Frontier Estimation and Extreme Values Theory

Abdelaati Daouia
Toulouse School of Economics (GREMAQ)

Jean-pierre Florens
Toulouse School of Economics (GREMAQ, IDEI)

Léopold Simar*
Université Catholique de Louvain (Institut de Statistique)

September, 2007

Abstract

The production/econometric frontier is the locus of the optimal combinations of inputs and outputs. From a statistical point of view, it can be viewed as the upper surface of the support of a random vector under shape constraints. In this paper we investigate the problem of nonparametric monotone frontier estimation from an extreme-values theory perspective. This allows to revisit the asymptotic theory of the popular FDH estimator, to derive new and asymptotically Gaussian estimators and to provide useful asymptotic confidence bands for the monotone boundary function. The study of the asymptotic properties of the resulting frontier estimators is carried out by relating them to an original dimensionless random sample and then applying standard extreme-values theory. The finite sample behavior of the suggested estimators is explored through Monte-Carlo experiments.

AMS 2000 subject classification: 62G20; 62G32; 91B38

Key words : Monotone boundary estimation, Extreme values, Productivity analysis.

*The authors acknowledge support from the "Interuniversity Attraction Pole", Phase V (No. P5/24) from the Belgian Science Policy is also acknowledged. A. Daouia also acknowledges support from the INRA.
E-mail addresses: daouia@cict.fr, florens@cict.fr, simar@stat.ucl.ac.be. Corresponding Author: Institut de Statistique, 20 voie du roman pays, B-1348 Louvain-la-Neuve, Belgium. Tel: +32 10 474308, Fax: +32 10 473032.

1 Introduction

In production theory and efficiency analysis (see *e.g.* Shephard, 1970), one is willing to estimate the boundary of a production set (the set of feasible combinations of inputs and outputs). This boundary (the production frontier) represents the set of optimal production plans so that the efficiency of a production unit (a firm, ...) is obtained by measuring the distance from this unit to the estimated production frontier. Parametric approaches rely on parametric models for the frontier and for the underlying stochastic process, whereas nonparametric approaches offer much more flexible models for the Data Generating Process (see *e.g.* Daraio and Simar 2007, for recent surveys on this topic).

Formally, we consider in this paper technologies where $x \in \mathbb{R}_+^p$, a vector of production factors (inputs) is used to produce a single quantity (output) $y \in \mathbb{R}_+$. The attainable production set is then defined, in standard microeconomic theory of the firm, as $\mathbb{T} = \{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+ \mid x \text{ can produce } y\}$. Assumptions are usually done on this set, such as free disposability of inputs and outputs, meaning that if $(x, y) \in \mathbb{T}$, then $(x', y') \in \mathbb{T}$, for any (x', y') such that $x' \geq x$ (with respect to the partial order) and $y' \leq y$. As far as efficiency of a firm is of concern, the boundary of $\mathbb{T}$ is of interest. The efficient boundary (production frontier) of $\mathbb{T}$ is the locus of optimal production plans (maximal achievable output for a given level of the inputs). In our setup, the production frontier is represented by the graph of the production function $\phi(x) = \sup\{y \,|\, (x, y) \in \mathbb{T}\}$. Then the Farrell-Debreu output efficiency score (Farrell 1957, Debreu 1951) of a firm operating at the level (x, y) is given by the ratio $\phi(x)/y$.

Cazals, Florens and Simar (2002) propose a probabilistic interpretation of the production frontier. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probability space on which the random variables X and Y are defined and let $\mathbb{T}$ be the support of the joint distribution of (X, Y). The distribution function of (X, Y) can be denoted $F(x, y)$ and $F(\cdot|x) = F(x, \cdot)/F_X(x)$ will be used to denote the conditional distribution function of Y given $X \leq x$, with $F_X(x) = F(x, \infty) > 0$. It has been proven in Cazals et al. (2002) that the function

$$\varphi(x) = \sup\{y \geq 0 | F(y|x) < 1\}$$

is monotone nondecreasing with x. So for all $x' \geq x$ with respect to the partial order, $\varphi(x') \geq \varphi(x)$. The graph of φ is the smallest nondecreasing surface which is larger than or equal to the upper boundary of $\mathbb{T}$. Further, it has been shown that under the free disposability assumption, $\varphi \equiv \phi$, *i.e.*, the graph of φ coincides with the production frontier.

Since $\mathbb{T}$ is unknown, it has to be estimated from a sample of i.i.d. firms $\mathcal{X}_n = \{(X_i, Y_i) | i = 1, \ldots, n\}$. The Free Disposal Hull (FDH) of $\mathcal{X}_n$, introduced by Deprins, Simar and Tulkens

(1984) is $\widehat{\mathbb{T}}_{FDH} = \left\{(x,y) \in \mathbb{R}_+^{p+1} \,|\, y \leq Y_i,\ x \geq X_i,\ i = 1,\ldots,n\right\}$. The resulting FDH estimator of the frontier function $\varphi(x)$ is defined as

$$\hat{\varphi}_1(x) = \sup\{y \geq 0 | \hat{F}(y|x) < 1\} = \max_{i:X_i \leq x} Y_i$$

where $\hat{F}(y|x) = \hat{F}_n(x,y)/\hat{F}_X(x)$ with $\hat{F}_n(x,y) = (1/n)\sum_{i=1}^n \mathbb{I}(X_i \leq x, Y_i \leq y)$ and $\hat{F}_X(x) = (1/n)\sum_{i=1}^n \mathbb{I}(X_i \leq x)$. This estimator represents the lowest monotone step function covering all the data points (X_i, Y_i). The asymptotic behavior of $\widehat{\varphi}_1(x)$ was first derived by Korostelev, Simar and Tsybakov (1995) for the consistency and by Park, Simar and Weiner (2000) and Hwang, Park and Ryu (2002) for the asymptotic sampling distribution. To summarize, under regularity conditions, the FDH estimator $\hat{\varphi}_1(x)$ is consistent and converges to a Weibull distribution with some unknown parameters. In Park et al (2000), the obtained convergence rate $n^{-1/(p+1)}$ requires that the joint density of (X,Y) has a jump at its support boundary. In addition, the estimation of the parameters of the Weibull requires the specification of smoothing parameters and the resulting procedure has very poor accuracy. In Hwang et al (2002), the convergence of $\hat{\varphi}_1(x)$ to the Weibull distribution has been established in a general case where the density of (X,Y) may decrease to zero or rise up to infinity at a speed of power β ($\beta > -1$) of the distance from the frontier. They obtain the convergence rate $n^{-1/(\beta+2)}$ and extend the particular result of Park et al (2000) where $\beta = 0$, but their result is only derived in the simple case of one-dimensional inputs ($p = 1$) which may be of less interest in practice.

In this paper we first analyze the properties of the FDH estimator from an extreme-value theory perspective. By doing so, we generalize and extend the results of Park *et al.* (2000) and Hwang *et al.* (2002) in at least three directions. First we provide the necessary and sufficient condition for the FDH estimator to converge in distribution and we specify the asymptotic distribution with the appropriate rate of convergence. We also provide a limit theorem of moments in a general framework. Second, we show how the unknown parameter $\rho_x > 0$ involved by the necessary and sufficient extreme-value condition, is linked to the dimension $p+1$ of the data and to the shape parameter $\beta > -1$ of the joint density: in the general setting where $p \geq 1$ and $\beta = \beta_x$ may depend on x, we obtain under a convenient regularity condition the general convergence rate $n^{-1/\rho_x} = n^{-1/(\beta_x+p+1)}$ of the FDH estimator $\hat{\varphi}_1(x)$. Third, we suggest a strongly consistent and asymptotically normal estimator of the unknown parameter ρ_x of the asymptotic Weibull distribution of $\hat{\varphi}_1(x)$. This also answers the important question of how to estimate the shape parameter β_x of the joint density of (X,Y) when it approaches to the frontier of the support $\mathbb{T}$.

By construction, the FDH estimator is very non-robust to extremes. Recently Aragon, Daouia and Thomas-Agnan (2005) have built an original estimator of $\varphi(x)$, which is more robust than $\hat{\varphi}_1(x)$ but it keeps the same limiting Weibull distribution as $\hat{\varphi}_1(x)$ under the restrictive condition $\beta = 0$. In this paper, we give more insights and generalize their main result. We also suggest attractive estimators of $\varphi(x)$ converging to a normal distribution and which appear to be robust to outliers. The study of the asymptotic properties of the different estimators considered in this paper, is carried out in a simple and clever way by relating them to an original dimensionless random sample and then applying standard extreme-values theory.

The paper is organized as follows. Section 2 presents the main results of the paper and Section 3 illustrates how the theoretical asymptotic results behave in finite sample situations and shows an example with a real data set on the production activity of the French post offices. Section 4 concludes and the proofs are reserved for the Appendix.

2 The Main Results

Throughout this paper we will denote by $\varphi_\alpha(x)$ and $\hat{\varphi}_\alpha(x)$, respectively, the α^{th} quantiles of the distribution function $F(\cdot|x)$ and its empirical version $\hat{F}(\cdot|x)$, with $\alpha \in]0,1]$,

$$\varphi_\alpha(x) = \inf\{y \geq 0 | F(y|x) \geq \alpha\} \quad \text{and} \quad \hat{\varphi}_\alpha(x) = \inf\{y \geq 0 | \hat{F}(y|x) \geq \alpha\}.$$

Asymptotic Weibull distribution

We first derive the following interesting results on the problem of convergence in distribution of suitably normalized maxima $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))$.

Theorem 2.1. *(i) If there exist constants $b_n > 0$ and some nondegenerate distribution function G_x such that*

$$b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G_x, \tag{2.1}$$

then $G_x(y)$ coincides with the Weibull distribution function

$$\Psi_{\rho_x}(y) = \begin{cases} \exp\{-(-y)^{\rho_x}\} & y < 0 \\ 1 & y \geq 0 \end{cases} \quad \textit{for some} \quad \rho_x > 0.$$

(ii) There exists $b_n > 0$ such that $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))$ converges in distribution if and only if

$$\lim_{t \uparrow \infty} \frac{1 - F(\varphi(x) - \frac{1}{tz}|x)}{1 - F(\varphi(x) - \frac{1}{t}|x)} = z^{-\rho_x} \quad \textit{for all} \quad z > 0 \tag{2.2}$$

[regular variation with exponent $-\rho_x$, notation $1 - F(\varphi(x) - \frac{1}{t}|x) \in RV_{-\rho_x}$]. In this case the norming constants b_n can be chosen as : $b_n = \varphi(x) - \varphi_{1-(1/nF_X(x))}(x)$.

(iii) Given (2.2), if $\mathbb{E}(Y^k|X \leq x) < \infty$ for some integer $k \geq 1$, then

$$\lim_{n \to \infty} \mathbb{E}\{b_n^{-1}(\varphi(x) - \hat{\varphi}_1(x))\}^k = \Gamma(1 + k\rho_x^{-1}),$$

where $\Gamma(\cdot)$ denotes the gamma function.

(iv) Given (2.2), if $\mathbb{E}(Y^2|X \leq x) < \infty$ then

$$\lim_{n \to \infty} \mathbb{P}\left[\frac{\hat{\varphi}_1(x) - \mathbb{E}(\hat{\varphi}_1(x))}{\{Var(\hat{\varphi}_1(x))\}^{1/2}} \leq y\right] = \Psi_{\rho_x}[\{\Gamma(1 + 2\rho_x^{-1}) - \Gamma^2(1 + \rho_x^{-1})\}^{1/2} y - \Gamma(1 + \rho_x^{-1})].$$

Remark 2.1. Since $F_X(x)[1 - F(\varphi(x) - \frac{1}{t}\,|\,x)] \in \mathrm{RV}_{-\rho_x}$ by (2.2), this function can be represented as $t^{-\rho_x} L_x(t)$ with $L_x(\cdot) \in \mathrm{RV}_0$ (L_x being slowly varying) and so, the extreme-value condition (2.2) holds if and only if we have the following representation

$$F_X(x)[1 - F(y\,|\,x)] = L_x\left(\frac{1}{\varphi(x) - y}\right)(\varphi(x) - y)^{\rho_x} \quad \text{as} \quad y \uparrow \varphi(x). \tag{2.3}$$

In the particular case where $L_x\left(\frac{1}{\varphi(x)-y}\right) = \ell_x$ is a strictly positive function in x, it is shown in the following corollary that $b_n \sim (n\ell_x)^{-1/\rho_x}$.

Corollary 2.1. *Given* (2.3) *or equivalently* (2.2) *with* $L_x\left(\frac{1}{\varphi(x)-y}\right) = \ell_x > 0$, *we have*

$$(n\ell_x)^{1/\rho_x}\big(\varphi(x) - \hat{\varphi}_1(x)\big) \xrightarrow{d} \text{Weibull}(1, \rho_x) \quad \text{as} \quad n \to \infty.$$

Remark 2.2. Consider the assumption that the joint density function of (X, Y) satisfies

$$f(x, y) = c_x \{\varphi(x) - y\}^{\beta} + o(\{\varphi(x) - y\}^{\beta}) \quad \text{as} \quad y \uparrow \varphi(x), \tag{2.4}$$

for some constant $\beta > -1$, with c_x being a strictly positive function in x. Under the restrictive condition that the density f is strictly positive on the frontier (*i.e.* $\beta = 0$) among others, Park et al (2000) have obtained the limiting Weibull distribution of the FDH estimator with the convergence rate $n^{-1/(p+1)}$. When β may be non nul, Hwang et al (2002) have obtained the asymptotic Weibull distribution with the convergence rate $n^{-1/(\beta+2)}$ in the simple case where $p = 1$ (here it is also assumed that (2.4) holds uniformly in a neighborhood of the point at which we want to estmate $\varphi(\cdot)$ and that this frontier function is strictly increasing in that neighborhood and satisfies a Lipschitz condition of order 1). In the general setting where $p \geq 1$ and $\beta = \beta_x > -1$ may depend on x, we have the following more general result which involves the link between the regular variation index ρ_x, the dimension $p+1$ of the data and the shape parameter β_x of the joint density near the boundary.

Corollary 2.2. *If the condition of Corollary 2.1 holds with the functions* $\ell_x > 0$, $\rho_x > p$ *and* $\varphi(x)$ *being differentiable and the partial derivatives of* $\varphi(x)$ *being strictly positive, then* (2.4) *holds with* $\beta_x = \rho_x - (p+1)$ *and we have*

$$(n\ell_x)^{1/(\beta_x+p+1)}\big(\varphi(x) - \hat{\varphi}_1(x)\big) \xrightarrow{d} \text{Weibull}(1, \beta_x + p + 1) \quad \text{as} \quad n \to \infty.$$

Remark 2.3. We assume the diffrentiability of the functions ℓ_x, ρ_x with $\rho_x > p$ and $\varphi(x)$ in order to ensure the existence of the joint density near its support boundary. We distinguich between three different behaviors of this density at the frontier point $(x, \varphi(x)) \in \mathbb{R}^{p+1}$ following the value of ρ_x compared with the dimension $(p+1)$: when $\rho_x > p+1$ the joint density decays to zero at a speed of power $\rho_x - (p+1)$ of the distance from the frontier; when $\rho_x = p+1$ the density has a sudden jump at the frontier; when $\rho_x < p+1$ the density rises up to infinity at a speed of power $\rho_x - (p+1)$ of the distance from the frontier. The case $\rho_x \leq p+1$ corresponds to sharp or fault-type frontiers.

As an immediate consequence of Corollary 2.2, when $p = 1$ and $\beta_x = \beta$ (or equivalently $\rho_x = \rho$) does not depend on x, we obtain the convergence in distribution of the FDH estimator as in Hwang et al (2002) with the same convergence rate $n^{-1/(\beta+2)}$ (in the notations of Theorem 1 in Hwang et al (2002), $\mu(x) = \ell_x(\beta+2)\varphi'(x) = \ell_x\rho_x\varphi'(x)$). In the other particular case where the joint density is strictly positive on the frontier, we achieve the best rate of convergence $n^{-1/(p+1)}$ as in Park et al, 2000 (in the notations of Theorem 3.1 in Park et al (2000), $\mu_{NW,0}/y = \ell_x^{1/(p+1)} = \ell_x^{1/\rho_x}$).

Note also that the condition (2.4) has been considered by Hardle, Park and Tsybakov (1995) and by Hall, Park and Stern (1998). In a next section (see Conditional tail index extimation) we answer the important question of how to estimate the shape parameter β_x in (2.4) or equivalently the regular variation exponent ρ_x in (2.2).

On the other hand, as an immediate consequence of Theorem 2.1 *(iii)* in conjunction with Corollary 2.2, we obtain

$$\mathbb{E}\{\varphi(x)-\hat{\varphi}_1(x)\}^k = k\{\beta_x+p+1\}^{-1}\{n\ell_x\}^{-k/(\beta_x+p+1)}\Gamma\left(\frac{k}{\beta_x+p+1}\right)+o(n^{-k/(\beta_x+p+1)}). \quad (2.5)$$

This extends the limit theorem of moments of Park et al (2000, Theorem 3.3) to the more general setting where β_x may be non nul. Likewise, Hwang et al (2002, see Remark 1) provides (2.5) only for $k \in \{1,2\}$, $p=1$ and $\beta_x = \beta$. The result (2.5) also reflects the well known curse of dimensionality from which suffers the FDH estimator $\hat{\varphi}_1(x)$ as the number p of inputs-usage increases, pointed out earlier by Park et al (2000) in the particular case where $\beta_x = 0$.

Robust frontier estimators

By an appropriate choice of the order α as a function of n, Aragon et al (2005) have shown that the empirical partial frontier $\hat{\varphi}_\alpha(x)$ estimates the full frontier $\varphi(x)$ itself and convereges to the same Weibull distribution as the FDH $\hat{\varphi}_1(x)$ under the restrictive conditions of Park et al (2000). The next theorem gives more insights and generalizes their main result.

Theorem 2.2. *(i) If $b_n^{-1}(\hat{\varphi}_1(x)-\varphi(x)) \xrightarrow{d} G_x$, then for each $k \in \{0,1,\ldots,n\hat{F}_X(x)-1\}$,*

$$b_n^{-1}\left(\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)-\varphi(x)\right) \xrightarrow{d} H_x$$

for the distribution function $H_x(y) = G_x(y)\sum_{i=0}^{k}(-\log G_x(y))^i/i!$.

(ii) Suppose the upper bound of the support of Y is finite. If $b_n^{-1}(\hat{\varphi}_1(x)-\varphi(x)) \xrightarrow{d} G_x$, then $b_n^{-1}(\hat{\varphi}_{\alpha_n}(x)-\varphi(x)) \xrightarrow{d} G_x$ for all sequences $\alpha_n \to 1$ satisfying $nb_n^{-1}(1-\alpha_n) \to 0$.

Remark 2.4. When the FDH $\hat{\varphi}_1(x)$ converges in distribution, the conditional quantile-based estimator $\hat{\varphi}_{\alpha(n)}(x)$, for $\alpha(n) = 1-k/n\hat{F}_X(x)$, estimates the frontier function $\varphi(x)$ itself and converges in distribution as well, with the same scaling but a different limit distribution (here $nb_n^{-1}(1-\alpha(n)) \to \infty$). To recover the same limit distribution as the FDH estimator, it suffices to choose $\alpha(n) \to 1$ rapidly so that $nb_n^{-1}(1-\alpha(n)) \to 0$. This extends the main result of Aragon *et al* (2005, Theorem 4.3) where the convergence rate achieves $n^{-1/(p+1)}$ under the restrictive assumption that the joint density of (X, Y) is strictly positive on the frontier. Note also that the estimate $\hat{\varphi}_{\alpha(n)}$ does not envelop all the data points providing a robust alternative to the FDH frontier $\hat{\varphi}_1$: see Daouia and Ruiz-Gazen (2006) for an analysis of its quantitative and qualitative robustness properties.

Conditional tail index extimation

An important question in the setup of the obtaind results, is how to estimate ρ_x from the multivariate random sample of production units (X_i, Y_i), $i = 1, \ldots, n$. This problem is very similar to that of the estimation of the so-called extreme value index based rather on a sample of *univariate* random variables (see, *e.g.*, Embrechts *et al.*, 1997 and the references therein). An attractive estimation method has been proposed by Pickands (1975). This procedure can be easily adapted to our approach: let $k = k_n$ be a sequence of integers tending to infinity and let $k/n \to 0$ as $n \to \infty$. A Pickands type estimate of ρ_x can be derived as:

$$\hat{\rho}_x = \log 2 \left(\log \frac{\hat{\varphi}_{1-\frac{2k-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{4k-1}{n\hat{F}_X(x)}}(x)}{\hat{\varphi}_{1-\frac{k-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k-1}{n\hat{F}_X(x)}}(x)} \right)^{-1}. \tag{2.6}$$

We show in the next theorem that this estimate is weakly consistent and that if k_n increases suitably rapidly, then there is strong consistency. We also give extreme-value conditions under which $\hat{\rho}_x$ is asymptotically normal. This result is particularly important since it allows to test the hypothesis $\rho_x > 0$ and will be employed in the next section (see Asymptotic Normal distribution) to derive confidence intervals for $\varphi(x)$.

Theorem 2.3. *(i) If* (2.2) *holds,* $k_n \to \infty$ *and* $k_n/n \to 0$*, then* $\hat{\rho}_x \xrightarrow{P} \rho_x$.

(ii) If (2.2) *holds,* $k_n/n \to 0$ *and* $k_n/\log\log n \to \infty$*, then* $\hat{\rho}_x \xrightarrow{a.s.} \rho_x$.

(iii) Assume that $U(t) := \varphi_{1-\frac{1}{tF_X(x)}}(x)$ *has a positive derivative and that there exists a positive function* $A(\cdot)$ *such that for* $z > 0$

$$\lim_{t\uparrow\infty} \frac{(tz)^{1+\frac{1}{\rho_x}}\, U'(tz) - t^{1+\frac{1}{\rho_x}}\, U'(t)}{A(t)} = \pm\log(z),$$

for either choice of the sign [Π-variation, notation $\pm t^{1+\frac{1}{\rho_x}} U'(t) \in \Pi(A)$ *]. Then*

$$\sqrt{k_n}(\hat{\rho}_x - \rho_x) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\rho_x)), \tag{2.7}$$

with asymptotic variance $\sigma^2(\rho_x) = (2^{1-\frac{2}{\rho_x}} + 1)/\{(2^{-\frac{1}{\rho_x}} - 1)\log 4\}^2$*, for* $k_n \to \infty$ *satisfying* $k_n = o(n/g^{-1}(n))$*, where* g^{-1} *is the generalized inverse function of* $g(t) = t^{3+\frac{2}{\rho_x}}\{U'(t)/A(t)\}^2$.

(iv) If for some $\kappa > 0$ *and* $\delta > 0$ *the function* $\left\{t^{\rho_x-1}F'(\varphi(x) - \frac{1}{t}|x) - \delta\right\} \in RV_{-\kappa}$*, then* $\sqrt{k_n}(\hat{\rho}_x - \rho_x) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\rho_x))$*, for* $k_n \to \infty$ *satisfying* $k_n = o(n/g^{-1}(n))$*, with* $g(t) = t^{3+\frac{2}{\rho_x}}\left\{U'(t)/\left(t^{1+\frac{1}{\rho_x}}U'(t) - [\delta F_X(x)]^{-1/\rho_x}(\rho_x)^{\frac{1}{\rho_x}-1}\right)\right\}^2$.

Remark 2.5. Note that the second-order regular variation conditions *(iii)* and *(iv)* of Theorem 2.3 are difficult to check in practice, which makes the theoretical choice of the sequence $\{k_n\}$ a hard problem. In practice, in order to choose a reasonable estimate $\hat{\rho}_x(k_n)$ of ρ_x, one can make the plot of $\hat{\rho}_x$ consisting of the points $\{(k, \hat{\rho}_x(k)), 1 \le k < n\hat{F}_X(x)/4\}$, and pick out a value of ρ_x at which the obtained graph looks stable (see also Remark 2.7). As it will be illustrated in Section 3, good plots of $\hat{\rho}_x$ may require a large sample of the order of several thousand.

Asymptotic Normal distribution

Another question of particular interest is how to derive asymptotically normal estimates of high partial frontiers $\varphi_\alpha(x)$, when $\alpha = \alpha(n) \uparrow 1$, and of the true full frontier $\varphi(x)$ itself.

Theorem 2.4. *(i) Assume that* $F(\cdot|x)$ *has a positive density* $F'(\cdot|x)$ *and that* $F'(\varphi(x) - \frac{1}{t}|x) \in RV_{1-\rho_x}$*. Then*

$$\sqrt{2k_n}\,\frac{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \varphi_{1-\frac{p_n}{F_X(x)}}(x)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} \mathcal{N}(0, V_1(\rho_x)) \tag{2.8}$$

where $V_1(\rho_x) = \rho_x^{-2}2^{1-\frac{2}{\rho_x}}/\left(2^{-1/\rho_x} - 1\right)^2$*, provided* $p_n \to 0$*,* $np_n \to \infty$ *and* $k_n = [np_n]$.

(ii) Suppose the conditions of Theorem 2.3 (iii) or (iv) hold and define

$$\hat{\varphi}_1^*(x) := \frac{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)}{2^{1/\hat{\rho}_x} - 1} + \hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x). \tag{2.9}$$

Then

$$\sqrt{2k_n}\,\frac{\hat{\varphi}_1^*(x) - \varphi(x)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} \mathcal{N}(0, V_2(\rho_x)), \tag{2.10}$$

where $V_2(\rho_x) = 3\rho_x^{-2}2^{-1-\frac{2}{\rho_x}}/\left(2^{-1/\rho_x} - 1\right)^6$.

Let us now consider the asymptotic behavior of $\hat{\varphi}_1^*(x)$ in two particular cases: the case when ρ_x is known (here we will denote the resulting estimator $\tilde{\varphi}_1^*(x)$) and the case where k is fixed (here we denote the estimator $\overline{\varphi}_1^*(x)$).

Theorem 2.5. *(i) Suppose the conditions of Theorem 2.3 (iii) or (iv) hold and define*

$$\tilde{\varphi}_1^*(x) := \frac{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)}{2^{1/\rho_x}-1} + \hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x). \tag{2.11}$$

Then

$$\sqrt{2k_n}\,\frac{\tilde{\varphi}_1^*(x)-\varphi(x)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} \mathcal{N}(0, V_3(\rho_x)), \tag{2.12}$$

where $V_3(\rho_x) = \rho_x^{-2}2^{-\frac{2}{\rho_x}} / \left(2^{-1/\rho_x}-1\right)^4$.

(ii) Assume that (2.2) *holds and define* $\overline{\varphi}_1^*(x) := \hat{\varphi}_1^*(x)$ *with* k *fixed. Then*

$$\frac{\overline{\varphi}_1^*(x)-\varphi(x)}{\hat{\varphi}_{1-\frac{k-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k-1}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} (1-2^{-1/\rho_x})^{-1} + \{e^{-H_k/\rho_x}-1\}^{-1}, \tag{2.13}$$

where H_k *is a random variable having a beta-type density function:*

$$f_k(z) = \frac{(2k)!}{k!(k-1)!}e^{-(k+1)z}(1-e^{-z})^{k-1}, \quad \text{for } z \geq 0.$$

Remark 2.6. Note that Theorem 2.4 *(ii)* is still valid if the estimate $\hat{\rho}_x$ is replaced by the true value ρ_x up to a change of the asymptotic variance.

Remark 2.7. Theorems 2.1-2.5 follow easily by applying the elegant devices of Dekkers and de Haan (1989), among others, in conjunction with the clever idea that $\varphi(x)$ coincides with the right endpoint of the common distribution of the *univariate* random variables $Z_i^x := Y_i 1\!\!1(X_i \leq x),\ i = 1, ..., n$. It is also clear from the proofs that $\hat{\rho}_x$ as well as the estimates of the partial and full frontiers can be easily computed in practice by employing the order statistics $Z_{(n-k)}^x = \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$. This identity can also be of some help in finding an optimal choice for the sequence k_n. Indeed various selection methods for k_n suggested in the context of univariate extreme value estimation (see *e.g.* Guillou and Hall 2001 and the references therein) could be adapted to our problem. Of course, the selected number k_n of extreme observations Z_i^x involved in the definition of the estimators $\hat{\rho}_x$, $\hat{\varphi}_1^*(x)$ and $\tilde{\varphi}_1^*(x)$ should depend on the fixed level $x \in \mathbb{R}^p$ of inputs-usage. In Section 3, we suggest a simple data driven method for selecting reasonable values of $k_n(x)$ for a set of grid of values for x.

Remark 2.8. In the particular case where $L_x\left(\frac{1}{\varphi(x)-y}\right) = \ell_x$ in (2.3), the condition of Theorem 2.4 *(i)* holds, that is $F'(\varphi(x) - \frac{1}{t}\,|\,x) = \frac{\ell_x\rho_x}{F_X(x)}\left(\frac{1}{t}\right)^{\rho_x-1} \in \mathrm{RV}_{1-\rho_x}$. But the conditions of Theorem 2.3 *(iii)* and *(iv)* do not hold in this particular case since both functions $t^{1+\frac{1}{\rho_x}}U'(t) = \frac{1}{\rho_x}\left(\frac{1}{\ell_x}\right)^{1/\rho_x}$ and $t^{\rho_x-1}F'(\varphi(x) - \frac{1}{t}\,|\,x) = \frac{\ell_x\rho_x}{F_X(x)}$ are constant in t. Nevertheless, we have $t^{1-\gamma_x}U'(t) \equiv$ constant (with the notation γ_x of our proofs), so that the left-hand side of Equation (2.3) in Dekkers and de Haan (1989) is identically zero. It follows that the conclusion of Theorem 2.3 *(iii)* or *(iv)* holds for all sequences $k_n \to \infty$ satisfying $\frac{k_n}{n} \to 0$. The same is true for the conclusion of Theorem 2.4 *(ii)*. The next theorem gives another variant of this result.

Theorem 2.6. *Suppose the condition of Corollary 2.1 holds. Then*

$$\frac{\rho_x k_n^{1/2}}{(k_n/n\ell_x)^{1/\rho_x}}\left[\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) + (k_n/n\ell_x)^{1/\rho_x} - \varphi(x)\right] \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty,$$

provided that $k_n \to \infty$ *and* $k_n/n \to 0$ *as* $n \to \infty$.

Alternatively, we have the following formulation.

Corollary 2.3. *Suppose* $\alpha_n \uparrow 1$ *and* $n(1-\alpha_n) \to \infty$ *as* $n \to \infty$. *Then, under the condition of Theorem 2.6,*

$$\frac{\rho_x k_n^{1/2}}{\mathcal{B}_n}\left[\hat{\varphi}_{\alpha_n}(x) - \varphi(x) + \mathcal{B}_n\right] \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty,$$

where $\mathcal{B}_n = (k_n/n\ell_x)^{1/\rho_x}$ *with* $k_n - 1$ *being the integral part of* $n(1-\alpha_n)\hat{F}_X(x)$.

Example 2.1. We consider the case where the monotone boundary of the support of (X,Y) is linear. We choose (X,Y) uniformly distributed over the region $D = \{(x,y)\,|\, \le x \le 1, 0 \le y \le x\}$. In this case (see Florens and Simar 2005, Daouia and Ruiz-Gazen 2006, among others), it can be easily seen that $\varphi(x) = x$ and $F_X(x)[1 - F(y\,|\,x)] = (\varphi(x) - y)^2$ for all $0 \le y \le \varphi(x)$. Thus $L_x(\cdot) = \ell_x = 1$ and $\rho_x = 2$ for all x. Therefore the conclusions of all Theorems 2.1-2.6 hold (see Remark 2.8).

Example 2.2. We now choose a non linear monotone upper boundary 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 $\lambda = 3$ (see, *e.g.*, Gijbels, Mammen, Park and Simar 1999). Here, the frontier function is $\varphi(x) = x^{1/2}$ and the conditional distribution function is $F(y|x) = 3x^{-1}y^2 - 2x^{-3/2}y^3$, for $0 < x \le 1$ and $0 \le y \le \varphi(x)$. It is then easy to see that the extreme-value condition (2.2), or equivalently (2.3), holds with $\rho_x = 3$ and $L_x(z) = F_X(x)[3\varphi(x) - \frac{2}{z}]/[\varphi(x)]^4$ for all $x \in]0,1]$ and $z > 0$.

3 Finite Sample Performance

The simulation experiments of this section illustrate how the convergence results work out in practice. We also apply our approach to a real data set.

3.1 Simulated samples

We will simulate a sample of size $n = 1000$ and one of size $n = 5000$ according the scenario of Example 2.1 above. Here $\varphi(x) = x$ and $\rho_x = 2$. By construction of the estimators $\hat{\rho}_x$, $\hat{\varphi}_1^*(x)$ and $\tilde{\varphi}_1^*(x)$, the threshold $k_n(x)$ can vary between 1 and $n\hat{F}_X(x)/4$. Denote by $N_x = n\hat{F}_X(x)$ the number of observations (X_i, Y_i) with $X_i \leq x$.

The number of extreme observations (X_i, Y_i), with $X_i \leq x$, used to estimate ρ_x and $\varphi(x)$ is $4k_n(x)$ for $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$, whereas it is $2k_n(x)$ for $\tilde{\varphi}_1^*(x)$. Then, as it can be expected, if $k_n(x)$ or N_x is too small, the variance of the estimator of $\hat{\rho}_x$ may be large because of the large variation of the few extreme observations (X_i, Y_i), with $X_i \leq x$, involved in the estimation of ρ_x: this large variation may result in negative or too large values of $\hat{\rho}_x$. It is easy to see that $\hat{\rho}_x \geq 0$ if and only if

$$\hat{\varphi}_{1-\frac{k_n(x)-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n(x)-1}{n\hat{F}_X(x)}}(x) \leq \hat{\varphi}_{1-\frac{2k_n(x)-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{4k_n(x)-1}{n\hat{F}_X(x)}}(x).$$

Likewise, the confidence bands $\hat{\rho}_x \pm 1.96\sigma(\hat{\rho}_x)/\sqrt{k_n(x)}$ of ρ_x obtained from (2.7) may be negative or too large. In particular, the use of small input valuex x may result in disappointing estimates $\hat{\rho}_x$ (and also for $\hat{\varphi}_1^*(x)$ that suffers from the vexing defects of $\hat{\rho}_x$) and corresponding confidence bands due to the conditioning by $X \leq x$ (this is a border effect).

This exactly what happens for the first case when our sample size $n = 1000$ and for values of x as small as 0.25 (see Table 1). On the contrary, the estimator $\tilde{\varphi}_1^*(x)$ computed with the true value of $\rho_x = 2$ provides more reasonable estimates of $\varphi(x)$ and is rather stable with respect to the choice of $k_n(x)$.

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
12	2.99	0.03	5.96	0.336	-0.414	1.087	0.264	0.179	0.348
11	2.48	-0.08	5.04	0.294	-0.221	0.808	0.261	0.178	0.343
10	1.68	-0.13	3.50	0.228	0.003	0.453	0.246	0.172	0.320
9	4.97	-0.77	10.71	0.480	-1.648	2.607	0.277	0.187	0.368
8	3.66	-0.80	8.12	0.397	-0.861	1.656	0.284	0.187	0.380
7	3.55	-1.07	8.16	0.379	-0.812	1.570	0.280	0.183	0.376
6	1.98	-0.78	4.74	0.262	-0.083	0.607	0.263	0.179	0.347
5	2.27	-1.20	5.75	0.262	-0.147	0.672	0.251	0.173	0.328

Table 1: *Estimation at $x = 0.25$, sample size $n = 1000$. Here $N_x = n\hat{F}_X(x) = 67$.*

Tables 2–4 show the results for larger values of x and, as expected, the estimators $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$ behaves better, at least for appropriate values of $k_n(x)$. Again $\tilde{\varphi}_1^*(x)$ performs rather well and is stable to the selected value of $k_n(x)$.

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
67	1.45	0.85	2.06	0.435	0.269	0.600	0.515	0.443	0.586
63	1.44	0.82	2.06	0.437	0.271	0.604	0.518	0.445	0.590
59	1.52	0.84	2.19	0.456	0.268	0.643	0.525	0.450	0.599
55	1.17	0.63	1.71	0.404	0.295	0.513	0.510	0.441	0.580
51	1.34	0.70	1.98	0.425	0.285	0.565	0.508	0.438	0.578
47	1.25	0.63	1.87	0.407	0.292	0.522	0.490	0.425	0.555
43	1.10	0.52	1.68	0.382	0.300	0.463	0.467	0.409	0.525
39	1.08	0.48	1.67	0.384	0.307	0.461	0.466	0.409	0.523

Table 2: *Estimation at $x = 0.50$, sample size $n = 1000$. Here $N_x = n\hat{F}_X(x) = 268$.*

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
121	1.52	1.05	1.99	0.615	0.448	0.782	0.702	0.636	0.768
113	1.49	1.01	1.96	0.596	0.445	0.747	0.680	0.618	0.743
105	1.41	0.94	1.88	0.581	0.447	0.716	0.673	0.612	0.734
97	1.61	1.05	2.16	0.619	0.443	0.796	0.680	0.617	0.743
89	1.77	1.13	2.42	0.654	0.438	0.871	0.689	0.624	0.753
81	2.32	1.44	3.20	0.744	0.375	1.113	0.696	0.629	0.763
73	2.15	1.29	3.02	0.714	0.400	1.028	0.693	0.627	0.758
65	2.23	1.29	3.18	0.738	0.393	1.083	0.706	0.639	0.774

Table 3: *Estimation at $x = 0.75$, sample size $n = 1000$. Here $N_x = n\hat{F}_X(x) = 582$.*

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
194	2.38	1.80	2.97	1.084	0.655	1.512	0.985	0.911	1.059
180	1.88	1.40	2.36	0.925	0.665	1.185	0.952	0.883	1.022
166	2.69	1.97	3.41	1.159	0.611	1.707	0.991	0.916	1.066
152	2.96	2.13	3.78	1.220	0.558	1.882	0.995	0.919	1.071
138	2.23	1.58	2.88	1.032	0.655	1.409	0.983	0.909	1.056
124	3.96	2.73	5.19	1.449	0.241	2.657	1.013	0.934	1.093
110	1.95	1.31	2.58	0.982	0.685	1.279	0.992	0.918	1.067
96	2.41	1.57	3.25	1.079	0.624	1.535	1.002	0.925	1.079

Table 4: *Estimation at $x = 1.00$, sample size $n = 1000$. Here $N_x = n\hat{F}_X(x) = 1000$.*

When the sample size increases, the estimators behave much better, even for moderate values of x. Tables 5–8 display the results for $n = 5000$. The improvements of $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$ are remarkable, although the convergence is rather slow.

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
40	2.29	1.05	3.53	0.292	0.085	0.500	0.275	0.236	0.313
37	1.74	0.77	2.72	0.253	0.135	0.371	0.267	0.231	0.304
34	1.86	0.77	2.94	0.257	0.127	0.386	0.264	0.229	0.299
31	1.42	0.55	2.29	0.231	0.157	0.305	0.258	0.224	0.291
28	1.65	0.59	2.72	0.243	0.145	0.341	0.258	0.225	0.291
25	1.27	0.40	2.13	0.224	0.168	0.280	0.252	0.221	0.283
22	1.35	0.37	2.33	0.233	0.167	0.299	0.258	0.225	0.290
19	1.79	0.39	3.19	0.251	0.137	0.365	0.259	0.225	0.292

Table 5: *Estimation at $x = 0.25$, sample size $n = 5000$. Here $N_x = n\hat{F}_X(x) = 306$.*

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
300	1.83	1.47	2.19	0.487	0.367	0.606	0.512	0.478	0.546
270	2.19	1.73	2.65	0.553	0.378	0.729	0.526	0.490	0.561
240	2.51	1.96	3.07	0.605	0.371	0.839	0.532	0.496	0.569
210	2.01	1.54	2.48	0.510	0.367	0.653	0.509	0.475	0.543
180	1.68	1.26	2.11	0.467	0.368	0.565	0.501	0.469	0.533
150	1.56	1.12	1.99	0.447	0.367	0.528	0.489	0.458	0.519
120	1.18	0.81	1.55	0.418	0.372	0.464	0.482	0.453	0.510
90	1.45	0.93	1.97	0.441	0.376	0.506	0.478	0.450	0.506

Table 6: *Estimation at $x = 0.50$, sample size $n = 5000$. Here $N_x = n\hat{F}_X(x) = 1261$.*

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
500	1.94	1.65	2.24	0.744	0.612	0.876	0.755	0.721	0.788
465	2.02	1.70	2.34	0.762	0.619	0.905	0.759	0.725	0.793
430	2.44	2.04	2.84	0.848	0.637	1.059	0.769	0.734	0.804
395	2.36	1.96	2.77	0.832	0.633	1.032	0.769	0.734	0.804
360	2.18	1.79	2.58	0.790	0.622	0.958	0.760	0.726	0.794
325	2.39	1.94	2.85	0.822	0.622	1.022	0.762	0.727	0.796
290	2.55	2.04	3.06	0.838	0.614	1.063	0.758	0.724	0.792
255	2.11	1.66	2.56	0.774	0.618	0.931	0.759	0.725	0.793

Table 7: *Estimation at $x = 0.75$, sample size $n = 5000$. Here $N_x = n\hat{F}_X(x) = 2843$.*

$k_n(x)$	$\hat{\rho}_x$	CI low	CI up	$\hat{\varphi}_1^*(x)$	CI low	CI up	$\tilde{\varphi}_1^*(x)$	CI low	CI up
1126	2.01	1.81	2.22	1.010	0.867	1.152	1.006	0.972	1.039
1064	2.04	1.82	2.25	1.015	0.869	1.161	1.005	0.971	1.038
1002	2.21	1.97	2.45	1.066	0.894	1.237	1.008	0.974	1.042
940	2.01	1.78	2.23	1.001	0.861	1.142	0.999	0.966	1.032
878	2.31	2.05	2.58	1.096	0.907	1.285	1.016	0.981	1.050
816	1.94	1.71	2.17	0.988	0.855	1.120	1.002	0.968	1.035
754	1.88	1.65	2.11	0.968	0.845	1.090	0.995	0.962	1.028
692	1.88	1.64	2.12	0.966	0.844	1.087	0.991	0.959	1.024

Table 8: *Estimation at $x = 1.00$, sample size $n = 5000$. Here $N_x = n\hat{F}_X(x) = 5000$.*

3.2 Monte-Carlo experiment

We simulated 5000 samples of size $n = 1000$ to analyze the bias and the mean squared error of the different estimators. Here $\bar{k}_n(x)$ and $\bar{N}_x$ are the average values observed over the 5000 Monte-Carlo replications. In the caption of each table, we also specify the bias and the MSE of the FDH estimator. For the case $n = 1000$ (see Tables 9-12), the comments above are confirmed: we have disappointing estimates $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$, in particular for small values of x . We see the improvement of $\tilde{\varphi}_1^*(x)$ over the FDH in terms of the bias, without increasing too much the MSE. Remember also that these extreme-value estimators have a normal limiting distribution and so that they are much more easy to handle for providing confidence intervals. The same exercise was done for the case $n = 5000$ (see Tables 13-16) confirming the excellent behavior of the various estimators, in particular for larger values of x.

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
15.2	-0.371250	2076.162199	-0.040702	17.936723	0.000981	0.001468
14.2	-3.755379	55378.220022	-0.322270	418.449614	0.001193	0.001464
13.2	-4.144351	92811.628558	-0.289035	474.677747	0.001219	0.001477
12.2	-3.104349	54401.318738	-0.287991	375.844276	0.001344	0.001477
11.2	-0.546109	3405.493177	-0.047624	21.182847	0.001541	0.001475
10.2	-2.331159	6147.906639	-0.176869	34.096812	0.001791	0.001446
9.2	-0.978040	4221.365859	-0.078110	17.370830	0.002526	0.001441

Table 9: *5000 Monte-Carlo simulations. Estimation at* $x = 0.25$*, sample size* $n = 1000$*. Here* $\bar{N}_x = 62.4$*. For the FDH estimator we have* $B_{\hat{\varphi}_1(x)} = -0.027853$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000988$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
51.5	0.625810	290.102767	0.103443	8.618906	0.000120	0.001447
48.0	0.388598	932.254086	0.063367	24.204135	0.000254	0.001445
44.4	-0.150934	4253.678818	-0.028357	104.883774	-0.000206	0.001441
40.9	4.530279	60967.002870	0.597043	1078.215861	-0.000202	0.001427
37.3	15.068476	986776.259939	2.103253	19385.905948	-0.000417	0.001445
33.8	0.225957	915.667792	0.035012	16.368848	-0.000057	0.001424
30.2	0.939559	1700.529419	0.110049	22.505675	-0.000082	0.001460
26.7	-1.318316	4942.662655	-0.138662	54.347023	-0.000961	0.001569

Table 10: *5000 Monte-Carlo simulations. Estimation at* $x = 0.50$*, sample size* $n = 1000$*. Here* $\bar{N}_x = 250.1$*,* $B_{\hat{\varphi}_1(x)} = -0.028417$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.001024$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
140.3	0.230635	4.049966	0.063638	0.302785	-0.000492	0.001463
131.7	0.377855	88.847625	0.101028	6.350934	-0.000508	0.001432
123.0	0.446401	104.217617	0.114126	6.651523	-0.000508	0.001432
114.3	0.877701	1327.752528	0.213208	76.116650	-0.000020	0.001438
105.7	-0.859698	7682.098671	-0.211983	455.487287	0.000388	0.001456
97.0	0.392573	59.642659	0.090922	3.382413	0.000781	0.001453
88.4	-0.080519	6226.584550	-0.001355	268.189614	0.001195	0.001473
79.7	0.154144	1486.535566	0.030968	60.918632	0.001474	0.001505

Table 11: *5000 Monte-Carlo simulations. Estimation at* $x = 0.75$*, sample size* $n = 1000$*. Here* $\bar{N}_x = 562.7$*,* $B_{\hat{\varphi}_1(x)} = -0.027926$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000993$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
250.0	0.114960	0.611321	0.042945	0.084115	0.000632	0.001462
234.3	0.140516	1.446132	0.050816	0.193949	0.001038	0.001469
218.5	0.149288	0.690302	0.051984	0.085653	0.001134	0.001460
202.8	0.191437	4.403866	0.063331	0.453027	0.000791	0.001467
187.0	0.171695	0.606067	0.054599	0.059936	0.000609	0.001436
171.3	0.189518	0.939087	0.057629	0.083464	0.000282	0.001420
155.6	0.211142	1.844034	0.061681	0.144070	0.000126	0.001429
139.8	0.245505	8.868598	0.067017	0.565693	0.000639	0.001464

Table 12: *5000 Monte-Carlo simulations. Estimation at* $x = 1.00$*, sample size* $n = 1000$*. Here* $\bar{N}_x = 1000.0$*,* $B_{\hat{\varphi}_1(x)} = -0.028166$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.001005$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
77.8	0.265806	363.990553	0.023532	3.175978	0.000379	0.000290
74.4	0.420226	43.632246	0.037947	0.372272	0.000427	0.000284
71.1	0.313375	282.300865	0.028278	2.210215	0.000401	0.000284
67.7	-1.326063	13787.214141	-0.117202	106.834595	0.000409	0.000283
64.4	0.860081	493.882481	0.071043	3.372632	0.000456	0.000287
61.0	0.459680	48.319968	0.038234	0.302038	0.000662	0.000290
57.7	-0.971909	13717.553317	-0.076138	78.373638	0.000581	0.000285
54.3	0.784383	846.926494	0.056539	4.762474	0.000641	0.000289

Table 13: *5000 Monte-Carlo simulations. Estimation at* $x = 0.25$*, sample size* $n = 5000$*. Here* $\bar{N}_x = 312.6$*,* $B_{\hat{\varphi}_1(x)} = -0.012577$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000201$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
312.1	0.086268	0.224352	0.015670	0.007356	-0.000087	0.000290
297.0	0.086851	0.248763	0.015382	0.007745	-0.000225	0.000286
281.8	0.090149	0.263425	0.015550	0.007781	-0.000278	0.000284
266.7	0.091899	0.286461	0.015474	0.008059	-0.000371	0.000284
251.6	0.092156	0.309629	0.015032	0.008238	-0.000612	0.000285
236.5	0.104809	0.340925	0.016595	0.008506	-0.000505	0.000288
221.4	0.125818	0.422421	0.019278	0.009943	-0.000388	0.000293
206.3	0.134020	0.513226	0.019892	0.011207	-0.000475	0.000296

Table 14: *5000 Monte-Carlo simulations. Estimation at* $x = 0.50$*, sample size* $n = 5000$*. Here* $\bar{N}_x = 1249.7$*,* $B_{\hat{\varphi}_1(x)} = -0.012617$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000202$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
702.8	0.035728	0.079383	0.009740	0.005874	-0.000052	0.000292
668.2	0.038507	0.085129	0.010421	0.005992	0.000183	0.000288
633.5	0.043920	0.092585	0.011551	0.006156	0.000316	0.000286
598.8	0.044422	0.095697	0.011334	0.006017	0.000239	0.000282
564.2	0.051037	0.102044	0.012665	0.006044	0.000459	0.000277
529.5	0.050885	0.109137	0.012209	0.006040	0.000287	0.000277
494.8	0.055984	0.120578	0.012913	0.006236	0.000266	0.000277
460.1	0.067703	0.139597	0.015060	0.006759	0.000489	0.000283

Table 15: *5000 Monte-Carlo simulations. Estimation at* $x = 0.75$*, sample size* $n = 5000$*. Here* $\bar{N}_x = 2812.9$*,* $B_{\hat{\varphi}_1(x)} = -0.012698$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000206$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
1250.0	0.022876	0.042506	0.008493	0.005594	0.000340	0.000290
1188.0	0.024300	0.046131	0.008811	0.005795	0.000335	0.000296
1126.0	0.023621	0.049323	0.008382	0.005870	0.000227	0.000296
1064.0	0.022926	0.052949	0.007950	0.005959	0.000102	0.000299
1002.0	0.024986	0.054950	0.008414	0.005816	0.000191	0.000297
940.0	0.025146	0.058081	0.008216	0.005764	0.000129	0.000297
878.0	0.027870	0.063172	0.008820	0.005854	0.000216	0.000297
816.0	0.031612	0.069223	0.009598	0.005932	0.000294	0.000289

Table 16: *5000 Monte-Carlo simulations. Estimation at* $x = 1.00$*, sample size* $n = 5000$*. Here* $\bar{N}_x = 5000.0$*,* $B_{\hat{\varphi}_1(x)} = -0.012661$ *and* $MSE_{\hat{\varphi}_1(x)} = 0.000203$

3.3 A data driven method for selecting $k_n(x)$

In a real data set situation, the question of selecting the optimal value of $k_n(x)$ is still an open issue and is not addressed here. We only suggest an empirical rule that turns out to give reasonable estimates of the frontier in the simulated samples above.

The idea is to select a set of grid of values for x, then for each x, we select a grid of values for $k_n(x) = [N_x/4] - k + 1$, where k is an integer varying between 1 and the integral part

$[N_x/4]$ of $N_x/4$. For each pair we evaluate $\hat{\rho}_x(k)$ and we select the k where the variation of the results is the smaller. We achieve this by computing the standard deviations of $\hat{\rho}_x(k)$ over a "window" of 5 successive values of k. For the sample generated with $n = 5000$ illustrated above (Tables 5–8) we obtain the results shown in Figure 1. This empirical rule seems to provide reasonable estimates. The estimator $\tilde{\varphi}_1^*(x)$ is computed with the true value $\rho_x = 2$.

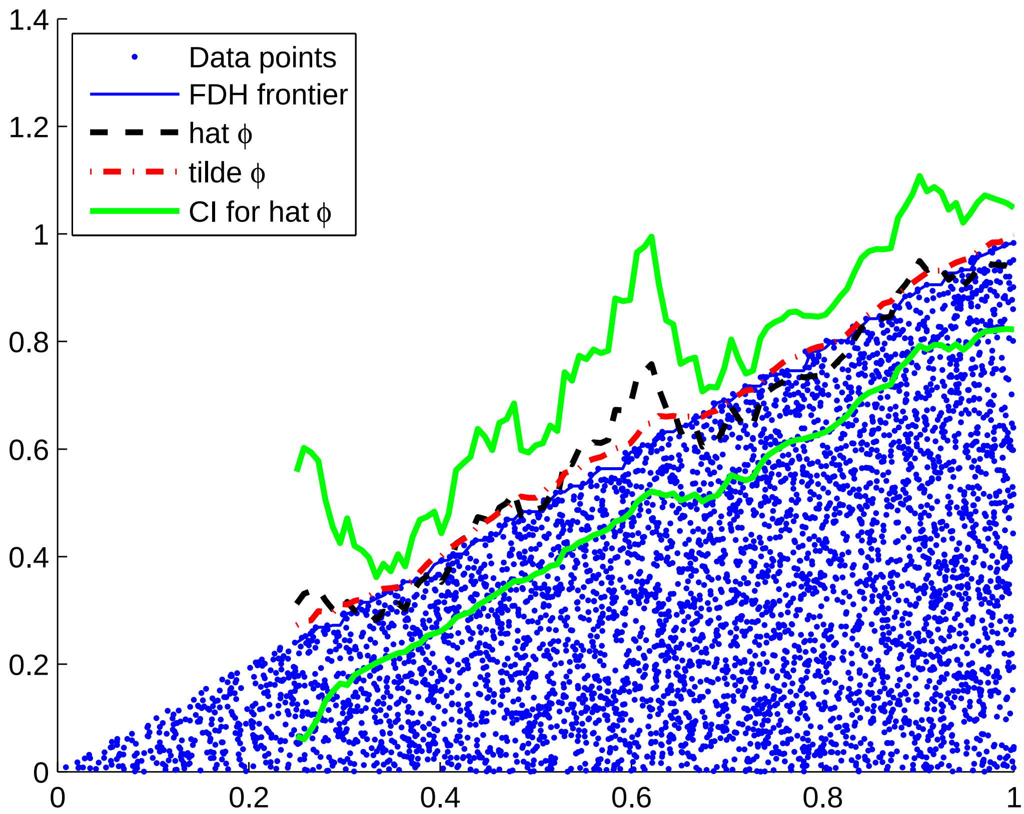

Figure 1: *One sample of size* $n = 5000$ *from Example 2.1.*

3.4 An application

We use the same real data example as in Cazals *et al.* (2002) and Daouia and Simar (2005) on the frontier analysis of 9521 French post offices observed in 1994, with X as the quantity of labor and Y as the volume of delivered mail. In this illustration, we only consider the $n = 4000$ observed post offices with the smallest levels x_i. We used the emprical rule explained above for selecting reasonable values for $k_n(x)$.

The cloud of points and the resulting estimates are provided in the top of Figure 2. The FDH estimator is clearly determined by only a few very extreme points. If we delete one extreme point from the sample (a clearly outlying point at the North-West of the cloud), we obtain the bottom picture: the FDH estimator changes drastically, whereas the extreme-values based estimators are very robust. We also display the values of $\tilde{\varphi}_1^*(x)$ computed with $\rho_x = \bar{\rho}_x = 1.0189$ being the average value of $\hat{\rho}_x$ computed over the 200 grid points of x: $\tilde{\varphi}_1^*(x)$ and $\hat{\varphi}_1^*(x)$ are very similar in this example.

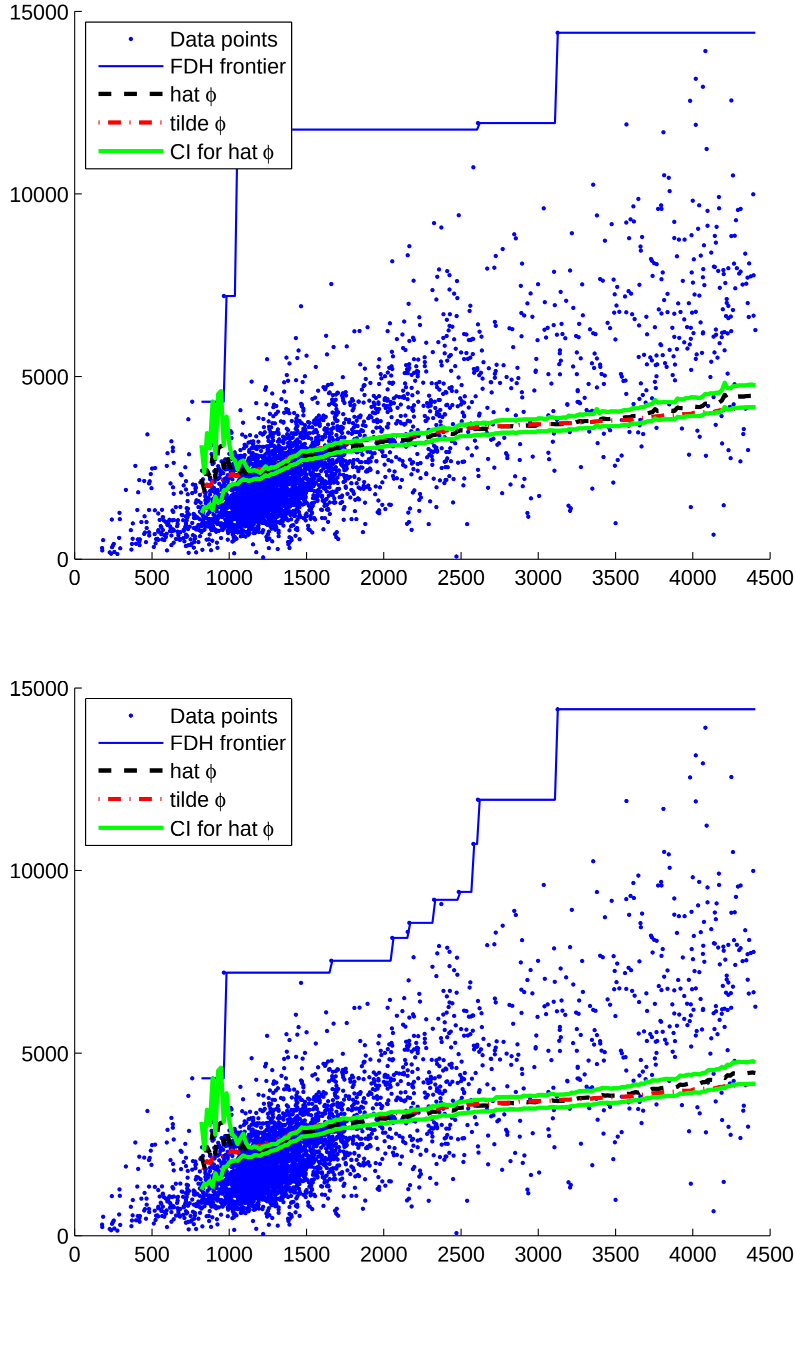

Figure 2: *French post offices data example.*

4 Conclusions

In our approach, we provide the necessary and sufficient condition for the FDH estimator to converge in distibution, we specify its asymptotic distribution with the appropriate convergence rate and provide a limit theorem of moments in a general framework. As an immediate consequence, we extend the previuos results of Park et al (2000) and Hwang et al (2002) to the general setting where $p \geq 1$ and $\beta = \beta_x$ may depend on x in the condition (2.4). We also give more insights and generalize the main result of Aragon et al (2005) on robust variants of the FDH estimator $\hat{\varphi}_1(x)$ and we provide a strongly consistent and asymptotically normal estimator of the unknown parameter ρ_x of the asymptotic Weibull distribution of $\hat{\varphi}_1(x)$. Moreover we answer the question of how ρ_x is linked to the dimension $p+1$ of the data and to the shape parameter β_x of the joint density of (X, Y) near its support boundary.

Compared to the frontier estimators defined by optimization problems (see, *e.g.*, Hall et al. 1998), our extreme value-based estimators $\hat{\varphi}_1(x)$, $\hat{\varphi}_{\alpha_n}(x)$, $\hat{\varphi}_1^*(x)$ and $\tilde{\varphi}_1^*(x)$ benefit from explicit and easy formulations. They also have the advantage to not be limited to a bidimensional support $\mathbb{T}$ since they do not require a partition of $\mathbb{T}$ as it is the case of the other extreme value-based estimators (see, *e.g.*, Gardes 2002) or piecewise polynomial estimators (see, *e.g.*, Korostelev and Tsybakov 1993, Hardle et al. 1995). Moreover, the new estimators $\hat{\varphi}_1^*(x)$ and $\tilde{\varphi}_1^*(x)$ are asymptotically normally distributed and provide useful asymptotic confidence bands for the monotone frontier function $\varphi(x)$. The study of extreme-value properties of the estimators considered in this paper, is easily carried out by relating them to a simple dimensionless random sample and then applying standard extreme-values theory (Dekkers and de Haan 1989,...).

We illustrate how the large sample theory applies in practice, by analyzing some simulated samples and by doing some Monte-Carlo experiment. Good estimates of the frontier and the conditional tail index may require a large sample of the order of several thousand. Selecting theoretically the optimal extreme conditional quantiles $\hat{\varphi}_{\alpha(k_n(x))}$ for estimating the frontier and/or the tail index is a difficult question that deserves for future work. Here, we suggest a simple data driven method that provides a reasonable choice of the sequence $\{k_n(x)\}$ for large samples.

Appendix: Proofs

Proof of Theorem 2.1 (i) Let $Z^x = Y 1\!\!1(X \leq x)$ and $F_x(\cdot) = 1 - F_X(x)[1 - F(\cdot|x)]$. It can be easily seen that $\mathbb{P}(Z^x \leq y) = F_x(y)$ for any $y \geq 0$. Therefore $\{Z_i^x = Y_i 1\!\!1(X_i \leq x), i = 1, ..., n\}$ is an iid sequence of random variables with common distribution function F_x. Moreover, it is easy to see that the right endpoint of F_x coincides with $\varphi(x)$ and that $\max_{i=1,\dots,n} Z_i^x$ coincides with $\hat{\varphi}_1(x)$. Thus according to the Fisher-Tippett Theorem, if there exists $b_n > 0$ such that $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G$ for a nondegenerate distribution function G, then $G(y) = e^{-(-y)^\rho}$ with support $]-\infty, 0]$ and $\rho > 0$ (see *e.g.* Embrechts et al. (1997), Theorem 3.2.3, p. 121).

(ii) On the other hand (see *e.g.* Embrechts et al. (1997), Theorem 3.3.12, p. 135), there exist norming constants b_n such that $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G$ (*i.e.*, F_x belongs to the domain of attraction of $G = \Psi_{\rho_x}$) if and only if

$$\bar{F}_x\left(\varphi(x) - \frac{1}{t}\right) \in \mathrm{RV}_{-\rho_x}, \tag{A.1}$$

where $\bar{F}_x = 1 - F_x$. This necessary and sufficient condition is equivalent to (2.2). In this case, b_n can be taken equal to $\varphi(x) - \inf\{y \geq 0 | F_x(y) \geq 1 - \frac{1}{n}\}$ which coincides with $\varphi(x) - \inf\{y \geq 0 | F(y|x) \geq 1 - \frac{1}{nF_X(x)}\}$.

(iii)-(iv) Under the given regularity conditions, we know that (A.1) holds and it is easy to see that $\int_{-\infty}^{\varphi(x)} |y|^k F_x(dy) < \infty$. Then it is immediate (see *e.g.* Resnick 1987, Proposition 2.1, p.77) that $\lim_{n\to\infty} \mathbb{E}\{b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))\}^k = (-1)^k \Gamma(1 + k/\rho_x)$. Likewise, the last result follows from *e.g.* Resnick (1987, Corollary 2.3, p.83). $\square$

Proof of Corollary 2.1 Following the proof of Theorem 2.1, we can set $b_n = \varphi(x) - F_x^{-1}(1 - \frac{1}{n})$ where $F_x^{-1}(t) = \inf\{y \in]0, \varphi(x)] : F_x(y) \geq t\}$ for all $t \in]0, 1]$. It follows from the regularity condition (2.3) that $F_x^{-1}(t) = \varphi(x) - ((1-t)/\ell_x)^{1/\rho_x}$ as $t \uparrow 1$. Whence $b_n = (1/n\ell_x)^{1/\rho_x}$ for all n sufficiently large. $\square$

Proof of Corollary 2.2 Under the given conditions, it can be easily seen from (2.3) that

$$f(x,y) = (\varphi(x) - y)^{\rho_x - (p+1)} \left[\ell_x \rho_x (\rho_x - 1) \cdots (\rho_x - p) \frac{\partial}{\partial x^1} \varphi(x) \cdots \frac{\partial}{\partial x^p} \varphi(x) + o(1)\right]$$

as $y \uparrow \varphi(x)$, where the term $o(1)$ depends on the partial derivatives of $x \mapsto \ell_x$, $x \mapsto \rho_x$ and $x \mapsto \varphi(x)$. $\square$

Proof of Theorem 2.2 (i) Let $Z_{(i)}^x$ be the i^{th} order statistic generated by the random variables $Z_1^x, ..., Z_n^x$. Following the standard extreme value theory (see *e.g.* van der Vaart

1998, Theorem 21.18, p. 313), if $b_n^{-1}(Z^x_{(n)} - \varphi(x)) \xrightarrow{d} G$, then $b_n^{-1}(Z^x_{(n-k)} - \varphi(x)) \xrightarrow{d} H$ for the distribution function $H(y) = G(y)\sum_{i=0}^{k}(-\log G(y))^i/i!$. The conclusion follows from

$$Z^x_{(n-k)} = \inf\left\{y \geq 0 | \frac{1}{n}\sum_{i=1}^{n} 1\!\!1_{(Z_i^x \leq y)} \geq 1 - \frac{k}{n}\right\} = \inf\left\{y \geq 0 | \hat{F}_n(x,y) + 1 - \hat{F}_X(x) \geq 1 - \frac{k}{n}\right\}$$

$$= \inf\left\{y \geq 0 | \hat{F}(x|y) \geq 1 - \frac{k}{n\hat{F}_X(x)}\right\} = \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$$

and from the fact that $\hat{\varphi}_1(x) = Z^x_{(n)}$.

(ii) Writting $b_n^{-1}(\hat{\varphi}_\alpha(x) - \varphi(x)) = b_n^{-1}(\hat{\varphi}_\alpha(x) - \hat{\varphi}_1(x)) + b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))$, it suffices to find an appropriate sequence $\alpha = \alpha_n \to 1$ so that $b_n^{-1}(\hat{\varphi}_{\alpha_n}(x) - \hat{\varphi}_1(x)) \xrightarrow{d} 0$. Aragon et al (2005, see the proof of Theorem 4.3, Equation (20)) showed that for any $\alpha > 0$:

$$|\hat{\varphi}_\alpha(x) - \hat{\varphi}_1(x)| \leq (1-\alpha)n\hat{F}_X(x)M \quad \text{with probability 1},$$

where $M < \infty$ is the upper bound of the support of Y. Thus it suffices to choose $\alpha = \alpha_n \to 1$ such that $nb_n^{-1}(1-\alpha_n) \to 0$. $\square$

Proof of Theorem 2.3 (i) Let us consider the random sample of univariate variables $Z_1^x, \ldots, Z_n^x$ introduced in the proof of Theorem 2.1, and let $\gamma_x = -1/\rho_x$ in (A.1). Then the Pickands estimate of the exponent of variation $\gamma_x < 0$ is given by:

$$\hat{\gamma}_x := (\log 2)^{-1} \log \frac{Z^x_{(n-k+1)} - Z^x_{(n-2k+1)}}{Z^x_{(n-2k+1)} - Z^x_{(n-4k+1)}} \equiv -\frac{1}{\hat{\rho}_x}.$$

Under (2.2), the condition (A.1) holds and so there exists $b_n > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\left[b_n^{-1}(Z^x_{(n)} - \varphi(x)) \leq y\right] = \Psi_{-1/\gamma_x}(y).$$

Since this limit is unique only up to affine transformations, we have

$$\lim_{n\to\infty} \mathbb{P}\left[c_n^{-1}(Z^x_{(n)} - d_n) \leq y\right] = \Psi_{-1/\gamma_x}(-\gamma_x y - 1) = \exp\left\{-(1+\gamma_x y)^{-1/\gamma_x}\right\},$$

for all $y \leq 0$, where $c_n = -\gamma_x b_n$ and $d_n = \varphi(x) - b_n$. Thus condition (1.1) in Dekkers and de Haan (1989) holds. Therefore $\hat{\gamma}_x \xrightarrow{P} \gamma_x$ if $k_n \to \infty$ and $\frac{k_n}{n} \to 0$ in view of Theorem 2.1 of Dekkers and de Haan (1989). This ends the proof of the weak consistency of $\hat{\rho}_x = -1/\hat{\gamma}_x$.

(ii) Likewise, if $\frac{k_n}{n} \to 0$ and $\frac{k_n}{\log\log n} \to \infty$, then $\hat{\gamma}_x \xrightarrow{a.s.} \gamma_x$ via Theorem 2.2 of Dekkers and de Haan (1989).

(iii) We have $U(t) = \inf\{y \geq 0 \,|\, \frac{1}{1-F_x(y)} \geq t\}$ which corresponds to the arrow means inverse function $(1/(1-F_x))^{-1}(t)$. Since $\pm t^{1-\gamma_x}U'(t) \in \Pi(A)$ with $\gamma_x = -1/\rho_x < 0$, it follows from Dekkers and de Haan (1989, Theorem 2.3) that $\sqrt{k_n}(\hat{\gamma}_x - \gamma_x) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\gamma_x))$ with $\sigma^2(\gamma_x) = \gamma_x^2(2^{2\gamma_x+1}+1)/\left\{2(2^{\gamma_x}-1)\log 2\right\}^2$ for $k_n \to \infty$ satisfying $k_n = o(n/g^{-1}(n))$, where $g(t) := t^{3-2\gamma_x}\left\{U'(t)/A(t)\right\}^2$. By the delta method we conclude that $\sqrt{k_n}(\hat{\rho}_x - \rho_x) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\rho_x))$, with asymptotic variance $\sigma^2(\rho_x) = \sigma^2(\gamma_x)/\gamma_x^2$.

(iv) Under the regularity condition, we have $\pm\left\{t^{-1-\frac{1}{\gamma_x}}F_x'(\varphi(x)-\frac{1}{t}) - \delta F_X(x)\right\} \in \mathrm{RV}_{-\kappa}$. The desired conclusion follows then immediately from Theorem 2.5 of Dekkers and de Haan (1989). $\square$

Proof of Theorem 2.4 (i) Under the regularity condition, the distribution function F_x of Z^x has a positive density $F_x'(y) = F_X(x)F'(y \mid x)$ and $F_x'(\varphi(x)-\frac{1}{t}) \in \mathrm{RV}_{1+\frac{1}{\gamma_x}}$. Therefore, according to Dekkers and de Haan (1989, Theorem 3.1), $\sqrt{2k_n}\,\dfrac{Z^x_{(n-k_n-1)} - F_x^{-1}(1-p_n)}{Z^x_{(n-k_n-1)} - Z^x_{(n-2k_n-1)}}$ is asymptotically normal with mean zero and variance $2^{2\gamma_x+1}\gamma_x^2/(2^{\gamma_x}-1)^2$. We conclude by using $Z^x_{(n-k)} = \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$ and $F_x^{-1}(1-p_n) = \varphi_{1-\frac{p_n}{F_X(x)}}(x)$.

(ii) We have $\hat{\varphi}_1^*(x) = \frac{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}}{2^{-\hat{\gamma}_x}-1} + Z^x_{(n-k_n+1)}$. Then following Dekkers and de Haan (1989, Theorem 3.2), $\sqrt{2k_n}\,\dfrac{\hat{\varphi}_1^*(x) - \varphi(x)}{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}}$ is asymptotically normal with mean zero and variance $3\gamma_x^2 2^{2\gamma_x-1}/(2^{\gamma_x}-1)^6$. This completes the proof. $\square$

Proof of Theorem 2.5 (i) We have

$$\sqrt{2k_n}\,\frac{\tilde{\varphi}_1^*(x) - \varphi(x)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} = \sqrt{2k_n}\left\{\frac{1}{2^{-\gamma_x}-1} + \frac{Z^x_{(n-k_n+1)} - \varphi(x)}{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}}\right\}.$$

Let $E_{(1)} \leq \ldots \leq E_{(n)}$ be the order statistics of iid exponential variables $E_1, \ldots, E_n$. Then $\{Z^x_{(n-k+1)}\}_{k=1}^n \stackrel{d}{=} \{U(e^{E_{(n-k+1)}})\}_{k=1}^n$. Writing $V(t) := U(e^t)$, we obtain

$$\begin{aligned}
\sqrt{2k_n}\,\frac{\tilde{\varphi}_1^*(x) - \varphi(x)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} &\stackrel{d}{=} \sqrt{2k_n}\left\{\frac{1}{2^{-\gamma_x}-1} + \frac{V(E_{(n-k_n+1)}) - \varphi(x)}{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})}\right\} \\
&= \left[-\sqrt{2k_n}\left\{\frac{V(\infty) - V(\log\frac{n}{2k_n})}{V'(\log\frac{n}{2k_n})} + \frac{1}{\gamma_x}\right\}\right. \\
&\quad + \sqrt{2k_n}\left\{\frac{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})}{2^{\gamma_x}V'(E_{(n-2k_n+1)})} - \frac{1-2^{-\gamma_x}}{\gamma_x}\right\}\frac{2^{\gamma_x}}{1-2^{\gamma_x}}\frac{V'(E_{(n-2k_n+1)})}{V'(\log\frac{n}{2k_n})} \\
&\quad \left. - \frac{\sqrt{2k_n}}{\gamma_x}\left\{\frac{V'(E_{(n-2k_n+1)})}{V'(\log\frac{n}{2k_n})} - 1 - \gamma_x\frac{V(E_{(n-k_n+1)}) - V(\log\frac{n}{2k_n})}{V'(\log\frac{n}{2k_n})}\right\}\right] \\
&\quad \times \frac{V'(\log\frac{n}{2k_n})}{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})}.
\end{aligned}$$

The first term at the right hand side tends to zero as established by Dekkers and de Haan (1989, proof of Theorem 3.2, p. 1809). The second term converges in distribution to $\mathcal{N}(0,1)\times \frac{2^{\gamma_x}}{1-2^{\gamma_x}}$ in view of Lemma 3.1 and Corollary 3.1 of Dekkers and de Haan (1989). The third term converges in probability to $\frac{\gamma_x}{2^{\gamma_x}-1}$ by the same Corollary 3.1. This ends the proof of (i).

(ii) As indicated in the proof of Theorem 2.3 (i), the extreme-value condition (1.1) in Dekkers and de Haan (1989) holds under (2.2). Therefore, it follows from Theorem 3.4 of Dekkers and de Haan (1989) that

$$\frac{\overline{\varphi}_1^*(x)-\varphi(x)}{\hat{\varphi}_{1-\frac{k-1}{n\hat{F}_X(x)}}(x)-\hat{\varphi}_{1-\frac{2k-1}{n\hat{F}_X(x)}}(x)}=\frac{\hat{\varphi}_1^*(x)-\varphi(x)}{Z^x_{(n-k+1)}-Z^x_{(n-2k+1)}}$$

converges in distribution to the random variable $(1-2^{\gamma_x})^{-1}+\{e^{\gamma_x H_k}-1\}^{-1}$ where H_k has the distribution of $\sum_{j=k+1}^{2k}E_j/j$ with $E_1,E_2,\ldots$ being iid standard exponential. The density of H_k is given in Remark 3.1 of Dekkers and de Haan (1989). This ends the proof of (ii). $\square$

Proof of Theorem 2.6 Write $\bar{F}_x(y):=F_X(x)[1-F(y|x)]$ and $F_x(y):=1-\bar{F}_x(y)$ for all $y\geq 0$. Let $R_x(y):=-\log\{\bar{F}_x(y)\}$ for all $y\in[0,\varphi(x)[$, and let $E_{(n-k_n+1)}$ be the $(n-k_n+1)^{\text{th}}$ order statistic generated by n independent standard exponential random variables. Then $Z^x_{(n-k_n+1)}$ has the same distribution as $R_x^{-1}[E_{(n-k_n+1)}]$, where

$$R_x^{-1}(t):=\inf\{y\geq 0|R_x(y)\geq t\}=\inf\{y\geq 0|F_x(y)\geq 1-e^{-t}\}:=F_x^{-1}(1-e^{-t}).$$

Hence

$$Z^x_{(n-k_n+1)}-F_x^{-1}\left(1-\frac{k_n}{n}\right)\stackrel{d}{=}R_x^{-1}[E_{(n-k_n+1)}]-R_x^{-1}\left[\log\left(\frac{n}{k_n}\right)\right]$$
$$=\left[E_{(n-k_n+1)}-\log\left(\frac{n}{k_n}\right)\right](R_x^{-1})'\left[\log\left(\frac{n}{k_n}\right)\right]+\frac{1}{2}\left[E_{(n-k_n+1)}-\log\left(\frac{n}{k_n}\right)\right]^2(R_x^{-1})''[\delta_n],$$

provided that $E_{(n-k_n+1)}\wedge\log(n/k_n)<\delta_n<E_{(n-k_n+1)}\vee\log(n/k_n)$. By the regularity condition (2.3), we have $R_x^{-1}(t)=\varphi(x)-(e^{-t}/\ell_x)^{1/\gamma_x}$ for all t large enough. Whence, for all n sufficiently large,

$$\frac{\rho_x k_n^{1/2}}{(k_n/n\ell_x)^{1/\rho_x}}\left[Z^x_{(n-k_n+1)}-F_x^{-1}\left(1-\frac{k_n}{n}\right)\right]\stackrel{d}{=}k_n^{1/2}\left[E_{(n-k_n+1)}-\log\left(\frac{n}{k_n}\right)\right]$$
$$-\frac{k_n^{1/2}}{2\rho_x}\left[E_{(n-k_n+1)}-\log\left(\frac{n}{k_n}\right)\right]^2\exp\left\{-\frac{1}{\rho_x}\left[\delta_n-\log\left(\frac{n}{k_n}\right)\right]\right\}.$$

Since $k_n^{1/2}[E_{(n-k_n+1)}-\log(n/k_n)]\xrightarrow{d}\mathcal{N}(0,1)$ and $|\delta_n-\log(n/k_n)|\leq|E_{(n-k_n+1)}-\log(n/k_n)|\xrightarrow{P}0$, as $n\to\infty$, we obtain

$$\frac{\rho_x k_n^{1/2}}{(k_n/n\ell_x)^{1/\rho_x}}\left[Z^x_{(n-k_n+1)}-F_x^{-1}(1-\frac{k_n}{n})\right]\xrightarrow{d}\mathcal{N}(0,1)\quad\text{as}\quad n\to\infty.$$

Since $F_x^{-1}(t) = \varphi(x) - ((1-t)/\ell_x)^{1/\rho_x}$ for all $t < 1$ large enough, we have $\varphi(x) - F_x^{-1}(1 - \frac{k_n}{n}) = (k_n/n\ell_x)^{1/\rho_x}$ for all n sufficiently large. Thus

$$\frac{\rho_x k_n^{1/2}}{(k_n/n\ell_x)^{1/\rho_x}} \left[Z^x_{(n-k_n+1)} + (k_n/n\ell_x)^{1/\rho_x} - \varphi(x) \right] \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty. \tag{A.2}$$

We conclude by using $Z^x_{(n-k_n+1)} = \hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x)$. $\square$

Proof of Corollary 2.3 We have $n - k_n + 1 \geq n - n(1-\alpha_n)\hat{F}_X(x) > n - k_n$, and so $\hat{\varphi}_{\alpha_n}(x) = \inf\{y \geq 0 | \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{(Z_i^x \leq y)} \geq 1 - (1-\alpha_n)\hat{F}_X(x)\} = Z^x_{(n-k_n+1)}$. We also have $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$, and so (A.2) holds. This ends the proof. $\square$

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] Cazals, C., J.P. Florens and L. Simar (2002), Nonparametric frontier estimation: a robust approach, *Journal of Econometrics*, 106, 1-25.

[3] Daouia, A. and A. Ruiz-Gazen (2006), Robust Nonparametric Frontier Estimators: Influence Function and Qualitative Robustness, *Statistica Sinica*, 16 (4), 1233–1253.

[4] Daouia, A. and L. Simar (2005), Robust Nonparametric Estimators of Monotone Boundaries, *Journal of Multivariate Analysis*, 96, 311–331.

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

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

[7] Dekkers, A.L.M. and L. de Haan (1989), On the estimation of extreme-value index and large quantiles estimation, *The Annals of Statistics*, 17(4), 1795–1832.

[8] 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.

[9] Embrechts, P., C. Klüppelberg and T. Mikosch (1997), *Modelling Extremal Events*, Springer.

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

[11] Gardes, L. (2002), Estimating the support of a Poisson process via the Faber-Shauder basis and extreme values, *Publications de l'Institut de Statistique de l'Université de Paris, XXXXVI*, 43–72.

[12] Gijbels, I. and Mammen, E. and Park, B. U. and Simar, L. (1999). On estimation of monotone and concave frontier functions. *J. Amer. Statistical Assoc.*, 94(445), 220–228.

[13] Guillou, A. and P. Hall (2001), A diagnostic for selecting the threshold in extreme-value analysis, *Journal of the Royal Statistical Society, Ser. B*, 63(2), 293–305.

[14] Hall, P. and Park, B.U. and Stern, S.E. (1998), On polynomial estimators of frontiers and boundaries, *Journal of Multivariate Analysis*, 66 (1), 71–98.

[15] Hardle, W. and Park, B.U. and Tsybakov, A.B. (1995), Estimation of non-sharp support boundaries, *Journal of Multivariate Analysis*, 43, 205–218.

[16] Hwang, J.H., B.U. Park and W. Ryu (2002), Limit theorems for boundary function estimators, *Statistics* & *Probability Letters*, 59, 353–360.

[17] Korostelev, A., Simar, L. and Tsybakov, A.B. (1995), Efficient estimation of monotone boundaries, *The Annals of Statistics*, 23, 476-489.

[18] Korostelev, A. and Tsybakov, A.B. (1993). *Minimax theory of image reconstruction.* Volume 82 of Lecture Notes in Statistics, Springer-verlag, New-York.

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

[20] Pickands, J. (1975), Statistical inference using extreme order statistics, *The Annals of Statistics*, 3, 119–131.

[21] Resnick, S.I. (1987), Extreme Values, Regular Variation, and Point Processes. Springer-Verlag.

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

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