TSE working paper series

10-195

Research Group: *Econometrics and Statistics*

September 30, 2010

Probabilistic Characterization of Directional Distances and their Robust Versions

LEOPOLD SIMAR AND ANNE VANHEMS

Ecole d'économie de Toulouse (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31000 TOULOUSE
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.TSE-fr.eu** - contact@TSE-fr.eu

Probabilistic Characterization of Directional Distances and their Robust versions

Léopold Simar* Anne Vanhems§

September 30, 2010

Abstract

In productivity analysis, the performance of production units is measured through the distance of the individual decision making units (DMU) to the technology which is defined as the frontier of the production set. Most of the existing methods, Farrell-Debreu and Shephard radial measures (input or output oriented) and hyperbolic distance functions, rely on multiplicative measures of the distance and so require to deal with strictly positive inputs and outputs. This can be critical when the data contain zero or negative values as in financial data bases for the measure of funds performances. Directional distance function is an alternative that can be viewed as an additive measure of efficiency. We show in this paper that using a probabilistic formulation of the production process, the directional distance can be expressed as simple radial or hyperbolic distance up to a simple transformation of the inputs/outputs space. This allows to propose simple methods of estimation but also to transfer easily most of the known properties of the estimators shared by the radial and hyperbolic distances. In addition, the formulation allows to define robust directional distances in the lines of α-quantile or order-m partial frontiers. Finally we can also define conditional directional distance functions, conditional to environmental factors. To illustrate the methodology, we show how it can be implemented using a Mutual Funds database.

Keywords: Directional distance function, partial frontier, conditional measures of efficiency, robust nonparametric estimation, mutual funds performances.

JEL Classification: Primary C13; secondary C14

*Institute of Statistics, Université Catholique de Louvain-la-Neuve, Belgium. L. Simar acknowledges support from the "Interuniversity Attraction Pole", Phase VI (No. P6/03) of the Belgian Science Policy and from the INRA-GREMAQ, Toulouse, France. `leopold.simar@uclouvain.be`

§Corresponding author. Toulouse Business School and Toulouse School of Economics, France. `a.vanhems@esc-toulouse.fr`

1 Introduction

In productivity analysis, the performance of production units is measured through efficiency scores. These efficiency measures are typically given by the distance of the individual decision making units (DMU) to the technology which is defined as the frontier of the production set. In the classical setting of productivity analysis and technical efficiency study, we consider a set of p inputs and q outputs used in the production process. The production set is the set of technically feasible combinations of inputs and outputs. It is defined as

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

and its efficient frontier (the technology) is defined by

$$\Psi^\partial = \left\{(x, y) \in \Psi \mid \left(\gamma^{-1}x, \gamma y\right) \notin \Psi \text{ for all } \gamma > 1\right\}. \tag{1.2}$$

The Farrell-Debreu input distance (see Debreu, 1951, Farrell, 1957) for a firm operating at the level (x, y) is then determined by the *radial* distance from (x, y) to the efficient frontier. It indicates how much all input quantities can be proportionately reduced so that the output levels y can still be produced. Formally, the Farrell input distance for a firm at a point (x, y) is given by

$$\theta(x, y) = \inf\{\theta > 0 | (\theta x, y) \in \Psi\}. \tag{1.3}$$

In the same spirit, the Farrell output distance of (x, y)

$$\lambda(x, y) = \sup\{\lambda > 0 | (x, \lambda y) \in \Psi\}, \tag{1.4}$$

is the maximum feasible equiproportionate expansion of all outputs attainable with the input level x. Note that the Shephard (1970) distances functions are the reciprocals of the Farrell distance functions.

The hyperbolic graph efficiency measure, proposed by Färe et al. (1985), Färe and Grosskopf (2000), provides an alternative measure of the distance of a firm (x, y) to the efficient frontier. It is given by

$$\gamma(x, y) = \sup\{\gamma > 0 | (\gamma^{-1}x, \gamma y) \in \Psi\}. \tag{1.5}$$

A use of hyperbolic path allows to avoid some of the ambiguity in choosing output or input orientation. In this case, input and output levels are adjusted simultaneously.

All the above methods rely on multiplicative measures of the distance and so require to deal with strictly positive inputs and outputs. This can be critical when the data contain zero or negative values as in financial data bases with the measure of funds performances.

A natural idea suggested by several authors is to translate the data to avoid negative values, but as pointed e.g. by Lovell and Pastor (1995), if multiplicative efficiency measures satisfy the unit invariant property they are not invariant from an affine translation of inputs or outputs, and only additive models can satisfy the translation invariant property. As these authors note, some specific DEA estimator can satisfy translation invariance for inputs or outputs, but not for both. Such a restriction may strongly constraint the choice of inputs and outputs (see Gregoriou, Sedzro and Zhu 2005).

Recently directional distance functions have been introduced (see Chambers, Chung and Färe, 1998, Färe and Grosskopf, 2004, Färe et al., 2008) that generalize both input and output distance functions. The directional distance function projects the input-output vector (x, y) onto the technology frontier in a direction given by vector $(-g_x, g_y)$ where $g = (g_x, g_y) \in \mathbb{R}_+^{p+q}$. It is defined as

$$D(x, y; g_x, g_y) = \sup\{\beta > 0 | (x - \beta g_x, y + \beta g_y) \in \Psi\}. \tag{1.6}$$

It encompasses input or output oriented radial distance measures as special cases: it can indeed be seen that if $g = (x, 0)$ (resp. $g = (0, y)$), the input (resp. output) radial distance introduced above can be recovered. The choice of the direction vector g is to be done by the researcher the only restriction is that it has to be positive and in the units chosen for the inputs and the outputs. For a discussion about this choice see Färe et al. (2008), where several natural candidates are discussed. This direction can be different for each point to be evaluated but it can also be the same for all the points (x, y) which would be like assuming all the firms face the same prices. Also, as pointed by Färe et al. (2008) "assuming a common direction would be a kind of egalitarian evaluation reflecting some underlying social welfare function". As discussed below, our methodology will admit very flexible choices of this directional vector.

This directional distance function can indeed be viewed as an additive definition of efficiency since to reach the frontier we substract portions of g_x from the input x and add g_y to the outputs y. So, the distance can be defined with negative inputs or outputs. We can see that $D(x, y; g_x, g_y) \geq 0$ if and only if $(x, y) \in \Psi$ and a point on the frontier is characterized by $D(x, y; g_x, g_y) = 0$. It benefits from translation property and is also independent of unit of measurement (such as radial distances). The translation property can be written as

$$D(x - \eta g_x, y + \eta g_y; g_x, g_y) = D(x, y; g_x, g_y) - \eta, \forall \eta \in \mathbb{R}, \tag{1.7}$$

and the independence of unit of measurement as

$$D(\theta. * x, \lambda. * y; \theta. * g_x, \lambda. * g_y) = D(x, y; g_x, g_y), \quad \text{where } \theta \in \mathbb{R}_+^p, \text{ and } \lambda \in \mathbb{R}_+^q, \tag{1.8}$$

where .$*$ refers to componentwise multiplication of vectors. See Färe et al. (2008), and the references there, for more discussions and properties.

To the best of our knowledge, the statistical properties of directional distances have never been established, robust versions (robust to outliers and extreme data points) have not been proposed and conditional directional distances, where conditioning is on environmental factor that may influence the production process and introduce possible heterogeneity, have not been proposed in the literature, so far. This is the objective of this paper.

The paper is organized as follows. We show in Section 2, that using the probabilistic formulation of the production process introduced by Cazals et al. (2002) and Daraio and Simar (2005), the directional distances can be expressed as simple radial distances up to a simple transformation of the inputs/outputs space. This property provides very simple methods of estimation of the directional distance, using standard tools available for radial measures. We show also that this perspective allows to define conditional directional distance functions and their nonparametric estimators, where conditioning is on environmental factors that may influence the production process (Cazals et al. 2002 and Daraio and Simar, 2005).

With this approach, we define in Section 3, robust versions of the directional distance functions by adapting the α-quantile (Daouia and Simar, 2007) and order-m partial frontiers (Cazals et al., 2002) introduced for radial measures. We will also show in Section 4 that most of the known statistical properties of the estimators of the radial distances can be translated for the directional distances, including tools for making inference. Finally, we propose in Section 5 an empirical illustration on a database of Mutual Funds. Section 6 concludes.

2 Probabilistic Formulations

2.1 General definitions

A probabilistic formulation of the production process was first introduced by Cazals, Florens, Simar (2002) and then developed in Daraio and Simar (2005). Consider the joint probability measure of (X, Y) and the associated probability function $H_{XY}(\cdot, \cdot)$ defined as

$$H_{XY}(x, y) = \text{Prob}(X \leq x, Y \geq y), \tag{2.1}$$

and Ψ is the support of (X, Y). It has been shown that if Ψ is a nonempty set with inputs and outputs freely disposable, Ψ can be identified with the support of $H_{XY}(\cdot, \cdot)$:

$$\Psi = \{(x, y) \in \mathbb{R}^{p+q} |\ H_{XY}(x, y) > 0\}, \tag{2.2}$$

where we notice that we allows for negative inputs and outputs[1]. If only positive inputs and outputs are considered, we can define the radial (multiplicative) efficiency scores in terms of the support of this probability function. For all $(x, y) \in \Psi$ such that $S_Y(y) = \text{Prob}(Y \geq y) > 0$ we have

$$\theta(x, y) = \inf\{\theta > 0 | H_{XY}(\theta x, y) > 0\}, \tag{2.3}$$

and similarly, for all $(x, y) \in \Psi$ such that $F_X(x) = \text{Prob}(X \leq x) > 0$ we have

$$\lambda(x, y) = \sup\{\lambda > 0 | H_{XY}(x, \lambda y) > 0\}. \tag{2.4}$$

For hyperbolic distance function we have (see Wheelock and Wilson, 2008)

$$\gamma(x, y) = \sup\{\gamma > 0 | H_{XY}(\gamma^{-1} x, \gamma y) > 0\}. \tag{2.5}$$

We now define, accordingly, the probabilistic version of the directional distance (1.6):

$$D(x, y; g_x, g_y) = \sup\{\beta > 0 | H_{XY}(x - \beta g_x, y + \beta g_y) > 0\}. \tag{2.6}$$

We note that only the latter formulation is valid when positive and/or negative inputs and outputs are considered as possible. We will show below that by a simple change of variable, a directional distance function can be viewed as a particular input/output/hyperbolic distance function in a transformed dataset. We can then benefit from the nice properties of directional efficiencies combined with simple tractable radial distance to compute appropriate estimators having known statistical properties.

2.2 Link between directional and hyperbolic distances

In what follows, we assume that inputs and outputs can take either positive or negative values. Such an assumption is rather unusual in classical production settings but is sensible in finance context and measure of funds performances. In this case, one input can be the lower skewness and one output the return, both taking either positive or negative values. So we consider a vector of inputs $x \in \mathbb{R}^p$ and a vector of outputs $y \in \mathbb{R}^q$. In order to define the directional distance function, we consider strictly positive directional vector $g = (g_x, g_y) \in \mathbb{R}^{p+q}_{*+}$.[2] Remember that this directional vector must have the same unit as the (input, output)

[1] Free disposability of inputs and outputs is a minimal natural assumption we can made on Ψ. It says that if $(x, y) \in \Psi$, any $(\tilde{x}, \tilde{y})$ such that $\tilde{x} \geq x$ and $\tilde{y} \leq y$ is such that $(\tilde{x}, \tilde{y}) \in \Psi$. Here inequalities between vectors are componentwise. In a sense, free disposablity assumes that wasting ressources is technically possible, although certainly not optimal.

[2] The case where one of the two components, g_x or g_y is set to zero is addressed in Section 2.3

vector; we will suggest below in our empirical illustration, some natural choice for g. So we repeat here the definition of the directional distance

$$\begin{aligned} D(x,y;g_x,g_y) &= \sup\{\beta>0 | (x-\beta g_x, y+\beta g_y)\in\Psi\} \\ &= \sup\{\beta>0 | H_{XY}(x-\beta g_x, y+\beta g_y)>0\}, \end{aligned} \tag{2.7}$$

where the latter equality implicitly assumes the free disposability of Ψ.

Consider now the following monotonic (increasing) transformation of the units for the inputs and for the outputs:

$$x^* = \exp(x./g_x) \text{ and } y^* = \exp(y./g_y), \tag{2.8}$$

where ./ refers to componentwise division of vectors. We will denote Ψ^* the attainable set in this new coordinate system. We have

$$\Psi^* = \{(x^*,y^*)\in\mathbb{R}_+^{p+q} \mid x^*=\exp(x./g_x), y^*=\exp(y./g_y) \text{ for some } (x,y)\in\Psi\}. \tag{2.9}$$

We have the following property which is trivially verified due to the monotonic (increasing) property of the transformation[3].

Lemma 2.1. *If Ψ is free disposal then Ψ^* is free disposal.*

Due to this Lemma we can also write Ψ^* as

$$\Psi^* = \{(x^*,y^*)\in\mathbb{R}_+^{p+q} \mid H_{X^*Y^*}(x^*,y^*)>0\}, \tag{2.10}$$

where $H_{X^*Y^*}(\cdot,\cdot)$ is the transformation of the probability $H_{XY}(\cdot,\cdot)$ for the random variables $X^*=\exp(X./g_x)$ and $Y^*=\exp(Y./g_y)$. By elementary algebraic manipulations we find that for all $(x^*,y^*)\in\mathbb{R}_+^{p+q}$ we have

$$\begin{aligned} H_{X^*Y^*}(x^*,y^*) &= \text{Prob}(X^*\le x^*, Y^*\ge y^*) \\ &= \text{Prob}\big(X\le g_x.*\log(x^*), Y\ge g_y.*\log(y^*)\big) \\ &= H_{XY}\big(g_x.*\log(x^*), g_y.*\log(y^*)\big) \end{aligned} \tag{2.11}$$

Equivalently we have for all $(x,y)\in\mathbb{R}^{p+q}$,

$$H_{XY}(x,y) = H_{X^*Y^*}\big(\exp(x./g_x), \exp(y./g_y)\big). \tag{2.12}$$

Now by using (2.7) and (2.12) we can write

$$\begin{aligned} D(x,y;g_x,g_y) &= \sup\{\beta>0 | H_{X^*Y^*}\Big(\exp\big((x-\beta g_x)./g_x\big), \exp\big((y+\beta g_y)./g_y\big)\Big)>0\} \\ &= \sup\{\beta>0 | H_{X^*Y^*}\big(x^*\exp(-\beta), y^*\exp(\beta)\big)>0\}. \end{aligned}$$

[3]Note that the eventual convexity assumption of Ψ is not preserved by the transformation, which is not an issue here because only the free disposability property will be used in defining our estimators below.

By the change of variable $\gamma = \exp(\beta) > 0$ we finally obtain

$$\begin{aligned} D(x, y; g_x, g_y) &= \log(\gamma(x^*, y^*)), \\ \text{where } \gamma(x^*, y^*) &= \sup\{\gamma > 0 \,|\, H_{X^*Y^*}(\gamma^{-1}x^*, \gamma y^*) > 0\}. \end{aligned} \tag{2.13}$$

We note that if $(x, y) \in \Psi$, $\beta > 0$ and so $\gamma(x^*, y^*) > 1$ and if $(x, y) \in \Psi^\partial$, $\beta = 0$ and $\gamma(x^*, y^*) = 1$.

The motivation of writing the directional distances in terms of a standard multiplicative hyperbolic distances, is multiple. First, they are very easy to estimate in a nonparametric setup. Second, as seen in Section 3, we can define partial versions of the frontier (order-m and order-α quantile frontiers) leading to robust estimators of the directional distance. Third, as shown in Section 4, the resulting nonparametric estimators will share similar statistical properties than the radial measures and inference is easy to implement. Fourth, in the lines of Daraio and Simar (2005), we can easily introduce environmental factors in the process. This latter point is rather immediate and we just give a sketch of the approach below.

Conditional directional distances: a summary

To summarize, if a set of environmental variables $Z \in \mathbb{R}^r$ is considered as eventually influencing the production process, Daraio and Simar (2005), extending results from Cazals et al. (2002), suggest that the production process cannot be only characterized by $H_{XY}(x, y)$ but rather by the conditional probability

$$H_{XY|Z}(x, y|z) = \text{Prob}(X \leq x, Y \geq y \,|\, Z = z). \tag{2.14}$$

Then all the concepts of efficiency developped above can be rewritten by replacing H_{XY} by $H_{XY|Z}$. This leads to

Definition 2.1. *For all $(x, y) \in \Psi$, the conditional distance function of (x, y), conditional to $Z = z$ and relative to the directional vector (g_x, g_y) is defined as*

$$D(x, y; g_x, g_y|z) = \sup\{\beta > 0 | H_{XY|Z}(x - \beta g_x, y + \beta g_y \,|\, Z = z) > 0\}. \tag{2.15}$$

It is easy to see that its link with hyperbolic measures is given by

$$\begin{aligned} D(x, y; g_x, g_y|z) &= \log(\gamma(x^*, y^*|z)), \\ \text{where } \gamma(x^*, y^*|z) &= \sup\{\gamma > 0 \,|\, H_{X^*Y^*|Z}(\gamma^{-1}x^*, \gamma y^* \,|\, Z = z) > 0\}. \end{aligned} \tag{2.16}$$

The nonparametric estimation involves smoothing (kernel) methods in the Z dimension, for estimating $H_{X^*Y^*|Z}(\cdot, \cdot|Z = z)$. For details on nonparametric estimation, statistical properties and procedure for bandwidths selection, see Daraio and Simar (2005, 2007), Jeong et al. (2010) and Badin et al. (2010). All what is presented in the rest of the paper could also be written in terms of this conditional directional distance with conditioning on $Z = z$. To save place we will not give the details and let the reader do the exercise.

2.3 Link between input/output directional and radial distances

The links between the input or output directional distances and the radial input or output distances follow the same scheme as above. For the input orientation, we consider the directional vector $g = (g_x, 0)$, with $g_x > 0$, so according (2.7) we have

$$\begin{aligned} D(x, y; g_x, 0) &= \sup\{\beta > 0 | (x - \beta g_x, y) \in \Psi\} \\ &= \sup\{\beta > 0 | H_{XY}(x - \beta g_x, y) > 0\}. \end{aligned} \tag{2.17}$$

This measure can be equivalently written as

$$\begin{aligned} &D(x, y; g_x, 0) = \log(\gamma(x^*, y^*)), \\ &\qquad \text{where } \gamma(x^*, y^*) = \sup\{\gamma > 0 \,|\, H_{X^*Y^*}(\gamma^{-1}x^*, y^*) > 0\}, \end{aligned} \tag{2.18}$$

where the new coordinates are defined as $x^* = \exp(x./g_x)$ and $y^* = \exp(y)$. We see that $\gamma(x^*, y^*)$ is the usual radial Shephard input distance (the inverse of the Farrell input distance, $\theta(x^*, y^*)$) in the (x^*, y^*) coordinates. Remember that the interest here is the possibility of handling positive and negative values of the inputs, and the flexibility in choosing the direction g_x.

For the output oriented case, the choice of the directional vector is here $g = (0, g_y)$ with $g_y > 0$ giving

$$\begin{aligned} D(x, y; 0, g_y) &= \sup\{\beta > 0 | (x, y + \beta g_y) \in \Psi\} \\ &= \sup\{\beta > 0 | H_{XY}(x, y + \beta g_y) > 0\}. \end{aligned} \tag{2.19}$$

The transformation of coordinates is here given by $x^* = \exp(x)$ and $y^* = \exp(y./g_y)$. Therefore

$$\begin{aligned} &D(x, y; 0, g_y) = \log(\gamma(x^*, y^*)), \\ &\qquad \text{where } \gamma(x^*, y^*) = \sup\{\gamma > 0 \,|\, H_{X^*Y^*}(x^*, \gamma y^*) > 0\}, \end{aligned} \tag{2.20}$$

here $\gamma(x^*, y^*)$ is the usual radial Farrell output distance in the (x^*, y^*) coordinates.

Remark 2.1. *It should be noticed that for the full frontier measures analyzed here, the equations* (2.18) *and* (2.20) *could be writen in terms of conditional distribution functions.*

For the input oriented case, if y^* is such that $S_{Y^*}(y^*) = \text{Prob}(Y^* \geq y^*) \geq 0$, the condition appearing in (2.18) *could be written $F_{X^*|Y^*}(\gamma^{-1}x^*|y^*) > 0$ where $F_{X^*|Y^*}(x^*|y^*) = \text{Prob}(X^* \leq x^*|Y^* \geq y^*)$. Similarly for the output oriented case, if x^* such that $F_{X^*}(x) = \text{Prob}(X^* \leq x^*) > 0$ the condition in* (2.20) *could be written as $S_{Y^*|X^*}(\gamma y^*|x^*) > 0$ where $S_{Y^*|X^*}(y^*|x^*) = \text{Prob}(Y^* \geq y^*|X^* \leq x^*)$.*

3 Robust Versions of Directional Distances

Cazals et al. (2002) introduced the concept of "partial" frontiers, i.e. defining a benchmark frontier which is less extreme than the full frontier Ψ^∂. They proposed the order-m frontier and its corresponding order-m efficiency distances. One of the advantages of these frontiers is that their nonparametric estimators have nice properties ($\sqrt{n}$-consistency and asymptotic normal distribution) but in addition the frontier estimates do not envelop all the data points and so are less sensitive to outliers and extreme observations. Aragon et al. (2005) and Daouia and Simar (2007) proposed an alternative approach to define partial frontier based on the concept of order-α quantiles frontiers. Wheelock and Wilson (2008) and Wilson (2010) generalize these results to hyperbolic distance function. Recently Daouia et al. (2009, 2010) have shown that these partial frontiers estimators can even be exploited to build new estimators of the full frontier Ψ^∂ being robust to extreme data points and having asymptotic normal distributions. So, these robust versions of distance functions seem to be very useful for the practitioners and we will show here that the idea can be extended rather easily to directional distances. We present the formulations for the general directional distance, and summarize in the appendix the presentation for the output oriented directional distances (the case of input oriented caseis left as an exercise for the reader).

3.1 Order-α quantile frontiers

In place of looking to the extreme support of $H_{XY}(\cdot,\cdot)$, the order-α frontier can be determined by points such that the order-α efficiency score is equal to 1. For instance for the Farrell input measure, for any $\alpha \in (0,1]$, we have (see Daouia and Simar, 2007 for details)

$$\begin{aligned}\theta_\alpha(x,y) &= \inf\{\theta > 0 | F_{X|Y}(\theta x, y) > 1-\alpha\}\\ &= \inf\{\theta > 0 | H_{XY}(\theta x, y) > (1-\alpha) S_Y(y)\},\end{aligned}$$

where $F_{X|Y}(x,y) = H_{XY}(x,y)/S_Y(y)$ under the condition that the survivor function of Y exists, $S_Y(y) = \text{Prob}(Y \geq y) > 0$. Remember that here (using radial distance) we restrict the input-output as being positive. Clearly when $\alpha \to 1$, $\theta_\alpha(x,y) \to \theta(x,y)$. Following the same argument, the order-α hyperbolic efficiency score is defined by:

$$\gamma_\alpha(x, y) = \sup\{\gamma > 0 | H_{XY}(\gamma^{-1}x, \gamma y) > 1 - \alpha\}.$$

The α-quantile frontier is characterized by units (x, y) having a order-α measure $\gamma_\alpha(x, y) = 1$. Units on the α-quantile frontier would thus have a probabilty $(1 - \alpha)$ of being dominated.

The natural extension to order-α distance functions in the general case (allowing negative inputs/outputs) is obtained from its probabilistic characterization given in (2.6). We have

Definition 3.1. *For all values $(x, y) \in \Psi$ and any $\alpha \in (0, 1]$, the order-α directional distance of (x, y) with directional vector (g_x, g_y) is defined as*

$$D_\alpha(x, y; g_x, g_y) = \sup\{\beta | H_{XY}(x - \beta g_x, y + \beta g_y) > 1 - \alpha\}, \tag{3.1}$$

It is easy to see that when $\alpha \to 1$, $D_\alpha(x, y; g_x, g_y) \to D(x, y; g_x, g_y)$. Here values of $D_\alpha(x, y; g_x, g_y) = 0$ characterize points (x, y) lying on the order-α quantile frontier, values $D_\alpha(x, y; g_x, g_y) > 0$ (resp < 0) correspond to points standing below (resp. above) the order-α frontier.

With the notations and the transformations introduced in Section 2, the conditional distance function of order-α can be equivalently written as

$$\begin{aligned} D_\alpha(x, y; g_x, g_y) &= \log(\gamma_\alpha(x^*, y^*)), \\ \text{where } \gamma_\alpha(x^*, y^*) &= \sup\{\gamma > 0 \,|\, H_{X^*Y^*}(\gamma^{-1}x^*, \gamma y^*) > 1 - \alpha\}, \end{aligned} \tag{3.2}$$

where we recognize in $\gamma_\alpha(x^*, y^*)$ the usual order-α hyperbolic distance in the (x^*, y^*) coordinates. This extension is rather immediate due to the monotonic (increasing) nature of the transformations into the coordinates (x^*, y^*) and because quantiles are stable to such monotonic transformations.

By the appropriate choices of the directional vector g and of the coordinates (x^*, y^*) we could also recover the input and output versions of these order-α distance functions (we omit the details, and give a summary in the Appendix).

3.2 Order-m partial frontiers

The order-m partial frontier, introduced by Cazals et al. (2002) and extended to hyperbolic distance functions in Wilson (2010) defines, in a sense, the benchmark frontier as the *expected optimal* level achieved by m peers drawn randomnly in the population, rather that the *real optimal* achievable level. We give the definition here in its most general (hyperbolic) version. Consider a set of m iid random variables $\mathcal{S}_m = \{(X_j, Y_j)\}_{j=1}^m$ drawn from the joint density of

(X, Y) (this density has support Ψ and is univocally determined by $H_{XY}(x, y)$). We define now the random set

$$\widetilde{\Psi}_m = \bigcup_{j=1}^{m} \{(x, y) \in \mathbb{R}_+^{p+q} \mid x \geq X_j, y \leq Y_j\}. \tag{3.3}$$

Then for any fixed $(x, y) \in \mathbb{R}_+^{p+q}$ we define the random hyperbolic distance

$$\widetilde{\gamma}_m(x, y) = \sup\{\gamma | (\gamma^{-1}x, \gamma y) \in \widetilde{\Psi}_m\}, \tag{3.4}$$

we remark that $\widetilde{\gamma}_m(x, y)$ is the FDH estimator of the hyperbolic distance computed with the random sample $\mathcal{S}_m$. The order-m hyperbolic distance is then defined as

$$\gamma_m(x, y) = \mathbb{E}(\widetilde{\gamma}_m(x, y)). \tag{3.5}$$

Explicit analytical expression involving H_{XY} are given in Cazals et al. (2002) for the radial measures (input/output) and in Wilson (2010) for the hyperbolic case described above. Note that all these multiplicative distance measures require that both inputs and outputs are positive. We also remark that the random distance given in (3.4) can be defined as

$$\widetilde{\gamma}_m(x, y) = \sup\{\gamma | \widetilde{H}_{m,XY}(\gamma^{-1}x, \gamma y) > 0\}, \tag{3.6}$$

where $\widetilde{H}_{m,XY}(x, y) = (1/m) \sum_{j=1}^{m} \mathbb{1}(X_j \leq x, Y_j \geq y)$ is the empirical version of the function H_{XY} defined from the random sample $\mathcal{S}_m$.

Now we can extend this concept to directional distance and define directional distance of order-m, that will be applicable when some of the inputs or outputs may be negative. As above we consider m iid random variables $\mathcal{S}_m = \{(X_j, Y_j)\}_{j=1}^{m}$ drawn from the joint density of (X, Y) defined on Ψ and we adapt the definition of the random set $\widetilde{\Psi}_m$ to the case where x and y are reals (not restricted to be positive)

$$\widetilde{\Psi}_m = \bigcup_{j=1}^{m} \{(x, y) \in \mathbb{R}^{p+q} \mid x \geq X_j, y \leq Y_j\}. \tag{3.7}$$

Now we define the random distance

$$\begin{aligned} \widetilde{D}_m(x, y; g_x, g_y) &= \sup\{\beta | (x - \beta g_x, y + \beta g_y) \in \widetilde{\Psi}_m\}, && (3.8) \\ &= \sup\{\beta | \widetilde{H}_{m,XY}(x - \beta g_x, y + \beta g_y) > 0\}. && (3.9) \end{aligned}$$

We can now define our order-m directional distance

Definition 3.2. *For any $(x, y) \in \Psi$ and for any integer $m \geq 1$, the order-m directional distance of (x, y) with directional vector (g_x, g_y) is defined as*

$$D_m(x, y; g_x, g_y) = \mathbb{E}(\widetilde{D}_m(x, y; g_x, g_y)), \tag{3.10}$$

where we assume the existence of the expectation.

Note that $\widetilde{D}_m(x, y; g_x, g_y)$ can be equivalently written in the transformed coordinate system introduced in Section 2 as the logarithm of an hyperbolic measure

$$\begin{aligned} \widetilde{D}_m(x, y; g_x, g_y) &= \log(\widetilde{\gamma}_m(x^*, y^*)), \\ \text{where } \widetilde{\gamma}_m(x^*, y^*) &= \sup\{\gamma > 0 \,|\, \widetilde{H}_{m,X^*Y^*}(\gamma^{-1}x^*, \gamma y^*) > 0\}, \end{aligned} \tag{3.11}$$

where $\widetilde{H}_{m,X^*Y^*}$ is the version of $\widetilde{H}_{m,XY}$ in the (x^*, y^*) coordinate system. In our notation here, by denoting W_m the random variable $\widetilde{\gamma}_m(x^*, y^*)$ and G_m its distribution function[4], we have:

$$\begin{aligned} G_m(w) &= \text{Prob}(W_m \leq w) = \text{Prob}\left[\max_{j=1,\ldots,m}\left\{\min_{k=1,\ldots,p,\,\ell=1,\ldots,q}\left(\frac{x^{*,k}}{X_j^{*,k}}, \frac{Y_j^{*,\ell}}{y^{*,\ell}}\right)\right\} \leq w\right] \\ &= \left[\text{Prob}\left\{\min_{k=1,\ldots,p,\,\ell=1,\ldots,q}\left(\frac{x^{*,k}}{X^{*,k}}, \frac{Y^{*,\ell}}{y^{*,\ell}}\right) \leq w\right\}\right]^m \\ &= [1 - \text{Prob}\,(x^* > w\,X^*, Y^* > w\,y^*)]^m \\ &= \left[1 - H_{X^*Y^*}(w^{-1}x^*, w\,y^*)\right]^m \end{aligned} \tag{3.12}$$

So, under the assumption that the expectation of $\log(W_m)$ exists for any $m \geq 1$, the explicit formulation of the order-m directional distance is given by

$$D_m(x, y; g_x, g_y) = \int_0^\infty \log(w)\, dG_m(w). \tag{3.13}$$

It should be noticed that the expectation of $\widetilde{\gamma}_m(x^*, y^*)$ is simply given by

$$\mathbb{E}(\widetilde{\gamma}_m(x^*, y^*)) = \int_0^\infty [1 - G_m(w)]\, dw,$$

but since in general, for non degenerate random variable W_m, $\mathbb{E}(\log(W_m)) \neq \log(\mathbb{E}(W_m))$, the latter result is not useful to recover $D_m(x, y; g_x, g_y)$. We show in the proposition below the convergence of the order-m directional distance to the "full" directional distance when $m \to \infty$.

Theorem 3.1. *For any values (x, y) in Ψ, we have:*

$$\lim_{m\to\infty} D_m(x, y; g_x, g_y) = D(x, y; g_x, g_y) \tag{3.14}$$

Proof. Note first that the integral in equation (3.13) can be rewritten:

$$D_m(x, y; g_x, g_y) = \int_0^{\gamma(x^*,y^*)} \log(w)\, dG_m(w) \tag{3.15}$$

[4]The distribution function $G_m(.)$ depends on the coordinates of the point of interest (x^*, y^*), but to simplify the notations, we omit to explicit this.

Indeed, by definition of $\gamma(x^*, y^*)$ given in (2.13), we have, for any values $(x^*, y^*) \in \Psi^*$:

$$\forall w > \gamma(x^*, y^*),\ H_{X^*Y^*}(w^{-1}x^*, w\,y^*) = 0 \text{ and } \forall m \geq 1,\, G_m(w) = 1.$$

Under the assumption that $\log(w)G_1(w) \to 0$ when $w \to 0$, we obtain, for any $m \geq 1$, by integration by parts:[5]

$$\begin{aligned} D_m(x, y; g_x, g_y) &= \log\left(\gamma(x^*, y^*)\right) - \int_0^{\gamma(x^*,y^*)} \frac{G_m(w)}{w}\, dw \\ &= D(x, y; g_x, g_y) - \int_0^{\gamma(x^*,y^*)} \frac{\left[1 - H_{X^*Y^*}(w^{-1}x^*, w\,y^*)\right]^m}{w}\, dw. \end{aligned}$$

Then, for all $w \leq \gamma(x^*, y^*)$ and $m \geq 1$, we have $\frac{G_m(w)}{w} \leq \frac{G_1(w)}{w}$ and $\int_0^{\gamma(x^*,y^*)} \frac{G_1(w)}{w} dw$ exists. Therefore, using the Dominated convergence theorem, the integral on the right-hand side converges to zero when $m \to \infty$ and $\lim_{m\to\infty} D_m(x, y; g_x, g_y) = D(x, y; g_x, g_y)$. □

4 Nonparametric Estimators

4.1 Full frontier estimation

Consider a sample of observations $\mathcal{X} = \{(X_i, Y_i) | i = 1, ..., n\}$ and a fixed value (x, y). For a given direction (g_x, g_y), a nonparametric estimator of the directional distance is given by

$$\widehat{D}(x, y, g_x, g_y) = \log\left(\widehat{\gamma}(x^*, y^*)\right) \tag{4.1}$$

where $\widehat{\gamma}(x^*, y^*)$ is the empirical analog of the hyperbolic distance computed in the transformed sample of observations $\mathcal{X}^* = \{(X_i^*, Y_i^*) = (\exp(X_i./g_x), \exp(Y_i./g_y)) | i = 1, ..., n\}$ at the fixed transformed value $(x^*, y^*) = (\exp(x./g_x), \exp(y./g_y))$. More precisely, define the set of observations which dominate (x^*, y^*), $D^* = \{i\ |(X_i^*, Y_i^*) \in \mathcal{X}^*, X_i^* \leq x^*, Y_i^* \geq y^*\}$. The estimation of $\widehat{H}_{X^*Y^*}$ is given by:

$$\widehat{H}_{X^*Y^*}(x^*, y^*) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i^* \leq x^*, Y_i^* \geq y^*) \tag{4.2}$$

and, following equation (3.2) and Wheelock and Wilson (2008), the estimation of $\widehat{\gamma}(x^*, y^*)$ is given by:

[5] The assumption $\log(w)G_1(w) \to 0$ when $w \to 0$ is not very restrictive. In particular, it is obtained if the support of (X, Y) is compact. In the latter case, the support of (X^*, Y^*) is also compact with $X^* > 0$ and $Y^* > 0$ and we have that for all $(x^*, y^*) \in \Psi^*$, $H_{X^*Y^*}(w^{-1}x^*, w\,y^*) = 1$ (and so, $G_1(w) = 0$) for all $w \leq w^*$ for some $w^* > 0$.

$$\widehat{\gamma}(x^*, y^*) = \sup\{\gamma > 0 \,|\, \widehat{H}_{X^*Y^*}(\gamma^{-1}x^*, \gamma y^*) > 0\}, \tag{4.3}$$

$$= \max_{j \in D^*} \left\{ \min_{k=1,\ldots,p,\, \ell=1,\ldots,q} \left(\frac{x^{*,k}}{X_j^{*,k}}, \frac{Y_j^{*,\ell}}{y^{*,\ell}} \right) \right\}. \tag{4.4}$$

So, the computation of $\widehat{\gamma}(x^*, y^*)$ can be derived from a standard FDH estimator by considering the extended sample of $p+q$ outputs $\{(1./X_j^*, Y_j^*), j = 1, ..., n\}$ and one single fixed input.

We note that the estimation of input/output directional distance can be derived similarly by computing the associated input/output oriented FDH estimator in the transformed set of data $\mathcal{X}^*$ and taking the logarithm of this value (see the appendix for more details).

The asymptotic properties of the estimator (4.1) are directly derived from the asymptotic properties of the FDH estimator given in (4.4) using the functional Delta method. These properties are obtained under standard regularity conditions, described in details in Park et al. (2000) (see also Wheelock and Wilson, 2008 and Wilson 2010). In summary we have:

– **Assumption A1.** The observations (X_i, Y_i) are independent and identically distributed random variables, with density function f_{XY} defined on the support Ψ. The support Ψ is assumed free disposal and compact.

– **Assumption A2.** At a frontier point $(x^\partial, y^\partial) \in \Psi^\partial$, the density f_{XY} is strictly positive and sequentially Lipschitz continuous, i.e., $f_{XY}(x^\partial, y^\partial) > 0$ and for all sequences $(x_k, y_k) \in \Psi$ converging to (x^∂, y^∂), $|f_{XY}(x_k, y_k) - f_{XY}(x^\partial, y^\partial)| \le c||(x_k, y_k) - (x^\partial, y^\partial)||$ for some positive constant c.

– **Assumption A3.** For all (x, y) in the interior of Ψ, $\theta(x, y)$, $\lambda(x, y)$ and $\gamma(x, y)$ are twice continuously differentiable in both arguments.

Note that the density f_{XY} appearing in Assumption A2 is determined by H_{XY} and conversely.

Theorem 4.1. *Under Assumptions $A1 - A3$ for any (x, y) in the interior of Ψ such that $\gamma(x^*, y^*) > 0$, we have:*

$$D(x, y, g_x, g_y) - \widehat{D}(x, y, g_x, g_y) = O_p(n^{-1/(p+q)}), \tag{4.5}$$

$$n^{-1/(p+q)} \left(D(x, y, g_x, g_y) - \widehat{D}(x, y, g_x, g_y) \right) \xrightarrow{\mathcal{L}} \text{Weibull} \left(\mu_{H,0}^{p+q}, \frac{p+q}{\gamma(x^*, y^*)} \right), \tag{4.6}$$

where $\mu_{H,0}$ is a constant.

An expression of the Weibull parameter $\mu_{H,0}$ is obtained by adapting Park et al. (2000) to hyperbolic measures. It is explicitely given in Wilson and Wheelock (2008) (see Definition A.2, in their Appendix).

Proof. The proposition follows from a straitghforward application of the functional delta method (see e.g. van der Vaart 1998) applied on $\gamma(x^*, y^*)$ with the functional log as continuously differentiable mapping. More precisely, we have:

$$n^{-1/(p+q)} \left(\log\left(\gamma(x^*, y^*)\right) - \log\left(\widehat{\gamma}(x^*, y^*)\right)\right) = \frac{1}{\gamma(x^*, y^*)} n^{-1/(p+q)} \left(\gamma(x^*, y^*) - \widehat{\gamma}(x^*, y^*)\right) + o_P(1).$$

We know from Park et al. (2000) and Wheelock and Wilson (2008) that under the above assumptions, $\gamma(x^*, y^*) - \widehat{\gamma}(x^*, y^*) = O_p(n^{-1/(p+q)})$ and that

$$n^{-1/(p+q)} \left(\gamma(x^*, y^*) - \widehat{\gamma}(x^*, y^*)\right) \xrightarrow{\mathcal{L}} \text{Weibull}(\mu_{H,0}^{p+q}, p+q).$$

Multiplying a Weibull distribution by a strictly positive constant preserves the Weibull distribution with a new scalar argument, which completes the proof. □

From a practical point of view, the result is difficult to use because the constant $\mu_{H,0}$ depends on many unknown parameters. Park et al. (2000) suggest a way to estimate this constant in the input (or output) oriented case but claim that it is only useful with large data sets (in particular if $p + q$ increases). So, the bootstrap seems to be an attractive way to provide statistical inference involving the measures $\log(\gamma(x^*, y^*))$. The consistency of the bootstrap using FDH estimators has been established in Jeong and Simar (2006), it is based on a subsampling approach. In Simar and Wilson (2009), it is shown how to determine the subsampling size by data driven methods.

4.2 Order-α frontier estimation

The estimation of *order-α directional distance* is driven from the estimation of α quantile hyperbolic estimator and the multivariate approach proposed by Daouia and Simar (2007). We simply have

$$\widehat{D}_\alpha(x, y, g_x, g_y) = \log(\widehat{\gamma}_\alpha(x^*, y^*)), \tag{4.7}$$

where the algorithm from Daouia and Simar is adapted as follows. Define:

$$\mathcal{X}_i^* = \min_{k=1,\ldots,p,\, \ell=1,\ldots,q} \left(\frac{x^{*,k}}{X_i^{*,k}}, \frac{Y_i^{*,\ell}}{y^{*,\ell}} \right), \quad i = 1, \ldots, n.$$

Let $M^* = \sum_{i=1}^n 1\!\!I(X_i^* \leq x^*, Y_i^* \geq y^*)$ be nonnull. For $j = 1, ..., M^*$, denote by $\mathcal{X}_{(j)}^*$ the j-th order statistic of the observations $\mathcal{X}_i^*$ such that $X_i^* \leq x^*$ and $Y_i^* \geq y^*$. Then it follows

$$\widehat{\gamma}_\alpha(x^*, y^*) = \begin{cases} \mathcal{X}_{((1-\alpha)M^*)}^* & \text{if } (1-\alpha)M^* \in \mathbb{N}, \\ \mathcal{X}_{([(1-\alpha)M^*]+1)}^* & \text{otherwise.} \end{cases} \tag{4.8}$$

As for the full frontier estimation, the asymptotic properties of (4.7) follows from the known asymptotic results for (4.8) using the functional Delta method and the log mapping. This provides the next theorem.

Theorem 4.2. *Under Assumptions A1 − A3, for $0 < \alpha < 1$, and for any (x, y) in the interior of Ψ such that $\gamma_\alpha(x^*, y^*) > 0$ let the function $\phi(w) = H_{X^*Y^*}(w^{-1}x^*, w\,y^*)$. If $\phi(w)$ is differentiable at $w = \gamma_\alpha(x^*, y^*)$, $\widehat{D}_\alpha(x, y, g_x, g_y)$ is a $\sqrt{n}$-consistent estimator of $D_\alpha(x, y, g_x, g_y)$ and*

$$\sqrt{n}\left(\widehat{D}_\alpha(x, y, g_x, g_y) - D_\alpha(x, y, g_x, g_y)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\alpha^2(x, y)), \tag{4.9}$$

where

$$\sigma_\alpha^2(x, y) = \frac{\alpha(1-\alpha)}{\gamma_\alpha(x^*, y^*)^2}\left[\phi'(\gamma_\alpha(x^*, y^*))\right]^{-2}. \tag{4.10}$$

Proof. The argument is similar to the proof in Theorem 4.1 above. Daouia and Simar (2007) derive the $\sqrt{n}$-consistency and the asymptotic normality for the α-quantile estimator (input or output orienttation). These asymptotic properties have been extended to the hyperbolic α-quantile estimator in Wheelock and Wilson (2008) which we need here. The delta method allows, in a straightforward way, to transfer these asymptotic properties to the directional estimator (4.7), which completes the proof. □

It is obvious that for a given n, if $\alpha \to 1$, the order-α estimator of the directional distance converges to the "full" directional distance: $\lim_{\alpha\to1} \widehat{D}_\alpha(x, y, g_x, g_y) = \widehat{D}(x, y, g_x, g_y)$. But we know (see Wheelock and Wilson, 2008) that if $\alpha(n)$ is a sequence in n with $\alpha(n) \to 1$ as $n \to \infty$ fast enough, the order-α estimator $\widehat{\gamma}_{\alpha(n)}(x^*, y^*)$ shares the same properties as the FDH estimator $\widehat{\gamma}(x^*, y^*)$. This result still holds for the directional distances.

Theorem 4.3. *Under Assumptions A1 − A3 for any (x, y) in the interior of Ψ such that $\gamma(x^*, y^*) > 0$, and with the order $\alpha(n) > 0$ such that $n^{(p+q+1)/(p+q)}(1 - \alpha(n)) \to 0$ as $n \to \infty$, $\widehat{D}_{\alpha(n)}(x, y, g_x, g_y)$ is a consistent estimator of $D(x, y, g_x, g_y)$ and*

$$n^{-1/(p+q)}\left(D(x, y, g_x, g_y) - \widehat{D}_{\alpha(n)}(x, y, g_x, g_y)\right) \xrightarrow{\mathcal{L}} Weibull\left(\mu_{H,0}^{p+q}, \frac{p+q}{\gamma(x^*, y^*)}\right) \tag{4.11}$$

where $\mu_{H,0}$ is the constant defined in Theorem 4.1.

Proof. The result derives immediately from the functional Delta method and from the Theorem 4.4 in Wheelock and Wilson (2008). □

The last results allows to define a robust estimator of the directional distance $D(x, y, g_x, g_y)$ since even if $\alpha(n) \to 1$ when $n \to \infty$, for finite n, the estimator $\widehat{D}_{\alpha(n)}(x, y, g_x, g_y)$ is not obtained by envelopping all the data points and so will be more robust to extreme data points.

Note that here, if inference on the order-α measures is desired (with fixed α), Theorem 4.2 allows to use the normal approximation, and in practice, the bootstrap is useful to approximate the variance term. Note that here, as explained in Florens and Simar (2005), the naive bootstrap can be used because we are not estimating a boundary. If Theorem 4.3 is advocated to make inference on the full frontier estimate, here the subsampling bootstrap has to be used, as explained above for the full frontier estimate.

4.3 Order-m frontier estimation

The computation of the *order-m directional distance estimator* $\widehat{D}_m(x, y, g_x, g_y)$ follows the same vein with a Monte Carlo procedure, adapted from Cazals et al. (2002), to estimate the expectation in (3.10). The estimator is defined by the empirical version of (3.10). We have

$$\widehat{D}_m(x, y, g_x, g_y) = \widehat{\mathbb{E}}\left(\log(\tilde{\gamma}_m(x^*, y^*))\right), \tag{4.12}$$

$$= \int_0^\infty \log(w)\, d\widehat{G_m}(w), \tag{4.13}$$

with $\widehat{G_m}(w) = \left[1 - \widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right]^m$. An simple algorithm to compute the expectation is as follows:

(1) For $b = 1, ..., B$, where B is large, redo the following steps (i)-(ii):

   (i) Draw a sample of size m with replacement from D^* and denote it by $\{(X^*_{i,b}, Y^*_{i,b}), i = 1, ..., m\}$

   (ii) Compute $\tilde{\gamma}^b_m(x^*, y^*) = \max_{j=1,\dots,m} \left\{ \min_{k=1,\dots,p,\, \ell=1,\dots,q} \left(\frac{x^{*,k}}{X^{*,k}_{j,b}}, \frac{Y^{*,\ell}_{j,b}}{y^{*,\ell}} \right) \right\}$.

(2) Finally, $\widehat{D}_m(x, y, g_x, g_y) \approx \frac{1}{B} \sum_{b=1}^{B} \log(\tilde{\gamma}^b_m(x^*, y^*))$.

For the quality of the Monte-Carlo approximation B should be large. The computations are so fast that $B = 1000$ is quite reasonable in practice.

Contrary to the previous estimators (full and order-α frontiers), the order-m directional distance estimator is not defined as a log transformation of the order-m hyperbolic estimator and its properties cannot be directly inferred using this previous estimator. However, we can still use the Delta method to derive asymptotic properties of $\widehat{D}_m(x, y, g_x, g_y)$. The first result shows the convergence of $\widehat{D}_m(x, y, g_x, g_y)$ to the FDH estimator $\widehat{D}(x, y, g_x, g_y)$ when $m \to \infty$.

Theorem 4.4. *Under Assumptions A1 − A3, for any (x, y) in the interior of Ψ, and for fixed n,*

$$\lim_{m\to\infty} \widehat{D}_m(x, y, g_x, g_y) = \widehat{D}(x, y, g_x, g_y) \tag{4.14}$$

Proof. Note first that the integral in equation (4.13) can be rewritten:

$$\widehat{D}_m(x, y; g_x, g_y) = \int_{k_w}^{\widehat{\gamma}(x^*,y^*)} \log(w)\, d\widehat{G}_m(w) \tag{4.15}$$

Indeed, by the definition of $\widehat{\gamma}(x^*, y^*)$, for any values $(x^*, y^*) \in \Psi^*$ we have that for all $w > \widehat{\gamma}(x^*, y^*)$, $\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*) = 0$ and so, $\widehat{G_m}(w) = 1$. In addition, under the assumption that (X, Y) have compact support, there exist a $w^* > 0$ such that for all $w \le w^*$, $\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*) = 1$ and therefore $\widehat{G_m}(w) = 0$ for all $w \le w^*$ (see footnote 5). Now we obtain by integration by parts:

$$\begin{aligned}
\widehat{D}_m(x, y; g_x, g_y) &= \log\left(\widehat{\gamma}(x^*, y^*)\right) - \int_{w^*}^{\widehat{\gamma}(x^*,y^*)} \frac{\widehat{G_m}(w)}{w}\, dw \\
&= \widehat{D}(x, y; g_x, g_y) - \int_{w^*}^{\widehat{\gamma}(x^*,y^*)} \frac{\left[1 - \widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right]^m}{w}\, dw \qquad (4.16)
\end{aligned}$$

The integrand of the right side term is bounded by $\left[\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right]^m / w^*$ which converges to zero as $m \to \infty$, since $\left[\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right] < 1$. This gives the desired result. □

Now we have to establish the convergence of the order-m estimator of the directional distance to the true order-m directional distance, as $n \to \infty$, with m fixed.

Theorem 4.5. *Under Assumptions A1 − A3, for any (x, y) in the interior of Ψ, and for any $m \ge 1$ we have, as $n \to \infty$*

$$\widehat{D}_m(x, y, g_x, g_y) \xrightarrow{a.s.} D_m(x, y, g_x, g_y), \tag{4.17}$$

$$\sqrt{n}\left(\widehat{D}_m(x, y, g_x, g_y) - D_m(x, y, g_x, g_y)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(x, y)), \tag{4.18}$$

with

$$\sigma^2(x, y) = \mathbb{E}\left[\int_0^\infty \frac{m}{w}\left(1 - H_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right)^{m-1} 1\!\!1(X_i^* \le w^{-1}x^*, Y_i^* \ge w\, y^*) dw\right]^2$$

Proof. We have to adapt the argument in Cazals et al.(2002) and use the functional delta method. Consider, for any (x, y) in the interior of Ψ and for any $m \geq 1$, the operator T defined as

$$T(H) = \int_0^\infty \log(w) dG_m(w),$$

where G_m is the function depending on H defined in (3.12). The operator T associates a real value to any probability function H. This operator is Frechet-differentiable w.r.t. the sup norm, that is:

$$T(H + R) - T(H) = DT_H(R) + o_P(\|R\|) \tag{4.19}$$

Before applying the Frechet differentiability of T, we first decompose it into two integrals and apply integration by parts on each of them, in order to have a more appropriate expression for $T(H)$ to define its Frechet derivative. Therefore, we have:

$$\begin{aligned} T(H) &= \int_0^1 \log(w) dG_m(w) + \int_1^\infty \log(w) dG_m(w) \\ &= \int_0^1 \log(w) dG_m(w) - \int_1^\infty \log(w)(-dG_m(w)) \\ &= (I) + (II) \end{aligned}$$

We have seen above in Section 3 that $G_m(w) = 0$ for all $w \leq w^*$ for some $w^* > 0$ and that $G_m(w) = 1$ for all $w \geq \gamma(x^*, y^*)$. Note that since (x^*, y^*) is in the interior of Ψ^*, $w^* < 1$ and $\gamma(x^*, y^*) > 1$. So the two integrals above simplifies and by integrating by parts, we obtain

$$\begin{aligned} (I) &= \int_{w^*}^1 \log(w) dG_m(w) \\ &= -\int_{w^*}^1 \frac{G_m(w)}{w} dw. \end{aligned}$$

In the same way,

$$\begin{aligned} (II) &= -\int_1^{\gamma(x^*,y^*)} \log(w)(-dG_m(w)) \\ &= -\left[\log(w)\left(1 - G_m(w)\right)\right]_1^{\gamma(x^*,y^*)} + \int_1^{\gamma(x^*,y^*)} \frac{1 - G_m(w)}{w} dw. \end{aligned}$$

Therefore, under our regularity conditions, the operator T can be rewritten as:

$$T(H) = -\int_0^1 \frac{G_m(w)}{w} dw + \int_1^\infty \frac{1 - G_m(w)}{w} dw \tag{4.20}$$

The Frechet derivative is then obtained by standard calculus:

$$\begin{aligned}
G_m(H+R) &= \left[1-(H+R)(w^{-1}x^*, w\,y^*)\right]^m \\
&= \sum_{k=0}^{m}\binom{m}{k}\left(1-H(w^{-1}x^*, w\,y^*)\right)^{m-k}\left(-R(w^{-1}x^*, w\,y^*)\right)^k \\
&= \left[1-H(w^{-1}x^*, w\,y^*)\right]^m \\
&\quad -m\left[1-H(w^{-1}x^*, w\,y^*)\right]^{m-1} R(w^{-1}x^*, w\,y^*) \\
&\quad +\ldots.
\end{aligned}$$

Therefore we get:

$$DT_H(R) = \int_0^\infty \frac{m}{w}\left\{\left[1-H(w^{-1}x^*, w\,y^*)\right]^{m-1} R(w^{-1}x^*, w\,y^*)\right\} dw$$

Now, applying equation (4.19) we have:

$$\begin{aligned}
\sqrt{n}\left[T(\widehat{H})-T(H)\right] &= \int_0^\infty \frac{m}{w}\left[1-H(w^{-1}x^*, w\,y^*)\right]^{m-1}\sqrt{n}(\widehat{H}-H)(w^{-1}x^*, w\,y^*)dw \\
&\quad +o_P(\sqrt{n}\|\widehat{H}-H\|) \qquad (4.21)
\end{aligned}$$

We are now back to the standard form (equation (3.10)) as in Cazals et al. (2002), our theorem then follows from the central limit theorem applied to the first term of (4.21) (see Cazals et al. for details). □

Our final result concerns the behavior of $\widehat{D}_{m(n)}(x, y, g_x, g_y)$ for a sequence $m(n)$ with $m(n) \to \infty$ fast enough, as $n \to \infty$. We know by Cazals et al. (2002) that under appropriate rates for $m(n)$, the asymptotic behavior of $\widehat{\gamma}_{m(n)}$ is similar to the FDH estimator as estimator of the full frontier measure γ, with the limiting Weibull distribution. The following theorem indicates that this property is preserved for directional distance estimators.

Theorem 4.6. *Under Assumptions A1−A3, for any (x, y) in the interior of Ψ, and let $m(n)$ be a sequence such that $m(n) \to \infty$ as $n \to \infty$, with $m(n) = O(\beta n \log(n))$ with $\beta > 1/(p+q)$. Then we have as $n \to \infty$*

$$n^{-1/(p+q)}\left(\widehat{D}_{m(n)}(x, y, g_x, g_y) - D(x, y, g_x, g_y)\right) \xrightarrow{\mathcal{L}} \mathit{Weibull}\left(\mu_{H,0}^{p+q}, \frac{p+q}{\gamma(x^*, y^*)}\right) \qquad (4.22)$$

where $\mu_{H,0}$ is the constant introduced in Theorem 4.1.

Proof. Now, using the decomposition (4.16) we obtain:

$$\begin{aligned}
n^{-1/(p+q)}\left(\widehat{D}_m(x, y, g_x, g_y) - D(x, y, g_x, g_y)\right) &= n^{-1/(p+q)}\left(\widehat{D}(x, y, g_x, g_y) - D(x, y, g_x, g_y)\right) \\
&\quad -n^{-1/(p+q)}\int_0^{\widehat{\gamma}(x^*,y^*)} \frac{\left[1-\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\,y^*)\right]^m}{w}dw. \qquad (4.23)
\end{aligned}$$

Following Cazals et al. (2002), the issue is to find $m(n)$ such that the last integral is $o_P(1)$ as $n \to \infty$. Note first that

$$\forall (x^*, y^*) \in \Psi^*,\ \forall w \in [0, \widehat{\gamma}(x^*, y^*)],\ \widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*) \geq \frac{1}{n},$$

and $\left[1 - \widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*)\right]^m \leq \left(1 - \frac{1}{n}\right)^m$. Under the assumption that X and Y have compact support, then for all $(x^*, y^*) \in \Psi^*$, $\widehat{H}_{X^*Y^*}(w^{-1}x^*, w\, y^*) = 1$ for all $w \leq w^*$ for some constant $w^* > 0$. Now, using a mean value theorem, we can bound the integral in (4.23) by $\widehat{\gamma}(x^*, y^*)\frac{1}{w^*}\left(1 - \frac{1}{n}\right)^m$. So, to achieve our goal, it is sufficient that $m(n)$ is such that $\left(1 - \frac{1}{n}\right)^{m(n)} = O_p(n^{-\beta})$ where $\beta > 1/(p+q)$. This completes the proof. □

Again this property indicates that the order-m directional distance estimator can be viewed as a robust estimator of the full directional distance measure, because even if $m(n) \to \infty$ fast enough as $n \to \infty$, for finite n, the corresponding estimator is not based on an estimator of Ψ that envelops all the data points and so $\widehat{D}_{m(n)}(x, y, g_x, g_y)$ will be more resistant to outliers and extreme data points.

For practical inference for deriving confidence intervals and/or testing issues, here again, the bootstrap can be used, with the same caveats carefully described for "full" and order-α directional distances above at the end of Sections 4.1 and 4.2.

5 Empirical Illustration

We illustrate our methodology using a cross-section database of US mutual funds collected from Morningstar and updated at May 2002, with one specific category of Mutual Funds (aggressive growth) and a final number of observations equal to 129. Note that a more complete version of this database was previously analyzed in Daraio and Simar (2006, 2007).[6].

In this setting the traditional output is Y, the *total annual return*, expressed in percentage terms and we consider three inputs: X_1 is a *measure of risk* given by the standard deviation of return, X_2 is the *expense ratio* which is a measure of transaction costs (operating expenses and management fees, administrative fees, and all other asset-based costs) and X_3, the *turnover ratio* that measures the fund's trading activity (funds with higher turnover incur greater brokerage fees for affecting the trades). Due to the aggressive strategies chosen by the funds, annual returns can take either positive or negative values and it is mandatory to take this specificity into account in the estimation of efficiencies. As a matter of fact, most of the returns in our data were negative over this period (spanning the 11 September 2001).

[6]We are grateful to Cinzia Daraio who kindly let us use her database

Table 1 below gives the directional distances obtained for full and partial frontiers. The table presents the different variants of the measure: input oriented $D^{in}(x, y) = D(x, y; g_x, 0)$, output oriented $D^{out}(x, y) = D(x, y; 0, g_y)$ and the more general directional distance $D(x, y) = D(x, y; g_x, g_y)$. The direction we have chosen for g_x is the average values for the inputs (all are positive in our case) and we obtain $g_x = (34.98, 155.19, 1.68)'$. For the output we choose the mean of the absolute values and obtain $g_y = 18.36$. Comparing the efficiency level along an average benchmark direction sounds meaningful, although from a theoretical view point, any other direction could be chosen.

Note that for the robust estimation methods, the orders α and m were chosen in order to fix a proportion of points staying outside the partial frontier nearly equal for both approaches. In the illustration below we choose $m = 40$ (we choose 1000 Monte-Carlo replications to compute the order-m estimates) and $\alpha = 0.98$ providing in both cases around 23% of points outside the robust estimated frontier. Among these points the most extremes may be warned as potential outliers. In order to save place, the tables below give the results only for 20 funds randomly chosen in the full set of 129 funds. The average values at the bottom line have been computed with the full sample.

Table 1 gives the results for the different estimates. Choosing the orientation is let to the analyst but of course, the advantage of the global measures ($D(x, y), D_\alpha(x, y)$ and $D_m(x, y)$) is that they do not force to privilege one particular orientation (input or output). We will thus mainly focus the comments on the 3 last columns of the table. When $\widehat{D}_n(x, y) = 0$, it means that the fund is estimated as being efficient. It is the case of the funds #3, #122, #123, #56 and #31 (we can see since they are on the efficient frontier, they are also efficient for the input oriented case (column 2) and for the output oriented case (column 5). Looking to the partial measures for these funds give an idea how far they are outside the partial frontier. For instance for the fund #3, we see that it stands, as expected, outside the partial frontiers ($\hat{D}_{\alpha,n} = -0.1361$ and $\hat{D}_{m,n} = -0.0904$). This means that it has to reduce its output (the annual return) by $0.0904 \times g_y = 1.66$ and simultaneously increase its inputs by an amount (in original units) of $0.0904 \times g_x = (3.16, 14.03, 0.15)'$ to be on the partial frontier of order $m = 40$. So, fund #3 seems not to be so far above the order-40 frontier.

As pointed above, these partial scores can be useful to detect the most extreme (efficient) funds and warn for potential outliers. For instance, the fund #56 seems to be very extreme to this respect (very high negative values of $\hat{D}_{\alpha,n}$ and of $\hat{D}_{m,n}$) . This idea has been suggested in Simar (2003), proposing to do this kind of analysis in a more systematic way. The procedure could be extended with our directional distances. On the other extreme the fund #99 and #91 seems particularly inefficient among the 20 selected funds. For instance fund #91 should increase its return by an amount of $0.7256 * g_y = 13.33$ and reduce its outputs by an amount of $0.7256 * g_x = (25.38, 112.61, 1, 22)'$ for being on the efficient frontier. We see that these inefficient funds would also be considered as being inefficient using other measures (input and output oriented); even the less extreme benchmark provided by the partial frontier seem very distant for these inefficient funds.

These distance of the data points to the frontier, in original units, are quite informative because they give, in the original units the size of the effort to become 'efficient'. These distances from the frontier are called "slacks" in the literature. Table 2 gives the slacks and the corresponding original data of the selected 20 funds. To save place we only give the slacks for the full directional measure $D(x, y)$ (estimated by the 8th column of Table 1.We see that of course for efficient funds the slacks are zero. The last row of the table gives the average slacks with the average levels of inputs and outputs in the full sample of 129 funds.

6 Conclusion

In this paper, we have shown that a probabilistic formulation of the production process allows to give an original characterization of the directional distances introduced by Chambers et al. (1998) and Färe and Grosskopf (2004). This is not only a simple theoretical exercise because by doing this we were able to define the robust versions of these directional distance function, introducing order-m and order-α quantile version of the distances. In addition we indicate that it is straightforward to extend the concept of conditional efficiency measures to the directional distances, which allows to introduce environmental factors or heterogeneity in the production process.

Finally we show that most of the known statistical properties of the nonparametric estimators of radial efficiency scores can be transferred to the nonparametric estimators of the directional distances. Even statistical inference is available by using the appropriate bootstrap algorithms. We provide also very simple and fast algorithms to compute these nonparametric estimators.

Directional distance are quite attractive, since they generalize in a sense the usual Farrell-Shephard radial distances when all the input and the outputs are all strictly positive. But when some of the inputs and/or the outputs could take negative values, there provide the most natural way to measure efficiencies. The procedure is illustrated in a real data example on Mutual Funds. This example illustrates how useful and informative are these directional measures to analyses the performance of the funds.

References

[1] Aragon, Y., A. Daouia and C. Thomas-Agnan (2005), Nonparametric Frontier Estimation: A Conditional Quantile-based Approach, *Econometric Theory*, 21, 358–389.

[2] Badin, M., Daraio, C. and L. Simar (2010), Optimal Bandwidth Selection for Conditional Efficiency Measures: a Data-driven Approach, *European Journal of Operational Research*, 201, 633–640

[3] Cazals, C. Florens, J.P. and L. Simar (2002), Nonparametric Frontier Estimation: a Robust Approach , in *Journal of Econometrics*, 106, 1–25.

[4] Chambers, R.G., Y.H. Chung,and R. Färe (1998), Profit, Directional Distance Functions and Nerlovian Efficiency, *Journal of Optimization Theory and Applications*, 98, 351–364.

[5] Daouia, A., Florens, J.P. and L. Simar (2009), Regularization of Non-parametric Frontier Estimators. Discussion paper 0922, Institut de Statistique, UCL.

[6] Daouia, A., Florens, J.P. and L. Simar (2010), Frontier estimation and extreme values theory. Discussion paper 0806, Institut de Statistique, UCL, in press *Bernoulli*.

[7] Daouia, A. and L. Simar (2007), Nonparametric efficiency analysis: a multivariate conditional quantile approach, *Journal of Econometrics*, 140, 375–400.

[8] Daraio, C. and L. Simar (2005), Introducing environmental variables in nonparametric frontier models: a probabilistic approach, *Journal of Productivity Analysis*, vol 24, 1, 93–121.

[9] Daraio, C. and L. Simar (2006), A Robust Nonparametric Approach to Evaluate and Explain the Performance of Mutual Funds, *European Journal of Operational Research*, vol 175, 1, 516–542.

[10] Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*, Springer, New-York.

[11] Debreu, G. (1951), The coefficient of resource utilization,*Econometrica*, 19:3, 273-292.

[12] Deprins, D., Simar, L. and H. Tulkens (1984), Measuring labor inefficiency in post offices, in *The Performance of Public Enterprises: Concepts and measurements*. M. Marchand, P. Pestieau and H. Tulkens (eds.), Amsterdam, North-Holland, 243-267.

[13] Färe, R., and S. Grosskopf (2000), Theory and application of dierectional distance functions, *Journal of Productivity Analysis*, 13, 93–103.

[14] Färe, R., and S. Grosskopf (2004), *New Directions: Efficiency and Productivity*, Boston: Kluwer Academic Publishers.

[15] Färe, R., Grosskopf, S. and C.A.K. Lovell (1985), *The Measurement of Efficiency of Production*. Boston, Kluwer-Nijhoff Publishing.

[16] Färe, R., S. Grosskopf and D. Margaritis (2008), Efficiency and Productivity: Malmquist and More, in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A. Knox Lovell and Shelton Schmidt, editors, Oxford University Press.

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

[18] Florens, J.P. and L. Simar, (2005), Parametric Approximations of Nonparametric Frontier, *Journal of Econometrics*, vol 124, 1, 91–116.

[19] Gregoriou, G.N., Sedzro, K. and J. Zhu (2005), Hedge fund performance appraisal using data envelopment analysis. *European Journal of Operational Research*, 164, 555–571.

[20] Jeong, S.O. , B. U. Park and L. Simar (2010), Nonparametric conditional efficiency measures: asymptotic properties. *Annals of Operations Research*, 173, 105–122.

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

[22] Kneip, A, L. Simar and P.W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory*, 24, 1663–1697.

[23] Lovell, K.C.A. and J. Pastor (1995), Units invariant and translation invariant DEA models, *Operations Research Letters*, 18, 147–151.

[24] Park, B. Simar, L. and Ch. Weiner (2000) : The FDH Estimator for Productivity Efficiency Scores : Asymptotic Properties, *Econometric Theory*, Vol 16, 855-877.

[25] Shephard, R.W. (1970). *Theory of Cost and Production Function.* Princeton University Press, Princeton, New-Jersey.

[26] Simar, L. (2003), Detecting Outliers in Frontiers Models: a Simple Approach, *Journal of Productivity Analysis*, 20, 391–424.

[27] Simar, L. and P.W. Wilson (2008), Statistical Inference in Nonparametric Frontier Models: recent developments and perspectives, in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A. Knox Lovell and Shelton Schmidt, editors, Oxford University Press.

[28] Simar, L. and P.W. Wilson (2009), Inference by Subsampling in Nonparametric Frontier Models. Discussion paper 0933, Institut de Statistique, UCL.

[29] van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics, 3, Cambridge University Press, Cambridge.

[30] Wheelock, D.C. and P.W. Wilson (2008), Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations, *Journal of Econometrics*, 145, 209–225.

[31] Wilson, P.W. (2008), FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R, *Socio-Economic Planning Sciences* 42, 247–254.

[32] Wilson, P.W. (2010), Asymptotic Properties of some Non-parametric Hyperbolic Efficiency Estimators, in *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, Van Keilegom, I. and P.W. Wilson, editors , Berlin Physica-Verlag, forthcoming.

A Appendix: Output-oriented Directional Distances

A.1 Full frontier estimation.

We have seen in Section 2.3 that for the output oriented case, the choice of the directional vector is $g = (0, g_y)$ with $g_y > 0$ provides $D(x, y; 0, g_y) = \sup\{\beta > 0 |\ H_{XY}(x, y + \beta g_x > 0\}$. The transformation of coordinates is here given by $x^* = \exp(x)$ and $y^* = \exp(y./g_y)$, so $D(x, y; 0, g_y) = \log(\lambda(x^*, y^*))$ where $\lambda(x^*, y^*)$ is the usual radial Farrell output distance in the (x^*, y^*) coordinates.

Consider a sample of observations $\mathcal{X} = \{(X_i, Y_i)|i = 1, ..., n\}$ and a fixed value (x, y). Given the direction $(0, g_y)$, a nonparametric estimator of the output directional distance is given by

$$\widehat{D}(x, y, 0, g_y) = \log\left(\widehat{\lambda}(x^*, y^*)\right) \tag{A.1}$$

where $\widehat{\lambda}(x^*, y^*)$ is the empirical analog of the output distance computed in the transformed sample of observations $\mathcal{X}^* = \{(X_i^*, Y_i^*) = (\exp(X_i), \exp(Y_i./g_y))|i = 1, ..., n\}$ at the fixed transformed value $(x^*, y^*) = (\exp(x), \exp(y./g_y))$. More precisely, define the set of observations which dominate (x^*, y^*), $D^* = \{i\ |(X_i^*, Y_i^*) \in \mathcal{X}^*, X_i^* \leq x^*, Y_i^* \geq y^*\}$. The estimation of $\widehat{\lambda}(x^*, y^*)$ is given by:

$$\widehat{\lambda}(x^*, y^*) = \sup\{\lambda > 0 \,|\, \widehat{H}_{X^*Y^*}(x^*, \lambda y^*) > 0\} = \max_{j \in D^*}\left\{\min_{\ell=1,\ldots,q}\left(\frac{Y_j^{*,\ell}}{y^{*,\ell}}\right)\right\}. \tag{A.2}$$

A.2 Order-α quantile estimation

For the Farrell output measure, for any $\alpha \in (0, 1]$ and if x such that $F_X(x) = \text{Prob}(X \leq x) > 0$ we have (Daouia and Simar, 2007)

$$\begin{aligned}
\lambda_\alpha(x, y) &= \sup\{\lambda > 0|S_{Y|X}(x, \lambda y) > 1 - \alpha\} \\
&= \sup\{\lambda > 0|H_{XY}(x, \lambda y) > (1 - \alpha)F_X(x)\},
\end{aligned}$$

where $S_{Y|X}(x, y) = H_{XY}(x, y)/F_X(x) = \text{Prob}(Y \geq y|X \leq x)$ and where (using radial distance) we restrict the input-output as being positive. Clearly when $\alpha \to 1$, $\lambda_\alpha(x, y) \to \lambda(x, y)$.

The natural extension to order-α distance functions is obtained from its probabilistic characterization. We can define

Definition A.1. *For any points $(x, y) \in \Psi$ and for all $\alpha \in (0, 1]$ the order-α output directional distance of (x, y) with respect to the output directional vector g_y is defined as*

$$D_\alpha^{out}(x, y; g_y) = \sup\{\beta|S_{Y|X}(x, y + \beta g_y) > 1 - \alpha\}, \tag{A.3}$$

It is easy to show that, here again, $D_\alpha^{out}(x,y;g_y) \to D(x,y;0,g_y)$ when $\alpha \to 1$. The order-α distance function can also be written as $D_\alpha^{out}(x,y;g_y) = \log(\lambda_\alpha(x^*,y^*))$ where $\lambda_\alpha(x^*,y^*) = \sup\{\lambda > 0 \,|\, S_{Y^*|X^*}(x^*,\lambda y^*) > 1-\alpha\}$ is the usual order-α output distance in the (x^*,y^*) coordinates. A nonparametric estimator of order-α directional distance is thus given by

$$\widehat{D}_\alpha(x,y,0,g_y) = \log(\widehat{\lambda}_\alpha(x^*,y^*)), \tag{A.4}$$

where $\widehat{\lambda}_\alpha(x^*,y^*)$ can be computed from the algorithm described in Daouia and Simar (2007) for the output-oriented case with reference sample $\mathcal{X}^*$.

A.3 Order-m partial estimation

For a given level of inputs x in the interior of the support of X, we consider a m i.i.d. random variables Y_i, $i = 1,\dots,m$ generated by the conditional distribution function $F_{Y|X}(y|x) = \text{Prob}(Y \le y|X \le x)$ and define the random set $\widetilde{\Psi}_m^x = \{(x',y) \in \Psi | x' \le x, y \le Y_i\}$. then we can define the random distances

$$\begin{aligned} \widetilde{\beta}_m(x,y;g_y) &= \sup\{\beta|(x,y+\beta g_y) \in \widetilde{\Psi}_m^x\}, \\ &= \sup\{\beta|\widetilde{F}_{m,Y|X}(y+\beta g_y|x) < 1\}, \end{aligned} \tag{A.5}$$

where $\widetilde{F}_{m,Y|X}(y|x)$ is the empirical version of $F_{Y|X}(y|x)$ computed from the random sample of size m. Now we can formally define our order-m output directional distance.

Definition A.2. *For any point $(x,y) \in \Psi$ and for all $m \ge 1$ the order-m output directional distance of (x,y) with respect to the output directional vector g_y is defined as*

$$D_m^{out}(x,y;g_y) = \mathbb{E}(\widetilde{\beta}_m(x,y;g_y)|X \le x), \tag{A.6}$$

where we assume the expectation exists.

Note that $\widetilde{\beta}_m(x,y;g_y)$ can be equivalently written as the logarithm of a radial Farrell output measure in the coordinates (X^*,Y^*): $\widetilde{\beta}_m(x,y;g_y) = \log(\widetilde{\lambda}_m(x^*,y^*))$, where $\widetilde{\lambda}_m(x^*,y^*) = \sup\{\lambda > 0 \,|\, \widetilde{F}_{m,Y^*|X^*}(\lambda y^*|x^*) < 1\}$, where $\widetilde{F}_{m,Y^*|X^*}$ is the version of $\widetilde{F}_{m,Y|X}$ in the (X^*,Y^*) coordinates. Then, by denoting W_m^{out} the random variable $\widetilde{\lambda}_m(x^*,y^*)$ and G_m^{out} its distribution function, we have:

$$\begin{aligned} G_m^{out}(w) &= \text{Prob}(W_m^{out} \le w|X^* \le x^*) = \text{Prob}\left[\max_{j=1,\dots,m}\left\{\min_{\ell=1,\dots,q}\left(\frac{Y_j^{*,\ell}}{y^{*,\ell}}\right)\right\} \le w|X^* \le x^*\right] \\ &= \left[1 - S_{Y^*|X^*}(w\,y^*|x^*)\right]^m, \end{aligned} \tag{A.7}$$

where $S_{Y^*|X^*}(y^*|x^*) = \text{Prob}(Y^* \geq y^*|X^* \leq x^*)$. Here it would be easy to obtain the $\mathbb{E}(\widetilde{\lambda}_m(x^*, y^*)|X^* \leq x^*)$ and its nonparametric estimator (see Theorem 2.2 in Daraio and Simar, 2005) but this would be of little interest for the directional distance $D_m^{out}(x, y; g_y)$ because $\mathbb{E}(\log(\cdot))$ is not equal to $\log(\mathbb{E}(\cdot))$. But, under the assumption that the conditional expectation of $\log(W_m^{out})$ exists for any $m \geq 1$, the explicit formulation of the order-m directional distance can be obtain trough

$$D_m^{out}(x, y; g_y) = \int_0^\infty \log(w)\, dG_m^{out}(w). \tag{A.8}$$

By adapting the proof of Theorem 3.1, we could also prove that as $m \to \infty$, $D_m^{out}(x, y; g_y) \to D(x, y; 0, g_y)$.

The computation of a nonparametric estimator of the order-m output directional distance $\widehat{D}_m^{out}(x, y; g_y)$ follows the same vein with a Monte Carlo procedure adapted from Cazals et al. (2002) and Daraio and Simar (2005).

$$\begin{aligned} \widehat{D}_m^{out}(x, y; g_y) &= \widehat{\mathbb{E}}\left(\log(\tilde{\lambda}_m(x^*, y^*)|X^* \leq x^*)\right), \\ &= \int_0^\infty \log(w)\, d\widehat{G_m^{out}}(w). \end{aligned} \tag{A.9}$$

where $\widehat{G_m^{out}}(w) = \left[1 - \widehat{S}_{Y^*|X^*}(w\, y^*|x^*)\right]^m$. A Monte-Carlo algorithm for approximating the expectation may be the following:

(1) For $b = 1, ..., B$, where B is large, redo the following steps (i)-(ii):

  (i) Draw a sample of size m with replacement from those Y_i^* such that $X_i^* \leq x^*$ and denote it by $\{Y_{j,b}^*,\ j = 1, ..., m\}$

  (ii) Compute $\widetilde{\lambda}_m^b(x^*, y^*) = \max_{j=1,\ldots,m}\left\{\min_{\ell=1,\ldots,q}\left(\frac{Y_{j,b}^{*,\ell}}{y^{*,\ell}}\right)\right\}$

(2) Finally we obtain $\widehat{D}_m^{out}(x, y; g_y) \approx \frac{1}{B}\sum_{b=1}^{B} \log(\tilde{\lambda}_m^b(x^*, y^*))$.

Table 1: Different estimates of Distance Measures for 20 randomnly selected funds.

Units	$\hat{D}_n^{in}(x,y)$	$\hat{D}_{\alpha,n}^{in}(x,y)$	$\hat{D}_{m,n}^{in}(x,y)$	$\hat{D}_n^{out}(x,y)$	$\hat{D}_{\alpha,n}^{out}(x,y)$	$\hat{D}_{m,n}^{out}(x,y)$	$\hat{D}_n(x,y)$	$\hat{D}_{\alpha,n}(x,y)$	$\hat{D}_{m,n}(x,y)$
3	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0211	0.0000	-0.1361	-0.0904
99	0.5026	0.3544	0.3646	1.3739	1.3739	1.3546	0.5026	0.3033	0.3262
107	0.1611	0.0000	0.0645	0.9829	0.9829	0.9829	0.1611	0.0000	0.0548
39	0.4551	0.4317	0.3459	1.0455	1.0455	1.0260	0.4317	0.1725	0.2327
51	0.1368	0.1289	0.0844	0.5358	0.5358	0.5358	0.1368	0.0000	0.0156
121	0.2771	0.1289	0.1876	1.4109	1.4109	1.4074	0.2771	0.1289	0.1531
122	0.0000	0.0000	-0.0358	0.0000	0.0000	0.0000	0.0000	-0.1476	-0.1614
15	0.4639	0.1546	0.3137	0.8963	0.8963	0.8683	0.1546	0.1318	0.1178
123	0.0000	0.0000	-0.0034	0.0000	0.0000	0.0000	0.0000	-0.1963	-0.1446
28	0.2706	0.1224	0.1450	1.1642	1.1642	1.1642	0.2706	0.0476	0.0913
65	0.3222	0.1611	0.2146	1.7681	1.7681	1.7548	0.3222	0.1611	0.2087
56	0.0000	0.0000	-0.0583	0.0000	0.0000	0.0000	0.0000	-0.3088	-0.2683
115	0.5429	0.5429	0.4680	0.8391	0.8391	0.8251	0.3568	0.2222	0.2126
27	0.4183	0.1784	0.2398	1.2170	1.2170	1.2140	0.4183	0.1368	0.1798
6	0.3628	0.0535	0.1531	1.1909	1.1909	1.1909	0.3628	0.0119	0.0845
31	0.0000	0.0000	-0.0008	0.0000	0.0000	0.0000	0.0000	-0.0022	-0.0370
61	0.4011	0.2736	0.2678	1.2513	1.2513	1.2509	0.4011	0.2319	0.2101
45	0.4683	0.4059	0.4276	0.9001	0.9001	0.8870	0.4059	0.3554	0.3232
91	0.7256	0.5174	0.5995	1.6015	1.4234	1.4330	0.7256	0.5174	0.5592
129	0.3673	0.1130	0.2048	1.0466	1.0466	1.0447	0.2821	0.1130	0.1141
$mean_n$	0.3165	0.2261	0.2272	0.8380	0.8258	0.8113	0.2697	0.1065	0.1204

Table 2: Original data and output- and input-slacks obtained with measure $\hat{D}_n(x, y)$.

Units	Y	S_Y	X_1	S_{X_1}	X_2	S_{X_2}	X_3	S_{X_3}
3	-3.92	0.00	39.69	0.00	250.00	0.00	1.15	0.00
99	-24.74	9.23	39.31	-17.58	105.00	-78.00	1.85	-0.85
107	-25.90	2.96	34.25	-5.63	44.00	-25.00	2.62	-0.27
39	-18.71	7.93	31.19	-15.10	228.00	-67.00	1.38	-0.73
51	-17.72	2.51	38.18	-4.78	181.00	-21.23	0.80	-0.23
121	-25.42	5.09	36.87	-9.69	70.00	-43.00	1.75	-0.47
122	-10.84	0.00	19.46	0.00	50.00	0.00	1.09	0.00
15	-15.97	2.84	34.74	-5.41	144.00	-24.00	1.35	-0.26
123	-6.42	0.00	32.66	0.00	78.00	0.00	1.01	0.00
28	-20.89	4.97	32.85	-9.47	69.00	-42.00	1.42	-0.46
65	-31.98	5.92	46.91	-11.27	69.00	-50.00	2.50	-0.54
56	-7.88	0.00	15.05	0.00	161.00	0.00	0.48	0.00
115	-14.92	6.55	34.26	-12.48	126.00	-55.38	1.76	-0.60
27	-21.86	7.68	29.90	-14.63	124.00	-64.91	1.39	-0.70
6	-21.38	6.66	29.72	-12.69	104.00	-56.30	1.18	-0.61
31	-7.82	0.00	40.16	0.00	50.00	0.00	2.05	0.00
61	-22.49	7.37	29.30	-14.03	105.00	-62.25	1.55	-0.67
45	-16.04	7.45	31.65	-14.20	224.00	-63.00	2.01	-0.68
91	-25.65	13.33	40.43	-25.38	305.00	-112.61	1.96	-1.22
129	-18.73	5.18	37.74	-9.87	84.00	-43.78	1.28	-0.47
$mean_n$	-18.17	4.95	34.98	-9.43	155.19	-41.86	1.68	-0.45