TSE WP working paper series

10-165

Research Group: *Econometrics and Statistics*

2009

Frontier Estimation and Extreme Values Theory

ABDELAATI DAOUIA,

JEAN-PIERRE FLORENS

AND LEOPOLD SIMAR

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

Toulouse School of Economics

Frontier Estimation and Extreme Values Theory

Abdelaati Daouia*
Toulouse School of Economics (GREMAQ, University of Toulouse)
Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France. E-mail: daouia@cict.fr

Jean-Pierre Florens
Toulouse School of Economics (GREMAQ/IDEI, University of Toulouse)
Manufacture des Tabacs, 21 Allée de Brienne, 31000 Toulouse, France. E-mail: florens@cict.fr

Léopold Simar
Université Catholique de Louvain (Institut de Statistique)
Voie du Roman Pays 20, B-1348, Louvain-la-Neuve, Belgium. E-mail: leopold.simar@uclouvain.be

A running title : Monotone Frontier Estimation
** Corresponding Author*

Abstract : 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 Free Disposal Hull estimator in a general setup, to derive new and asymptotically Gaussian estimators and to provide useful asymptotic confidence bands for the monotone boundary function. The finite sample behavior of the suggested estimators is explored through Monte-Carlo experiments. We also apply our approach to a real data set on the production activity of the French postal services.

Key words : conditional quantile; extreme values; monotone boundary; production frontier

1 Introduction

In production theory and efficiency analysis, 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.* [4] 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 $x' \geq x$ (this inequality has to be understood componentwise) 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 $\phi(x) = \sup\{y \mid (x,y) \in \mathbb{T}\}$. Then the economic efficiency score of a firm operating at the level (x,y) is given by the ratio $\phi(x)/y$.

Cazals et al [2] proposed a probabilistic interpretation of the production frontier. Let $\mathbb{T}$ be the support of the joint distribution of a random vector $(X,Y) \in \mathbb{R}_+^p \times \mathbb{R}_+$ and let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probability space on which the vector of inputs X and the output Y are defined. 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 [2] that

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

is a monotone nondecreasing function 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) $\widehat{\mathbb{T}}_{FDH} = \left\{(x,y) \in \mathbb{R}_+^{p+1} | y \leq Y_i,\ x \geq X_i,\ i = 1,\ldots,n\right\}$ of $\mathcal{X}_n$ has been introduced by [7]. The resulting FDH estimator of $\varphi(x)$ is

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

where $\hat{F}(y|x) = \hat{F}_n(x,y)/\hat{F}_X(x)$ with $\hat{F}_n(x,y) = (1/n)\sum_{i=1}^n \mathbb{1}(X_i \leq x, Y_i \leq y)$ and $\hat{F}_X(x) = \hat{F}_n(x,\infty)$. 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 [13] for the consistency and by [14, 12] 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 [14], 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 distribution requires the specification of smoothing parameters and the resulting procedure has very poor accuracy. In Hwang et al [12], 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 [14] 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 [14] and Hwang et al [12] 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 et al [1] 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 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

From now on we assume that $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$ and will denote by $\varphi_\alpha(x)$ and $\hat{\varphi}_\alpha(x)$, respectively, the αth quantiles of the distribution function $F(\cdot|x)$ and its empirical version $\hat{F}(\cdot|x)$,

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

with $\alpha \in]0,1]$. When $\alpha \uparrow 1$, the conditional quantile $\varphi_\alpha(x)$ tends to $\varphi_1(x)$ which coincides with the frontier function $\varphi(x)$. Likewise, $\hat{\varphi}_\alpha(x)$ tends to the FDH estimator $\hat{\varphi}_1(x)$ of $\varphi(x)$ as $\alpha \uparrow 1$.

2.1 Asymptotic Weibull distribution

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

Theorem 2.1. *(i) If there exist $b_n > 0$ and some non-degenerate 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 $\Psi_{\rho_x}(y) = \exp\{-(-y)^{\rho_x}\}$ with support $]-\infty, 0]$ for some $\rho_x > 0$.

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

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

[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), $\lim_{n\to\infty}\mathbb{E}\{b_n^{-1}(\varphi(x) - \hat{\varphi}_1(x))\}^k = \Gamma(1 + k\rho_x^{-1})$ *for all integer $k \geq 1$, and*

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

Remark 2.1. Since the function $t \mapsto F_X(x)[1 - F(\varphi(x) - \frac{1}{t}|x)] \in \mathrm{RV}_{-\rho_x}$ (regularly varying in $t \to \infty$) 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

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

In the particular case where $L_x\left(\{\varphi(x) - y\}^{-1}\right) = \ell_x$ is a strictly positive function in x, it is shown in the next corollary that $b_n \sim (n\ell_x)^{-1/\rho_x}$. From now on, a random variable W is said to follow the distribution Weibull$(1, \rho_x)$ if W^{ρ_x} is Exponential with parameter 1.

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

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

Remark 2.2. Park et al [14] and Hwang et al [12] have obtained similar results under more restrictive conditions. Indeed, a unified formulation of the assumptions used in [14, 12] can be expressed as

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

where $f(x,y)$ is the joint density of (X,Y), β is a constant satisfying $\beta>-1$, and c_x is a strictly positive function in x. Under the restrictive condition that f is strictly positive on the frontier (*i.e.* $\beta=0$) among others, Park et al [14] have obtained the limiting Weibull distribution of the FDH estimator with the convergence rate $n^{-1/(p+1)}$. When β may be non null, Hwang et al [12] have obtained the asymptotic Weibull distribution with the convergence rate $n^{-1/(\beta+2)}$ in the simple case $p=1$ (here it is also assumed that (2.4) holds uniformly in a neighborhood of the point at which we want to estimate $\varphi(\cdot)$ 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 tail index ρ_x, the data dimension $p+1$ 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* $F(x,y)$ *being differentiable near the frontier (i.e.* $\ell_x>0$, $\rho_x>p$ *and* $\varphi(x)$ *are differentiable in* x *with first partial derivatives of* $\varphi(x)$ *being strictly positive), then* (2.4) *holds with* $\beta=\beta_x=\rho_x-(p+1)$ *and we have*

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

Remark 2.3. We assume the differentiability 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 distinguish 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.

Remark 2.4. 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 [12] (see Remark 2.2) with the same convergence rate $n^{-1/(\beta+2)}$ (in the notations of Theorem 1 in [12], $\mu(x)=\ell_x(\beta+2)\varphi'(x)=\ell_x\rho_x\varphi'(x)$). 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 [14] (in the notations of Theorem 3.1 in [14], $\mu_{NW,0}/y=\ell_x^{1/(p+1)}=\ell_x^{1/\rho_x}$).

Note also that the condition (2.4) with $\beta=\beta_x>-1$ (as in Corollary 2.2) has been considered by [11, 10, 8]. In Section 2.3 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).

As an immediate consequence of Theorem 2.1 (iii) in conjunction with Corollary 2.2, we obtain

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

This extends the limit theorem of moments of Park et al ([14], Theorem 3.3) to the more general setting where β_x may be non null. Likewise, Hwang et al ([12], see Remark 1) provide (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 [14] in the particular case where $\beta_x=0$.

2.2 Robust frontier estimators

By an appropriate choice of α as a function of n, Aragon et al [1] have shown that $\hat{\varphi}_\alpha(x)$ estimates the full frontier $\varphi(x)$ itself and converges to the same Weibull distribution as the FDH $\hat{\varphi}_1(x)$ under the restrictive conditions of [14]. 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 any fixed integer $k \geq 0$,*

$$b_n^{-1}\left(\hat{\varphi}_{1-k/(n\hat{F}_X(x))}(x)-\varphi(x)\right) \xrightarrow{d} H_x \quad as \quad n\to\infty,$$

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 2.5. When $\hat{\varphi}_1(x)$ converges in distribution, the estimator $\hat{\varphi}_{\alpha_n}(x)$, for $\alpha_n := 1-k/n\hat{F}_X(x) < 1$ (*i.e.* $k = 1,2,\ldots$ in Theorem 2.2 (i)), estimates $\varphi(x)$ itself and converges in distribution as well, with the same scaling but a different limit distribution (here $nb_n^{-1}(1-\alpha_n) \overset{a.s.}{\to} \infty$). 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 ([1], Theorem 4.3) where the convergence rate achieves $n^{-1/(p+1)}$ under the restrictive assumption that the 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 [3] for an analysis of its quantitative and qualitative robustness properties.

2.3 Conditional tail index estimation

The important question of how to estimate ρ_x from the multivariate random sample $\mathcal{X}_n$ is very similar to the problem of estimation of the so-called extreme value index based rather on a sample of *univariate* random variables. An attractive estimation method has been proposed by [15] which can be easily adapted to our conditional 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

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

The following result is particularly important since it allows to test the hypothesis $\rho_x > 0$ and will be employed in a next section to derive asymptotic 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 $U(t) := \varphi_{1-\frac{1}{tF_X(x)}}(x)$, $t > \frac{1}{F_X(x)}$, has a positive derivative and there exists a positive function $A(\cdot)$ such that, for $z > 0$, $\lim_{t\to\infty}\left\{(tz)^{1+1/\rho_x}U'(tz)-t^{1+1/\rho_x}U'(t)\right\}/A(t) = \pm\log(z)$, for either choice of the sign [Π-variation, notation $\pm t^{1+1/\rho_x}U'(t) \in \Pi(A)$]. Then

$$\sqrt{k_n}(\hat{\rho}_x-\rho_x) \xrightarrow{d} \mathcal{N}(0,\sigma^2(\rho_x)), \tag{2.6}$$

with asymptotic variance $\sigma^2(\rho_x) = \rho_x^2(2^{1-\frac{2}{\rho_x}}+1)/\{(2^{-\frac{1}{\rho_x}}-1)\log 4\}^2$, 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 (2.6) *holds 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 2.6. 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. This technique is known as the Pickands plot in the univariate extreme-value literature (see *e.g.* [17] and the references therein, Section 4.5, p.93-96). This is this kind of idea which guides the automatic data driven rule we suggest in Section 3.

We also can easily adapt the well-known moment estimator for the index of a univariate extreme-value distribution (Dekkers et al [6]) to our conditional setup. Define

$$M_n^{(j)} = \frac{1}{k}\sum_{i=0}^{k-1}\left(\log\hat{\varphi}_{1-\frac{i}{n\hat{F}_X(x)}}(x) - \log\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)\right)^j \quad \text{for each} \quad j = 1,2 \quad \text{and} \quad k = k_n < n.$$

Then one can define the moment type estimator for the conditional regular-variation exponent ρ_x as

$$\tilde{\rho}_x = -\left\{M_n^{(1)} + 1 - \frac{1}{2}\left[1 - \left(M_n^{(1)}\right)^2/M_n^{(2)}\right]^{-1}\right\}^{-1}.$$

Theorem 2.4. *(i) If* (2.2) *holds, $k_n/n \to 0$ and $k_n \to \infty$, then $\tilde{\rho}_x \xrightarrow{p} \rho_x$.*

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

(iii) Suppose $\pm t^{1/\rho_x}\{\varphi(x) - U(t)\} \in \Pi(B)$ for some positive function B. Then $\sqrt{k_n}(\tilde{\rho}_x - \rho_x)$ has asymptotically a normal distribution with mean 0 *and variance*

$$\rho_x(2+\rho_x)(1+\rho_x)^2\left\{4 - 8\frac{(2+\rho_x)}{(3+\rho_x)} + \frac{(11+5\rho_x)(2+\rho_x)}{(3+\rho_x)(4+\rho_x)}\right\},$$

for $k_n \to \infty$ satisfying $k_n = o(n/g^{-1}(n))$, where $g(t) = t^{1+\frac{2}{\rho_x}}\left[\{\log\varphi(x) - \log U(t)\}/B(t)\right]^2$.

Remark 2.7. Note that the Π-variation condition $\pm t^{1+\frac{1}{\rho_x}}U'(t) \in \Pi$ of Theorem 2.3 (iii) is equivalent to $\pm(t^{1/\rho_x}\{\varphi(x) - U(t)\})' \in \mathrm{RV}_{-1}$ following Theorem A.3 in [5] and that this equivalent regular-variation condition implies $\pm t^{1/\rho_x}\{\varphi(x) - U(t)\} \in \Pi$ according to Proposition 0.11(a) in [16], with auxiliary function $B(t) = \pm t(t^{1/\rho_x}\{\varphi(x) - U(t)\})'$. Hence the condition of Theorem 2.3 (iii) implies that of Theorem 2.4 (iii). Note also that a similar result to Theorem 2.4 (iii) can be given under the conditions of Theorem 2.3 (iv).

2.4 Asymptotic confidence intervals

The next theorem enables one to construct confidence intervals for $\varphi(x)$ and for high quantile-type frontiers $\varphi_{1-p_n/F_X(x)}(x)$ when $p_n \to 0$ and $np_n \to \infty$.

Theorem 2.5. *(i) Suppose $F(\cdot|x)$ has a positive density $F'(\cdot|x)$ such that $F'(\varphi(x) - \frac{1}{t}|x) \in RV_{1-\rho_x}$.*

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

where $V_1(\rho_x) = \rho_x^{-2}2^{1-2/\rho_x}/(2^{-1/\rho_x} - 1)^2$, 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

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

Then, putting $V_2(\rho_x) = 3\rho_x^{-2} 2^{-1-2/\rho_x}/(2^{-1/\rho_x}-1)^6$, *we have*

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

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

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

Then, putting $V_3(\rho_x) = \rho_x^{-2} 2^{-2/\rho_x}/(2^{-1/\rho_x}-1)^4$, *we have*

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

$$\{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)\} / \{\frac{n}{2k_n} U'(\frac{n}{2k_n})\} \xrightarrow{p} \rho_x(1 - 2^{-1/\rho_x}). \tag{2.7}$$

Remark 2.8. Note that Theorem 2.5 *(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. It is easy to see that $V_2(\rho_x) \geq V_3(\rho_x)$ and so the estimator $\tilde{\varphi}_1^*(x)$ of $\varphi(x)$ is asymptotically more efficient than $\hat{\varphi}_1^*(x)$. We also conclude from (2.7) that both $\tilde{\varphi}_1^*(x)$ and $\hat{\varphi}_1^*(x)$ have the same rate of convergence, namely $nU'(\frac{n}{2k_n})/(2k_n)^{3/2}$. In the particular case where $L_x\left(\{\varphi(x)-y\}^{-1}\right) = \ell_x$ in (2.3), we have $U'(\frac{n}{2k_n}) = \frac{1}{\rho_x}(\frac{1}{\ell_x})^{1/\rho_x}(\frac{2k_n}{n})^{1+1/\rho_x}$. Note also that in this particular case, the condition of Theorem 2.5 *(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 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, the conclusions of Theorem 2.3 (iii) and (iv) hold in this case for all sequences $k_n \to \infty$ satisfying $\frac{k_n}{n} \to 0$. The same is true for the conclusion of Theorem 2.5 (ii).

Theorem 2.6. *If the condition of Corollary 2.1 holds,* $k_n \to \infty$ *and* $k_n/n \to 0$ *as* $n \to \infty$, *then*

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

Remark 2.9. The optimization of the asymptotic mean squared error of $\hat{\varphi}_{1-(k_n-1)/n\hat{F}_X(x)}(x)$ is not an appropriate criteria for selecting the optimal k_n since the resulting value of k_n does not depend on n.

We shall now construct asymptotic confidence intervals for both $\varphi(x)$ and $\varphi_{1-p_n/F_X(x)}(x)$ using the sums $M_n^{(1)}$ and $M_n^{(2)}$.

Theorem 2.7. *(i) Under the conditions of Theorem 2.5 (i),*

$$\sqrt{k_n}\, \frac{\hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x) - \varphi_{1-\frac{p_n}{F_X(x)}}(x)}{M_n^{(1)} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} \mathcal{N}(0, V_4(\rho_x))$$

where $V_4(\rho_x) = (1 + 1/\rho_x)^2$, *provided* $p_n \to 0$, $np_n \to \infty$ *and* $k_n = [np_n]$.

(ii) Suppose the conditions of Theorem 2.4 (iii) hold and that $U(\cdot)$ has a regularly varying derivative $U' \in RV_{-\rho_x}$. Define the moment estimator $\hat{\varphi}(x) = \hat{\varphi}_{1-k_n/n\hat{F}_X(x)}(x)\left\{1+M_n^{(1)}(1+\tilde{\rho}_x)\right\}$. Then

$$\sqrt{k_n}\,\frac{\hat{\varphi}(x)-\varphi(x)}{M_n^{(1)}(1+1/\tilde{\rho}_x)\hat{\varphi}_{1-k_n/n\hat{F}_X(x)}(x)} \xrightarrow{d} \mathcal{N}(0,V_5(\rho_x)),$$

$$V_5(\rho_x) = \rho_x^2\left[\frac{\rho_x}{(2+\rho_x)}+\rho_x(2+\rho_x)\left\{4-8\frac{(2+\rho_x)}{(3+\rho_x)}+\frac{(11+5\rho_x)(2+\rho_x)}{(3+\rho_x)(4+\rho_x)}\right\}-\frac{4\rho_x}{(3+\rho_x)}\right].$$

2.5 Examples

Example 2.1. We consider the case where the support frontier is linear. We choose (X,Y) uniformly distributed over the region $D = \{(x,y)\,|\,0 \le x \le 1,\, 0 \le y \le x\}$. In this case (see *e.g.* [3]), it is easy to see 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.* [3]). 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 easily seen that the extreme-value condition (2.2), or equivalently (2.3), holds with $\rho_x = 2$ and $L_x(z) = F_X(x)[3\varphi(x) - \frac{2}{z}]/[\varphi(x)]^3$ 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 on the production activity of the French postal services.

3.1 Monte-Carlo experiment

We will simulate 2000 samples of size $n = 1000$ and of size $n = 5000$ according the scenario of Example 2.1 above. Here $\varphi(x) = x$ and $\rho_x = 2$. Denote by $N_x = n\hat{F}_X(x)$ the number of observations (X_i,Y_i) with $X_i \le x$. By construction of the estimators $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$, the threshold $k_n(x)$ can vary between 1 and $N_x/4$. For the estimator with known ρ_x, $\tilde{\varphi}_1^*(x)$, $k_n(x)$ is bounded by $N_x/2$ and finally, for the moment estimators $\tilde{\rho}_x$ and $\hat{\varphi}(x)$, the upper bound for $k_n(x)$ is given by $N_x - 1$. So, in our Monte-Carlo experiments for the Pickands estimator, $k_n(x)$ was selected on a grid of values determined by the observed value of N_x. We choose $k_n(x) = [N_x/4] - k + 1$, where k is an integer varying between 1 and $[N_x/4]$. In the tables below, $\overline{N}_x$ is the average value observed over the 2000 Monte-Carlo replications, the tables display the values of $\overline{k}_n(x)$ which is the average of the Monte-Carlo values of $k_n(x)$ obtained for a fixed selection of values of k. For the moment estimators, the upper values of $k_n(x)$ were chosen as $N_x - 1$. The Tables display only a part of the results to save place, but typically we choose, in each case, a set of values of k that includes not only the most favourable cases but also covering a wide range of values for $k_n(x)$. These tables provide the Monte-Carlo estimates of the Bias and the Mean Squared Error (MSE) of the various estimators computed over the 2000 random replications, as well as the average lengths and the achieved coverages of the corresponding 95% asymptotic confidence intervals. They display only the results for x ranging over $\{0.25, 0.75\}$ to save place.

We will first comment the results obtained for the Pickands estimators and for the estimator of $\varphi(x)$ obtained by knowing that $\rho_x = p+1 = 2$ (jump of the joint density of (X,Y) on the frontier). We observe the disappointing behavior of the Pickands estimates when the sample size is $n = 1000$ and for values of x as small as 0.25 (see the first top block of Tables 1, 2). 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 $k_n(x)$. We see the improvement of $\tilde{\varphi}_1^*(x)$ over the FDH in terms of the bias, without increasing too much the MSE and this even with sample sizes as small as $N_x = 62$. The achieved coverages of the normal confidence intervals obtained from $\tilde{\varphi}_1^*(x)$ are also quite satisfactory, and much more easy to derive than those obtained from the FDH estimator (assuming also $\rho_x = 2$).

Table 1: *Pickands and known ρ_x cases. Bias and Mean Squared Error, sample size $n = 1000$*

$x = 0.25$ $\quad \bar{N}_x = 62$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.028136$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.001005$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
12.0	-0.48504	906.91451	-0.03127	6.63766	0.00148	0.00142
11.4	-0.53609	9149.56965	-0.06785	36.77153	0.00168	0.00139
10.7	-1.26568	2095.81240	-0.12033	18.01733	0.00190	0.00142
10.1	-1.34925	2727.05598	-0.09043	13.39646	0.00165	0.00141
9.4	-1.01093	887.86044	-0.06853	4.08058	0.00213	0.00142
8.8	-0.99741	836.96814	-0.06174	3.82524	0.00220	0.00138
8.2	-1.43421	1084.83722	-0.07957	4.19400	0.00302	0.00135
7.5	-1.37656	1070.81436	-0.06913	4.36908	0.00340	0.00139
6.9	-1.09290	994.97474	-0.05734	3.45696	0.00446	0.00144
6.3	-0.40340	1406.03721	-0.01298	4.61059	0.00431	0.00137

$x = 0.75$ $\quad \bar{N}_x = 562$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.028080$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.001002$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
140.2	0.26635	6.32441	0.07343	0.47926	0.00030	0.00140
131.3	0.23266	1.28492	0.06191	0.09050	-0.00070	0.00138
122.4	0.25461	1.29701	0.06549	0.08546	-0.00065	0.00144
113.4	-0.09004	344.07913	-0.02658	22.67641	-0.00034	0.00142
104.5	0.42033	7.63112	0.09925	0.41662	0.00014	0.00145
95.6	0.33652	8.45253	0.07712	0.44647	-0.00004	0.00145
86.7	-9.40572	167972.74166	-2.13352	8553.19136	0.00036	0.00144
77.7	0.55786	22.85975	0.11535	0.99713	-0.00007	0.00148
68.8	0.25662	265.60614	0.04855	10.49201	-0.00008	0.00155
59.9	4.52123	23061.37346	0.82289	753.52315	0.00049	0.00151

Table 2: *Pickands and known ρ_x cases. Average Lengths (avl) and Coverages (cov) of the 95% confidence intervals, sample size $n = 1000$*

$x = 0.25$ $\quad \bar{N}_x = 62$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
12.0	1881.0192	0.8160	159.5440	0.7965	0.1504	0.9180
11.4	20972.8304	0.8185	1306.2047	0.7970	0.1507	0.9195
10.7	5065.5884	0.8035	467.0065	0.7810	0.1510	0.9190
10.1	6725.7862	0.8010	465.4399	0.7780	0.1508	0.9165
9.4	2061.6130	0.7960	132.1592	0.7735	0.1514	0.9130
8.8	2156.7584	0.7850	134.9646	0.7630	0.1514	0.9085
8.2	3305.2779	0.7780	182.7162	0.7545	0.1526	0.9085
7.5	3404.4945	0.7610	194.7502	0.7335	0.1534	0.8990
6.9	3559.2686	0.7335	170.6059	0.7065	0.1555	0.8975
6.3	4439.2558	0.6990	225.3314	0.6690	0.1557	0.8825

$x = 0.75$ $\quad \bar{N}_x = 562$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
140.2	6.6631	0.9190	1.8299	0.9150	0.1496	0.9520
131.3	3.7299	0.9130	0.9875	0.9055	0.1493	0.9520
122.4	3.9269	0.9020	1.0045	0.8985	0.1493	0.9420
113.4	231.0248	0.9045	59.2685	0.9025	0.1494	0.9430
104.5	9.1233	0.9150	2.1431	0.9030	0.1496	0.9445
95.6	9.8572	0.9115	2.2522	0.9040	0.1495	0.9485
86.7	127039.0252	0.9065	28640.0512	0.9010	0.1497	0.9540
77.7	22.9894	0.8990	4.7819	0.8950	0.1495	0.9470
68.8	230.8260	0.8910	45.8299	0.8805	0.1495	0.9325
59.9	20400.0683	0.8950	3687.5438	0.8825	0.1498	0.9390

For the larger value $x = 0.75$, as expected, $\hat{\rho}_x$ and $\hat{\varphi}_1^*(x)$ behave better, at least for appropriate

values of $k_n(x)$. Again $\tilde{\varphi}_1^*(x)$ performs rather well and is again stable to the selected value of $k_n(x)$. The achieved coverages of the confidence intervals are almost equal to the nominal level of 95%.

When the sample size increases, the Pickands estimators behave much better, even for moderate values of x. Tables 3 and 4 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. Here, as soon as N_x is larger than 1000, all the estimators provide reasonably good confidence intervals of the corresponding unknown, with quite good achieved coverages. In these cases ($N_x \geq 1000$), we observe also some stability of the results with respect to the choice of $k_n(x)$.

Table 3: *Pickands and known* ρ_x *cases. Bias and Mean Squared Error, sample size* $n = 5000$

$x = 0.25$ $\quad \bar{N}_x = 312$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.012591$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.000203$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
77.7	-0.25757	784.19539	-0.02585	6.93961	0.00021	0.00028
74.4	0.41215	17.20703	0.03723	0.14471	0.00024	0.00028
71.0	0.42344	105.75775	0.03830	0.89895	0.00016	0.00028
67.7	0.44401	16.30552	0.03877	0.11468	0.00030	0.00028
64.4	0.30552	145.08207	0.02564	1.01166	0.00031	0.00029
61.0	0.68905	35.13730	0.05654	0.24012	0.00053	0.00029
57.7	0.82177	15489.98302	0.05929	89.02353	0.00053	0.00029
54.3	1.17914	1780.66037	0.08527	9.90370	0.00055	0.00029
51.0	-4.41384	13169.38480	-0.33207	74.80129	0.00046	0.00030
47.6	0.03147	3204.61688	-0.00179	14.27123	0.00064	0.00029

$x = 0.75$ $\quad \bar{N}_x = 2813$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.012627$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.000201$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}_1^*(x)}$	$MSE_{\hat{\varphi}_1^*(x)}$	$B_{\tilde{\varphi}_1^*(x)}$	$MSE_{\tilde{\varphi}_1^*(x)}$
702.9	0.03859	0.08296	0.01034	0.00614	-0.00016	0.00030
668.2	0.04106	0.08652	0.01096	0.00610	0.00014	0.00029
633.6	0.04436	0.09402	0.01146	0.00622	0.00010	0.00029
598.9	0.04647	0.09685	0.01170	0.00606	0.00017	0.00028
564.2	0.05097	0.10266	0.01251	0.00605	0.00033	0.00027
529.5	0.05241	0.11087	0.01247	0.00614	0.00022	0.00028
494.8	0.05749	0.11876	0.01314	0.00614	0.00024	0.00027
460.2	0.07181	0.13817	0.01581	0.00668	0.00054	0.00028
425.5	0.06895	0.14227	0.01470	0.00635	0.00039	0.00028
390.8	0.07308	0.16153	0.01506	0.00660	0.00041	0.00028

Table 4: *Pickands and known* ρ_x *cases. Average Lengths and Coverages , sample size* $n = 5000$

$x = 0.25$ $\quad \bar{N}_x = 312$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
77.7	630.9019	0.9040	59.3041	0.8925	0.0670	0.9455
74.4	18.4635	0.9060	1.6821	0.8970	0.0670	0.9505
71.0	92.5814	0.9000	8.5104	0.8960	0.0670	0.9480
67.7	18.6125	0.8990	1.5673	0.8910	0.0670	0.9485
64.4	131.0169	0.8910	10.9372	0.8845	0.0670	0.9525
61.0	37.9315	0.8960	3.1260	0.8840	0.0671	0.9465
57.7	14491.7449	0.8965	1098.2578	0.8850	0.0671	0.9470
54.3	1735.9675	0.8930	129.3070	0.8820	0.0671	0.9430
51.0	13077.3352	0.8910	981.3170	0.8805	0.0671	0.9440
47.6	3374.6016	0.8925	224.7041	0.8735	0.0672	0.9410

$x = 0.75$ $\quad \bar{N}_x = 2813$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
702.9	1.0921	0.9460	0.2970	0.9430	0.0669	0.9445
668.2	1.1237	0.9480	0.2981	0.9435	0.0669	0.9490
633.6	1.1598	0.9445	0.2996	0.9410	0.0669	0.9495
598.9	1.1961	0.9485	0.3004	0.9455	0.0669	0.9500
564.2	1.2392	0.9485	0.3022	0.9430	0.0670	0.9555
529.5	1.2834	0.9415	0.3032	0.9425	0.0670	0.9560
494.8	1.3365	0.9470	0.3052	0.9460	0.0670	0.9525
460.2	1.4106	0.9475	0.3109	0.9490	0.0670	0.9555
425.5	1.4646	0.9450	0.3103	0.9415	0.0670	0.9550
390.8	1.5408	0.9380	0.3130	0.9355	0.0670	0.9560

We now turn to the performances of the moment estimators $\tilde{\rho}_x$ and $\widehat{\varphi}(x)$. The results are displayed in Table 5 for $n = 1000$ and Table 6 for $n = 5000$. Note that we used the same seed in the Monte-Carlo experiments than the one used for the preceding tables. We observe here much more reasonable results, in terms of the Bias and MSE of the estimators $\tilde{\rho}_x$ and $\hat{\varphi}(x)$, as soon as N_x is larger

than, say, 200. In addition, when N_x increases, the results are much less sensitive to the choice of $k_n(x)$ than for the Pickands estimators. We also observe that the most favorable values of $k_n(x)$ for estimating ρ_x or $\varphi(x)$ are not necessarily in the same range of values.

Table 5: *Moment Estimators. Bias, MSE, Average Lengths and Coverages, sample size* $n = 1000$

$x = 0.25$ $\bar{N}_x = 62$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}(x)}$	$MSE_{\hat{\varphi}(x)}$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}(x)}$	$cov_{\hat{\varphi}(x)}$
31.4	7.69194	98852.85196	0.18856	102.10294	69618.3092	0.8105	2237.8909	0.4845
28.5	0.78155	603.95223	-0.02837	0.61147	465.2116	0.8075	14.7657	0.5210
25.3	2.91920	6022.50946	0.04939	6.39901	4536.7150	0.8105	147.9476	0.5535
22.3	5.14393	21118.10510	0.12234	21.75798	18862.3079	0.8285	605.2293	0.5940
19.2	-0.13751	1402.87695	-0.03458	1.38802	1249.5570	0.8225	39.1572	0.6020
16.0	-0.57398	3611.92685	-0.03721	2.06825	3643.9352	0.7910	86.5993	0.6235
12.9	-2.87575	5952.16812	-0.09824	4.32304	6474.0510	0.8150	173.0064	0.6455
9.8	-0.69028	2209.06514	-0.02620	1.17234	3140.6753	0.7690	71.6783	0.6310
6.7	154.77576	48461004.94093	2.22488	10164.75120	77554551.1229	0.7280	1123213.7963	0.6190
3.6	-1.21190	1912.09995	-0.03132	0.58698	4166.0973	0.6175	71.9080	0.5080
2.0	-0.87003	2394.13723	-0.03639	0.31533	6640.2573	0.4625	68.3937	0.3635

$x = 0.75$ $\bar{N}_x = 562$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}(x)}$	$MSE_{\hat{\varphi}(x)}$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}(x)}$	$cov_{\hat{\varphi}(x)}$
281.5	0.22963	0.43342	-0.14881	0.02651	1.5512	0.8845	0.1537	0.1820
253.6	0.24167	0.45421	-0.12336	0.01954	1.6506	0.9190	0.1623	0.2535
225.4	0.24100	0.48387	-0.10137	0.01476	1.7570	0.9225	0.1698	0.3310
197.3	0.22582	0.49760	-0.08310	0.01121	1.8650	0.9255	0.1749	0.3985
169.2	0.21128	0.55801	-0.06660	0.00872	2.0150	0.9210	0.1808	0.4900
141.0	0.21154	0.54369	-0.05033	0.00625	2.2000	0.9205	0.1863	0.5900
112.9	0.22414	0.74955	-0.03492	0.00563	2.5452	0.9015	0.1993	0.6400
84.8	0.23220	1.02117	-0.02156	0.00544	3.0558	0.9120	0.2148	0.7115
56.7	0.29779	3.60304	-0.00729	0.01205	5.1691	0.8835	0.3054	0.7475
28.6	-0.47319	1765.30827	-0.03043	2.80568	1288.1794	0.8750	51.3417	0.7915
14.5	1.06058	508.21548	0.02489	0.47542	533.3150	0.8130	16.2354	0.7430

Table 6: *Moment Estimators. Bias, MSE, Average Lengths and Coverages, sample size* $n = 5000$

$x = 0.25$ $\bar{N}_x = 312$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}(x)}$	$MSE_{\hat{\varphi}(x)}$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}(x)}$	$cov_{\hat{\varphi}(x)}$
150.4	0.36520	1.47278	-0.04187	0.00339	2.5969	0.8900	0.0869	0.3350
137.9	0.35077	1.86333	-0.03615	0.00337	2.8243	0.8905	0.0939	0.3765
125.3	0.33799	1.26492	-0.03080	0.00226	2.7378	0.8990	0.0893	0.4435
112.9	0.30315	1.02334	-0.02670	0.00173	2.7495	0.9005	0.0874	0.4840
100.4	0.27374	0.93872	-0.02284	0.00139	2.8414	0.8930	0.0873	0.5495
87.9	0.28569	1.22921	-0.01810	0.00137	3.1695	0.8965	0.0936	0.5860
75.4	0.30500	9.96907	-0.01330	0.00806	7.3693	0.8865	0.2075	0.6340
62.9	0.26381	29.37920	-0.01097	0.02156	17.2434	0.8880	0.4629	0.6740
50.5	0.51850	18.67121	-0.00130	0.01090	14.4349	0.8780	0.3524	0.7020
38.0	0.53418	21.11753	0.00124	0.00956	18.2022	0.8645	0.3897	0.7225
19.2	0.62323	267.28452	0.00481	0.06789	246.3768	0.8430	3.8848	0.7525
12.9	-0.30491	1266.44113	-0.00977	0.30730	1431.7282	0.8150	22.2514	0.7315

$x = 0.75$ $\bar{N}_x = 2813$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\hat{\varphi}(x)}$	$MSE_{\hat{\varphi}(x)}$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}(x)}$	$cov_{\hat{\varphi}(x)}$
1125.7	0.14910	0.08588	-0.10940	0.01264	0.7039	0.8355	0.0674	0.0235
1013.2	0.14041	0.08293	-0.09393	0.00945	0.7374	0.8605	0.0690	0.0430
900.7	0.12149	0.07648	-0.08060	0.00707	0.7716	0.8890	0.0700	0.0720
788.2	0.11754	0.08188	-0.06686	0.00504	0.8233	0.9025	0.0718	0.1525
675.7	0.10905	0.08467	-0.05454	0.00352	0.8845	0.9250	0.0732	0.2565
563.0	0.10191	0.09542	-0.04300	0.00239	0.9658	0.9255	0.0749	0.3910
450.6	0.09008	0.11126	-0.03272	0.00163	1.0734	0.9310	0.0763	0.5145
338.1	0.08654	0.13468	-0.02274	0.00104	1.2404	0.9405	0.0783	0.6520
225.5	0.08933	0.19885	-0.01341	0.00071	1.5356	0.9420	0.0812	0.7665
113.0	0.10900	0.40414	-0.00468	0.00059	2.2621	0.9255	0.0875	0.8445
84.9	0.15855	0.61982	-0.00131	0.00065	2.7736	0.9170	0.0941	0.8515
56.7	0.08492	16.31728	-0.00208	0.01225	11.4038	0.8900	0.3139	0.8305

We note that the confidence intervals for ρ_x achieve quite reasonable coverage as soon as N_x is greater than, say, 500. However, the results for the confidence intervals of $\varphi(x)$ obtained from the moment estimator $\hat{\varphi}(x)$ are very poor even when N_x is as large as 5000. A more detailed analysis of the Monte-Carlo results allows us to conclude that this comes from an under evaluation of the asymptotic variance of $\hat{\varphi}(x)$ given in Theorem 2.7. Indeed, in most of the cases, the Monte-Carlo standard deviation of $\hat{\varphi}(x)$ was larger than the asymptotic theoretical expression by a factor of the order 2 to 5 when $N_x = 1250$ and by a factor of 1.3 to 1.7 when $N_x = 5000$. So the poor behavior

seems to improve slightly when N_x increases but at a very slow rate.

To summarize, we could say that using the Pickands estimators $\hat{\rho}_x$ and $\widehat{\varphi}_1^*(x)$, is only reasonable in our set-up when N_x is larger than, say, 1000. These estimators are highly sensitive to the choice of $k_n(x)$. The moment estimators $\tilde{\rho}_x$ and $\hat{\varphi}(x)$ have a much better behavior in terms of bias and MSE and a greater stability with respect to the choice of $k_n(x)$ even for moderate sample sizes. When N_x is very large ($N_x = 5000$), $\hat{\rho}_x$ and $\widehat{\varphi}_1^*(x)$ become more accurate than the moment estimators. On the other hand, inference on the value of ρ_x built from the asymptotic distribution of $\tilde{\rho}_x$ shows quite good coverage of the corresponding confidence intervals as soon as $N_x \geq 500$. However the confidence intervals derived from the Pickands estimator $\hat{\rho}_x$ provide more satisfactory results for large values of N_x, say, $N_x \geq 1000$. For inference purpose on the frontier function itself, the estimate of the asymptotic variance of the moment estimator $\hat{\varphi}(x)$ does not provide reliable confidence intervals even for relatively large values of N_x. It would be better to use in the latter case the confidence intervals obtained from the asymptotic distribution of the Pickands estimator $\hat{\varphi}_1^*(x)$.

So, in terms of bias and MSE computed over the 2000 random replications, as well as the average lengths and the achieved coverages of the 95% asymptotic confidence intervals, the moment estimators of ρ_x and $\varphi(x)$ sometimes are preferred over the Pickands estimators and sometimes not. It is difficult to imagine one procedure being preferred in all contexts. Hence a sensible practice is not to restrict the frontier analysis to one procedure but rather to check that both Pickands and moment estimators point toward similar conclusions. However when ρ_x is known, we have remarkable results for $\widetilde{\varphi}_1^*(x)$ even when N_x is small with remarkable properties of the resulting normal confidence intervals with a great stability with respect to the choice of $k_n(x)$. Remember that in most situations described so far in the econometric literature on frontier analysis, this tail index ρ_x is supposed to be known and equal to $p+1$ (here $\rho_x = 2$): this corresponds to the common assumption that there is a jump of the joint density of (X,Y) at the frontier.

This might suggest the following strategy with a real data set: either ρ_x is known (typically equal to $p+1$ if the assumption of a jump at the frontier is reasonable) and so we can use the estimator $\widetilde{\varphi}_1^*(x)$, or ρ_x is unknown, in this case we could suggest to use the following two-step estimator: first estimate ρ_x (the moment estimator of ρ_x seems the more appropriate, unless N_x is large enough) and second use the estimator $\widetilde{\varphi}_1^*(x)$, as if ρ_x was known, by plugging the estimated value $\tilde{\rho}_x$ or $\hat{\rho}_x$ at the place of ρ_x. In a real data set situation, the best prescription is not to favor the moment or the Pickands estimator of ρ_x in the first step, but to compute $\widetilde{\varphi}_1^*(x)$ by plugging both of them and then hope that the two resulting values of $\widetilde{\varphi}_1^*(x)$ point toward similar conclusions.

It should be clear that the two-step estimator $\widetilde{\varphi}_1^*(x)$, obtained by plugging $\hat{\rho}_x$, does not coincide necessarily with the Pickands estimator $\widehat{\varphi}_1^*(x)$ which is rather obtained by a simultaneous estimation of ρ_x and $\varphi(x)$. Indeed, we have observed in our Monte-Carlo exercise that the most favorable values of $k_n(x)$ for estimating ρ_x and $\varphi(x)$ are not necessarily in the same range of values. Thus nothing guarantees that the selected value $k_n(x)$ when computing $\hat{\rho}_x$ in the first step is the same as the one selected when computing $\widehat{\varphi}_1^*(x)$. Of course, when N_x is huge, the two values of $k_n(x)$ are expected to be similar, but the idea in the two-step procedure is to use the asymptotic results of the more efficient estimator $\tilde{\varphi}_1^*(x)$ and not those of $\widehat{\varphi}_1^*(x)$. In the next section, we suggest some *ad hoc* procedure for determining appropriate values of $k_n(x)$ with a real data set.

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

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.

First we have observed in our Monte-Carlo exercise that the optimal value for selecting $k_n(x)$ when estimating the index ρ_x is not necessarily the same than the value for estimating $\varphi(x)$. The idea is thus to select first, for each x (in a chosen grid of values), a grid of values for $k_n(x)$ for estimating ρ_x. For the Pickands estimator $\hat{\rho}_x$, we choose $k_n(x) = [N_x/4] - k + 1$, where k is an integer varying between 1 and $[N_x/4]$ and for the moment estimator $\tilde{\rho}_x$ we choose $k_n(x) = N_x - k$, where k is an integer varying between 1 and N_x. Then we evaluate the estimator $\hat{\rho}_x(k)$ (resp. $\tilde{\rho}_x(k)$) and we select the k where the variation of the results is the smaller. We achieve this by computing the standard deviations of $\hat{\rho}_x(k)$ (resp. $\tilde{\rho}_x(k)$) over a "window" of $2 \times [\sqrt{N_x/4}]$ (resp. $2 \times [\sqrt{N_x}]$) successive values of k. The value of k where this standard deviation is minimal defines the value of $k_n(x)$.

We follow the same idea for selecting a value for $k_n(x)$ for estimating the frontier $\varphi(x)$ itself. Here, in all the cases, we choose a grid of values for $k_n(x)$ given by $k = 1, \ldots, [\sqrt{N_x}]$ and select the k where the variation of the results is the smaller. To achieve this here, we compute the standard deviations of $\tilde{\varphi}_1^*(x)$ (resp. $\hat{\varphi}_1^*(x)$ and $\hat{\varphi}(x)$) over a "window" of size $2 \times \max(3, [\sqrt{N_x}/20])$ (this corresponds to have a window large enough to cover around 10% of the possible values of k in the selected range of values for $k_n(x)$). From now on, we only present illustrations for $\tilde{\varphi}_1^*(x)$ to save place.

For one sample generated with $n = 1000$ in the uniform case of our Monte-Carlo exercise above, we obtain the results shown in Figure 1.

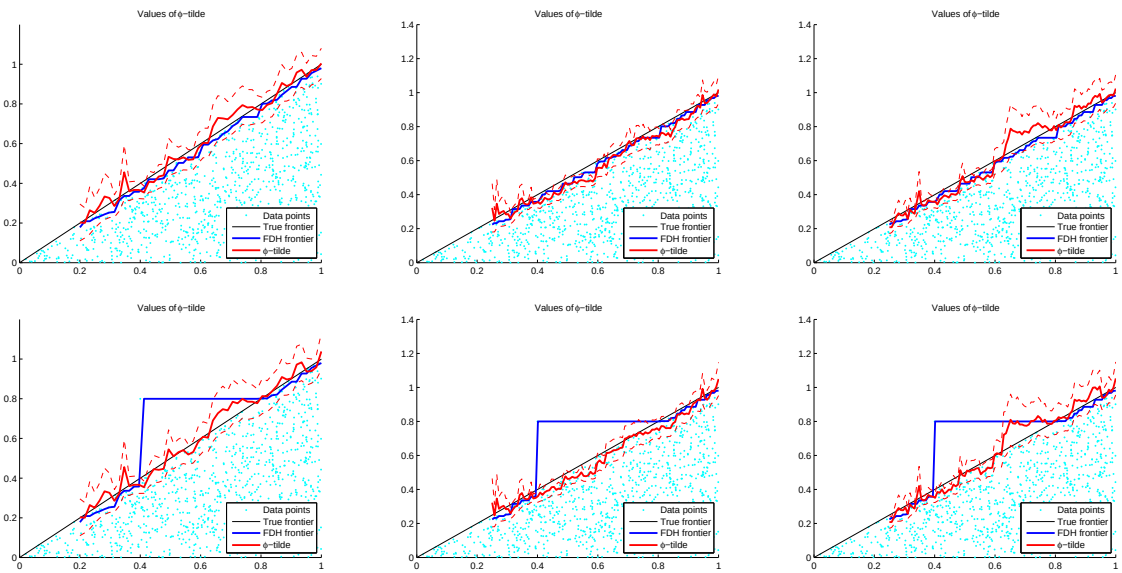

Figure 1: *Resulting estimator $\tilde{\varphi}_1^*(x)$ for a uniform data set of size $n = 1000$ (plus one outlier for the bottom panels), from left to right, we have the cases $\rho_x = 2$, plugging $\hat{\rho}_x$, plugging $\tilde{\rho}_x$.*

In this figure the estimator $\tilde{\varphi}_1^*(x)$ is first computed with the true value $\rho_x = 2$ (left panel of the figure) and then with a plug-in value of ρ_x estimated by the Pickands estimator (middle panel) and for the moment estimator $\tilde{\rho}_x$ (right panel). The pointwise confidence intervals are also displayed. The three bottom panels correspond to the same data set plus one outlier. This allows to illustrate how our robust estimators behave in the presence of outlying points, in contrast with the FDH estimator. In particular, due to the remarkable behavior of $\tilde{\varphi}_1^*(x)$ in the Monte-Carlo experiment, if we know that $\rho_x = 2$, we should use the left panel results and according our suggestion at the end of the preceding section, if ρ_x is unknown, we should use in this particular example the right panel results, where we

replace ρ_x by its moment estimator $\tilde{\rho}_x$ (since here $N_x \leq 1000$) and continue as if ρ_x was known. It is quite admirable that both panels are very similar.

3.3 An application

We use the same real data example as in [2] 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 empirical rules explained above for selecting reasonable values for $k_n(x)$. The cloud of points and the resulting estimates are provided in Figure 2. The FDH estimator is clearly determined by only a few very extreme points. If we delete 4 extreme points from the sample (represented by circles in the figure), we obtain the pictures of the top panels: the FDH estimator changes drastically, whereas the extreme-values based estimator $\tilde{\varphi}_1^*(x)$ is very robust to the presence of these 4 extreme points. We also note the great stability of the various forms of the estimator $\tilde{\varphi}_1^*(x)$, when ρ_x is supposed to be equal to 2 or when it is estimated by the Pickands or the moment estimator.

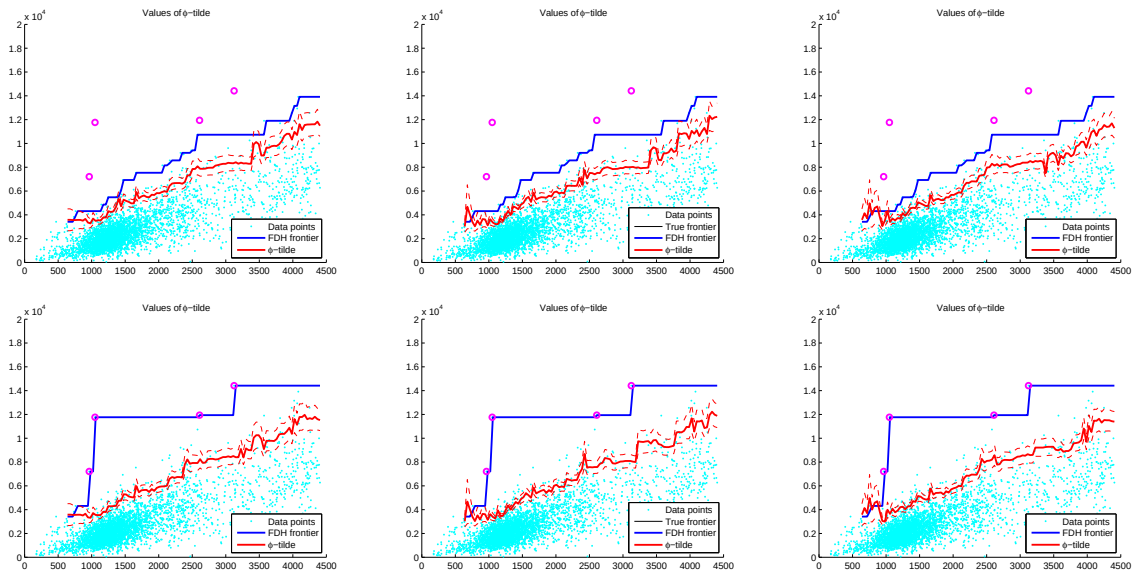

Figure 2: *Resulting estimator $\tilde{\varphi}_1^*(x)$ for the French post offices. We include 4 extreme data points (circles) for the bottom panels. From left to right, we have the cases $\rho_x = 2$, plugging $\hat{\rho}_x$, plugging $\tilde{\rho}_x$.*

4 Concluding Remarks

In our approach, we provide the necessary and sufficient condition for the FDH estimator $\hat{\varphi}_1(x)$ to converge in distribution, we specify its asymptotic distribution with the appropriate convergence rate and provide a limit theorem of moments in a general framework. We also give more insights and generalize the main result of [1] on robust variants of the FDH estimator and provide strongly consistent and asymptotically normal estimators $\hat{\rho}_x$ and $\tilde{\rho}_x$ of the unknown conditional tail index ρ_x involved in the limit law of $\hat{\varphi}_1(x)$. Moreover when the joint density of (X, Y) decreases to zero or rises up to infinity at a speed of power $\beta_x > -1$ of the distance from the boundary, as it is often assumed

in the literature, we answer the question of how ρ_x is linked to the data dimension $p+1$ and to the shape parameter β_x. The quantity $\beta_x \neq 0$ describes the rate at which the density tends to infinity (in case $\beta_x < 0$) or to 0 (in case $\beta_x > 0$) at the boundary. When $\beta_x = 0$, the joint density is strictly positive on the frontier. We establish that $\rho_x = \beta_x + (p+1)$. As an immediate consequence, we extend the previous results of [12, 14] to the general setting where $p \geq 1$ and $\beta = \beta_x$ may depend on x.

We propose new extreme-value based frontier estimators $\hat{\varphi}_1^*(x)$, $\tilde{\varphi}_1^*(x)$ and $\hat{\varphi}(x)$ which are asymptotically normally distributed and provide useful asymptotic confidence bands for the monotone frontier function $\varphi(x)$. These estimators have the advantage to not be limited to a bi-dimensional support and benefit from their explicit and easy formulations which is not the case of estimators defined by optimization problems such as local polynomial estimators (see *e.g.* [10]). Their asymptotic normality is derived under quite natural and general extreme-value conditions, without Lipschitz conditions on the boundary and without recourse to assumptions neither on the marginal distribution of X nor on the conditional distribution of Y given $X = x$ as it is often the case in both statistical and econometrics literature on frontier estimation. The study of the asymptotic properties of the different estimators considered in the present paper, is easily carried out by relating them to a simple dimensionless random sample and then applying standard extreme-values theory ([5], [6],...).

A closely related work in boundary estimation via extreme-values theory includes [9] in which the estimation of the frontier function at a point x is based on an increasing number of upper order statistics generated by the Y_i observations falling into a strip around x, and [8] in which estimators are rather based on a fixed number of upper order statistics. The main difference with the present approach is that Hall et al [9] only focus on estimation of the support curve of a bivariate density (*i.e.* $p = 1$) in the case $\beta_x > 1$ (*i.e.* the decrease in density is no more than algebraically fast), where it is known that estimators based on an increasing number of upper order statistics give optimal convergence rates. In contrast, Gijbels and Peng [8] consider the maximum of all Y_i observations falling into a strip around x and an endpoint type of estimator based on three large order statistics of the Y_i's in the strip. This methodology is closely related and comparable with our estimation method using the Pickands type estimator but, like the procedure of [9], it is only provided in the simple case $p = 1$ and involves in addition to the sequence k_n an extra smoothing parameter (bandwidth of the strip) which also needs to be selected. Moreover the asymptotic results in [8] are provided for densities of (X,Y) decreasing as a power of the distance from the boundary, whereas the setup in our approach is a general one. Note also that our transformed dimensionless data set $(Z_1^x, \ldots, Z_n^x)$ is constructed in such a way to take into account the monotonicity of the frontier (the endpoint of the common distribution of the Z_i^x's coincides with the frontier function $\varphi(x)$), the univariate random variables Z_i^x do not depend on the sample size and allow to employ easily the available results from the standard extreme-values theory, which is not the case for both [8, 9].

It should be clear that the monotonicity constraint on the frontier is the main difference with most of the existing approaches in the statistical literature. Indeed, the joint support of a random vector (X,Y) is often described in the literature as the set $\{(x,y)|y \leq \phi(x)\}$ where the graph of ϕ is interpreted as its upper boundary. As a matter of fact, the function of interest φ in our approach is the smallest monotone nondecreasing function which is larger than or equal to the frontier function ϕ. To our knowledge, only the estimators FDH and DEA estimate the quantity φ. Of course ϕ coincides with φ when the boundary curve is monotone, but the construction of estimators of the endpoint $\phi(x)$ of the conditional distribution of Y given $X = x$ requires a smoothing procedure which is not the case when the distribution of Y is conditioned by $X \leq x$.

We illustrate how the large sample theory applies in practice by doing some Monte-Carlo experiment. Good estimates of $\varphi(x)$ and ρ_x 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 $\varphi(x)$ and/or ρ_x is a difficult question that deserves for future work. Here, we suggest a simple automatic data driven method that provides a reasonable choice of the sequence $\{k_n(x)\}$ for large samples.

The empirical study reveals that the simultaneous estimation of the tail index and of the frontier function requires large sample sizes to provide sensible results. The moment estimators of ρ_x and of $\varphi(x)$ sometimes provide better estimations than the Pickands estimates and sometimes not. When considering bias and MSE, $\hat{\varphi}(x)$ and $\tilde{\rho}_x$ provide more accurate estimations, but when the sample size is large enough, $\hat{\varphi}_1^*(x)$ and $\hat{\rho}_x$ improve a lot and even seem to outperform the moment estimators. As far as the inference on ρ_x is concerned, $\tilde{\rho}_x$ provides also quite reliable confidence intervals, but $\hat{\rho}_x$ provides more satisfactory results for sufficiently large samples. However, when inference about the frontier function itself is concerned, the moment estimator provides very poor results compared with the Pickands estimator.

On the other hand, the performance of the estimator $\tilde{\varphi}_1^*(x)$, computed when ρ_x is known, is quite remarkable even compared with the benchmarked FDH. The confidence intervals for $\varphi(x)$ are very easy to compute and have quite good coverages. In addition, the results are quite stable with respect to the choice of the "smoothing" parameter $k_n(x)$. As shown in our illustrations, the estimates have also the merit of being robust to extreme values. This suggests, even if ρ_x is unknown, to use a plug-in version of $\tilde{\varphi}_1^*(x)$ for making inference on $\varphi(x)$: here, in a first step we estimate ρ_x (by the moment estimator unless N_x is large enough), then we use the asymptotic results for $\tilde{\varphi}_1^*(x)$, as if ρ_x was known. A sensible practice is not to restrict the first step to one procedure but rather to check that both Pickands and moment estimators point toward similar conclusions.

Appendix: Proofs

Proof of Theorem 2.1 Let $Z^x = Y\mathbb{1}(X \le x)$ and $F_x(\cdot) = \{1 - F_X(x)[1 - F(\cdot|x)]\}\mathbb{1}(\cdot \ge 0)$. It can be easily seen that $\mathbb{P}(Z^x \le y) = F_x(y)$ for any $y \in \mathbb{R}$. Therefore $\{Z_i^x = Y_i\mathbb{1}(X_i \le x), i = 1, ..., n\}$ 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 $\max_{i=1,...,n} Z_i^x$ coincides with $\hat{\varphi}_1(x)$. Thus Assertion (i) follows from the Fisher-Tippett Theorem. It is well known that the normalized maxima $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G$ (*i.e.* F_x belongs to the domain of attraction of $G = \Psi_{\rho_x}$) if and only if

$$\bar{F_x}(\varphi(x) - 1/t) \in \mathrm{RV}_{-\rho_x}, \tag{A.1}$$

where $\bar{F_x} = 1 - F_x$. This necessary and sufficient condition is equivalent to (2.2). In this case, the norming constant b_n can be taken equal to $\varphi(x) - \inf\{y \ge 0 | F_x(y) \ge 1 - \frac{1}{n}\} = \varphi(x) - \inf\{y \ge 0 | F(y|x) \ge 1 - \frac{1}{nF_X(x)}\}$, which gives Assertion (ii). For Assertion (iii), since (A.1) holds and $\mathbb{E}[|Z^x|^k] = F_X(x)\mathbb{E}(Y^k|X \le x) \le \varphi(x)^k$, it is immediate (see [16], 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 [16] (Corollary 2.3, p.83). □

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 $F_x^{-1}(t) = \inf\{y \in]0, \varphi(x)] : F_x(y) \ge t\}$ for all $t \in]0, 1]$. It follows from (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. □

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

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

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

For the next proofs we need the following lemma whose proof is quite easy and so is omitted.

Lemma .1. *Let $Z^x_{(1)} \leq \cdots \leq Z^x_{(n)}$ be the order statistics generated by the random variables $Z^x_1, \ldots, Z^x_n$.*

(i) If $\hat{F}_X(x) > 0$, then $\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k)}$ for each $k \in \{0, 1, \ldots, n\hat{F}_X(x) - 1\}$.

(ii) For any fixed integer $k \geq 0$, we have $\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k)}$ as $n \to \infty$, with probability 1.

(iii) For any sequence of integers $k_n \geq 0$ such that $k_n/n \to 0$ as $n \to \infty$, we have

$$\hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k_n)} \quad \text{as} \quad n \to \infty, \quad \text{with probability } 1.$$

Proof of Theorem 2.2 (i) Since $\varphi(x) = F_x^{-1}(1)$ and $\hat{\varphi}_1(x) = Z^x_{(n)}$ for all $n \geq 1$, we have $(\hat{\varphi}_1(x) - \varphi(x)) = (Z^x_{(n)} - F_x^{-1}(1))$. Hence, if $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G_x$, then $b_n^{-1}(Z^x_{(n)} - F_x^{-1}(1))$ converges to the same distribution G_x. Therefore, following [18] (Theorem 21.18, p. 313), $b_n^{-1}(Z^x_{(n-k)} - F_x^{-1}(1)) \xrightarrow{d} H_x$ for any integer $k \geq 0$, where $H_x(y) = G_x(y)\sum_{i=0}^{k}(-\log G(y))^i/i!$. Finally since $Z^x_{(n-k)} \overset{a.s.}{=} \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$, as $n \to \infty$, in view of Lemma .1 (ii), we obtain $b_n^{-1}(\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x) - F_x^{-1}(1)) \xrightarrow{d} H_x$.

(ii) Writing $b_n^{-1}(\hat{\varphi}_\alpha(x) - \varphi(x)) = b_n^{-1}(\hat{\varphi}_\alpha(x) - \hat{\varphi}_1(x)) + b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))$, 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 [1] (see Equation (20)) showed that $|\hat{\varphi}_\alpha(x) - \hat{\varphi}_1(x)| \leq (1-\alpha)n\hat{F}_X(x)F_Y^{-1}(1)$ with probability 1, for any $\alpha > 0$. 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 $\gamma_x = -1/\rho_x$ in (A.1). Then the Pickands [15] estimate of the exponent of variation $\gamma_x < 0$ is given by $\hat{\gamma}_x := (\log 2)^{-1}\log\{(Z^x_{(n-k+1)} - Z^x_{(n-2k+1)})/(Z^x_{(n-2k+1)} - Z^x_{(n-4k+1)})\}$. Under (2.2), Condition (A.1) holds and so there exists $b_n > 0$ such that $\lim_{n\to\infty}\mathbb{P}[b_n^{-1}(Z^x_{(n)} - \varphi(x)) \leq y] = \Psi_{-1/\gamma_x}(y)$. 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 the condition (1.1) in Dekkers and de Haan [5] 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 in [5]. This gives the weak consistency of $\hat{\rho}_x$ since $\hat{\gamma}_x \overset{a.s.}{=} -1/\hat{\rho}_x$, as $n \to \infty$, in view of Lemma .1 (iii).

(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 in [5] and so $\hat{\rho}_x \xrightarrow{a.s.} \rho_x$.

(iii) We have $U(t) = \inf\{y \geq 0 \,|\, \frac{1}{1-F_x(y)} \geq t\}$ which corresponds to the 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 [5] (see Theorem 2.3) that $\sqrt{k_n}(\hat{\gamma}_x - \gamma_x) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\gamma_x))$ with $\sigma^2(\gamma_x) = \gamma_x^2(2^{2\gamma_x+1}+1)/\{2(2^{\gamma_x}-1)\log 2\}^2$ for $k_n \to \infty$ satisfying $k_n = o(n/g^{-1}(n))$, where $g(t) := t^{3-2\gamma_x}\{U'(t)/A(t)\}^2$. By using the fact that $\sqrt{k_n}(\hat{\rho}_x - \rho_x) \overset{a.s.}{=} \sqrt{k_n}(-\frac{1}{\hat{\gamma}_x} + \frac{1}{\gamma_x})$, as $n \to \infty$, in view of Lemma .1 (iii) and applying 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 $\sigma^2(\rho_x) = \sigma^2(\gamma_x)/\gamma_x^4$.

(iv) Under the regularity condition, we have $\pm\left\{t^{-1-\frac{1}{\gamma_x}}F_x'(\varphi(x) - \frac{1}{t}) - \delta F_X(x)\right\} \in \mathrm{RV}_{-\kappa}$. Then the conclusion follows immediately from Theorem 2.5 of [5] in conjunction with Lemma .1 (iii). $\square$

Proof of Theorem 2.4 We have by Lemma .1 (iii), for each $j = 1, 2$,

$$M_n^{(j)} = (1/k) \sum_{i=0}^{k-1} \left(\log Z^x_{(n-i)} - \log Z^x_{(n-k)} \right)^j \quad \text{as} \quad n \to \infty, \quad \text{with probability 1.} \tag{A.2}$$

Then $-1/\tilde{\rho}_x$ coincides almost surely, for all n large enough, with the well-known moment estimator $\tilde{\gamma}_x$ (given by Equation (1.7) in [6]) of the index defined in (A.1) by $\gamma_x = -1/\rho_x$. Hence Theorem 2.4 (i) and (ii) follow from the weak and strong consistency of $\tilde{\gamma}_x$ proved in Theorem 2.1 of [6]. Likewise, Theorem 2.4 (iii) follows by applying Corollary 3.2 of [6] in conjunction with the delta method. □

Proof of Theorem 2.5 (i) Under the regularity condition, the distribution function F_x of Z^x has a positive derivative $F_x'(y) = F_X(x)F'(y|x)$ for all $y > 0$ such that $F_x'(\varphi(x) - \frac{1}{t}) \in \mathrm{RV}_{1+\frac{1}{\gamma_x}}$. Therefore, according to [5] (see 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 $F_x^{-1}(1-p_n) = \varphi_{1-\frac{p_n}{F_X(x)}}(x)$ and

$$\sqrt{2k_n} \frac{Z^x_{(n-k_n+1)} - F_x^{-1}(1-p_n)}{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}} \overset{a.s.}{=} \sqrt{2k_n} \frac{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - F_x^{-1}(1-p_n)}{\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x) - \hat{\varphi}_{1-\frac{2k_n-1}{n\hat{F}_X(x)}}(x)} \quad \text{as} \quad n \to \infty.$$

(ii) We have $\hat{\varphi}_1^*(x) \overset{a.s.}{=} \frac{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}}{2^{-\tilde{\gamma}_x}-1} + Z^x_{(n-k_n+1)}$ as $n \to \infty$. Then following Theorem 3.2 in [5], $\dfrac{\sqrt{2k_n}(\hat{\varphi}_1^*(x) - \varphi(x))}{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}}$ is asymptotically normal with mean 0 and variance $3\gamma_x^2 2^{2\gamma_x-1}/(2^{\gamma_x}-1)^6$.

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

$$\begin{aligned}
&\sqrt{2k_n} \left\{ \frac{1}{2^{-\gamma_x}-1} + \frac{Z^x_{(n-k_n+1)} - \varphi(x)}{Z^x_{(n-k_n+1)} - Z^x_{(n-2k_n+1)}} \right\} \overset{d}{=} \sqrt{2k_n} \left\{ \frac{1}{2^{-\gamma_x}-1} + \frac{V(E_{(n-k_n+1)}) - \varphi(x)}{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})} \right\} \\
&= \Bigg[-\sqrt{2k_n} \left\{ \frac{V(\infty) - V(\log \frac{n}{2k_n})}{V'(\log \frac{n}{2k_n})} + \frac{1}{\gamma_x} \right\} \\
&+ \sqrt{2k_n} \left\{ \frac{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})}{2^{\gamma_x} V'(E_{(n-2k_n+1)})} - \frac{1-2^{-\gamma_x}}{\gamma_x} \right\} \frac{2^{\gamma_x}}{1-2^{\gamma_x}} \frac{V'(E_{(n-2k_n+1)})}{V'(\log \frac{n}{2k_n})} \\
&- \frac{\sqrt{2k_n}}{\gamma_x} \left\{ \frac{V'(E_{(n-2k_n+1)})}{V'(\log \frac{n}{2k_n})} - 1 - \gamma_x \frac{V(E_{(n-k_n+1)}) - V(\log \frac{n}{2k_n})}{V'(\log \frac{n}{2k_n})} \right\} \Bigg] \frac{V'(\log \frac{n}{2k_n})}{V(E_{(n-k_n+1)}) - V(E_{(n-2k_n+1)})}.
\end{aligned}$$

The first term at the right hand side tends to zero as established by Dekkers and de Haan ([5], 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 [5]. 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 (iii) in conjunction with the fact that

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

with probability 1. □

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,

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

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 $\{\rho_x k_n^{1/2}/(k_n/n\ell_x)^{1/\rho_x}\}[Z^x_{(n-k_n+1)} - F_x^{-1}(1-k_n/n)] \xrightarrow{d} \mathcal{N}(0,1)$ as $n \to \infty$. 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 $\{\rho_x k_n^{1/2}/(k_n/n\ell_x)^{1/\rho_x}\}[Z^x_{(n-k_n+1)} + (k_n/n\ell_x)^{1/\rho_x} - \varphi(x)] \xrightarrow{d} \mathcal{N}(0,1)$ as $n \to \infty$. We conclude by using $Z^x_{(n-k_n+1)} \stackrel{a.s.}{=} \hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x)$ as $n \to \infty$. $\square$

Proof of Theorem 2.7 (i) As shown in the proof of Theorem 2.5 (i), we have $F_x'(\varphi(x) - \frac{1}{t}) \in \mathrm{RV}_{1+1/\gamma_x}$. Then by applying Theorem 5.1 in Dekkers et al [6] in conjunction with (A.2), we get

$$\sqrt{k_n}\{Z^x_{(n-k_n)} - F_x^{-1}(1-p_n)\}/M_n^{(1)} Z^x_{(n-k_n)} \xrightarrow{d} \mathcal{N}(0, V_4(-1/\gamma_x)).$$

This ends the proof by using simply $F_x^{-1}(1-p_n) = \varphi_{1-\frac{p_n}{F_X(x)}}(x)$ and $Z^x_{(n-k_n)} \stackrel{a.s.}{=} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)$ as $n \to \infty$.

(ii) Since $Z^x_{(n-k_n)} \stackrel{a.s.}{=} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)$ and $\tilde{\gamma}_x \stackrel{a.s.}{=} -1/\tilde{\rho}_x$ as $n \to \infty$, we have $\hat{\varphi}(x) \stackrel{a.s.}{=} Z^x_{(n-k_n)} M_n^{(1)}(1 - 1/\tilde{\gamma}_x) + Z^x_{(n-k_n)}$, $n \to \infty$. It is then easy to see from (A.2) that $\hat{\varphi}(x)$ coincides almost surely, for all n large enough, with the endpoint estimator $\hat{x}_n^*$ of $F_x^{-1}(1)$ introduced by [6] in Equation (4.8). It is also easy to check that $U(t) = (1/(1-F_x))^{-1}(t)$ satisfies the conditions of Theorem 3.1 in [6] with $\gamma_x = -1/\rho_x < 0$. Then according to Theorem 5.2 in [6], we have $\sqrt{k_n}\{\hat{x}_n^* - F_x^{-1}(1)\}/M_n^{(1)} Z^x_{(n-k_n)}(1 - \tilde{\gamma}_x) \xrightarrow{d} \mathcal{N}(0, V_5(-1/\gamma_x))$ which gives the desired convergence in distribution of Theorem 2.7 (ii) since $F_x^{-1}(1) = \varphi(x)$, $\hat{x}_n^* \stackrel{a.s.}{=} \hat{\varphi}(x)$, $\tilde{\gamma}_x \stackrel{a.s.}{=} -1/\tilde{\rho}_x$ and $Z^x_{(n-k_n)} \stackrel{a.s.}{=} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)$ as $n \to \infty$. $\square$

Acknowledgement

The authors acknowledge support from the "Interuniversity Attraction Pole", Phase VI (No. P6/03) from the Belgian Science Policy.

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

[5] 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.

[6] Dekkers, A.L.M., Einmahl, J.H.J. and L. de Haan (1989), A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, 17(4), 1833–1855.

[7] 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.

[8] Gijbels, I. and Peng, L. (2000), Estimation of a Support Curve via Order Statistics. *Extremes*, 3(3), 251–277.

[9] Hall, P. and Nussbaum, M. and Stern, S.E. (1997), On the estimation of a support curve of indeterminate sharpness, *Journal of Multivariate Analysis*, 62, 204–232.

[10] 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.

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

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

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

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

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

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

[17] Resnick, S.I. (2007), Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering.

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