies. We consider in this section where $d_x = x_0$ and $d_y = y_0$ and we also consider the "pure" input and the "pure" output particular cases. Since these directional distances are specific, we use an adapted notation. Thus we are interested in the directional distance defined as

$$\alpha(x_0, y_0) = \sup\{\alpha \,|\, ((1-\alpha)x_0, y_0) \in \Psi\}, \text{ input radial }, \tag{2.16}$$

$$\beta(x_0, y_0) = \sup\{\beta \,|\, (x_0, (1+\beta)y_0) \in \Psi\}, \text{ output radial }, \tag{2.17}$$

$$\gamma(x_0, y_0) = \sup\{\gamma \,|\, ((1-\gamma)x_0, (1+\gamma)y_0) \in \Psi\}, \text{ general radial} \tag{2.18}$$

and their conical versions where Ψ is replaced by $C(\Psi)$. For obvious reasons, these distances are often called proportional directional distances; the input radial is sometimes referred to as "cost minimization", the output radial as "revenue maximization" and the general radial as "profit maximization" (see e.g. Boussemart et al. (2003) and Epure et al. (2011)). In the one-dimensional case ($p = q = 1$), at a scale factor, all directions are radial, so Fig. 2 can still be used to illustrate the three measures. In more general multidimensional cases, the radial feature allows us to simplify the calculations.

First, as noted above, the input (output) radial allows to recovery of the Farrell–Debreu input (output) efficiency scores defined for the conical case in Kneip et al. (2023). We have

$$1 - \alpha_C(x_0, y_0) = \theta_C(x_0, y_0) = \inf_{(x,y)\in\Psi} \frac{\max_{j=1,\dots,p}[x_j/x_{0,j}]}{\min_{k=1,\dots,q}[y_k/y_{0,k}]}, \tag{2.19}$$

$$1 + \beta_C(x_0, y_0) = \lambda_C(x_0, y_0) = \sup_{(x,y)\in\Psi} \frac{\min_{k=1,\dots,q}[y_k/y_{0,k}]}{\max_{j=1,\dots,p}[x_j/x_{0,j}]} \tag{2.20}$$

As shown in the proof of Lemma 3.2 in Kneip et al. (2021), and adapted to possible non-convex Ψ by Lemma 3.1 in Kneip et al. (2023), the three projected points on $C^\partial(\Psi)$ are on the same ray. This appears as obvious in the case $p = q = 1$ of Fig. 2, it is still true but less obvious for multidimensional radial cases. However, it is not true for general directions $d = (-d_x, d_y)$. For the radial cases, it comes from the properties of θ_C and λ_C, we have for any $a, b > 0$ $\lambda_C(ax, y) = a\lambda_C(x, y)$ and $\lambda_C(x, by) = b^{-1}\lambda_C(x, y)$ and since (see Lemma 3.1 of Kneip et al. (2023)) $\theta_C(x, y)\lambda_C(x, y) = 1$, we have analogue (inverse) properties for $\theta_C(x, y)$.

For the general radial case (not considered in Kneip et al. (2023)), the proof is as follows: the point (x_0, y_0) projected on $C^\partial(\Psi)$ in the direction $(-x_0, y_0)$, is also "output" efficient, so we have

$$\lambda_C((1 - \gamma_C(x_0, y_0))x_0, (1 + \gamma_C(x_0, y_0))y_0) = \frac{1 - \gamma_C(x_0, y_0)}{1 + \gamma_C(x_0, y_0)}\lambda_C(x_0, y_0) = 1$$

So the point with coordinates $((1-\gamma_C(x_0, y_0))x_0, (1+\gamma_C(x_0, y_0))y_0)$ can be written as $((1 - \gamma_C(x_0, y_0))x_0, (1 - \gamma_C(x_0, y_0))\lambda_C(x_0, y_0)y_0)$, which in turn has the form $(ax_0, a\lambda_C(x_0, y_0)y_0)$ for some $a > 0$ and it belongs to the efficient ray defined by the radial output measure. The same is true for the efficient ray defined by the radial input measure. To summarize, even in large dimensions, the three frontier points, A_i^∂, A_g^∂ and A_o^∂ of Fig. 2 are on the same ray. So we have the following Lemma.

Lemma 2.1. *When using radial directions we have for any point* $(x_0, y_0) \in \mathbb{R}_+^{p+q}$ *the relations*

$$\big(1 - \alpha_C(x_0, y_0)\big)\big(1 + \beta_C(x_0, y_0)\big) = 1 \tag{2.21}$$

$$\gamma_C(x_0, y_0) = \frac{\alpha_C(x_0, y_0)\beta_C(x_0, y_0)}{\alpha_C(x_0, y_0) + \beta_C(x_0, y_0)}. \tag{2.22}$$

The first relation is already proven in Kneip et al. (2023). To the best of our knowledge, the second relation has never been stated in the literature. We left its proof as an exercise by playing with properties of similar triangles, due to the fact, explained above, that the three projected points are on the same ray (this is crucial, so the Lemma is not valid for general directions). These explicit relations will be useful to simplify the computation of the nonparametric estimators when radial orientations have been chosen.

3. Nonparametric estimators of $\delta_C(x, y)$

In practice, we do not know the attainable set Ψ, but we can estimate this set based on a random sample of n units $\mathcal{X}_n = \{(X_i, Y_i)\}_{i=1}^n$, where X_i is the vector of the observed p inputs and Y_i is the vector of the q outputs, $i = 1, \dots, n$. For general technologies, where we do not impose the convexity of Ψ, we can use the FDH estimator of Ψ, suggested by Deprins et al. (1984). It is the free disposal hull of the cloud of data points

$$\widehat{\Psi}_{FDH} = \{(x, y) \in \mathbb{R}_+^{p+q} \mid y \le Y_i, x \ge X_i, (X_i, Y_i) \in \mathcal{X}_n\}, \tag{3.1}$$

i.e. the union of all the orthants (positive in the inputs x and negative in the outputs y) having a vertex at the observed points. Then, as in Kneip et al. (2023), $C(\Psi)$ can be estimated by the cone spanned by $\widehat{\Psi}_{FDH}$. We have

$$C(\widehat{\Psi}_{FDH}) = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}_+^{p+q} \mid \tilde{x} = ax, \tilde{y} = ay, \forall a \in \mathbb{R}_+ \text{ and } \forall (x, y) \in \widehat{\Psi}_{FDH}\} \tag{3.2}$$

Now the FDH estimators of the directional distances are obtained by plugging $\widehat{\Psi}_{FDH}$ in place of Ψ in the definitions given in the preceding section. To be explicit, see e.g. Daraio et al. (2020), we have from (2.2)

$$\hat{\delta}(x_0, y_0; d_x, d_y) = \max_{i=1,\dots,n}\Big\{\min_{\substack{j=1,\dots,p\\k=1,\dots,q}}[x_{0,j}^\star - X_{i,j}^\star, Y_{i,k}^\star - y_{0,k}^\star]\Big\}, \tag{3.3}$$

with the notation $X_i^\star = X_i \oslash d_x$ and $Y_i^\star = Y_i \oslash d_y$, $i = 1, \dots, n$.

For the Conical FDH (CFDH) estimator of $\delta_C(x_0, y_0; d_x, d_y)$, we have the general formulation from (2.4)

$$\hat{\delta}_C(x_0, y_0; d_x, d_y) = \sup_{a>0}\Big\{\frac{1}{a}\max_{i=1,\dots,n}\Big\{\min_{\substack{j=1,\dots,p\\k=1,\dots,q}}[ax_{0,j}^\star - X_{i,j}^\star, Y_{i,k}^\star - ay_{0,k}^\star]\Big\}\Big\}. \tag{3.4}$$

In the next section, we will see how the latter simplifies according to the chosen direction d.

3.1. Radial orientations (proportional directional distances)

As seen in Section 2.3 when $d_x = x_0$ and $d_y = y_0$ the definition of the directional distances is greatly simplified and due to the radial nature of the directions, some relations can be established between the various versions (input oriented, output oriented, simultaneous input and output orientations). These three versions of the directional radial distances were denoted $\alpha_C(x_0, y_0)$, $\beta_C(x_0, y_0)$ and $\gamma_C(x_0, y_0)$ and some useful relations were given in Lemma 2.1. For the input and the output orientations, we can use, by (2.19) and (2.20), the results of Kneip et al. (2023). So we obtain the explicit CFDH estimators

$$\hat{\alpha}_C(x_0, y_0) = 1 - \hat{\theta}_C(x_0, y_0) = 1 - \min_{i=1,\dots,n}\frac{\max_{j=1,\dots,p}[X_{i,j}/x_{0,j}]}{\min_{k=1,\dots,q}[Y_{i,k}/y_{0,k}]}, \tag{3.5}$$

$$\hat{\beta}_C(x_0, y_0) = \hat{\lambda}_C(x_0, y_0) - 1 = \max_{i=1,\dots,n}\frac{\min_{k=1,\dots,q}[Y_{i,k}/y_{0,k}]}{\max_{j=1,\dots,p}[X_{i,j}/x_{0,j}]} - 1, \tag{3.6}$$

which solve explicitly (3.4) for the cases $d = (-x_0, 0_q)$ and $d = (0_p, y_0)$, respectively.

For the general case, where we use both the input and the output radial directions simultaneously, i.e. $d = (-x_0, y_0)$, by Lemma 2.1 we obtain

$$\hat{\gamma}_C(x_0, y_0) = \frac{\hat{\alpha}_C(x_0, y_0)\hat{\beta}_C(x_0, y_0)}{\hat{\alpha}_C(x_0, y_0) + \hat{\beta}_C(x_0, y_0)}. \tag{3.7}$$

3.2. General directions

3.2.1. Input and output oriented cases

Here we consider cases where $d = (-d_x, 0_q)$ and cases where $d = (0_p, d_y)$ for general directions d_x and d_y. In these cases, due to the results of Section 2.2 we obtain again explicit expressions of the FDH estimators. By plugging $\widehat{\Psi}_{FDH}$ in place of Ψ in (2.10) we have the CFDH estimator

$$\hat{\delta}_C^{\text{inp}}(x_0, y_0) = \max_{i=1,\dots,n}\Bigg\{\frac{1}{\min_{k=1,\dots,q}[Y_{i,k}/y_{0,k}]} \times \Big\{\min_{j=1,\dots,p}\big[\min_{k=1,\dots,q}[Y_{i,k}/y_{0,k}]x_{0,j}^\star - X_{i,j}^\star\big]\Big\}\Bigg\}. \tag{3.8}$$

Similarly, for the output orientation, we have (2.15)

$$\hat{\delta}_C^{\text{out}}(x_0, y_0) = \max_{i=1,\dots,n}\Bigg\{\frac{1}{\max_{j=1,\dots,p}[X_{i,j}/x_{0,j}]} \times \Big\{\min_{k=1,\dots,q}\big[Y_{i,k}^\star - \max_{j=1,\dots,p}[X_{i,j}/x_{0,j}]y_{0,k}^\star\big]\Big\}\Bigg\}. \tag{3.9}$$

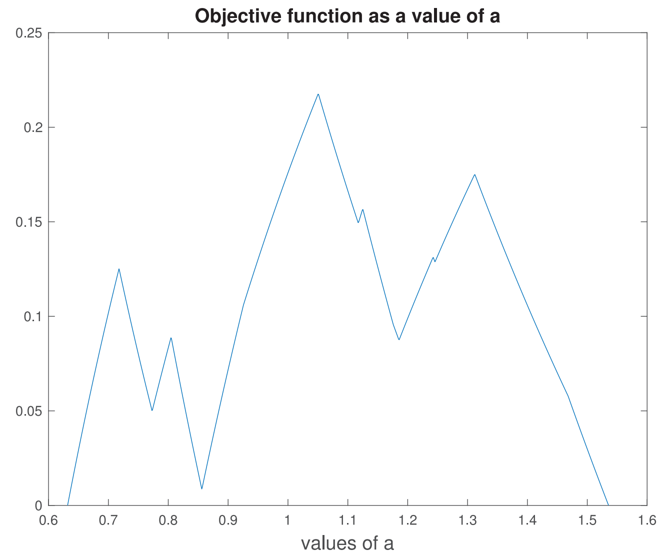

Fig. 3. Typical shape of the objective function as a function of a over a grid of equally spaced $M = 1001$ points, between a_{min} and a_{max} defined in (3.11).

3.2.2. General cases

In this case, when $d = (-d_x, d_y)$, we cannot avoid the (one dimensional) numerical optimization problem coming from (2.4). By plugging $\widehat{\Psi}_{FDH}$ in place of Ψ we obtain the FDH estimator for this general case

$$\widehat{\delta}_C(x_0, y_0) = \max_{a>0}\left\{ \frac{1}{a} \max_{i=1,\dots,n}\left\{ \min_{\substack{j=1,\dots,p\\ k=1,\dots,q}} [a x^\star_{0,j} - X^\star_{i,j}, Y^\star_{i,k} - a y^\star_{0,k}]\right\}\right\}. \quad (3.10)$$

This one-dimensional optimization problem (in a) can be easily solved, but care should be taken to avoid the multiple local maxima in the objective function to maximize, due to the "staircase" shape of the frontier of $\widehat{\Psi}_{FDH}$. One way to overcome this difficulty is to compute the objective function in a on a fine grid of values of a, $a = a_1, \dots, a_M$, where M could be very large (say, $M = 1000$ or 2000, the objective function is very fast to compute) in the interval $[a_{min}, a_{max}]$. Fig. 3 displays a typical graph of the objective function, as a function of a in one of the simulated examples used below in Section 5, with $n = 100$ and $p = q = 2$.

It is indeed easy to show, by using the developments in Section 2.1, that if $(x_0, y_0) \in \widehat{\Psi}_{FDH}$, the optimal value for a must satisfy (2.5) where Ψ is replaced by $\widehat{\Psi}_{FDH}$. So we have

$$a \geq a_{min} = \min_{i=1,\dots,n}\left\{ \max_{j=1,\dots,p} \frac{X_{i,j}}{x_{0,j}} \right\} \text{ and } a \leq a_{max} = \max_{i=1,\dots,n}\left\{ \min_{k=1,\dots,q} \frac{Y_{i,k}}{y_{0,k}} \right\}. \quad (3.11)$$

Then we detect the optimal value of a in this grid, say a_k, and then we can search for the maximum of the objective function using any numerical optimizer, searching for the maximum of a univariate function with starting value a_k. Note that due to the non-smooth nature of the objective function, a simple robust algorithm should be used, e.g. the Nelder–Mead simplex method which does not use derivatives. It should be noticed that if $(x_0, y_0) \notin \widehat{\Psi}_{FDH}$, the bounds in (2.5) are not valid, in this case we only know that $a \in (0, \infty)$.

3.3. Statistical properties

The statistical properties of directional distances are well known. According to Simar and Vanhems (2012) and Simar et al. (2012), we know that the statistical properties of the radial oriented measures can be extended to the directional distances for any directional vector d. By Kneip et al. (2023), these properties have been extended to FDH conical estimators. To summarize, we have for the CFDH estimators of $\delta_C(x_0, y_0)$ and for any direction vector d, as $n \to \infty$

$$n^\kappa\left(\delta_C(x_0, y_0) - \widehat{\delta}_C(x_0, y_0)\right) \xrightarrow{\mathcal{L}} \text{Weibull}(\eta_{x_0,y_0}), \quad (3.12)$$

where η_{x_0,y_0} is an unknown constant depending on the Data Generating Process (DGP). This result is sufficient to provide a valid inference about $\delta_C(x_0, y_0)$ for a given unit (x_0, y_0) using the subsampling methods described by Simar and Wilson (2011). As shown in Kneip et al. (2023) the rate of convergence κ depends on the returns to scale assumed on Ψ: under the CRS assumption ($\Psi = C(\Psi)$) we have $\kappa = 1/(p+q-1)$, otherwise, $\kappa = 1/(p+q-0.5)$. Note the difference with the rate of $\widehat{\delta}(x_0, y_0)$, the traditional FDH estimators of $\delta(x_0, y_0)$, the directional distance to the boundary of Ψ, where, as shown by Simar and Vanhems (2012), $\kappa = 1/(p+q)$, under CRS or not. As explained in Simar and Wilson (2011) using the correct rate is crucial for using subsampling techniques, and this rate depends on the chosen estimator and the assumptions on the shape of Ψ. We will use these results in the next section to provide confidence intervals on the productivity indices and their decompositions.

4. Luenberger productivity indices

4.1. Definition

As explained in the literature (see Färe et al. (1992, 1985, 2008, 1994, 1998)), the correct definition of the MPI and the LPI requires the use of measures of the distance to the boundary of $C(\Psi)$, the cone spanned by Ψ, which is identical to Ψ if and only if we have CRS (an often assumed hypothesis in this literature, but often rejected by the appropriate test, see e.g. Kneip et al. (2023)). Denote the attainable set at time t by

$$\Psi^t = \{(x, y) \mid x \text{ can produce } y \text{ at time } t\}. \quad (4.1)$$

Consider now the two periods t_1 and t_2. The "input" oriented MPI for a DMU moving from (x^1, y^1) in period 1 to (x^2, y^2) in period 2 is defined as

$$\mathcal{M}^{\text{inp}} = \left(\frac{\theta(x^2, y^2 \mid C(\Psi^1))}{\theta(x^1, y^1 \mid C(\Psi^1))} \times \frac{\theta(x^2, y^2 \mid C(\Psi^2))}{\theta(x^1, y^1 \mid C(\Psi^2))} \right)^{1/2}. \quad (4.2)$$

The boundaries of $C(\Psi^1)$ and of $C(\Psi^2)$ serve as benchmarks against which changes in productivity are measured and the MPI is the geometric mean of these changes. An "output" version is obtained by using rather $\lambda(x^s, y^s \mid C(\Psi^t))$ in place of the $\theta(x^s, y^s \mid C(\Psi^t))$, for $s, t = 1, 2$. Note that Kneip et al. (2021, 2023)) use also hyperbolic graph measures of efficiency. $\mathcal{M}^{\text{inp}} > (= \text{ or } <)\ 1$ when the productivity increases (remains unchanged or decreases) for this DMU between the two periods.

The LPI for the same DMU may be defined as follows (using the notation introduced in (1.7))

$$\mathcal{L} = \frac{1}{2}\left\{ \left[\delta(x^1, y^1 | C(\Psi^1)) - \delta(x^2, y^2 | C(\Psi^1))\right] \right.$$
$$\left. + \left[\delta(x^1, y^1 | C(\Psi^2)) - \delta(x^2, y^2 | C(\Psi^2))\right] \right\}, \quad (4.3)$$

with a similar interpretation of the MPI. In particular $\mathcal{L} > (= \text{ or } <)\ 0$ indicates the growth (unchanged or declining) of productivity for this DMU between the two periods. It is important to note that in the literature, the assumption of CRS is often made to define this index (that is, $C(\Psi^t) = \Psi^t$ in both periods, $t = 1, 2$). It is important to point out that if the assumption of CRS is true we have better rates of convergence, and hence better estimation, for all the indices, as explained above in Section 3.3. On the contrary, if the assumption of CRS is wrong, we cannot use the CRS estimator of the production set because it is inconsistent. But in our approach, we do not need this assumption, nor the convexity assumption.

4.2. Decompositions

Several decompositions of the MPIs have been proposed in the literature to investigate the sources of the productivity changes (see e.g. Simar and Wilson (2019, 2023) and the references therein). As shown e.g. in Epure et al. (2011), similar decompositions can be made for the LPIs. In our general formulation here, in (4.3), which does not assume CRS, we suggest a decomposition similar to the one proposed for MPIs by Simar and Wilson (2023). By simple arithmetic, it is easy to show that we can decompose $\mathcal{L}$ in an additive way as follows

$$\mathcal{L} = \mathcal{E} + \mathcal{T} + \mathcal{S}_1 + \mathcal{S}_2, \tag{4.4}$$

where the elements are defined as follows. We have:

$$\mathcal{E} = \delta((x^1, y^1)|\Psi^1) - \delta((x^2, y^2)|\Psi^2), \tag{4.5}$$

$$\mathcal{T} = \frac{1}{2}\left\{\left[\delta((x^1, y^1)|\Psi^2) - \delta((x^1, y^1)|\Psi^1)\right] + \left[\delta((x^2, y^2)|\Psi^2) - \delta((x^2, y^2)|\Psi^1)\right]\right\}, \tag{4.6}$$

$$\mathcal{S}_1 = \left[\delta((x^1, y^1)|C(\Psi^1)) - \delta((x^1, y^1)|\Psi^1)\right] - \left[\delta((x^2, y^2)|C(\Psi^2)) - \delta((x^2, y^2)|\Psi^2)\right], \tag{4.7}$$

$$\mathcal{S}_2 = \frac{1}{2}\left\{\left[\left(\delta((x^1, y^1)|\Psi^1) - \delta((x^1, y^1)|C(\Psi^1))\right) - \left(\delta((x^1, y^1)|\Psi^2) - \delta((x^1, y^1)|C(\Psi^2))\right)\right] + \left[\left(\delta((x^2, y^2)|\Psi^1) - \delta((x^2, y^2)|C(\Psi^1))\right) - \left(\delta((x^2, y^2)|\Psi^2) - \delta((x^2, y^2)|C(\Psi^2))\right)\right]\right\}. \tag{4.8}$$

The interpretation of these elements is the same as in Simar and Wilson (2019, 2023)) but measured in terms of directional distances. For instance, $\mathcal{E}$ measures a *change in efficiency*, $\mathcal{T}$ is the mean of two terms, each measuring the *shift of the frontier* of Ψ^t between the two periods. As explained in Simar and Wilson (2019, 2023)), $\mathcal{S}_1$ measures the *change in scale efficiency* of the DMU between the two periods under consideration. Specifically, it quantifies the difference in the scale efficiency of the DMU when comparing its performance relative to the production possibility frontier (or the attainable set) in both periods. Scale efficiency refers to how effectively a DMU is operating relative to an optimal production scale. If the unit is operating at increasing or decreasing returns to scale, $\mathcal{S}_1$ captures how this scale efficiency changes over time. $\mathcal{S}_2$ may be viewed as a *residual* that makes the equality of the right term in (4.4) to the value of $\mathcal{L}$ defined in (4.3). $\mathcal{S}_2$ can also be viewed as capturing the *change in the scale of technology* across the two periods. In a way, $\mathcal{S}_2$ acts as a term that accounts for differences between the actual production possibility frontier and a convexified or optimal version of that frontier. These components ($\mathcal{S}_1$ and $\mathcal{S}_2$) together help to provide a deeper understanding of how both efficiency and scale affect the productivity of a DMU across different time periods, without relying on assumptions of convexity or constant returns to scale.

We remark that we use the true possibility sets Ψ^t as reference sets for defining $\mathcal{E}$ and $\mathcal{T}$ as we should. The scale terms $\mathcal{S}_1$ and $\mathcal{S}_2$ provide measures of the differences between Ψ^t and $C(\Psi^t)$. Note again that this does not require convexity of Ψ^t nor the CRS assumptions, but clearly under CRS on both periods, $\mathcal{S}_1 = \mathcal{S}_2 = 0$.

4.3. Estimation and inference

We consider the data over two time periods: $\{(X_i^1, Y_i^1)\}_{i=1}^{n_1}$ and $\{(X_i^2, Y_i^2)\}_{i=1}^{n_2}$. Estimation and inference about MPIs and their components have been analyzed by Kneip et al. (2023) and Simar and Wilson (2019). We can follow the same route and use the appropriate FDH and CFDH estimators with their properties described in Section 3 to make inferences about the LPIs. These estimators are obtained by plugging the FDH estimator $\widehat{\Psi}^t_{n_t}$ in the place of Ψ^t for $t = 1, 2$ in all the expressions above, providing $\widehat{\mathcal{L}}, \widehat{\mathcal{E}}, \widehat{\mathcal{T}}, \widehat{\mathcal{S}}_1$ and $\widehat{\mathcal{S}}_2$.

To simplify the notations suppose that $n_1 = n_2 = n$, we can estimate from the full sample $\mathcal{X}_n = \{(X_i^1, Y_i^1, X_i^2, Y_i^2)\}_{i=1}^n$ the n individual LPIs, for $i = 1, \ldots, n$

$$\widehat{\mathcal{L}}_i = \frac{1}{2}\left\{\left[\delta(X_i^1, Y_i^1|C(\widehat{\Psi}_n^1)) - \delta(X_i^2, Y_i^2|C(\widehat{\Psi}_n^1))\right] + \left[\delta(X_i^1, Y_i^1|C(\widehat{\Psi}_n^2)) - \delta(X_i^2, Y_i^2|C(\widehat{\Psi}_n^2))\right]\right\}, \tag{4.9}$$

where $\delta(X_i^s, Y_i^s|C(\widehat{\Psi}_n^t))$ are the estimators $\widehat{\delta}_C(\cdot)$ defined in Section 3 with $s, t = 1, 2$. We describe in Appendix how to make inferences on the value of $\mathcal{L}$ (and any of its components in (4.4)) for an individual DMU.

If overall measures of productivity change and their components are of interest, researchers focus their attention on the geometric mean of MPIs. Here with the more general directional distance measures, we focus more simply, due to the additive nature of the indices, on arithmetic means of the individual indices computed for all DMUs in the sample. For instance, for the LPI we may want to make an inference on $\mu_{\mathcal{L}} = \mathbb{E}(\mathcal{L})$ where the expectation is over the whole population of DMUs. Following the arguments of Kneip et al. (2021, 2023)), it can be shown that

$$\overline{\mathcal{L}}_n = n^{-1} \sum_{i=1}^{n} \widehat{\mathcal{L}}_i \tag{4.10}$$

can be used to make inferences about $\mu_{\mathcal{L}}$, by using an appropriate Central Limit Theorem (CLT). As shown in Appendix, the adaptation to LPIs, $\mathcal{L}$ and its components appearing in (4.4) is straightforward and even easier due to the additive form of the LPIs. In particular, care should be taken to correct the inherent bias of the FDH and CFDH estimators (by using Jackknife methods) and, depending on the value of κ, to compute the mean in (4.10) of the estimators over a subsample of observations. We present in Appendix, all the practical details to make inferences on $\mu_{\mathcal{L}}$ and for the mean of any of its components in (4.4). The procedure is illustrated with a real data example in Section 5.2.

5. Illustratives examples

5.1. A simulated example

We first illustrate the various radial and non-radial estimators following Model II of Park and Weiner (2000) which defines a non-convex Ψ with 2 inputs and 2 outputs. This DGP denoted hereafter as PSW, was also used in Jeong and Simar (2006) and can be defined as follows. We define the function

$$g(x_1, x_2) = 2^{-1.4} x_1^{1.8} x_2^{0.6} \tag{5.1}$$

and select the points of the frontier $(\tilde{y}_1, \tilde{y}_2)$ such that

$$\tilde{y}_1 \tilde{y}_2 = g(x_1, x_2). \tag{5.2}$$

This defined a non-convex Ψ. Now the inputs are simulated according to $X_{i,j} \sim \text{Unif}(1, 2)$ independently for $j = 1, 2$. Then we generate $Y_{i,1}^* \sim \text{Unif}(0.2, 5)$ and define $Y_{i,2}^* = 1/Y_{i,1}^*$, hence the generated frontier points are given by

$$\widetilde{Y}_{i,k} = \sqrt{g(X_{i,1}, X_{i,2})} Y_{i,k}^*, \text{ for } k = 1, 2. \tag{5.3}$$

The points $(X_{i,1}, X_{i,2}, \widetilde{Y}_{i,2}, \widetilde{Y}_{i,2})$ are uniformly distributed on the frontier. Finally, as in Park and Weiner (2000), we generate inefficiencies through radial output inefficiencies producing the simulated data points $(X_{i,1}, X_{i,2}, Y_{i,2}, Y_{i,2})$, where

$$Y_{i,k} = \widetilde{Y}_{i,k} \exp(-\xi_i), \text{ for } k = 1, 2, \tag{5.4}$$

with $\xi_i \sim \text{Exp}(3)$, i.e. an exponential with mean 1/3. In the example below we select $n = 200$ and present some results for the first 5 simulated units.

In this analysis, we begin by using individual radial distance vectors, specifically $d_x = x$ and $d_y = y$, in Table 1. We calculate three cases: pure input case, pure output case, and a more general case that optimizes in both orientations (inputs and outputs) simultaneously. We

Table 1
PSW example: point estimates of proportional directional distances.

Unit	δ_i^{inp}	$\delta_{C,i}^{inp}$	δ_i^{out}	$\delta_{C,i}^{out}$	δ_i^{both}	$\delta_{C,i}^{both}$
1	0.0042	0.0647	0.0647	0.0692	0.0042	0.0335
2	0.1934	0.3087	0.4444	0.4466	0.0521	0.1825
3	0.2612	0.6070	1.1630	1.5447	0.2612	0.4358
4	0.0000	0.0660	0.0000	0.0707	0.0000	0.0341
5	0.0000	0.1991	0.0000	0.2487	0.0000	0.1106

Table 2
PSW example: point estimates of directional distances with common direction (median of the sample).

Unit	δ_i^{inp}	$\delta_{C,i}^{inp}$	δ_i^{out}	$\delta_{C,i}^{out}$	δ_i^{both}	$\delta_{C,i}^{both}$
1	0.0042	0.0650	0.0558	0.1338	0.0042	0.0439
2	0.2452	0.3717	0.3737	0.4051	0.0320	0.2244
3	0.2588	0.6443	0.5886	0.8912	0.2473	0.3987
4	0.0000	0.0723	0.0000	0.1017	0.0000	0.0423
5	0.0000	0.1487	0.0000	0.1840	0.0000	0.0822

Table 3
PSW example with $n = 200$: bootstrap 95% Confidence Intervals (CIs), left proportional and right common (median) directions. LB stands for Lower Bound and UB for Upper Bound of the intervals.

Unit	$\delta_{C,i}^{prop}$	LB	UB	$\delta_{C,i}^{median}$	LB	UB
1	0.0335	0.0335	0.0404	0.0439	0.0439	0.0653
2	0.1825	0.1825	0.2541	0.2244	0.2244	0.3640
3	0.4358	0.4358	0.4696	0.3987	0.3987	0.4653
4	0.0341	0.0341	0.0916	0.0423	0.0423	0.1172
5	0.1106	0.1106	0.1626	0.0822	0.0822	0.1227

use $d = (-d_x, 0_2)$ to measure the proportional reduction of inputs to reach the boundaries in the pure input case, $d = (0_p, d_y)$ to measure the proportional increase of outputs in the pure output case, and $d = (-d_x, d_y)$ to optimize in both orientations. To compare the estimates of the distance of each unit (X_i, Y_i) to the attainable set Ψ and $\mathcal{C}(\Psi)$, the cone spanned by Ψ, we present the estimates of both $\delta(X_i, Y_i)$ and $\delta_C(X_i, Y_i)$ for each case in Table 1. We observe that the differences between these estimates can be substantial in several cases, indicating that we are far from a CRS case (as can be seen from (5.1)–(5.2)). We also observe that the chosen orientation plays an important role in measuring efficiencies, which is consistent with our expectations.

Table 2 displays the results for the same sample and the same units but by selecting a common distance vector for all the units. We have chosen $d_x = \text{Median}(X)$ and $d_y = \text{Median}(Y)$. As for Table 1, in Table 2 we present the pure input, the pure output, and the case where we use both orientations simultaneously. Again we display the distances to Ψ and $\mathcal{C}(\Psi)$. We observe a similar general pattern as in Table 1. Note however that except for observations lying on the frontiers, the measures of inefficiency depend on the chosen direction vector. This is a well-known feature of directional distances (see e.g. the discussion in Remark 2.1 in Daraio and Simar (2016)).

To appreciate the quality of the point estimates with only $n = 200$ in this $p + q = 4$ dimensional space, we compute for the same sample and the same units the 95% confidence intervals of the individual distances to the cone $\mathcal{C}(\Psi)$, by the subsampling techniques described in Simar and Wilson (2011). We use $B = 2000$ bootstrap replications. The results are displayed in Table 3. In some cases, these intervals are rather wide indicating that $n = 200$ is not large enough with this dimension: the rate for CFDH is $n^{1/(p+q-0.5)} = n^{2/7}$, far below the parametric rate $n^{1/2}$. In Table 3, the column headed $\delta_{C,i}^{prop}$ is thus the same as the column headed $\delta_{C,i}^{both}$ in Table 1 and the column headed $\delta_{C,i}^{median}$ is the same as the column headed $\delta_{C,i}^{both}$ in Table 2.

Finally, to illustrate the role of the sample size, we redo the same exercise with $n = 800$. Since the units in Table 4 are not the same as the one selected for Table 3, we cannot make a comparison unit by unit. However we observe, in general, narrower confidence intervals (by a factor on the average equal to $(800/200)^{2/7} = 1.49$), illustrating the role of n in the asymptotic behavior of the CFDH estimators. The average of the width of the 5 CIs of Table 3, for $n = 200$ is 0.0443 and 0.0686 (for the proportional and common directional cases, respectively); these values over the 5 (different) units in Table 4, for $n = 800$ are 0.0399 and 0.0330, respectively.

Table 4
PSW example with $n = 800$: bootstrap 95% CIs, left proportional and right common (median) directions. LB stands for Lower Bound and UB for Upper Bound of the intervals.

Unit	δ_i^{prop}	LB	UB	δ_i^{median}	LB	UB
1	0.2095	0.2095	0.2343	0.2063	0.2063	0.2449
2	0.2530	0.2530	0.3065	0.2034	0.2034	0.2211
3	0.0000	0.0000	0.0525	−0.0000	−0.0000	0.0829
4	0.0382	0.0382	0.0590	0.0343	0.0343	0.0447
5	0.3272	0.3272	0.3751	0.1874	0.1874	0.2027

5.2. A real data illustration of LPIs

We illustrate how inference can be conducted on the means of the LPIs and all its components in the population. We use the same real data set used by Simar and Wilson (2019) (hereafter SW2019) where they impose convexity and focus on the MPIs.

The data come from Färe et al. (1992) who examine productivity changes among $n = 42$ Swedish pharmacies over the period 1980–1989. As described in SW2019, the original inputs are (i) labor input for pharmacists; (ii) labor input for technical staff; (iii) building services; and (iv) equipment services. The original outputs are (i) drug deliveries to hospitals; (ii) prescription drugs for outpatient care; (iii) medical appliances for the handicapped; and (iv) over the counter goods. See Färe et al. (1992) for further details. We are grateful to the authors for making the data available. As noticed by SW2019, there are only $n = 42$ observations for each year, so there is no hope of making reasonable inferences when working with $p + q = 8$ dimensions. This is a case in which we have the so called "curse of dimensionality". As explained in e.g. Daraio and Simar (2007), p. 26, footnote 10) the curse of dimensionality, shared by many nonparametric methods, means that to avoid large variances and wide confidence interval estimates a large quantity of data is needed. To appreciate the curse of dimensionality, Wilson (2018) introduces the "Effective parametric sample size" m, which for the FDH case is $m = [n^{2\kappa}]$ with $\kappa = 1/(p + q - 0.5)$. In our case, $n = 42$ and $p + q = 8$, $m \approx 3$. Hence, keeping all the variables in the nonparametric analysis, would be similar to providing inference in a parametric model (e.g. in a regression analysis) with only 3 observations. It is necessary then to reduce the dimension and this is possible in our example due to the correlations among the inputs and among some of the outputs. Hopefully, the inputs are highly correlated and the 3 last outputs also, so using the method described in detail in Chapter 6 of Daraio and Simar (2007), we can reduce the dimensions to one input factor (sharing 95.4% of the total inertia of the moment matrix $X'X$) and two outputs, i.e. the first original output and one output factor for the 3 other outputs (sharing 96.5% of the total inertia of the corresponding moment matrix). Hence, by doing so we do not lose so much information and as illustrated by Wilson (2018), we can gain a lot of precision in the statistical inference. SW2019 use the same dimension reduction, so we end up with $p = 1$ input and $q = 2$ outputs.

We compute for each of the 9 pairs of years from 1980 to 1989 the mean of the LPIs over the 42 pharmacies and we use the CLT described in Appendix to check which changes are significant at various levels. The results are displayed in Table 5 where we add a final row to indicate the changes in the various indices between the year 1980 and the year 1989.

We can compare with the results displayed in Table 2 in SW2019 where they use MPIs, with the following analogues: our $\mathcal{L}$ is the

Table 5
Productivity Change and its components for Swedish pharmacies, 1980–1989 (p = 1 and q = 2). Each entry is the mean of the changes in the indices in the pharmacies n = 42 between the two periods described in the first column.

Period	$\hat{\mathcal{L}}$	$\hat{\mathcal{E}}$	$\hat{\mathcal{T}}$	$\hat{\mathcal{S}}_1$	$\hat{\mathcal{S}}_2$
1980–1981	0.0027	0.0201***	−0.0068***	−0.0004	−0.0102
1981–1982	0.0304	−0.0612***	0.1004***	−0.0337***	0.0250***
1982–1983	0.0100	0.0620***	−0.0471***	0.0192	−0.0242**
1983–1984	−0.0282	−0.0081*	−0.0292	−0.0009*	0.0100
1984–1985	0.0159	0.0036	0.0148	0.0195***	−0.0219***
1985–1986	−0.0011	−0.0088**	0.0155***	−0.0086***	0.0009
1986–1987	0.0280***	0.0017	0.0287***	0.0085***	−0.0108***
1987–1988	0.0155***	0.0002	0.0204***	−0.0002	−0.0049
1988–1989	0.0156	0.0025	0.0034	−0.0001	0.0099
1980–1989	0.0974***	0.0118	0.0873***	0.0032	−0.0049

* Statistical significant difference from 0 at level .1.
** Statistical significant difference from 0 at level .05.
*** Statistical significant difference from 0 at level .01.

analogue of $\mathcal{M}$, our $\mathcal{E}$ is the analogue of $\mathcal{E}_2$, our $\mathcal{T}$ is the analogue of $\mathcal{T}_2$, our $\mathcal{S}_1$ is the analogue of $\mathcal{S}_1$ and our $\mathcal{S}_2$ is the analogue of $\mathcal{S}_3$. SW2019 impose convexity of all the attainable sets and so, use DEA estimators of hyperbolic distances to Ψ^t and to $\mathcal{C}(\Psi^t)$. Our approach has a less restrictive setup (without imposing convexity and using proportional directional distances) and we observe some differences between the two global pictures.

We can share some major qualitative conclusions as those of SW2019. We observe a global gain in productivity from 1980 to 1989, but mainly due to gains in technologies and not efficiencies. However, we find some notable differences. We do not observe so many significant productivity changes in one year during the 10 years (we have only 2 significant changes for $\mathcal{L}$ over the 9 pairs, where SW2019 have significant changes for almost all the years), but globally, at the end a highly significant change from 1980 to 1989, which agrees with SW2019. The main reason for these changes is indeed due to the technological progress during these years and it seems that in some years, the significant gain of the technology was even waved out by a significant loss of efficiency (see e.g. the periods 1981–1982 or 1985–1986). The opposite happens in the period 1980–1981. Looking at Table 5 in detail reveals some substantial differences with the approach of SW2019, e.g. they observe a global significant gain of scale efficiency over the period 1980–1989, which is not our case. So, this may question the validity of imposing the convexity of the attainable sets that modify significantly the results. Of course, due to the small number of DMUs ($n = 42$) with a dimension $p + q = 3$, we should remain careful (the rate of convergence is $n^{1/3}$ for the FDH estimators, far below the parametric rate $n^{1/2}$) and avoid definite conclusions from our exercise. Our objective was mainly to illustrate our approach and show that by being less restrictive than the existing ones, we may obtain different results and this may be of interest.

6. Conclusions

Directional distances provide a flexible way to measure the distance from a DMU to the efficient frontier of the attainable set in the input–output space, Ψ, in the case of general technologies, i.e. without imposing convexity of Ψ. In this paper, we describe how to define and characterize these distances to the boundary of $\mathcal{C}(\Psi)$, the cone spanned by Ψ and we derive nonparametric estimators with their statistical properties. This allows us to define and estimate directional distances under the Constant Returns to Scale (CRS) assumption for general directional vectors and general technologies.

In addition, we can define and estimate Luenberger Productivity Indices (LPI) and their decompositions in these general situations to analyze the possible source of changes in productivity. We finally derive Central Limit Theorems (CLTs) allowing statistical inference on these indices making available the statistical tools to assess the significance and the confidence intervals on the LPI and their components.

All this represents a major step forward in the literature, introducing a general and flexible tool for estimating directional distances for general technologies that do not impose convexity. In addition, the application of this tool allows us to estimate Luenberger productivity indices and their decompositions for general technologies. Finally, the development of statistical inference in this context makes available for the first time information on the statistical significance and confidence intervals of LPI and its decompositions.

This approach is very promising and can be used in a variety of application areas. To show how it works in practice, we present an application on simulated data generated based on a classical Data Generating Process from the literature. Finally, an application on real and well-known Swedish pharmacy data from the literature shows the added value of the proposed approach by providing for the first time estimates on the significance and confidence intervals of LPI and its components.

The main limitation of this approach is that it refers to deterministic frontiers. The main extensions to be addressed in further developments are the inclusion of noise in stochastic frontier models and the extension of this framework to take account of external or environmental variables that are neither inputs nor outputs of the production process but which may affect the performance of the units analyzed. All these are left to further work.

CRediT authorship contribution statement

Cinzia Daraio: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Simone Di Leo:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Léopold Simar:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Acknowledgments

The financial support of the Sapienza University of Rome, Italy, research award no. RM11916B8853C925 and no. RM12117A8A5DBD18, is gratefully acknowledged.

Appendix. Inference on LPIs

A.1. Inference on individual $\mathcal{L}$

Looking at the definition of the LPI for an individual DMU, in (4.3), we see that introducing the notation $z^t = (x^t, y^t)$ for $t = 1, 2$, we could rewrite it as $\mathcal{L}(z^1, z^2 \mid \mathcal{C}(\Psi^1), \mathcal{C}(\Psi^2))$, to explicit the ingredients needed to define the LPI for a DMU moving from z^1 to z^2 between periods 1 and 2. The full sample of available observations (we assume $n = n_1 = n_2$) can be written $\mathcal{X}_n = \{(Z_i^1, Z_i^2)\}_{i=1}^n$. For a particular DMU, the CFDH estimator provides

$$\hat{\mathcal{L}}(z^1, z^2 \mid \mathcal{C}(\hat{\Psi}_n^1), \mathcal{C}(\hat{\Psi}_n^2)) = \frac{1}{2}\Big\{\big[\delta(z^1|\mathcal{C}(\hat{\Psi}_n^1)) - \delta(z^2|\mathcal{C}(\hat{\Psi}_n^1))\big] + \big[\delta(z^1|\mathcal{C}(\hat{\Psi}_n^2)) - \delta(z^2|\mathcal{C}(\hat{\Psi}_n^2))\big]\Big\}, \quad (A.1)$$

where for $t = 1, 2$, $\hat{\Psi}_n^t$ is the FDH estimator of Ψ^t computed with the sample of period t taken form $\mathcal{X}_n$, i.e., $\{(Z_i^t)\}_{i=1}^n$.

By applying the theory described in Kneip et al. (2021, 2023)) and in Simar and Wilson (2019) we can derive the limiting distribution of $\hat{\mathcal{L}}$ for a given DMU. Under mild regularity assumptions, and as a direct consequence of (3.12), we have, as $n \to \infty$,

$$n^{\kappa}\Big(\hat{\mathcal{L}}(z^1, z^2 \mid \mathcal{C}(\hat{\Psi}_n^1), \mathcal{C}(\hat{\Psi}_n^2)) - \mathcal{L}(z^1, z^2 \mid \mathcal{C}(\Psi^1), \mathcal{C}(\Psi^2))\Big) \xrightarrow{\mathcal{L}} \mathrm{Q}_{\mathcal{L}, z^1, z^2}, \quad (A.2)$$

where $\mathrm{Q}_{\mathcal{L}, z^1, z^2}$ is a non-degenerate distribution and κ was defined in (3.12). This result is sufficient to enable valid inference about $\mathcal{L}(z^1, z^2 \mid \mathcal{C}(\Psi^1), \mathcal{C}(\Psi^2))$ for a single DMU, using the subsampling methods described in Simar and Wilson (2011).

A.2. Inference on $\mu_{\mathcal{L}} = \mathbb{E}(\mathcal{L})$

We summarize and particularize here for the LPIs, the developments described in Simar and Wilson (2019) for MPIs. Given the sample $\mathcal{X}_n = \{(Z_i^1, Z_i^2)\}_{i=1}^n$, one may obtain n estimates $\widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_n^1), \mathcal{C}(\widehat{\Psi}_n^2))$, $i = 1, \dots, n$. We have

$$\mu_{\mathcal{L}} = \mathbb{E}\left(\mathcal{L}(Z_i^1, Z_i^2 \mid \mathcal{C}(\Psi^1), \mathcal{C}(\Psi^2))\right), \tag{A.3}$$

and a natural estimator is given by

$$\widehat{\mu}_{\mathcal{L},n} = n^{-1}\sum_{i=1}^n \widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_n^1), \mathcal{C}(\widehat{\Psi}_n^2)). \tag{A.4}$$

Due to the inherent bias of FDH estimators, usual Lindeberg–Levy CLT cannot be used, without correcting for the bias (see Kneip et al. (2015)). In fact, under regularity assumptions described e.g. in Simar and Wilson (2019), we have the following CLT

$$n^{1/2}\left(\widehat{\mu}_{\mathcal{L},n} - \mu_{\mathcal{L}} - C_{\mathcal{L}}n^{-\kappa} + \xi_{n,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_{\mathcal{L}}^2), \tag{A.5}$$

where $\sigma_{\mathcal{L}}^2 = \mathbb{V}(\mathcal{L})$ and the remainder $\xi_{n,\kappa} = o(n^{-\kappa})$. The bias term $C_{\mathcal{L}}n^{-\kappa}$ can be eliminated by adapting a generalized Jackknife method that can be summarized, in our setup, as follows. We randomly split $\mathcal{X}_n = \{(Z_i^1, Z_i^2)\}_{i=1}^n$ into two parts $\mathcal{X}_{m_1}^{(1)}$ and $\mathcal{X}_{m_2}^{(2)}$, with $m_1 = \lfloor n/2 \rfloor$ and $m_2 = n - m_1$ (note that if n is even, $m_1 = m_2$). Then we define the mean over the two subsamples, for $\ell = 1, 2$

$$\widehat{\mu}_{\mathcal{L},m_\ell} = m_\ell^{-1} \sum_{\{i \mid (Z_i^1, Z_i^2) \in \mathcal{X}_{m_\ell}^{(\ell)}\}} \widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_{m_\ell}^1), \mathcal{C}(\widehat{\Psi}_{m_\ell}^2)), \tag{A.6}$$

where for $t = 1, 2$, $\widehat{\Psi}_{m_\ell}^t$ is the FDH estimator of Ψ^t obtained from the observations for period t in the sample $\mathcal{X}_{m_\ell}^{(\ell)}$. Now we set

$$\widetilde{\mu}_{\mathcal{L},n/2} = \frac{1}{2}(\widehat{\mu}_{\mathcal{L},m_1} + \widehat{\mu}_{\mathcal{L},m_2}). \tag{A.7}$$

Using a similar argument as Kneip et al. (2015) we can show that

$$\widehat{B}_{\mathcal{L},n,\kappa} = (2^\kappa - 1)^{-1}(\widetilde{\mu}_{\mathcal{L},n/2} - \widehat{\mu}_{\mathcal{L},n}) = C_{\mathcal{L}}n^{-\kappa} + \tilde{\xi}_{n,\kappa} + o_p(n^{-1/2}), \tag{A.8}$$

where the remainder $\tilde{\xi}_{n,\kappa} = o(n^{-\kappa})$.

To reduce the variance of the biased estimator $\widehat{B}_{\mathcal{L},n,\kappa}$, we can randomly split the original sample $\mathcal{X}_n$, a large number of times, say K. For each split, $k = 1, \dots, K$, we obtain an estimate $\widehat{B}_{\mathcal{L},n,\kappa,k}$. Then we use as bias estimate

$$\widehat{B}_{\mathcal{L},n,\kappa} = K^{-1}\sum_{k=1}^K \widehat{B}_{\mathcal{L},n,\kappa,k}. \tag{A.9}$$

In practice, K should be as large as say 100. In the CLT described in (A.5), we see that the factor $n^{1/2}$ can still be too large to neglect the remainder, even after the bias correction if $\kappa < 1/2$. This leads to the following Theorem (For the regularity conditions, see e.g. Theorems 4.6 and 4.7 in Kneip et al. (2023), or Theorem 4.4 in Simar and Wilson (2019)),

Theorem A.1. *Under mild regularity conditions, as $n \to \infty$,*

$$n^{1/2}\left(\widehat{\mu}_{\mathcal{L},n} - \mu_{\mathcal{L}} - \widehat{B}_{\mathcal{L},n,\kappa} + \xi_{n,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_{\mathcal{L}}^2), \tag{A.10}$$

provided $\kappa \geq 1/2$. If $\kappa < 1/2$,

$$n^{\kappa}\left(\widehat{\mu}_{\mathcal{L},n_\kappa} - \mu_{\mathcal{L}} - \widehat{B}_{\mathcal{L},n,\kappa} + \xi_{n,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_{\mathcal{L}}^2), \tag{A.11}$$

where $\widehat{\mu}_{\mathcal{L},n_\kappa} = n_\kappa^{-1}\sum_{i=1}^{n_\kappa} \widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_n^1), \mathcal{C}(\widehat{\Psi}_n^2))$, with $n_\kappa = \min(\lfloor n^{2\kappa} \rfloor, n)$.

In fact $\widehat{\mu}_{\mathcal{L},n_\kappa}$ is the average of a random subsample of size $n_\kappa \leq n$ drawn from the full sample $\widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_n^1), \mathcal{C}(\widehat{\Psi}_n^2))$, $i = 1, \dots, n$, where the notation explicit that they are computed with the full sample $\mathcal{X}_n$. Now it is shown in Kneip et al. (2023) that

$$\widehat{\sigma}_{\mathcal{L},n}^2 = n^{-1}\sum_{i=1}^n \left[\widehat{\mathcal{L}}(Z_i^1, Z_i^2 \mid \mathcal{C}(\widehat{\Psi}_n^1), \mathcal{C}(\widehat{\Psi}_n^2)) - \widehat{\mu}_{\mathcal{L},n}\right]^2 \xrightarrow{p} \sigma_{\mathcal{L}}^2. \tag{A.12}$$

The latter can be used to make practical inferences about $\mu_{\mathcal{L}}$ by using the quantiles of the standard normal distribution. For instance, if $\kappa \geq 1/2$, an asymptotically correct $(1-\alpha)$ confidence interval for $\mu_{\mathcal{L}}$ is obtained by

$$\left[\widehat{\mu}_{\mathcal{L},n} - \widehat{B}_{\mathcal{L},n,\kappa} \pm z_{1-\alpha/2}\frac{\widehat{\sigma}_{\mathcal{L},n}}{n^{1/2}}\right], \tag{A.13}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution. When $\kappa < 1/2$ we should rather use the interval

$$\left[\widehat{\mu}_{\mathcal{L},n_\kappa} - \widehat{B}_{\mathcal{L},n,\kappa} \pm z_{1-\alpha/2}\frac{\widehat{\sigma}_{\mathcal{L},n}}{n^{\kappa}}\right]. \tag{A.14}$$

Remark A.1. As explained in Simar and Wilson (2019) for the case of MPIs, the same reasoning can be applied to any component of the decomposition of the LPI described in (4.4). Note that for $\mathcal{L}$, the rate κ is the rate of the CFDH estimator (i.e. $1/(p+q-1)$ or $1/(p+q-0.5)$ according to we assume CRS or not), but for the other components, the rate κ is always governed by the FDH estimators, i.e. $1/(p+q)$.

References

Boussemart, J. P., Briec, W., Kerstens, K., & Poutineau, J. C. (2003). Luenberger and Malmquist productivity indices: Theoretical comparisons andempirical illustration. *Bulletin of Economic Research*, *55*(4), 391–405.

Caves, D. W., Christensen, L. R., & Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output and productivity. *Econometrica*, *50*, 436–505.

Chambers, R. G. (2002). Exact nonradial input, output, and productivity measurement. *Economic Theory*, *20*(4), 751–765.

Chambers, R. G., Chung, Y., & Färe, R. (1996). Benefit and distance functions. *Journal of Economic Theory*, *70*, 407–419.

Chambers, R. G., Chung, Y., & Färe, R. (1998). Profit, directional distance functions, and nerlovian efficiency. *Journal of Optimization Theory and Applications*, *98*, 351–364.

Daraio, C., & Simar, L. (2007). *Advanced robust and nonparametric methods in efficiency analysis: methodology and applications*. New-York: Springer.

Daraio, C., & Simar, L. (2016). Efficiency and benchmarking with directional distances: a data-driven approach. *Journal of the Operational Research Society*, *67*(7), 928–944.

Daraio, C., Simar, L., & Wilson, P. W. (2020). Fast and efficient computations of directional distances estimators. *Annals of Operations Research*, *288*, 805–835.

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

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor inefficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: concepts and measurements* (pp. 243–267). Amsterdam: North-Holland.

Epure, M., Kerstens, K., & Prior, D. (2011). Bank productivity and performance groups: A decomposition approach based upon the luenberger productivity indicator. *European Journal of Operational Research*, *211*, 630–641.

Färe, R., & Grosskopf, S. (2000). Theory and application of directional distance functions. *Journal of Productivity Analysis*, *13*(2), 93–103.

Färe, R., Grosskopf, S., Lindgren, B., & Roos, P. (1992). Productivity changes in Swedish pharmacies 1980–1989: A non-parametric Malmquistapproach. *Journal of Productivity Analysis*, *3*, 85–101.

Färe, R., Grosskopf, S., & Lovell, C. A. K. (1985). *The measurement of efficiency of production*. Boston: Kluwer-Nijhoff Publishing.

Färe, R., Grosskopf, S., & Margaritis, D. (2008). Efficiency and productivity: Malmquist and more. In Harold Fried, C. A. Knox Lovell, & Shelton Schmidt (Eds.), *The measurement of productive efficiency* (2nd ed.). Oxford University Press.

Färe, R., Grosskopf, S., & Margaritis, D. (2015). Directional distance functions revisited: Selective overview and update. *1*, (pp. 57–79).

Färe, R., Grosskopf, S., Norris, M., & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review*, *84*, 66–83.

Färe, R., Grosskopf, S., & Roos, P. (1998). Malmquist productivity indexes: a survey of theory and practice. In *Index numbers: essays in honour of sten Malmquist* (pp. 127–190). Dordrecht: Springer Netherlands.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society, A*, (120), 253–281.

Jeong, S. O., & Simar, L. (2006). Linearly interpolated FDH efficiency score for nonconvex frontiers. *Journal of Multivariate Analysis*, *97*, 2141–2161.

Kevork, I. S., Pange, J., Tzeremes, P., & Tzeremes, N. G. (2017). Estimating Malmquist productivity indexes using probabilistic directional distances: An application to the European banking sector. *European Journal of Operational Research*, *261*(3), 1125–1140.

Kneip, A., Simar, L., & Wilson, P. W. (2015). When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores. *Econometric Theory*, *31*, 394–422.

Kneip, A., Simar, L., & Wilson, P. W. (2021). Inference in dynamic, nonparametric models of production: Central limit theorems for Malmquist indices. *Econometric Theory*, *37*, 537–572.

Kneip, A., Simar, L., & Wilson, P. W. (2023). Conical FDH estimators for testing returns to scale and making inference about change of productivity. submitted manuscript, UCL.

Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estatistica y de Investigacion Operativa*, *4*, 209–242.

Park, L., & Weiner, Ch. (2000). The FDH estimator for productivity efficiency scores: Asymptotic properties. *Econometric Theory*, *16*, 855–877.

Pastor, J. T., Lovell, C. K., & Aparicio, J. (2020). Defining a new graph inefficiency measure for the proportional directional distance function and introducing a new Malmquist productivity index. *European Journal of Operational Research*, *281*(1), 222–230.

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

Shephard, R. W. (1970). *Theory of cost and production function*. Princeton, New-Jersey: Princeton University Press.

Simar, L., & Vanhems, A. (2012). Probabilistic characterization of directional distances and their robust versions. *Journal of Econometrics*, *166*, 342–354.

Simar, L., Vanhems, A., & Wilson, P. W. (2012). Statistical inference with DEA estimators of directional distances. *European Journal of Operational Research*, *220*, 853–864.

Simar, L., & Wilson, P. W. (2011). Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*, *36*, 33–53.

Simar, L., & Wilson, P. W. (2019). Central limit theorems and inference for sources of productivity change measured by nonparametric Malmquist indices. *European Journal of Operational Research*, *277*, 756–769.

Simar, L., & Wilson, P. W. (2023). Another look at productivity growth in industrialized countries. *Journal of Productivity Analysis*, *60*, 257–272.

Simar, L., & Zelenyuk, V. (2018). Central limit theorems for aggregate efficiency. *Operations Research*, *66*(1), 137–149.

Wilson, P. W. (2018). Dimension reduction in nonparametric models of production. *European Journal of Operational Research*, *267*, 349–367.