ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 140 (2007) 375–400
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Nonparametric efficiency analysis: A multivariate conditional quantile approach ☆

Abdelaati Daouia[a], Léopold Simar[b,*]

[a]*GREMAQ, Université de Toulouse I and LSP, Université de Toulouse III, France*
[b]*Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium*

Available online 25 September 2006

Abstract

This paper focuses on nonparametric efficiency analysis based on robust estimation of partial frontiers in a complete multivariate setup (multiple inputs and multiple outputs). It introduces α-quantile efficiency scores. A nonparametric estimator is proposed achieving strong consistency and asymptotic normality. Then if α increases to one as a function of the sample size we recover the properties of the FDH estimator. But our estimator is more robust to the perturbations in data, since it attains a finite gross-error sensitivity. Environmental variables can be introduced to evaluate efficiencies and a consistent estimator is proposed. Numerical examples illustrate the usefulness of the approach.

JEL classification: C13; C14; D20

Keywords: Efficiency; Frontier estimation; Robust nonparametric estimators; Conditional quantiles

1. Introduction and basic concepts

Foundations of the economic theory on productivity and efficiency analysis date back to the works of Koopmans (1951) on activity analysis. Shephard (1970) proposes a modern formulation of the problem. Following these lines, we consider a production technology

☆This work was supported in part by "Actions Thématiques de l'Université Paul Sabatier", Toulouse III. Research support from the "Interuniversity Attraction Pole", Phase V (No. P5/24) from the Belgian Science Policy is acknowledged.
*Corresponding author. Tel.: +32 10 47 43 08; fax: +32 10 47 30 32.
E-mail addresses: daouia@cict.fr (A. Daouia), simar@stat.ucl.ac.be (L. Simar).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2006.07.002

where the activity of the production units is characterized by a set of inputs $x \in \mathbb{R}_+^p$ used to produce a set of outputs $y \in \mathbb{R}_+^q$. In this framework the production set is the set of technically feasible combinations of (x, y). It is defined as

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

Assumptions are usually done on this set, such as free disposability of inputs and outputs, meaning that if $(x, y) \in \Psi$, then $(x', y') \in \Psi$, as soon as[1] $x' \geqslant x$ and $y' \leqslant y$. Often convexity of Ψ is also assumed, and no free lunches (if $y \geqslant 0$ with $y \neq 0$, $(0, y) \notin \Psi$, see Shephard, 1970, for more details). The production set can be described in terms of its sections

$$\text{Input requirement sets:} \quad \forall y \in \mathbb{R}_+^q, \quad X(y) = \{x \in \mathbb{R}_+^p \mid (x, y) \in \Psi\},$$
$$\text{Output requirement sets:} \quad \forall x \in \mathbb{R}_+^p, \quad Y(x) = \{y \in \mathbb{R}_+^q \mid (x, y) \in \Psi\}.$$

As far as efficiency is of concern, the boundaries of Ψ are of interest. The efficient boundary (frontier) of Ψ is the locus of optimal production scenarios (minimal achievable input level for a given output or maximal achievable output given the input). The Farrell efficient frontier is defined in a "radial sense". It can be described in the output space:[2]

$$\begin{aligned}\forall x \in \mathbb{R}_+^p,\ Y^\partial(x) &= \{(x, y^\partial(x)) \mid y^\partial(x) \in Y(x) : \lambda y^\partial(x) \notin Y(x),\ \forall \lambda > 1\} \\ &= \{(x, y^\partial(x)) \mid y^\partial(x) \in Y(x) : (x, \lambda y^\partial(x)) \notin \Psi,\ \forall \lambda > 1\},\end{aligned}$$

where the points $y^\partial(x)$ are the maximal outputs a unit operating at the level x can produce. Finally the Farrell efficiency scores for a given production unit $(x, y) \in \Psi$, are defined as

$$\lambda(x, y) = \sup\{\lambda \mid (x, \lambda y) \in \Psi\} = \sup\{\lambda \mid \lambda y \in Y(x)\}. \tag{1}$$

We have $\lambda(x, y) \geqslant 1$ represents the proportionate increase of outputs the unit operating at level (x, y) should attain to be considered as being efficient. Here, $y^\partial(x) = \lambda(x, y)\, y$ is the radial projection of (x, y) on the frontier, in the output direction (orthogonal to the vector x). Note that for the input orientation, the Farrell efficiency scores are defined as

$$\theta(x, y) = \inf\{\theta \mid (\theta x, y) \in \Psi\} = \inf\{\theta \mid \theta x \in X(y)\} \tag{2}$$

with an analog interpretation.

In practice Ψ is unknown and so has to be estimated from a random sample of production units $\{(X_i, Y_i) \mid i = 1, \ldots, n\}$, where we assume that $\text{Prob}((X_i, Y_i) \in \Psi) = 1$ (refereed in the literature as deterministic frontier models). So the problem is related to the problem of estimating the support of the random variable (X, Y) where, for mathematical convenience, we will assume that Ψ is compact. The most popular nonparametric estimators are based on the envelopment ideas: we search for estimators of Ψ which envelops at best the observed data points. The statistical properties of these estimators are now well established (see e.g. Simar and Wilson, 2000, for a recent survey).

The most flexible nonparametric estimator, initiated by Deprins et al. (1984), is the free disposal hull (FDH) estimator. It is provided by the FDH of the sample points

$$\widehat{\Psi}_{\text{FDH}} = \{(x, y) \in \mathbb{R}_+^{p+q} \mid y \leqslant Y_i,\ x \geqslant X_i,\ i = 1, \ldots, n\}.$$

[1] From here and below inequalities between vectors $a, b \in \mathbb{R}^k$ have to be understood element by element. Writing $a \leqslant b$ means $a_i \leqslant b_i$, for $i = 1, \ldots, k$.

[2] In what follows, we will make the presentation in the output orientation and we will only give the main results for the input orientation.

The FDH efficiency scores are obtained by plugging $\widehat{\Psi}_{\mathrm{FDH}}$ in Eqs. (1) and (2) in place of the unknown Ψ. The asymptotic properties of the resulting estimators are provided by Park et al. (2000). In summary, the error of estimation converges at a rate $n^{1/(p+q)}$ to a limiting Weibull distribution. If we assume that Ψ is convex, the convex hull of $\widehat{\Psi}_{\mathrm{FDH}}$ provides the data envelopment analysis (DEA) estimator of Ψ introduced by Farrell (1957). Its error of estimation converges at a rate $n^{2/(p+q+1)}$ to a nondegenerate distribution.

The FDH/DEA estimators envelop all the data points and so are very sensitive to outliers and/or to extreme values. Cazals et al. (2002) have introduced the concept of partial frontiers (order-m frontiers) with a nonparametric estimator which does not envelop all the data points. It is shown that by selecting the value of m as an appropriate function of n, the estimator of the partial order-m efficiency scores provides a robust estimator of the full Farrell efficiency scores sharing the same asymptotic properties as the FDH estimators but being less sensitive to outliers and/or extreme values. These properties have been investigated from the robustness perspective by Daouia and Ruiz Gazen (2006).

Recently Aragon et al. (2005) have proposed an alternative to order-m partial frontiers by introducing quantile based partial frontiers. The idea is to replace this concept of "discrete" order-m partial frontier by a "continuous" order-α partial frontier where $\alpha \in [0, 1]$ corresponds to the level of an appropriate non-standard conditional quantile frontier. A nonparametric estimator of the frontier is proposed which shares similar properties than the order-m estimators. As pointed out in Aragon et al. (2005) and in Daouia and Ruiz Gazen (2006), partial frontiers based on α-quantile estimators have better robustness properties than the ones based on the order-m estimators.

Unlike the order-m partial frontiers, due to the absence of natural ordering of Euclidean spaces for dimension greater than one, the α-quantile approach is limited to one-dimensional input for the input oriented frontier and to one-dimensional output for the output oriented frontier. In this paper, we overcome this difficulty and we propose an extension to the full multivariate case, introducing the concept of α-quantile efficiency scores and the corresponding α-quantile frontier set. We provide the asymptotic properties of our estimator, we investigate its robustness characteristics and show how to introduce environmental factors.

The paper is organized as follows: in the next section, we reformulate the concept of frontier and efficiency in a probabilistic framework, in the lines of Daraio and Simar (2006). We characterize the monotonicity properties of the resulting efficiency scores and show that in the case of free disposability of Ψ, these scores coincide with the Farrell efficiency scores defined above. Due to this unifying presentation it is then easy to define a concept of α-quantile efficiency scores in a full multivariate framework and to investigate its properties. Section 3 analyzes the asymptotic properties of the corresponding nonparametric estimators and Section 4 investigates their reliability from the robustness theory point of view. Then Section 5 shows how environmental variables can be introduced to evaluate efficiencies and analyzes the properties of the resulting estimator. Section 6 illustrates with some numerical examples and Section 7 concludes.

2. Multivariate quantile-type efficiency scores

2.1. Probabilistic formulation

Daraio and Simar (2006), extending previous works of Cazals et al. (2002), propose a probabilistic formulation of efficiency concepts. The data generating process (DGP) of

(X, Y) is completely characterized by the knowledge of the probability function $H_{XY}(\cdot,\cdot)$ defined as

$$H_{XY}(x,y) = \text{Prob}(X \leqslant x,\, Y \geqslant y).$$

The support of $H_{XY}(\cdot,\cdot)$ is Ψ and $H_{XY}(x,y)$ can be interpreted as the probability for a unit operating at the level (x,y) to be dominated. Note that this function is a nonstandard distribution function, having a cumulative distribution form for X and a survival form for Y. So $H_{XY}(\cdot,\cdot)$ is monotone nondecreasing with x and monotone nonincreasing[3] with y.

This joint probability can be decomposed as follows:

$$\begin{aligned} H_{XY}(x,y) &= \text{Prob}(X \leqslant x \mid Y \geqslant y)\,\text{Prob}(Y \geqslant y) = F_{X|Y}(x|y)\,S_Y(y) \\ &= \text{Prob}(Y \geqslant y \mid X \leqslant x)\,\text{Prob}(X \leqslant x) = S_{Y|X}(y|x)\,F_X(x), \end{aligned}$$

where we suppose the conditional probabilities exits (i.e., when needed, $F_X(x)>0$ or $S_Y(y)>0$). Note that the conditional distribution $F_{X|Y}$ and the conditional survival $S_{Y|X}$ are nonstandard due to the event describing the condition. We can now define, as in Daraio and Simar (2006) efficiency scores in terms of the support of these probabilities. For the output oriented case, for all x such that $F_X(x)>0$, we define the output efficiency score as

$$\widetilde{\lambda}(x,y) = \sup\{\lambda \mid S_{Y|X}(\lambda y|x)>0\} = \sup\{\lambda \mid H_{XY}(x,\lambda y)>0\}. \tag{3}$$

This output efficiency score can be interpreted as the proportionate increase of outputs a unit working at the level (x,y) should perform to be dominated with probability zero. These efficiency scores share the following properties.

Proposition 2.1. *Whenever defined, $\widetilde{\lambda}(x,y)$ is monotone nondecreasing with x and monotone nonincreasing with y.*

If we now define $\widetilde{y}^{\partial}(x) = \widetilde{\lambda}(x,y)\,y$, for fixed y, this is monotone nondecreasing with x. This result can be seen as a multivariate extension of Theorem 2.1 of Cazals et al. (2002). The efficient frontier, according to this probabilistic definition of efficiency, can be described for all x such that $F_X(x)>0$ by the set $\{(x, \widetilde{\lambda}(x,y)\,y) \mid (x,y) \in \Psi\}$. By construction and by Proposition 2.1, the set bounded by this frontier is the FDH of Ψ. If Ψ is free disposal, the two sets coincide and $\widetilde{\lambda}(x,y) = \lambda(x,y)$. The same could be done in the input orientation. From now on, we will assume that Ψ is free disposal.

Natural nonparametric estimators of $\theta(x,y)$ and of $\lambda(x,y)$ are obtained by plugging $\widehat{H}_{XY,n}$ in place of H_{XY} in the definition of the efficiency scores. Defining

$$\widehat{H}_{XY,n}(x,y) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \leqslant x,\, Y_i \geqslant y)$$

the most natural nonparametric estimators of the efficiency scores are given for the input orientation,[4] by

$$\widehat{\theta}_n(x,y) = \inf\{\theta \mid \widehat{F}_{X|Y,n}(\theta x|y)>0\} = \min_{i|Y_i \geqslant y}\ \max_{j=1,\ldots,p} X_i^j/x^j,$$

[3]A function f from $\mathbb{R}^k$ to $\mathbb{R}$ is monotone nonincreasing if $a_1 \leqslant a_2$ implies $f(a_1) \geqslant f(a_2)$. We will say that f is monotone decreasing, if $a_1 \leqslant a_2$ and $a_1 \neq a_2$ implies $f(a_1) > f(a_2)$.

[4]For a vector $a \in \mathbb{R}^k$, we denote by a^j its jth component.

where $\widehat{F}_{X|Y,n}(x|y) = \widehat{H}_{XY,n}(x,y)/\widehat{H}_{XY,n}(\infty,y)$. For the output oriented case we have

$$\widehat{\lambda}_n(x,y) = \sup\{\lambda \mid \widehat{S}_{Y|X,n}(\lambda y|x) > 0\} = \max_{i|X_i \leqslant x} \min_{j=1,\ldots,q} Y_i^j/y^j,$$

where $\widehat{S}_{Y|X,n}(y|x) = \widehat{H}_{XY,n}(x,y)/\widehat{H}_{XY,n}(x,0)$. As pointed out in Daraio and Simar (2006), these estimators are the FDH estimators of the Farrell efficiency scores. Note also that the FDH efficiency scores share the properties of Proposition 2.1.

2.2. Conditional quantile-based efficiency scores

Aragon et al. (2005) have introduced the conditional quantile frontier function for a production (output) function when the output is unidimensional and for a cost (input) function when the input is one-dimensional. We extend the ideas to a full multivariate setup. Since a natural ordering of Euclidean spaces of dimension greater than one does not exist, we overcome the difficulty by defining α-quantile efficiency scores as follows.

Definition 2.1. For all y such that $S_Y(y) > 0$ and for $\alpha \in]0,1]$, the α-quantile input efficiency score for the unit $(x,y) \in \Psi$ is defined as

$$\theta_\alpha(x,y) = \inf\{\theta \mid F_{X|Y}(\theta x|y) > 1-\alpha\}.$$

For all x such that $F_X(x) > 0$ and for $\alpha \in]0,1]$, the α-quantile output efficiency score for the unit $(x,y) \in \Psi$ is defined as

$$\lambda_\alpha(x,y) = \sup\{\lambda \mid S_{Y|X}(\lambda y|x) > 1-\alpha\}. \tag{4}$$

For instance, in the output direction, $\lambda_\alpha(x,y)$ is the proportionate reduction (if <1) or increase (if >1) of outputs, a unit working at the level (x,y) should perform to be dominated by firms using less input than the level x with probability $1-\alpha$. Clearly when $\alpha = 1$, this is, under free disposability of Ψ, the Farrell output efficiency score and roughly said, $\lambda_\alpha(x,y)$ is the output efficiency score of (x,y) at the level $\alpha \times 100\%$.

We can now for all x such that $F_X(x) > 0$ define the α-quantile efficient frontier in the output direction as the set

$$Y_\alpha^\partial(x) = \{(x, \lambda_\alpha(x,y)y) \mid (x,y) \in \Psi\}.$$

The points $y_\alpha^\partial(x) = \lambda_\alpha(x,y)y$ represent the efficient outputs for the input x at the level $\alpha \times 100\%$ and the pairs $(x, y_\alpha^\partial(x))$ have a probability $H_{XY}(x, y_\alpha^\partial(x)) = (1-\alpha)F_X(x) \leqslant 1-\alpha$ of being dominated if $H_{XY}(\cdot,\cdot)$ is continuous on Ψ.

Note that in the particular case of $q = 1$, for any x such that $F_X(x) > 0$, the output efficient frontier at the level $\alpha \times 100\%$ may be described as the set $Y_\alpha^\partial(x) = \{(x, \varphi(x))\}$ where $\varphi_\alpha(x) = \lambda_\alpha(x,1)$ is the conditional quantile production function of order α of Aragon et al. (2005).

As shown below, the α-quantile efficiency scores share most of the properties of their univariate correspondent. We remind here that Ψ is assumed to be free disposal. We define $\Psi^* = \{(x,y) \in \Psi \mid 0 < H_{XY}(x,y) < S_Y(y) \wedge F_X(x)\}$ as being the interior of Ψ.

Proposition 2.2. *Assume that $F_{X|Y}$ is continuous and monotone increasing in x and that $S_{Y|X}$ is continuous and monotone decreasing in y. Then, for all $(x,y) \in \Psi^*$, there exist α and β*

in $]0,1]$ *such that*

$$\theta_\alpha(x,y)=1 \quad \textit{where } \alpha=1-F_{X|Y}(x|y),$$
$$\lambda_\beta(x,y)=1 \quad \textit{where } \beta=1-S_{Y|X}(y|x).$$

Proposition 2.2 shows that any point (x,y) in the interior of Ψ, belongs to an appropriate α-quantile efficient frontier in both directions (input and output). For instance in the output orientation, it can be described as the set $Y_\alpha^{\partial}(x)$ where $\alpha=1-S_{Y|X}(y|x)$. Since for any (x,y) belonging to the efficient frontier of Ψ, $\theta(x,y)=\theta_1(x,y)=1$ and $\lambda(x,y)=\lambda_1(x,y)=1$, we can use the value of $\alpha=\alpha(x,y)$ and of $\beta=\beta(x,y)$ of the proposition to define a new concept of input and output efficiency score. This is in the same spirit as in Aragon et al. (2003), this will not be pursued here.

Proposition 2.3. *For all y such that* $S_Y(y)>0$, *we have* $\lim_{\alpha\to 1}\searrow \theta_\alpha(x,y)=\theta(x,y)$ *and for all x such that* $F_X(x)>0$, $\lim_{\alpha\to 1}\nearrow \lambda_\alpha(x,y)=\lambda(x,y)$.

The α-quantile input efficiency score $\theta_\alpha(x,y)$ is clearly monotone nonincreasing with x but it is in general not monotone in y, unless we add an assumption on $F_{X|Y}$.

Proposition 2.4. *Assume that* $F_{X|Y}(\cdot|y)$ *is continuous for any y. Then, for points* (x,y) *such that* $F_{X|Y}(x|y)<1$, *the two following properties are equivalent.*

$$F_{X|Y}(x|y) \textit{ is monotone nonincreasing with } y, \tag{5}$$
$$\theta_\alpha(x,y) \textit{ is monotone nondecreasing with } y \textit{ for all } \alpha. \tag{6}$$

Note that both conditions of the proposition are quite reasonable in production analysis. The first relation (5) says that there is less probability to observe a level of input lower than a fixed value x for firms producing more than a level y_2, than for firms producing more than a level $y_1\leqslant y_2$. It is more difficult to reduce inputs when producing higher level of outputs. Whereas (6) states that everything else being kept constant, the α-quantile input efficiency score cannot decrease when the output increases.

Of course, *mutatis mutandis*, we have the same property in the output direction. $\lambda_\alpha(x,y)$ is monotone nonincreasing with y, but for the monotonicity with respect to x, we have:

Proposition 2.5. *Assume that* $S_{Y|X}(\cdot|x)$ *is continuous for any x. Then, for points* (x,y) *such that* $S_{Y|X}(y|x)<1$, *the two following properties are equivalent.*

$$S_{Y|X}(y|x) \textit{ is monotone nondecreasing with } x,$$
$$\lambda_\alpha(x,y) \textit{ is monotone nondecreasing with } x \textit{ for all } \alpha.$$

3. Nonparametric estimator

A natural nonparametric estimator of the α-quantile efficiency scores is obtained by plugging the empirical $\widehat{H}_{XY,n}(x,y)$ in the above formulas so we have:

$$\widehat{\theta}_{\alpha,n}(x,y)=\inf\{\theta \mid \widehat{F}_{X|Y,n}(\theta x|y)>1-\alpha\},$$
$$\widehat{\lambda}_{\alpha,n}(x,y)=\sup\{\lambda \mid \widehat{S}_{Y|X,n}(\lambda y|x)>1-\alpha\},$$

where $\widehat{F}_{X|Y,n}$ and $\widehat{S}_{Y|X,n}$ were defined in Section 2.

These nonparametric estimators can be computed very easily. Indeed, define

$$\mathscr{Y}_i = \min_{k=1,\ldots,q} \frac{Y_i^k}{y^k}, \quad i = 1, \ldots, n$$

and let $N_x = n\widehat{H}_{XY,n}(x, 0)$ be nonnull. For $j = 1, \ldots, N_x$, denote by $\mathscr{Y}^x_{(j)}$ the jth order statistic of the observations $\mathscr{Y}_i$ such that $X_i \leqslant x$: $\mathscr{Y}^x_{(1)} \leqslant \mathscr{Y}^x_{(2)} \leqslant \cdots \leqslant \mathscr{Y}^x_{(N_x)}$. Then

$$\widehat{\lambda}_{\alpha,n}(x, y) = \begin{cases} \mathscr{Y}^x_{(\alpha N_x)} & \text{if } \alpha N_x \in \mathbb{N}^*, \\ \mathscr{Y}^x_{([\alpha N_x]+1)} & \text{otherwise,} \end{cases} \tag{7}$$

where $\mathbb{N}^*$ denotes the set of positive integers and $[\alpha N_x]$ denotes the integral part of αN_x.

Likewise, let $M_y = n\widehat{H}_{XY,n}(\infty, y) > 0$, and define

$$\mathscr{X}_i = \max_{k=1,\ldots,p} \frac{X_i^k}{x^k}, \quad i = 1, \ldots, n.$$

For $j = 1, \ldots, M_y$, denoted by $\mathscr{X}^y_{(j)}$ the jth order statistic of the observations $\mathscr{X}_i$ such that $Y_i \geqslant y$: $\mathscr{X}^y_{(1)} \leqslant \mathscr{X}^y_{(2)} \leqslant \cdots \leqslant \mathscr{X}^y_{(M_y)}$. Then

$$\widehat{\theta}_{\alpha,n}(x, y) = \begin{cases} \mathscr{X}^y_{((1-\alpha)M_y)} & \text{if } (1-\alpha)M_y \in \mathbb{N}, \\ \mathscr{X}^y_{([(1-\alpha)M_y]+1)} & \text{otherwise,} \end{cases}$$

where $\mathbb{N}$ denotes the set of all nonnegative integers.

The nonparametric α-quantile efficiency scores $\widehat{\theta}_{\alpha,n}(x, y)$ and $\widehat{\lambda}_{\alpha,n}(x, y)$ share the following properties:

Proposition 3.1. *For all y such that $\widehat{H}_{XY,n}(\infty, y) > 0$, we have $\lim_{\alpha\to 1} \searrow \widehat{\theta}_{\alpha,n}(x, y) = \widehat{\theta}_n(x, y)$ and for all x such that $\widehat{H}_{XY,n}(x, 0) > 0$, $\lim_{\alpha\to 1} \nearrow \widehat{\lambda}_{\alpha,n}(x, y) = \widehat{\lambda}_n(x, y)$.*

Now we investigate some of the asymptotic properties of our estimators. In what follows, we limit the presentation for the nonparametric estimator in the output oriented case. The same properties hold for the input oriented case.

Theorem 3.1. *Let $(x, y) \in \Psi$ be such that $F_X(x) > 0$ and let $0 < \alpha < 1$. Assume that $\lambda \mapsto S_{Y|X}(\lambda y|x)$ is decreasing in a neighborhood of $\lambda_\alpha(x, y)$. Then, for every $\varepsilon > 0$,*

$$\text{Prob}(|\widehat{\lambda}_{\alpha,n}(x, y) - \lambda_\alpha(x, y)| > \varepsilon) \leqslant 2\mathrm{e}^{-2n\delta^2_{\alpha,\varepsilon,x,y}} \quad \textit{for all } n \geqslant 1,$$

where

$$\delta_{\alpha,\varepsilon,x,y} = \frac{F_X(x)}{2-\alpha} \min\{(1-\alpha) - S_{Y|X}((\lambda_\alpha(x, y) + \varepsilon)y|x); \\ S_{Y|X}((\lambda_\alpha(x, y) - \varepsilon)y|x) - (1-\alpha)\}.$$

Proof. Let $\varepsilon > 0$. We have $\text{Prob}(|\widehat{\lambda}_{\alpha,n}(x, y) - \lambda_\alpha(x, y)| > \varepsilon) = \text{Prob}(\widehat{\lambda}_{\alpha,n}(x, y) > \lambda_\alpha(x, y) + \varepsilon) + \text{Prob}(\widehat{\lambda}_{\alpha,n}(x, y) < \lambda_\alpha(x, y) - \varepsilon)$. By applying the fact that $\widehat{\lambda}_{\alpha,n}(x, y) > \lambda$ implies $\widehat{S}_{Y|X,n}(\lambda y|x) > 1 - \alpha$, we get $\text{Prob}(\widehat{\lambda}_{\alpha,n}(x, y) > \lambda_\alpha(x, y) + \varepsilon) \leqslant \text{Prob}(\sum_{i=1}^n V_i - \sum_{i=1}^n E(V_i) > n\delta_1)$, where $V_i = \mathbb{1}(X_i \leqslant x, Y_i \geqslant (\lambda_\alpha(x, y) + \varepsilon)y) - (1-\alpha)\mathbb{1}(X_i \leqslant x)$ and $\delta_1 = -E(V_1) = F_X(x)[(1-\alpha) - S_{Y|X}((\lambda_\alpha(x, y) + \varepsilon)y|x)] > 0$. Since $\text{Prob}(\alpha - 1 \leqslant V_i \leqslant 1) = 1$, for each i, we obtain by

applying Hoeffding's Inequality (Hoeffding, 1963),

$$\mathrm{Prob}(\widehat{\lambda}_{\alpha,n}(x,y) > \lambda_\alpha(x,y) + \varepsilon) \leqslant \mathrm{e}^{-2n\delta_1^2/(2-\alpha)^2}.$$

Likewise, by applying the fact that $\widehat{\lambda}_{\alpha,n}(x,y) < \lambda$ implies $\widehat{S}_{Y|X,n}(\lambda y|x) \leqslant 1-\alpha$, we get $\mathrm{Prob}(\widehat{\lambda}_{\alpha,n}(x,y) < \lambda_\alpha(x,y) - \varepsilon) \leqslant \mathrm{e}^{-2n\delta_2^2/(2-\alpha)^2}$, where $\delta_2 = F_X(x)[S_{Y|X}((\lambda_\alpha(x,y)-\varepsilon)y|x) - (1-\alpha)] > 0$. Putting $\delta_{\alpha,\varepsilon,x,y} = \min\{\delta_1;\delta_2\}/(2-\alpha)$, the proof is complete. □

Thus $\mathrm{Prob}(|\widehat{\lambda}_{\alpha,n}(x,y) - \lambda_\alpha(x,y)| > \varepsilon) \to 0$ exponentially fast, which implies that $\widehat{\lambda}_{\alpha,n}(x,y)$ converges completely to $\lambda_\alpha(x,y)$. This generalizes the exponential probability inequality obtained in Daouia (2005, see Theorem 2.3) for the nonparametric α-quantile frontier $\widehat{y}^{\partial}_{\alpha,n}(x) = \widehat{\lambda}_{\alpha,n}(x,y)y$ in the univariate case where $p \geqslant 1$ and $q = 1$.

We also obtain the following asymptotic normality result which extends the one established in Aragon et al. (2005, see Theorem 4.1) to the more general case where $p, q \geqslant 1$.

Theorem 3.2. *Let $0<\alpha<1$ be a fixed order and let $(x,y) \in \Psi$ be a fixed unit such that $F_X(x) > 0$. Assume that $G(\lambda) = S_{Y|X}(\lambda y|x)$ is differentiable at $\lambda_\alpha(x,y)$ with negative derivative $G'(\lambda_\alpha(x,y))$. Then,*

$$\sqrt{n}(\widehat{\lambda}_{\alpha,n}(x,y) - \lambda_\alpha(x,y)) \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_\alpha^2(x,y)) \quad \textit{as } n \to \infty,$$

where $\sigma_\alpha^2(x,y) = \alpha(1-\alpha)/[G'(\lambda_\alpha(x,y))]^2 F_X(x)$.

Note that this theorem requires a slightly stronger hypothesis on the function $G(\lambda) = S_{Y|X}(\lambda y|x)$ than in the preceding theorem. This assumption is standard in quantile theory for the generalized inverse of the cdf $S_{Y|X}(\lambda y|x)$ to coincide with the reciprocal.

Proof. Let $V_n = \sqrt{n}(\widehat{\lambda}_{\alpha,n}(x,y) - \lambda_\alpha(x,y))$ and write

$$W_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{(1-\alpha)\mathbb{1}(X_i \leqslant x) - \mathbb{1}(X_i \leqslant x, Y_i \geqslant \lambda_\alpha(x,y)y)}{G'(\lambda_\alpha(x,y))F_X(x)}.$$

By using the fact that $H_{XY}(x, \lambda_\alpha(x,y)y) = (1-\alpha)F_X(x)$, we obtain in view of the central limit theorem that $W_n \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_\alpha^2(x,y))$. To show that V_n has the same asymptotic normal distribution, it suffices to prove that

$$R_n = V_n - W_n \xrightarrow{\mathrm{p}} 0. \tag{8}$$

Using $\widehat{\lambda}_{\alpha,n}(x,y) \geqslant \lambda \Leftrightarrow \widehat{S}_{Y|X,n}(\lambda y|x) > 1-\alpha$ (this can be easily proved from the definition of $\widehat{\lambda}_{\alpha,n}$ and by using the left-continuity of $\lambda \mapsto \widehat{S}_{Y|X,n}(\lambda y|x)$), we get for any real t

$$\{V_n \geqslant t\} = \left\{ \widehat{S}_{Y|X,n}\left(\left(\lambda_\alpha(x,y) + \frac{t}{\sqrt{n}} \right) y|x \right) > 1-\alpha \right\} = \{Z_{t,n} > T_n\}, \tag{9}$$

where

$$Z_{t,n} = \frac{\sqrt{n}\widehat{H}_{XY,n}(x,0)}{G'(\lambda_\alpha(x,y))F_X(x)} \left[G\left(\lambda_\alpha(x,y) + \frac{t}{\sqrt{n}} \right) - \widehat{S}_{Y|X,n}\left(\left(\lambda_\alpha(x,y) + \frac{t}{\sqrt{n}} \right) y|x \right) \right],$$

$$T_n = \frac{\sqrt{n}\widehat{H}_{XY,n}(x,0)}{G'(\lambda_\alpha(x,y))F_X(x)} \left[G\left(\lambda_\alpha(x,y) + \frac{t}{\sqrt{n}} \right) - (1-\alpha) \right].$$

Since $G\left(\lambda_\alpha(x,y)+\frac{t}{\sqrt{n}}\right)-(1-\alpha)=\frac{t}{\sqrt{n}}G'(\lambda_\alpha(x,y))+\frac{t}{\sqrt{n}}o(1)$ and $\widehat{H}_{XY,n}(x,0)\xrightarrow{\mathrm{p}}F_X(x)$, as $n\to\infty$, we obtain

$$T_n\xrightarrow{\mathrm{p}}t \quad \text{as } n\to\infty. \tag{10}$$

We also have

$$Z_{t,n}-W_n=\frac{\sqrt{n}\widehat{H}_{XY,n}(x,0)}{G'(\lambda_\alpha(x,y))F_X(x)}\left[\left(G\left(\lambda_\alpha(x,y)+\frac{t}{\sqrt{n}}\right)-\widehat{S}_{Y|X,n}\left(\left(\lambda_\alpha(x,y)+\frac{t}{\sqrt{n}}\right)y|x\right)\right)\right.$$
$$\left.-((1-\alpha)-\widehat{S}_{Y|X,n}((\lambda_\alpha(x,y)y|x))\right].$$

An easy computation shows that

$$\begin{aligned}&\mathrm{E}[(Z_{t,n}-W_n)^2]\{G'(\lambda_\alpha(x,y))F_X(x)\}^2\\&\quad=(1-\alpha)F_X(x)-(1-\alpha)^2F_X(x)\\&\qquad-F_X(x)G^2\left(\lambda_\alpha(x,y)+\frac{t}{\sqrt{n}}\right)+(1+2(1-\alpha))F_X(x)G\left(\lambda_\alpha(x,y)+\frac{t}{\sqrt{n}}\right)\\&\qquad-2F_X(x)G\left(\lambda_\alpha(x,y)+\left(\frac{t}{\sqrt{n}}\vee 0\right)\right).\end{aligned}$$

Using the continuity of $G(\cdot)$ in $\lambda_\alpha(x,y)$, we then obtain $\mathrm{E}[(Z_{t,n}-W_n)^2]\to 0$, and so

$$Z_{t,n}-W_n\xrightarrow{\mathrm{p}}0 \quad \text{as } n\to\infty. \tag{11}$$

To prove (8), it suffices to show that $\{V_n\}$ and $\{W_n\}$ satisfy the two conditions of Ghosh (1971, Lemma 1, p. 1958). Since W_n converges in law in view of the central limit theorem, it is uniformly tight and thus the first Ghosh's condition is satisfied. On the other hand, for any k and any $\varepsilon>0$, putting $t=k+\varepsilon$, we obtain in view of (9),

$$\begin{aligned}\mathrm{Prob}(V_n\geqslant k+\varepsilon, W_n\leqslant k)&=\mathrm{Prob}(Z_{t,n}>T_n, W_n\leqslant t-\varepsilon)\\&\leqslant\mathrm{Prob}(|(Z_{t,n}-W_n)-(T_n-t)|\geqslant\varepsilon).\end{aligned}$$

Then it follows immediately from (10) to (11) that $\lim_{n\to\infty}\mathrm{Prob}(V_n\geqslant k+\varepsilon, W_n\leqslant k)=0$. Similarly, by applying (9) to $t=k$, we get

$$\mathrm{Prob}(V_n<k, W_n\geqslant k+\varepsilon)\leqslant\mathrm{Prob}(|(W_n-Z_{t,n})+(T_n-t)|\geqslant\varepsilon)\longrightarrow 0 \quad \text{as } n\to\infty,$$

and so the second Ghosh's condition is also satisfied, which completes the proof. □

Note that this convergence result can be extended to the analysis of the asymptotic properties of a vector $(\sqrt{n}(\widehat{\lambda}_{\alpha,n}(x^1,y^1)-\lambda_\alpha(x^1,y^1)),\ldots,\sqrt{n}(\widehat{\lambda}_{\alpha,n}(x^r,y^r)-\lambda_\alpha(x^r,y^r)))$. We still have the asymptotic r-variate normal distribution with asymptotic covariances given by $\Sigma_{k,l}=\mathrm{E}[\Gamma(x^k,y^k,X,Y)\Gamma(x^l,y^l,X,Y)]$, where

$$\Gamma(x,y,X,Y)=\frac{(1-\alpha)\mathbb{1}(X\leqslant x)-\mathbb{1}(X\leqslant x,\ Y\geqslant\lambda_\alpha(x,y)y)}{G'(\lambda_\alpha(x,y))F_X(x)}. \tag{12}$$

The expression of the variance factors can be used to derive asymptotic confidence intervals for the order-α efficiency scores. For instance, consistent estimators for the factors $\sigma_\alpha^2(x,y)$ and $\Sigma_{k,l}$ can be obtained by plugging nonparametric estimators for $G'(\lambda)$ and $F_X(x)$ and taking the empirical mean for the expectation.

A more robust estimator of the Farrell efficiency scores $\lambda(x,y)$ than the standard FDH estimator $\widehat{\lambda}_n(x,y)$, which however shares the same asymptotic distribution with this later one, can be derived as follows.

Lemma 3.1. *Assume that the support of Y is bounded. Then, for any $(x,y)\in\Psi$,*

$$n^{1/(p+q)}(\widehat{\lambda}_n(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y))\xrightarrow{\text{a.s.}}0 \quad \text{as } n\to\infty,$$

where the order $\alpha(n)>0$ is such that: $n^{(p+q+1)/(p+q)}(1-\alpha(n))\to 0$ as $n\to\infty$.

Proof. We have from (7),

$$\begin{aligned}\widehat{\lambda}_n(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y) &= (\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{(\alpha(n)N_x)})\mathbb{1}(\alpha(n)N_x\in\mathbb{N}^*)\\ &\quad+(\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{([\alpha(n)N_x]+1)})\mathbb{1}(\alpha(n)N_x\notin\mathbb{N}^*).\end{aligned}$$

Let $C_{x,y,k}(n)=\frac{\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{(k)}}{1-k/N_x}$ for $k\in\{1,\ldots,N_x-1\}$, and let $C_{x,y}(n)=\max_{1\leqslant k\leqslant N_x-1}C_{x,y,k}(n)$. It can be then easily seen that

$$\begin{aligned}&(\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{(\alpha(n)N_x)})\mathbb{1}(\alpha(n)N_x\in\mathbb{N}^*)\leqslant C_{x,y}(n)(1-\alpha(n))\mathbb{1}(\alpha(n)N_x\in\mathbb{N}^*),\\ &(\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{([\alpha(n)N_x]+1)})\mathbb{1}(\alpha(n)N_x\notin\mathbb{N}^*)\leqslant C_{x,y}(n)(1-\alpha(n))\mathbb{1}(\alpha(n)N_x\notin\mathbb{N}^*)\end{aligned}$$

which gives $n^{1/(p+q)}(\widehat{\lambda}_n(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y))\leqslant n^{1/(p+q)}C_{x,y}(n)(1-\alpha(n))$. Since the support of Y is bounded, there exists a constant $M_y>0$ (depending on y) such that $\mathscr{Y}_i\leqslant M_y$ almost surely, for any $i=1,\ldots,n$. Hence, $\mathscr{Y}^x_{(N_x)}-\mathscr{Y}^x_{(k)}\leqslant\mathscr{Y}^x_{(N_x)}\leqslant M_y$ almost surely, for any $k=1,\ldots,N_x-1$. Using the fact that $\frac{1}{1-k/N_x}\leqslant N_x$, we therefore obtain $C_{x,y}(n)\leqslant M_yN_x$ almost surely, and so

$$\begin{aligned}n^{1/(p+q)}(\widehat{\lambda}_n(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y)) &\leqslant n^{1/(p+q)}M_yN_x(1-\alpha(n))\\ &= M_y\widehat{H}_{XY,n}(x,0)n^{(p+q+1)/(p+q)}(1-\alpha(n))\end{aligned}$$

almost surely. The conclusion follows by applying the strong law of large numbers. □

Making use of this lemma and the following decomposition

$$\begin{aligned}&n^{1/(p+q)}(\lambda(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y))\\ &\quad= n^{1/(p+q)}(\lambda(x,y)-\widehat{\lambda}_n(x,y))+n^{1/(p+q)}(\widehat{\lambda}_n(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y))\end{aligned}$$

we get immediately from Corollary 3.2 of Park et al. (2000) the following result.

Theorem 3.3. *Under Assumptions* AI–AIII *of* Park et al. (2000), *we have for any (x,y) interior to Ψ,*

$$n^{1/(p+q)}(\lambda(x,y)-\widehat{\lambda}_{\alpha(n),n}(x,y))\xrightarrow{\mathscr{L}}\text{Weibull}(\mu^{p+q}_{\text{NW},0},p+q) \quad \text{as } n\to\infty,$$

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

An explicit expression of the Weibull parameter $\mu_{\text{NW},0}$ is given in Park et al. (2000, see Definition A.2 of the appendix). A consistent estimator of this unknown parameter is also provided (see Park et al., 2000, Theorem 3.4).

The nonparametric conditional quantile efficiency scores $\widehat{\lambda}_{\alpha(n),n}(x,y)$ lead to an estimator of the full frontier $Y^\partial(x)$. For all x such that $\widehat{H}_{XY,n}(x,0)>0$, we have

$$\widehat{Y}^\partial_{\alpha(n),n}(x)=\{(x,\widehat{\lambda}_{\alpha(n),n}(x,y)\,y)\mid(x,y)\in\widehat{\Psi}_{FDH}\}.$$

This estimator does not envelop all the observed data points and so, is more resistant to extremes than the usual nonparametric envelopment estimators (FDH, DEA).

4. Robustness properties

The most popular nonparametric estimators (FDH,DEA) of the Farrell technical efficiency are not robust to the perturbations in data. A robust estimator has been suggested recently by Cazals et al. (2002). It is based on a concept of expected order-m efficiency scores, where m is a positive integer. For instance, in the output direction, in place of looking for the full upper boundary of $Y(x)$, as it is the case for defining $\lambda(x,y)$, the partial order-m efficiency score $\lambda_m(x,y)$ can roughly be viewed as the expectation of the maximal output efficiency score of the unit (x,y), when compared to m units randomly drawn from the population of units using less inputs than the level x (see, e.g., Definition 2.2 and Theorem 2.2 of Daraio and Simar, 2006). In this section, we analyze and compare the reliability of the nonparametric estimators based on the two concepts of order-m and order-α efficiency measures from a theoretical point of view.

Let us start by the robustness properties of the order-α efficiency scores. From now on we only focus on the output oriented case to save place. The same presentation can be done in the input direction. The estimators $\widehat{\lambda}_{\alpha,n}(x,y)$ are representable as a functional T^{α}_{xy} of the empirical version of the probability function H_{XY} which characterizes completely the DGP:

$$\lambda_\alpha(x,y) = \sup\left\{\lambda \,\middle|\, \frac{H_{XY}(x,\lambda y)}{H_{XY}(x,0)} > 1-\alpha\right\} = T^{\alpha}_{xy}(H_{XY}),$$

$$\widehat{\lambda}_{\alpha,n}(x,y) = \sup\left\{\lambda \,\middle|\, \frac{\widehat{H}_{XY,n}(x,\lambda y)}{\widehat{H}_{XY,n}(x,0)} > 1-\alpha\right\} = T^{\alpha}_{xy}(\widehat{H}_{XY,n}).$$

Therefore, the reliability of $\{\widehat{\lambda}_{\alpha,n}(x,y)\}_n$ in estimating $\lambda_\alpha(x,y)$ can be analyzed from a robustness theory point of view. The richest robustness information is provided by the influence function $(X_i, Y_i) \mapsto IF((X_i, Y_i); T^{\alpha}_{xy}, H_{XY})$ of T^{α}_{xy} at H_{XY} (Hampel, 1974). It is defined as the first Gâteaux derivative of T^{α}_{xy} at H_{XY} in the direction of $\Delta_{X_iY_i}(\cdot,\cdot) = \mathbb{1}(X_i \leqslant \cdot, Y_i \geqslant \cdot)$. Formally, $IF((X_i,Y_i); T^{\alpha}_{xy}, H_{XY}) = (\partial/\partial\delta)\, T^{\alpha}_{xy}(H_{XY} + \delta(\Delta_{X_iY_i} - H_{XY}))|_{\delta=0+}$.

The importance of the IF lies in the fact that it allows to assess the relative influence of individual observations towards the value of the estimate. If it is unbounded, even a single outlier may cause trouble. Its maximum absolute value $\gamma^*(T^{\alpha}_{xy}, H_{XY}) = \sup_{u\in\mathbb{R}^{p+q}} |IF(u; T^{\alpha}_{xy}, H_{XY})|$ defines the gross-error sensitivity of T^{α}_{xy} at H_{XY}. It measures the effect of contamination of the data by gross-errors, whereby some of the observations (X_i, Y_i) may have a distribution grossly different from H_{XY}. Specifically, γ^* is interpreted as the worst possible influence which a fixed amount of contamination can have upon the estimator.

Proposition 4.1. *Under the same conditions of Theorem* 3.2, *the gross-error sensitivity of the sequence of estimators* $\{\widehat{\lambda}_{\alpha,n}(x,y)\}_n$ *is given by*

$$\gamma^*(T^{\alpha}_{xy}, H_{XY}) = \frac{\min(-\alpha, \alpha-1)}{G'(\lambda_\alpha(x,y))F_X(x)}.$$

We have almost surely $IF((X_i, Y_i); T^{\alpha}_{xy}, H_{XY}) = \Gamma(x, y, X_i, Y_i)$, *where* Γ *is described in* (12).

Proof. We have

$$\begin{aligned}
&T^{\alpha}_{xy}(H_{XY}+\delta[\Delta_{X_iY_i}-H_{XY}])\\
&\quad=\sup\left\{\lambda|\, S_{Y|X}(\lambda y|x)>(1-\alpha)+\frac{\delta(1-\alpha)\mathbb{1}(X_i\leqslant x)-\delta\Delta_{X_iY_i}(x,\lambda y)}{(1-\delta)F_X(x)}\right\}\\
&\quad=\sup\left\{\lambda\leqslant\mathscr{Y}_i|\, G(\lambda)>(1-\alpha)-\frac{\alpha\delta\mathbb{1}(X_i\leqslant x)}{(1-\delta)F_X(x)}\right\}\\
&\quad\vee\sup\left\{\lambda>\mathscr{Y}_i|\, G(\lambda)>(1-\alpha)+\frac{\delta(1-\alpha)\mathbb{1}(X_i\leqslant x)}{(1-\delta)F_X(x)}\right\}.
\end{aligned}$$

The last equality is obtained by using the fact that $Y_i\geqslant\lambda y\Leftrightarrow\mathscr{Y}_i\geqslant\lambda$. For δ small enough, if $\lambda_\alpha(x,y)\leqslant\mathscr{Y}_i$ then the second supremum on the right-hand side of this equality is $-\infty$ and, if $\lambda_\alpha(x,y)>\mathscr{Y}_i$, then the first supremum is $\mathscr{Y}_i$. Therefore, for δ sufficiently small,

$$\begin{aligned}
T^{\alpha}_{xy}(H_{XY}+\delta[\Delta_{X_iY_i}-H_{XY}])&=G^{-1}\left((1-\alpha)-\frac{\alpha\delta\mathbb{1}(X_i\leqslant x)}{(1-\delta)F_X(x)}\right)\mathbb{1}(\lambda_\alpha(x,y)<\mathscr{Y}_i)\\
&\quad+\lambda_\alpha(x,y)\mathbb{1}(\lambda_\alpha(x,y)=\mathscr{Y}_i)\\
&\quad+G^{-1}\left((1-\alpha)+\frac{\delta(1-\alpha)\mathbb{1}(X_i\leqslant x)}{(1-\delta)F_X(x)}\right)\mathbb{1}(\lambda_\alpha(x,y)>\mathscr{Y}_i).
\end{aligned}$$

Hence

$$\begin{aligned}
&IF((X_i,Y_i);T^{\alpha}_{xy},H_{XY})\\
&\quad=\frac{-\alpha\mathbb{1}(X_i\leqslant x)\mathbb{1}(\lambda_\alpha(x,y)<\mathscr{Y}_i)+(1-\alpha)\mathbb{1}(X_i\leqslant x)\mathbb{1}(\lambda_\alpha(x,y)>\mathscr{Y}_i)}{G'(\lambda_\alpha(x,y))F_X(x)}.
\end{aligned}$$

More generally, we have for any $(x_0,y_0)\in\mathbb{R}^{p+q}$

$$\begin{aligned}
&IF((x_0,y_0);T^{\alpha}_{xy},H_{XY})\\
&\quad=\frac{\mathbb{1}(x_0\leqslant x)}{G'(\lambda_\alpha(x,y))F_X(x)}\left\{-\alpha\mathbb{1}\left(\lambda_\alpha(x,y)<\min_{1\leqslant k\leqslant q}\frac{y_0^k}{y^k}\right)\right.\\
&\qquad\left.+(1-\alpha)\mathbb{1}\left(\lambda_\alpha(x,y)>\min_{1\leqslant k\leqslant q}\frac{y_0^k}{y^k}\right)\right\}.
\end{aligned}$$

A simple computation gives then the desired conclusion. □

Thus the sequence of estimators $\{\widehat{\lambda}_{\alpha,n}(x,y)\}$ is bias-robust (Rousseeuw, 1981) in estimating $\lambda_\alpha(x,y)$ since it possesses a finite gross-error sensitivity. The approximate influence of the observations (X_i,Y_i) toward the error of estimation is described by the following asymptotic bias representation which follows from (8):

$$\begin{aligned}
\widehat{\lambda}_{\alpha,n}(x,y)-\lambda_\alpha(x,y)&=\frac{1}{n}\sum_{i=1}^{n}IF((X_i,Y_i);T^{\alpha}_{xy},H_{XY})+\frac{1}{\sqrt{n}}R_n,\\
&=\frac{\widehat{H}_{XY,n}(x,0)}{G'(\lambda_\alpha(x,y))F_X(x)}[(1-\alpha)-\widehat{S}_{Y|X,n}(\lambda_\alpha(x,y)y|x)]+R_n^{\alpha,xy},
\end{aligned}$$

where $\sqrt{n}R_n^{\alpha,xy}=\mathrm{o_p}(1)$ as $n\to\infty$. The fact that $IF((X_i,Y_i);T^{\alpha}_{xy},H_{XY})$ is zero when $\mathbb{1}(X_i\leqslant x)=0$ ensures that the efficiency scores $\widehat{\lambda}_{\alpha,n}(x,y)$ are not influenced by outlying

production units (X_i, Y_i) whose inputs $X_i \not\leqslant x$. Because of the irregularity of the *IF* due to the discontinuity of the indicator functions, the local-shift sensitivity of $\{\widehat{\lambda}_{\alpha,n}(x,y)\}$ defined as the smallest Lipschitz constant the influence function obeys, i.e.,

$$\lambda^*(T^\alpha_{xy}, H_{XY}) = \sup_{s \neq t \in \mathbb{R}^{p+q}} |IF(s; T^\alpha_{xy}, H_{XY}) - IF(t; T^\alpha_{xy}, H_{XY})| / \|s - t\|$$

is infinite. By $\|\cdot\|$ we denote the usual Euclidean norm on $\mathbb{R}^{p+q}$. This means that the estimators $\{\widehat{\lambda}_{\alpha,n}(x,y)\}$ may be sensitive to rounding errors. But this is much less important than the fact that $\gamma^*(T^\alpha_{xy}, H_{XY})$ is finite. Note also that, if $G(\lambda) = S_{Y|X}(\lambda y|x)$ is continuously differentiable on its support with negative derivative $G'(\lambda)$, then

$$\lim_{\alpha \nearrow 1} \gamma^*(T^\alpha_{xy}, H_{XY}) = \frac{1}{|G'(\lambda(x,y))| F_X(x)}. \tag{13}$$

This implies that $\widehat{\lambda}_{\alpha,n}(x,y)$ can be resistant to outliers even for large values of α.

Let us now turn to robustness characteristics of the expected order-m efficiency scores:

$$\lambda_m(x,y) = \int_0^\infty \left[1 - \left(1 - \frac{H_{XY}(x, \lambda y)}{H_{XY}(x,0)}\right)^m\right] \mathrm{d}\lambda = S^m_{xy}(H_{XY}),$$

$$\widehat{\lambda}_{m,n}(x,y) = \int_0^\infty \left[1 - \left(1 - \frac{\widehat{H}_{XY,n}(x, \lambda y)}{\widehat{H}_{XY,n}(x,0)}\right)^m\right] \mathrm{d}\lambda = S^m_{xy}(\widehat{H}_{XY,n}),$$

where the integrands are identically zero, respectively, for $\lambda \geqslant \lambda(x,y)$ and $\lambda \geqslant \widehat{\lambda}_n(x,y)$ (see Definition 2.2 and Theorem 2.2 of Daraio and Simar, 2006). It can be then easily seen that the influence function $(x_0, y_0) \in \mathbb{R}^{p+q}_+ \mapsto (\partial/\partial\delta) S^m_{xy}(H_{XY} + \delta(\Delta_{X_i Y_i} - H_{XY}))|_{\delta=0+}$ of the sequence of estimators $\{\widehat{\lambda}_{m,n}(x,y)\}$ is given by

$$\begin{aligned} &IF((x_0, y_0); S^m_{xy}, H_{XY}) \\ &\quad = \frac{m}{F_X(x)} \mathbb{1}(x_0 \leqslant x) \int_0^\infty [1 - S_{Y|X}(\lambda y|x)]^{m-1} (\mathbb{1}(y_0 \geqslant \lambda y) - S_{Y|X}(\lambda y|x)) \, \mathrm{d}\lambda. \end{aligned}$$

Like the order-α efficiency measure, $\widehat{\lambda}_{m,n}(x,y)$ rejects outlying production units using inputs $X_i \not\leqslant x$, for any sample size n. But unlike $\widehat{\lambda}_{\alpha,n}(x,y)$, the order-$m$ efficiency measure is not bias-robust since its gross-error sensitivity equals infinity. Indeed,

$$\begin{aligned} \gamma^*(S^m_{xy}, H_{XY}) &= \frac{m}{F_X(x)} \sup_{y_0 \in \mathbb{R}^q_+} \left| \int_0^\infty [1 - S_{Y|X}(\lambda y|x)]^{m-1} (\mathbb{1}(y_0 \geqslant \lambda y) - S_{Y|X}(\lambda y|x)) \, \mathrm{d}\lambda \right| \\ &\geqslant \frac{m}{F_X(x)} \sup_{y_0 \in \mathbb{R}^q_+ \,|\, \min_{1 \leqslant k \leqslant q} (y_0^k / y^k) > \lambda(x,y)} \int_0^{\min_{1 \leqslant k \leqslant q} \frac{y_0^k}{y^k}} [1 - S_{Y|X}(\lambda y|x)]^m \, \mathrm{d}\lambda \\ &\geqslant \frac{m}{F_X(x)} \sup_{y_0 \in \mathbb{R}^q_+ \,|\, \min_{1 \leqslant k \leqslant q} (y_0^k / y^k) > \lambda(x,y)} \left(\min_{1 \leqslant k \leqslant q} \frac{y_0^k}{y^k} - \lambda(x,y) \right) = \infty. \end{aligned}$$

This reflects the fact that even a single outlier (X_i, Y_i) with a level of inputs $X_i \leqslant x$, if it is far enough from the cloud of data points in the direction of Y, can attract $(x, \widehat{\lambda}_{m,n}(x,y) y)$ nearly to its outlying output Y_i. Besides this deficiency, the local-shift sensitivity is infinite too due to the discontinuity of the indicator function $x_0 \mapsto \mathbb{1}(x_0 \leqslant x)$.

However, if $\lambda(x,y)$ is majorized by a finite constant, then $\gamma^*(S^m_{xy}, H_{XY})$ is finite.

Proposition 4.2. *Let $m \geqslant 1$ be a fixed order and let $(x,y) \in \Psi$ such that $F_X(x) > 0$. If $\lambda(x,y) < \rho$, a finite positive constant, then the gross-error sensitivity of the sequence of estimators $\{\widehat{\lambda}_{m,n}(x,y)\}_n$ is such that*

$$\frac{m}{F_X(x)}(\rho - \lambda(x,y)) \leqslant \gamma^*(S^m_{xy}, H_{XY}) \leqslant \frac{m\rho}{F_X(x)}.$$

Proof. Let $(x_0, y_0) \in \mathbb{R}^{p+q}_+$ and put $h_\delta = H_{XY} + \delta(\Delta_{X_i Y_i} - H_{XY})$. Then $IF((x_0,y_0); S^m_{xy}, H_{XY}) = \frac{\partial}{\partial\delta} S^m_{xy}(h_\delta)|_{\delta=0+}$, where

$$S^m_{xy}(h_\delta) = \int_0^{\lambda_{h_\delta}(x,y)} \left[1 - \left(1 - \frac{h_\delta(x,\lambda y)}{h_\delta(x,0)}\right)^m\right] \mathrm{d}\lambda$$

with $\lambda_{h_\delta}(x,y) = \sup\{\lambda \mid \frac{h_\delta(x,\lambda y)}{h_\delta(x,0)} > 0\}$. Here $h_\delta(x,0) > 0$ for all δ small enough since $\lim_{\delta \searrow 0} h_\delta(x,0) = F_X(x) > 0$. It can be easily seen that $\lim_{\delta \searrow 0} \lambda_{h_\delta}(x,y) = \lambda_(x,y)$, which gives $\lambda_{h_\delta}(x,y) < \rho$ for δ sufficiently small. Hence $S^m_{xy}(h_\delta) = \int_0^\rho [1 - (1 - \frac{h_\delta(x,\lambda y)}{h_\delta(x,0)})^m] \mathrm{d}\lambda$ for δ small enough. Therefore

$$\begin{aligned} &IF((x_0,y_0); S^m_{xy}, H_{XY}) \\ &\quad = \frac{m}{F_X(x)} \mathbb{1}(x_0 \leqslant x) \int_0^\rho [1 - S_{Y|X}(\lambda y|x)]^{m-1} (\mathbb{1}(y_0 \geqslant \lambda y) - S_{Y|X}(\lambda y|x)) \,\mathrm{d}\lambda. \end{aligned}$$

It is then immediate that $\gamma^*(S^m_{xy}, H_{XY}) \geqslant \frac{m}{F_X(x)} \int_{\lambda(x,y)}^\rho [1 - S_{Y|X}(\lambda y|x)]^m \,\mathrm{d}\lambda = \frac{m}{F_X(x)}(\rho - \lambda(x,y))$ and $\gamma^*(S^m_{xy}, H_{XY}) \leqslant \frac{m\rho}{F_X(x)}$. □

The lower and upper bounds of $\gamma^*(S^m_{xy}, H_{XY})$ indicate that the nonparametric expected order-m efficiency scores are all the more sensitive to extreme values as the order m is large:

$$\lim_{m \nearrow \infty} \gamma^*(S^m_{xy}, H_{XY}) = \infty. \tag{14}$$

This means in particular that $\widehat{\lambda}_{m,n}(x,y)$, when considered as an estimator of the Farrell efficiency score $\lambda(x,y) = \lim_{m \nearrow \infty} \lambda_m(x,y)$, may be very sensitive to extreme values.

It should be clear that the partial efficiency scores $\{S^m_{xy}(\widehat{H}_{XY,n})\}$ and $\{T^\alpha_{xy}(\widehat{H}_{XY,n})\}$ do not estimate the same quantity. But in the limiting case where m tends to infinity and α to one, both sequences coincide with $\{\widehat{\lambda}_n(x,y)\}$ and can be then viewed as estimators of the full efficiency measure $\lambda(x,y)$. Results (13) and (14) indicate then that extreme order-α efficiencies are more robust than extreme order-m measures for estimating $\lambda(x,y)$. It is also important to note that by an appropriate choice of m and α as functions of n, the functionals $\{S^{m(n)}_{xy}(\widehat{H}_{XY,n})\}$ and $\{T^{\alpha(n)}_{xy}(\widehat{H}_{XY,n})\}$ estimate the true efficiency measure $\lambda(x,y)$. The advantage of quantile-type efficiency scores can be then clearly showed by comparing $\lim_{n\to\infty} \gamma^*(\cdot, H_{XY})$ of both estimators $S^{m(n)}_{xy}(\widehat{H}_{XY,n})$ and $T^{\alpha(n)}_{xy}(\widehat{H}_{XY,n})$ of $\lambda(x,y)$.

Theorem 4.1. *Let $(x,y) \in \Psi$ such that $F_X(x) > 0$ and let $\{\alpha(n)\}_{n\geqslant 1}$ and $\{m(n)\}_{n\geqslant 1}$ be nondecreasing sequences such that $0 < \alpha(n) < 1$, $\lim_{n\to\infty} \alpha(n) = 1$, $m(n) \geqslant 1$ and $\lim_{n\to\infty} m(n) = \infty$.*

1. *$\gamma^*(S^{m(n)}_{xy}, H_{XY})$ is infinite for any n unless the condition of Proposition 4.2 holds. Furthermore, $\lim_{n\nearrow\infty} \gamma^*(S^{m(n)}_{xy}, H_{XY}) = \lim_{m\nearrow\infty} \gamma^*(S^m_{xy}, H_{XY}) = \infty$.*
2. *If $G(\lambda) = S_{Y|X}(\lambda y|x)$ is differentiable at $\lambda_{\alpha(n)}(x,y)$ with derivative $G'(\lambda_{\alpha(n)}(x,y)) < 0$, then $\gamma^*(T^{\alpha(n)}_{xy}, H_{XY}) = \min(-\alpha(n), \alpha(n) - 1)/G'(\lambda_{\alpha(n)}(x,y)) F_X(x)$. If furthermore $G(\cdot)$ is*

continuously differentiable on its support with negative derivative $G'(\cdot)$, *then* $\lim_{n\nearrow\infty} \gamma^*(T_{xy}^{\alpha(n)}, H_{XY}) = \lim_{\alpha\nearrow 1} \gamma^*(T_{xy}^{\alpha}, H_{XY}) < \infty$.

In both cases the local-shift sensitivity equals infinity.

5. Introducing environmental variables

The analysis of the preceding section can easily be extended to the case where additional information is provided by other variables $Z \in \mathbb{R}^r$, exogenous to the production process itself, but which may explain a part of it. The basic idea for introducing this additional information in the model is to condition the production process to a given value of $Z = z$. Inspired from Cazals et al. (2002), Daraio and Simar (2006) introduce the concepts of conditional efficiency measure and of partial conditional efficiency measure of discrete order $m \geqslant 1$. Similarly, we propose below the idea for conditional quantile efficiency measure of continuous order $\alpha \in [0, 1]$.

If the joint distribution of (X, Y) conditional on $Z = z$ defines the production process, the efficiency measure $\lambda(x, y)$ defined in (1) and (3) has to be adapted to the condition $Z = z$ as follows:

$$\lambda(x, y|z) = \sup\{\lambda \mid S_{Y|X,Z}(\lambda y|x, z) > 0\},$$

where $S_{Y|X,Z}(y|x, z) = \text{Prob}(Y \geqslant y|X \leqslant x, Z = z)$.

A nonparametric estimator of the conditional full-frontier efficiency $\lambda(x, y|z)$ is given by plugging in its formula a nonparametric estimator of $S_{Y|X,Z}(y|x, z)$. We can use the following smoothed estimator:

$$\widehat{S}_{Y|X,Z,n}(y|x, z) = \frac{\sum_{i=1}^{n} \mathbb{1}(X_i \leqslant x, Y_i \geqslant y) K((z - Z_i)/h_n)}{\sum_{i=1}^{n} \mathbb{1}(X_i \leqslant x) K((z - Z_i)/h_n)},$$

where K is the kernel and h_n is the bandwidth of appropriate size. Practical bandwidth selection issues, based on a k-nearest neighbor method, are addressed in Section 4 of Daraio and Simar (2006) in the input oriented framework. As also pointed out there, the estimate of the conditional full-frontier efficiency for kernels with unbounded support is unable to detect any influence of the environmental factors. Therefore, kernels with compact support have to be used. Let the observations $(X_i, Y_i, Z_i) \in \mathbb{R}^{p+q+r}$, $i = 1, \ldots, n$, be independent with the same distribution as (X, Y, Z).

Lemma 5.1. *If the kernel K is of bounded variation with bounded support and the band sequence $h_n \to 0$ in such a way that $\sum_{n\geqslant 1} \exp[-\rho n h_n^r] < \infty$ for all $\rho > 0$, then*

$$\widehat{S}_{Y|X,Z,n}(\lambda y|x, z) \xrightarrow{\text{a.s.}} S_{Y|X,Z}(\lambda y|x, z) \quad \text{as } n \to \infty$$

for any λ, y and any x interior to the support of X and for almost all z, i.e., for all $z \notin N$ where N is such that $\text{Prob}(Z \in N) = 0$.

Proof. Since the indicator functions $\mathbb{1}(X \leqslant x)$ and $\mathbb{1}(X \leqslant x, Y \geqslant \lambda y)$ are bounded, Theorem 1 of Stute (1986a) immediately implies almost sure convergence of the Nadaraya–Watson estimates $\sum_{i=1}^{n} \mathbb{1}(X_i \leqslant x, Y_i \geqslant \lambda y) K((z - Z_i)/h_n) / \sum_{i=1}^{n} K((z - Z_i)/h_n)$ and $\sum_{i=1}^{n} \mathbb{1}(X_i \leqslant x) K((z - Z_i)/h_n) / \sum_{i=1}^{n} K((z - Z_i)/h_n)$ to $\text{E}[\mathbb{1}(X \leqslant x, Y \geqslant \lambda y)|Z = z]$ and $\text{E}[\mathbb{1}(X \leqslant x)|Z = z]$, respectively, as $n \to \infty$, for any λ, x, y and for all $z \notin N$. Which implies the desired almost sure convergence of $\widehat{S}_{Y|X,Z,n}(\lambda y|x, z)$ to $S_{Y|X,Z}(\lambda y|x, z)$. $\square$

By following the idea of Definition (4), we can define the conditional order-α output efficiency measure as follows:

Definition 5.1. For any $y \in \mathbb{R}_+^q$, the conditional order-α output efficiency measure given that $Z = z$, denoted by $\lambda_\alpha(x, y|z)$ is defined for all x in the interior of the support of X as

$$\lambda_\alpha(x, y|z) = \sup\{\lambda \mid S_{Y|X,Z}(\lambda y|x, z) > 1 - \alpha\}.$$

Therefore, for any $y \in \mathbb{R}_+^q$, the conditional order-α quantile frontier given that $Z = z$, is defined as the set of points $y_\alpha^\partial(x|z) = \lambda_\alpha(x, y|z)y$, $y \in \mathbb{R}_+^q$. As above we have immediately the following result.

Proposition 5.1. *For any $y \in \mathbb{R}_+^q$ and for all x in the interior of the support of X,*

$$\lim_{\alpha \to 1} \nearrow \lambda_\alpha(x, y|z) = \lambda(x, y|z).$$

A nonparametric estimator of $\lambda_\alpha(x, y|z)$ is provided by plugging in its formula the nonparametric estimator of $S_{Y|X,Z}(y|x, z)$. Formally, it is defined as

$$\widehat{\lambda}_{\alpha,n}(x, y|z) = \sup\{\lambda \mid \widehat{S}_{Y|X,Z,n}(\lambda y|x, z) > 1 - \alpha\}.$$

Here also we have

$$\lim_{\alpha \to 1} \nearrow \widehat{\lambda}_{\alpha,n}(x, y|z) = \widehat{\lambda}_n(x, y|z) = \sup\{\lambda \mid \widehat{S}_{Y|X,Z,n}(\lambda y|x, z) > 0\}.$$

These conditional nonparametric estimators are very easy to implement and very fast to compute in practice. Indeed, for $j = 1, \ldots, N_x$, denote by $Z_{[j]}^x$ the observation Z_i corresponding to the order statistic $\mathscr{Y}_{(j)}^x$, and let $R_{x,z} = \sum_{i=1}^n \mathbb{1}(X_i \leqslant x) K(\frac{z - Z_i}{h_n}) > 0$. Then,

$$\begin{aligned}\widehat{S}_{Y|X,Z,n}(\lambda y|x, z) &= \frac{1}{R_{x,z}} \sum_{j=1}^{N_x} \mathbb{1}(\lambda \leqslant \mathscr{Y}_{(j)}^x) K((z - Z_{[j]}^x)/h_n) \\ &= \begin{cases} 1 & \text{if } \lambda \leqslant \mathscr{Y}_{(1)}^x, \\ L_{k+1} & \text{if } \mathscr{Y}_{(k)}^x < \lambda \leqslant \mathscr{Y}_{(k+1)}^x, \quad k = 1, \ldots, N_x - 1, \\ 0 & \text{if } \lambda > \mathscr{Y}_{(N_x)}^x, \end{cases}\end{aligned}$$

where $L_{k+1} = (1/R_{x,z}) \sum_{j=k+1}^{N_x} K((z - Z_{[j]}^x)/h_n)$. It follows

$$\widehat{\lambda}_{\alpha,n}(x, y|z) = \begin{cases} \mathscr{Y}_{(k)}^x & \text{if } L_{k+1} \leqslant 1 - \alpha < L_k, \ k = 1, \ldots, N_x - 1, \\ \mathscr{Y}_{(N_x)}^x & \text{if } 0 \leqslant 1 - \alpha < L_{N_x}. \end{cases}$$

Theorem 5.1. *Let $x \in \mathbb{R}_+^p$ be an interior point of the support of X, $y \in \mathbb{R}_+^q$, $z \notin N$ and $\alpha \in]0, 1[$. If $\lambda \mapsto S_{Y|X,Z}(\lambda y|x, z)$ is decreasing in a neighborhood of $\lambda_\alpha(x, y|z)$, then*

$$\widehat{\lambda}_{\alpha,n}(x, y|z) \xrightarrow{\text{a.s.}} \lambda_\alpha(x, y|z) \quad \text{as } n \to \infty.$$

Proof. Let $\varepsilon > 0$ and $G(\lambda) = S_{Y|X,Z}(\lambda y|x, z)$. We have in view of the definition of $\lambda_\alpha(x, y|z)$ and the regularity condition

$$G(\lambda_\alpha(x, y|z) + \varepsilon) < 1 - \alpha < G(\lambda_\alpha(x, y|z) - \varepsilon).$$

Let $\widehat{G}_n(\lambda) = \widehat{S}_{Y|X,Z,n}(\lambda y|x, z)$. It follows from Lemma 5.1 that $\widehat{G}_n(\lambda_\alpha(x, y|z) + \varepsilon) \xrightarrow{\text{a.s.}} G(\lambda_\alpha(x, y|z) + \varepsilon)$ and $\widehat{G}_n(\lambda_\alpha(x, y|z) - \varepsilon) \xrightarrow{\text{a.s.}} G(\lambda_\alpha(x, y|z) - \varepsilon)$ as $n \to \infty$. This yields

(see, e.g., Serfling (1980, p. 6), for an equivalent condition for the almost sure convergence)

$$\forall \eta > 0: \ \text{Prob}\left\{\sup_{m \geqslant n} |\widehat{G}_m(\lambda_\alpha(x,y|z)+\varepsilon) - G(\lambda_\alpha(x,y|z)+\varepsilon)| > \eta\right\} \to 0, \quad n \to \infty,$$

$$\forall \delta > 0: \ \text{Prob}\left\{\sup_{m \geqslant n} |\widehat{G}_m(\lambda_\alpha(x,y|z)-\varepsilon) - G(\lambda_\alpha(x,y|z)-\varepsilon)| > \delta\right\} \to 0, \quad n \to \infty.$$

Putting $\eta = (1-\alpha) - G(\lambda_\alpha(x,y|z)+\varepsilon)$ and $0 < \delta < G(\lambda_\alpha(x,y|z)-\varepsilon) - (1-\alpha)$, we get

$$\text{Prob}[\widehat{G}_m(\lambda_\alpha(x,y|z)+\varepsilon) \leqslant 1-\alpha < \widehat{G}_m(\lambda_\alpha(x,y|z)-\varepsilon) \text{ for all } m \geqslant n] \to 1, \ \ n \to \infty.$$

Hence, by using the fact that $\widehat{\lambda}_{\alpha,m}(x,y|z) \geqslant \lambda$ if and only if $\widehat{G}_m(\lambda) > (1-\alpha)$, we obtain

$$\text{Prob}[\lambda_\alpha(x,y|z) - \varepsilon \leqslant \widehat{\lambda}_{\alpha,m}(x,y|z) < \lambda_\alpha(x,y|z) + \varepsilon \text{ for all } m \geqslant n] \to 1, \ \ n \to \infty.$$

That is $\text{Prob}(\sup_{m \geqslant n} |\widehat{\lambda}_{\alpha,m}(x,y|z) - \lambda_\alpha(x,y|z)| > \varepsilon) \to 0$ as $n \to \infty$. This is equivalent to the almost sure convergence of $\widehat{\lambda}_{\alpha,n}(x,y|z)$ to $\lambda_\alpha(x,y|z)$ as $n \to \infty$. □

Note that the asymptotic properties of $\widehat{\lambda}_n(x,y|z)$ have not yet been derived in the literature. Its weak consistency can be easily derived from Theorem 5.1 as follows.

Corollary 5.1. *Let $x \in \mathbb{R}_+^p$ be in the interior of the support of X, $y \in \mathbb{R}_+^q$, $z \notin N$ and let $\kappa(x,y|z) = \sup\{\lambda \mid S_{Y|X,Z}(\lambda y|x,z) = 1\}$. If $\lambda \mapsto S_{Y|X,Z}(\lambda y|x,z)$ is decreasing on $[\kappa(x,y|z), \lambda_\alpha(x,y|z)]$, then*

$$\widehat{\lambda}_n(x,y|z) \xrightarrow{\text{p}} \lambda(x,y|z) \quad \text{as } n \to \infty.$$

Proof. Let $\varepsilon > 0$. Because $\lim_{\alpha \nearrow 1} \lambda_\alpha(x,y|z) = \lambda(x,y|z)$, there exists $0 < \alpha_\varepsilon < 1$ such that $|\lambda_{\alpha_\varepsilon}(x,y|z) - \lambda(x,y|z)| < \varepsilon/2$. Since $\widehat{\lambda}_{\alpha_\varepsilon,n}(x,y|z) \leqslant \widehat{\lambda}_n(x,y|z) \leqslant \lambda(x,y|z)$ with probability 1, we obtain $|\widehat{\lambda}_n(x,y|z) - \lambda(x,y|z)| \leqslant |\widehat{\lambda}_{\alpha_\varepsilon,n}(x,y|z) - \lambda(x,y|z)| < |\widehat{\lambda}_{\alpha_\varepsilon,n}(x,y|z) - \lambda_{\alpha_\varepsilon}(x,y|z)| + \varepsilon/2$ with probability 1. Whence $\text{Prob}[|\widehat{\lambda}_n(x,y|z) - \lambda(x,y|z)| > \varepsilon] \leqslant \text{Prob}[|\widehat{\lambda}_{\alpha_\varepsilon,n}(x,y|z) - \lambda_{\alpha_\varepsilon}(x,y|z)| > \varepsilon/2]$. Since $\widehat{\lambda}_{\alpha,n}(x,y|z) \xrightarrow{\text{a.s.}} \lambda_\alpha(x,y|z)$ for all $0 < \alpha < 1$, $\text{Prob}[|\widehat{\lambda}_{\alpha_\varepsilon,n}(x,y|z) - \lambda_{\alpha_\varepsilon}(x,y|z)| > \varepsilon/2] \to 0$ as $n \to \infty$. This ends the proof. □

A slightly different version of the estimator $\widehat{S}_{Y|X,Z,n}(\lambda y|x,z)$ can be adapted from a proposal by Yang (1981) and is given by

$$\widetilde{S}_{Y|X,Z,n}(\lambda y|x,z) = \frac{\sum_{i=1}^n \mathbb{1}(X_i \leqslant x, Y_i \geqslant \lambda y) K((\widehat{F}_{Z,n}(z) - \widehat{F}_{Z,n}(Z_i))/h_n)}{\sum_{i=1}^n \mathbb{1}(X_i \leqslant x) K((\widehat{F}_{Z,n}(z) - \widehat{F}_{Z,n}(Z_i))/h_n)},$$

where $\widehat{F}_{Z,n}(\cdot)$ denotes the empirical distribution function of values of Z. It turns out that this estimator is more efficient than $\widehat{S}_{Y|X,Z,n}(\lambda y|x,z)$ if there are few observations in neighborhoods of z (see, e.g., Stute, 1984). Consistency of the resulting estimator $\widetilde{\lambda}_{\alpha,n}(x,y|z) = \sup\{\lambda \mid \widetilde{S}_{Y|X,Z,n}(\lambda y|x,z) > 1-\alpha\}$ can be easily derived from consistency of $\widetilde{S}_{Y|X,Z,n}(\lambda y|x,z)$. Note also that the asymptotic normality of $\widehat{\lambda}_{\alpha,n}(x,y|z)$ and $\widetilde{\lambda}_{\alpha,n}(x,y|z)$ as estimators of $\lambda_\alpha(x,y|z)$ can be established by using a similar method as for showing asymptotic normality of unconditional quantiles (see, e.g., Wretman, 1978). Compared with Wretman's proof, the uniform tightness of $\{\sqrt{nh_n^r}(\widehat{S}_{Y|X,Z,n}(\lambda y|x,z) - S_{Y|X,Z}(\lambda y|x,z))\}$ as a process indexed by λ should be proved and used rather than Chebyshev's inequality. This prescription can be found, e.g., in Stute (1986b).

5.1. Stressing the influence of Z on the production process

The comparison of $\widehat{\lambda}_n(x, y|z)$ with $\widehat{\lambda}_n(x, y)$ is certainly of interest for analyzing the global influence of Z on the production process. When Z is univariate, Daraio and Simar (2006) suggest that the use of a scatter plot of the ratios $\widehat{\lambda}_n(x, y|z)/\widehat{\lambda}_n(x, y)$ against Z and its smoothed nonparametric regression line would be helpful to describe the influence of Z on efficiency. An increasing regression corresponds to favorable environmental factor and a decreasing regression indicates an unfavorable factor. The correspondent α-quantile efficiency scores provide a more robust analysis, robust to extremes or outliers. Of course, we do not propose any inference here, but only an easy and useful descriptive diagnostic tool.

In the input oriented case, the interpretation of the shape of the regression line of the ratios $\widehat{\theta}_n(x, y|z)/\widehat{\theta}_n(x, y)$ against Z (and of their correspondent α-quantile based measures), is in the opposite direction.

6. Numerical illustrations

We illustrate first the estimation of the α-quantile frontiers in some simulated data sets and compare with the FDH and the order-m frontiers and even with some traditional parametric estimators. Then in a second part, we will show through a multivariate simulated data set, the behavior of the estimators of the α-quantile efficiency scores, conditional to environmental variables and compare with alternative nonparametric estimators (FDH and order-m). Finally, we illustrate how our procedure works with a real data set.

6.1. Estimation of the frontier functions

We will compare the estimators with two simulated data sets used in Florens and Simar (2005) and one proposed in Simar (2003).

Example 1 (*Cobb–Douglas with exponential*). We choose here a concave frontier given by the Cobb–Douglas model $Y = X^{1/2}\exp(-U)$, where X is uniform on $[0, 1]$ and U, independent of X, is exponential with parameter $\lambda = 3$. Here the true frontier function is given by $\varphi(x) = x^{1/2}$. In this particular model, it can be shown that the true α-quantile frontier is given by $\varphi_\alpha(x) = x^{1/2}\lambda_\alpha$, where $\lambda_\alpha = \cos(\frac{\arccos(1-2\alpha)+4\pi}{3}) + \frac{1}{2}$, whereas the true order-m frontier is given by $\varphi_m(x) = x^{1/2}\lambda_m$, where $\lambda_m = 1 - B_m$ with $B_m = \sum_{j=0}^{m}\binom{m}{j}(-2)^{m-j}3^j/(3m - j + 1)$. In general, these partial order frontiers are different, except for the limiting case where m tends to infinity and α to one. But in this particular example, if $\alpha = \frac{1}{2}(1 - \cos[3\arccos(\frac{1}{2} - B_m) - 4\pi])$, then $\lambda_\alpha = \lambda_m$. We choose here $m = 20$ and $\alpha = 0.9612$ to compare the two estimates of the same object. We will then compare with the FDH frontier, the deterministic Cobb–Douglas fit (shifted-OLS, so that all the residuals are negative) and a stochastic parametric model with a correct specification for the frontier and for the efficiency distribution (exponential) plus, as usual in these models, a normal noise.

The results are illustrated in Fig. 1, for one sample of 100 observations generated by the model above where we add three outliers. For this case, when ones estimates the true full frontier $\varphi(x) = x^{1/2}$, there is no way to obtain better results than the stochastic parametric

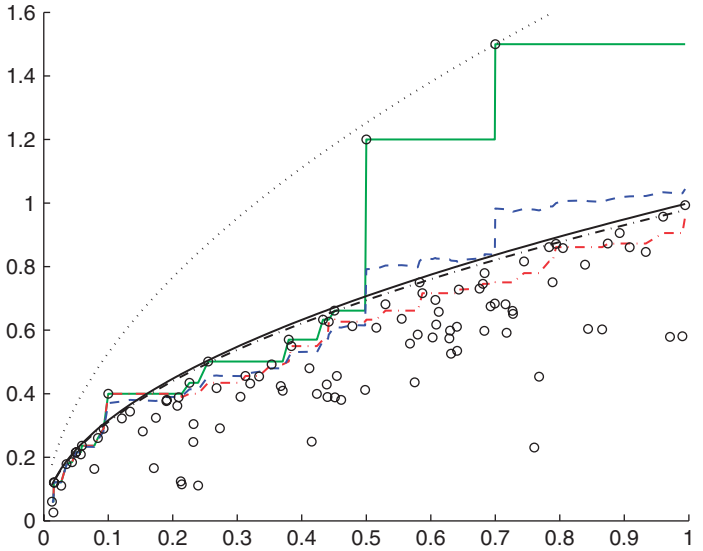

Fig. 1. Results for Example 1. In solid black line, the true frontier $y = x^{0.5}$. In cyan solid, the FDH frontier estimate, in blue dashed the estimated order-m frontier and in dash-dot red the estimate of the order-α frontier. In black dotted, the shifted OLS estimate and in dash-dot black, the parametric stochastic fit.

fit which is able to handle the three outliers by the normal noise term (estimated variance of the noise is 0.0308). Note however, as well known in stochastic frontier models with cross-sections, it is very difficult in these models to estimate the individual inefficiencies. Remember that the two estimators of the partial frontier estimate another object, situated below the full frontier ($\varphi_{m=20}(x) = \varphi_{\alpha=0.9612}(x) = 0.8815x^{1/2}$ is not reproduced on the figure for clarity) and we see, as expected by the theoretical properties developed above, that the order-α frontier estimate shows much more resistance to the outliers than the order-m. It is also interesting to see how both partial order frontiers behave pretty well even for estimating the full frontier in spite of not using any parametric assumption. As expected, the parametric deterministic estimate and the nonparametric full frontier estimate (FDH) are too sensitive to the outliers and miss the target.

Example 2 (*Cobb–Douglas with uniform*). We first consider a slightly different case where the frontier is linear (particular case of a Cobb–Douglas) but the stochastic scenario is different. We choose (X, Y) uniformly distributed over the region $D = \{(x, y)|0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant x\}$. Here the true frontier $\varphi(x) = x$. The true conditional α-quantile frontier is $\varphi_\alpha(x) = x(1 - \sqrt{1-\alpha})$, for $0 < x \leqslant 1$, and the order-m frontier can be computed as $\varphi_m(x) = x(1 - A_m)$, where $A_m = \sum_{j=0}^{m} \binom{m}{j} (-1)^{m-j} 2^j / (2m - j + 1)$. Again, in this particular example, if $\alpha = 1 - A_m^2$, then both partial frontiers coincide and we choose $m = 20$ and $\alpha = 0.9622$ to achieve this and facilitate the comparison of both estimators. We generate a sample of $n = 100$ observations and we add here four outliers, we use here for the parametric estimators the same specification as in Example 1 and we do the same exercise as above. The results are displayed in Fig. 2.

The comments are very similar to the preceding example: deterministic (parametric or FDH) break down and the stochastic parametric estimate (same specification as in

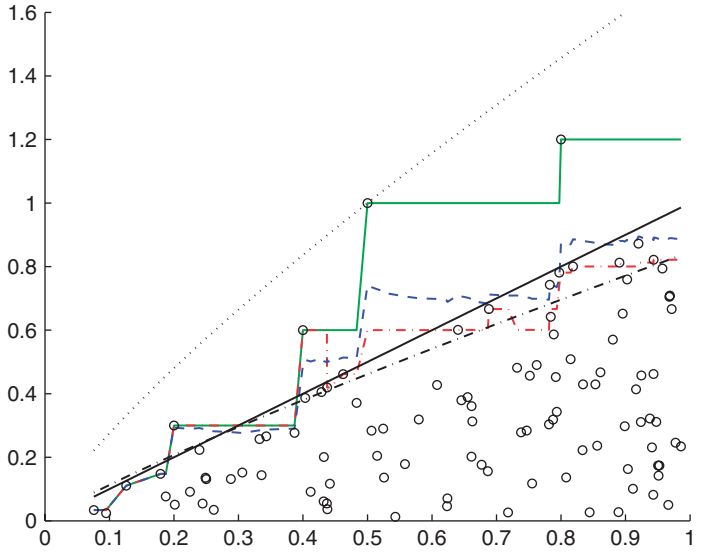

Fig. 2. Results for Example 2. In solid black line, the true frontier $y = x$. In cyan solid, the FDH frontier estimate, in blue dashed the estimated order-m frontier and in dash-dot red the estimate of the order-α frontier. In black dotted, the shifted OLS estimate and in dash-dot black, the parametric stochastic fit.

Example 1), behaves well but, as expected, not so well as above. This is due to the fact that the parametric specification of the model is correct (Cobb–Douglas) but the stochastic specification of the model is incorrect (as above we specified a normal noise minus an exponential in the log scale). So the estimated stochastic parametric frontier underestimates the true frontier. Remember that both partial frontier estimates estimate a different object, slightly below the full frontier. However, note again that even without any parametric assumptions, the partial frontiers behave not so badly when considered as estimates of the full frontier $\varphi(x) = x$, with again a clear advantage to the order-α frontier. This is the main advantage of these robust nonparametric estimators.

Example 3 (*Logit with exponential*). As just noticed, the advantages of the nonparametric estimators rely on the fact that no particular parametric specification is assumed for the frontier and for the efficiency term. To illustrate further, we did the same exercise as the two preceding examples but here the parametric specifications for the parametric models will be wrong. We use a Cobb–Douglas with exponential inefficiency term specification, as above, whereas the true model is $Y = \exp(10X - 5)/(1 + \exp(10X - 5))\exp(-U)$, with X and U as in Example 1. For the robust nonparametric estimators, we choose values of $\alpha = 0.95$ and $m = 20$, so that both order-α and order-m efficiency scores are very close. With little surprise, the nonparametric estimators clearly show their superiority as shown in Fig. 3.

6.2. A multivariate simulated example

In this set-up, we cannot produce pictures of the frontier surfaces, so we will focus on the detection of the effect of environmental variables on the efficiency scores

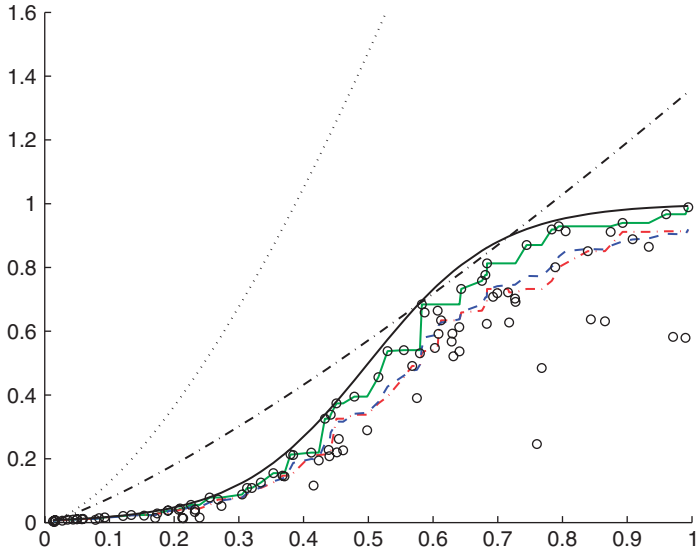

Fig. 3. Results for Example 3. In solid black line, the true logit frontier. In cyan solid, the FDH frontier estimate, in blue dashed the estimated order-m frontier and in dash-dot red the estimate of the order-α frontier. In black dotted, the shifted OLS estimate and in dash-dot black, the parametric stochastic fit.

and focusing on the comparison between the nonparametric estimators (full and partial frontiers).

Here a multi-input ($p = 2$) and multi-output ($q = 2$) data set is simulated and the function describing the efficient frontier is (as in Park et al., 2000)

$$y^{(2)} = 1.0845(x^{(1)})^{0.3}(x^{(2)})^{0.4} - y^{(1)},$$

where $y^{(j)}$, $(x^{(j)})$, denotes the jth component of y, (of x), for $j = 1, 2$. We draw $X_i^{(j)}$ independent uniforms on $(1, 2)$ and $\tilde{Y}_i^{(j)}$ independent uniforms on $(0.2, 5)$. Then the generated random rays in the output space are characterized by the slopes $S_i = \tilde{Y}_i^{(2)} / \tilde{Y}_i^{(1)}$. Finally, the generated random points on the frontier are defined by

$$Y_{i,\text{eff}}^{(1)} = \frac{1.0845(X_i^{(1)})^{0.3}(X_i^{(2)})^{0.4}}{S_i + 1},$$
$$Y_{i,\text{eff}}^{(2)} = 1.0845(X_i^{(1)})^{0.3}(X_i^{(2)})^{0.4} - Y_{i,\text{eff}}^{(1)}.$$

The efficiencies are generated by $\exp(-U_i)$ where U_i are drawn from an exponential with mean $\mu = \frac{1}{3}$. Finally, in a standard setup (without environmental factors), we define $Y_i = Y_{i,\text{eff}} * \exp(-U_i)$.

Now we introduce the dependency on an favorable environmental factor Z (we adapt Case 1 of Daraio and Simar, 2006): Z is uniform on $(0.5, 1.5)$ and

$$Y_i^{(1)} = Z * Y_{i,\text{eff}}^{(1)} * \exp(-U_i),$$
$$Y_i^{(2)} = Z * Y_{i,\text{eff}}^{(2)} * \exp(-U_i).$$

For the nonparametric estimation, we have chosen a triangle kernel for the smoothing: the results are very stable with respect to other choice of the kernel with compact support. For the partial frontier we have chosen $m = 20$ and $\alpha = 0.90$, just to illustrate the procedures. In practice, the choice of these two "tunning" parameters may be governed by their economic interpretation (a benchmark against the best of m virtual competitors, or against a level of production with a probability $(1-\alpha) \times 100\%$ of being dominated) but these estimators are so fast to compute that in practice, a sensitivity analysis could be performed to select an appropriate level for m and for α (see Simar, 2003 for more details). The values chosen here are such that both order-α and order-m efficiency scores are very close in absence of outliers.

We simulate $n = 100$ observations according to this scenario and we will compare our results with those obtained when adding five outliers. The latter are introduced at the following values of X: $(1.25, 1.5), (1.25, 1.75), (1.5, 1.5), (1.75, 1.25)$ and $(1.5, 1.25)$, the corresponding values for the slopes in the Y space are $(0.25, 0.75, 1, 3, 5)$. The corresponding values of Z have been chosen in the range of Z as $(0.6, 0.8, 1, 1.2, 1.4)$. Finally the outliers in the output direction were projected outside the true frontier by a factor 3 for the first three points and a factor of 2 for the remaining two outliers.

The results are displayed in Fig. 4. On the left panels, we have the results for the regular sample: we see that all the ratios allow to detect the favorable "linear" effect of Z on the production process. On the right panels, we see the results when the five outliers have been added. The comparison of the left to the right panel allows to appreciate the robustness of the measures to the outliers. The FDH estimator fails to detect the correct effect when the outliers are added and even shows a negative slope at the right. The order-m resists better to the outliers, although after $Z = 1.1$ indicate a flat slope. The α-quantile measures being the more robust to extreme values, give the best picture: the difference between the left and the right panels is the weakest; in both cases they indicate correctly the positive influence of Z on the production process.

6.3. Mutual funds data

We also illustrate our methodology analyzing US Mutual Funds data. We use a cross-section data set, collected by the reputed Morningstar, which consists of the US Mutual Funds universe updated at 05-31-2002. Among this universe we select the aggressive-growth (AG) category of mutual funds. These are funds that seek rapid growth of capital and that may invest in emerging market growth companies. For details about the data, the variables and references to this literature, see Daraio and Simar (2006), where the analysis is also fully motivated.

We have a sample of 129 mutual funds and we apply an input oriented framework. The traditional output in this framework is the total return of funds (the annual return at the 05-31-2002, expressed in percentage terms). Most returns were negative in this period, hence we shift them to get all positive returns by adding 100. This does not change our input oriented analysis. The inputs are risk (standard deviation, or volatility of the return), expense ratio (the percentage of fund assets paid for operating expenses, management fees, administrative fees, and all other asset-based costs) and turnover ratio (a measure of the fund's trading activity). In our illustration we use the market risks of mutual funds (the percentage of fund's movements that can be explained

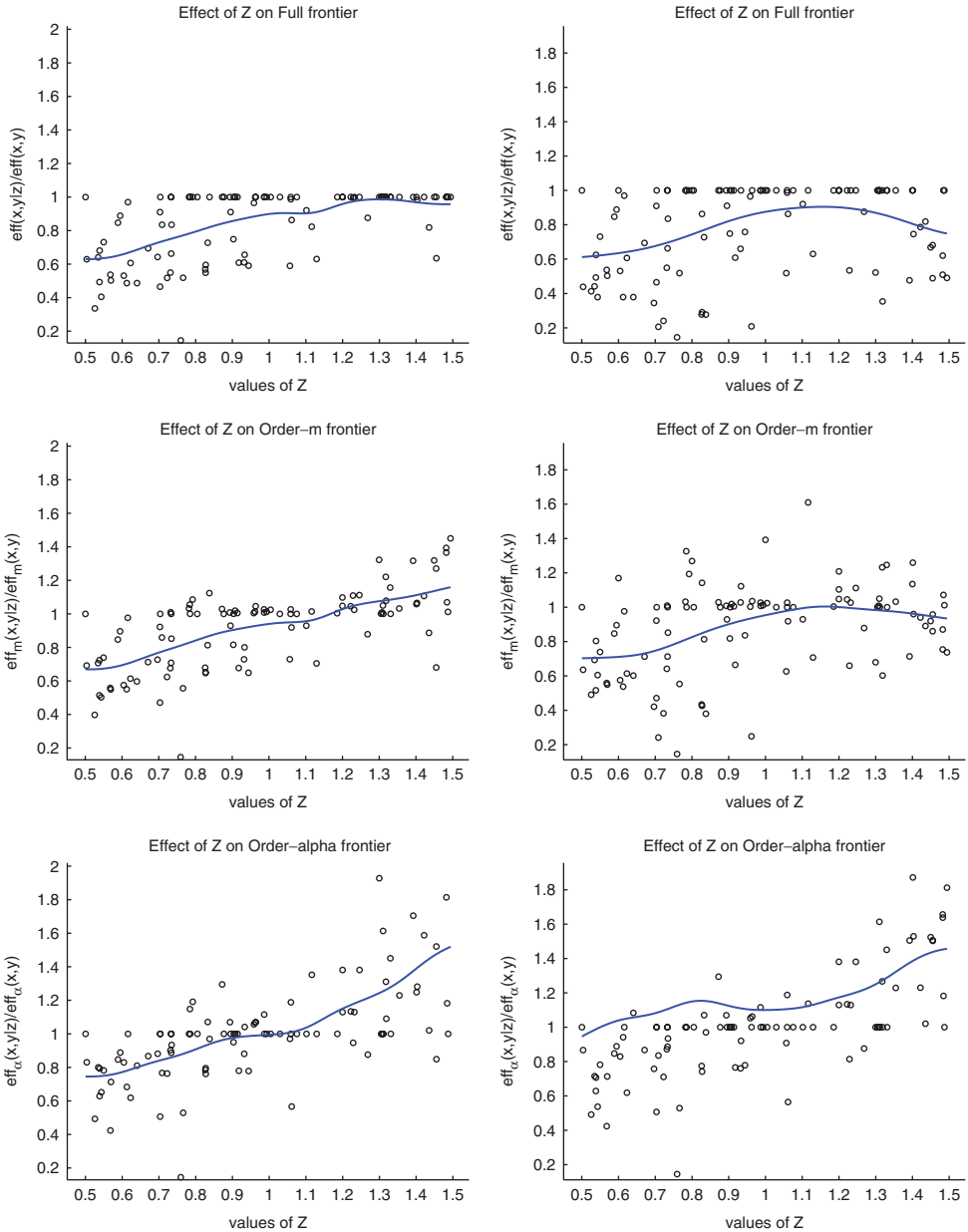

Fig. 4. Simulated example, $n = 100$: "positive" effect of Z on production efficiency (output oriented framework). Left panels, regular data and right panels, same data plus with 5 outliers. Scatterplot and smoothed regression of the ratios $\hat{\lambda}_n(x, y|z)/\hat{\lambda}_n(x, y)$ on Z (top panels), of $\hat{\lambda}_{m,n}(x, y|z)/\hat{\lambda}_{m,n}(x, y)$ on Z (middle panels) and of $\hat{\lambda}_{\alpha,n}(x, y|z)/\hat{\lambda}_{\alpha,n}(x, y)$ (bottom panels) on Z.

by movements in its benchmark index) as environmental variable, to investigate its effect on our data, i.e. if it is detrimental or favorable to the performance of mutual funds in the period under consideration.

We compare FDH, order-m (with $m = 25$, chosen in Daraio and Simar, 2006) and α-quantile efficiency scores, unconditional and conditional to Z. We have chosen four values of α, from 0.80 to 0.975 for showing the sensitivity of the procedure to this choice. The results are displayed in Fig. 5. We see indeed that for the order-α measures, the choice of α is not so important (the case $\alpha = 0.99$ is not reproduced here but is very similar, as it should be, to the FDH case). All the pictures confirm the global positive effect of the risk market on the performance of the funds, as expected from the literature (Sengupta, 2000 used this variable as additional input in this framework), but here this interpretation is a result of our analysis. We note also that the FDH measures fail to give this global interpretation, but we know this data set contains a lot of outliers (see Daraio and Simar, 2006). The effect appears more clearly with the order-α measures, because they are less sensitive to extreme values, and the robustness properties developed in Section 4 lead us to favor this measure. We see also that the results are rather stable when choosing "reasonable" values for α not far from the standard level of 95% of the statistical literature.

7. Conclusions

In this paper we develop a generalized concept of efficiency measure, the α-quantile efficient scores, related to a nonstandard conditional α-quantile frontier in a full multivariate set-up. The approach can be viewed as an alternative to the order-m efficiency scores and order-m efficient frontier developed by Cazals et al. (2002) and Daraio and Simar (2006).

Both approaches provide nonparametric estimators of the efficient frontier which are more robust than the usual envelopment estimators (like FDH/DEA estimators). The α-quantile approach is more easy to interpret since the parameter α is just the selected level of the quantile. The choice of m in order-m efficient frontier is more delicate although it can be interpreted as the number of potential firms against which the benchmark is done to determine the efficiency score of a particular firm. The choice of m can also be indirectly piloted by the percentage of observed firms staying above the frontier for a given m, but the α-quantile approach seems to be more direct.

The asymptotic normality of our estimator is provided for a fixed order $\alpha < 1$. An exponential probability inequality yielding the complete convergence of the estimator is also established. Then, by letting the order α increase to 1 as a function of the sample size, we derive an estimator of the full true Farrell efficiency scores which converges to a limiting Weibull distribution with the same rate as the FDH estimator.

The estimation procedure is robust to the perturbations in data, which attains a bounded influence function. A theoretical analysis shows that the multivariate quantile-type efficiency scores are more robust to extremes than the nonparametric order-m efficiency scores. Moreover, as for the order-m frontiers, the α-quantile frontiers can also be used to detect outliers in the spirit of Simar (2003).

In this framework, it is also easy to introduce environmental factors and we propose useful tools for detecting their influence on efficiencies. Here also, we derive some asymptotic properties for the resulting estimator and we show how it can be easily calculated. Numerical examples (with simulated data and with mutual funds data) illustrate the usefulness of the approach.

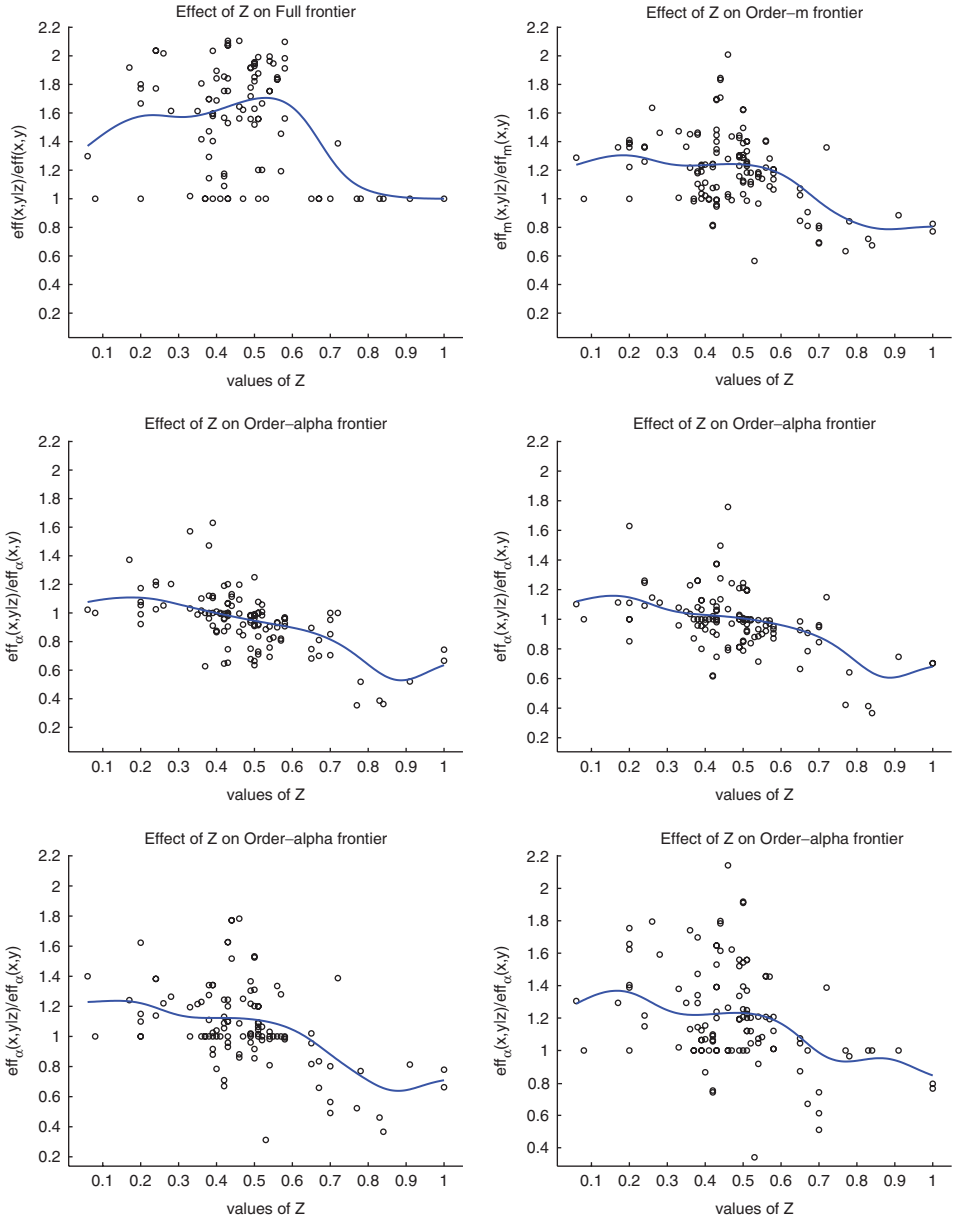

Fig. 5. Aggressive-growth US mutual funds. Scatterplot and smoothed regression of the ratios $\hat{\theta}_n(x,y|z)/\hat{\theta}_n(x,y)$ on Z (top left), of $\hat{\theta}_{m,n}(x,y|z)/\hat{\theta}_{m,n}(x,y)$ on Z (top right, with $m=25$) and of $\hat{\theta}_{\alpha,n}(x,y|z)/\hat{\theta}_{\alpha,n}(x,y)$ on Z (middle panel, left $\alpha=0.80$ and right $\alpha=0.90$ and bottom panel, left $\alpha=0.95$ and right $\alpha=0.975$).

Finally, since for every attainable point (x_i, y_i), there exists a α such that $\widehat{\theta}_{\alpha,n}=1$ (or $\widehat{\lambda}_{\alpha,n}=1$), this α could serve as an alternative measure of input (or output) efficiency. In other words, one may set the performance measure for the unit (x_i, y_i) to be the order α of the quantile frontier which passes through this unit. For instance, in the output orientation, it can be easily seen that $\alpha(x_i, y_i) = 1 - \widehat{S}_{Y|X,n}(y_i|x_i) + (1/n\widehat{H}_{XY,n}(x_i, 0))$.

Acknowledgments

Research support from the "Interuniversity Attraction Pole", Phase V (No. P5/24) from the Belgian Science Policy is acknowledged.

References

Aragon, Y., Daouia, A., Thomas-Agnan, C., 2003. Efficiency measurement: a nonparametric approach. Discussion Paper, GREMAQ et LSP, Université de Toulouse ⟨http://www.univ-tlse1.fr/GREMAQ/Statistique/mesure.pdf⟩.

Aragon, Y., Daouia, A., Thomas-Agnan, C., 2005. Nonparametric frontier estimation: a conditional quantile-based approach. Econometric Theory 21, 358–389.

Cazals, C., Florens, J.P., Simar, L., 2002. Nonparametric frontier estimation: a robust approach. J. Econometrics 106, 1–25.

Daouia, A., 2005. Asymptotic representation theory for nonstandard conditional quantiles. J. Nonparametric Statist. 17 (2), 253–268.

Daouia, A., Ruiz Gazen, A., 2006. Robust nonparametric frontier estimators: influence function and qualitative robustness. Statistica Sinica 16(4), to appear.

Daraio, C., Simar, L., 2006. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J. Productiv. Anal. 24 (1), 93–121.

Daraio, C., Simar, L., 2006, A robust nonparametric approach to evaluate and explain the performance of mutual funds. European J. Oper. Res. 175(1), 516–542.

Deprins, D., Simar, L., Tulkens, H., 1984. Measuring labor inefficiency in post offices. In: Marchand, M., Pestieau, P., Tulkens, H. (Eds.), The Performance of Public Enterprises: Concepts and Measurements. Amsterdam, North-Holland, pp. 243–267.

Farrell, M.J., 1957. The measurement of productive efficiency. J. Roy. Statistical Society Ser. A 120, 253–281.

Florens, J.P., Simar, L., 2005. Parametric approximations of nonparametric frontier. J. Econometrics 124 (1), 91–116.

Ghosh, J.K., 1971. A new proof of the Bahadur representation of quantiles and an application. Ann. Math. Statist. 42, 1957–1961.

Hampel, F.R., 1974. The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. 69, 383–393.

Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. 58, 13–30.

Koopmans, T.C., 1951. An analysis of production as an efficient combination of activities. In: Koopmans, T.C. (Ed.), Activity Analysis of Production and Allocation, Cowles Commission for Research in Economics, Monograph 13. Wiley, New York.

Park, B., Simar, L., Weiner, Ch., 2000. The FDH estimator for productivity efficiency scores: asymptotic properties. Econometric Theory 16, 855–877.

Rousseeuw, P.J., 1981. A new infinitesimal approach to robust estimation. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 56, 127–132.

Sengupta, J.K., 2000. Dynamic and stochastic efficiency analysis. Economics of Data Envelopment Analysis. World Scientific, Singapore.

Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.

Simar, L., 2003. Detecting outliers in frontiers models: a simple approach. J. Productiv. Anal. 20, 391–424.

Simar, L., Wilson, P.W., 2000. Statistical inference in nonparametric frontier models: the state of the art. J. Productiv. Anal. 13, 49–78.

Stute, W., 1984. Asymptotic normality of nearest neighbor regression function estimates. Ann. Statist. 12 (3), 917–926.

Stute, W., 1986a. On almost sure convergence of conditional empirical distribution functions. Ann. Probab. 14 (3), 891–901.

Stute, W., 1986b. Conditional empirical processes. Ann. Statist. 14 (2), 638–647.

Wretman, J., 1978. A simple derivation of the asymptotic distribution of a sample quantile. Scand. J. Statist. 5 (2), 123–124.

Yang, S.S., 1981. Linear functions of concomitants of order statistics with applications to nonparametric estimation of a regression function. J. Amer. Statist. Assoc. 76, 658–662.