I N S T I T U T D E
S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

0806

FRONTIER ESTIMATION AND EXTREME VALUES THEORY

DAOUIA, A., FLORENS, J.-P. and L. SIMAR

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 : 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 in a general setup, 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. We also apply our approach to a real data set.

Key words : conditional quantiles; extreme values; monotone boundaries; nonparametric estimation; production frontier

1 Introduction

In production theory and efficiency analysis (see *e.g.* Shephard [31]), 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 [5] 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 $\phi(x) = \sup\{y \,|\, (x,y) \in \mathbb{T}\}$. Then the Farrell-Debreu output efficiency score ([11], [6]) of a firm operating at the level (x,y) is given by the ratio $\phi(x)/y$.

Cazals, Florens and Simar [2] 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 [2] 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 [9] 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 $\hat{F}_X(x) = \hat{F}_n(x,\infty)$. 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 Korostelev, Simar and Tsybakov [25] for the consistency and by Park, Simar and Weiner [27] and Hwang, Park and Ryu [22] 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 [27], 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 [22], 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 [27] 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 [27] and Hwang et al [22] 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 [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 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

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 α^{th} 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 $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x))$.

Theorem 2.1. *(i) If there exist constants $b_n > 0$ 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 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* $\mathbb{E}(Y^k|X \le x) < \infty$ *for some integer* $k \ge 1$, *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* $\mathbb{E}(Y^2|X \le x) < \infty$ *then*

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

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*

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

Remark 2.2. Park et al [27] and Hwang et al [22] have obtained similar results under more restrictive conditions. Indeed, a unified formulation of the assumptions used in [27] and [22] can be expressed as follows:

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

where $f(x,y)$ is the joint density function 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 the density f is strictly positive on the frontier (*i.e.* $\beta = 0$) among others, Park et al [27] 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 [22] 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 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 $p \ge 1$ 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* $F(x,y)$ *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*

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

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 [22] (see Remark 2.2) with the same convergence rate $n^{-1/(\beta+2)}$ (in the notations of Theorem 1 in [22], $\mu(x) = \ell_x(\beta+2)\varphi'(x) = \ell_x\rho_x\varphi'(x)$). 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 [27] (in the notations of Theorem 3.1 in [27], $\mu_{NW,0}/y = \ell_x^{1/(p+1)} = \ell_x^{1/\rho_x}$).

Note also that the condition (2.4) with $\beta = \beta_x > -1$ (as in Corollary 2.2) has been considered by Hardle, Park and Tsybakov [21], Hall, Park and Stern [20] and by Gijbels and Peng [14]. In a next section (see Conditional tail index estimation) 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 ([27], Theorem 3.3) to the more general setting where β_x may be non null. Likewise, Hwang et al ([22], 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 [27] in the particular case where $\beta_x = 0$.

2.2 Robust frontier estimators

By an appropriate choice of the order α as a function of n, Aragon et al [1] have shown that the empirical partial frontier $\hat{\varphi}_\alpha(x)$ 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 Park et al [27]. 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* $k \geq 0$,

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

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 the FDH $\hat{\varphi}_1(x)$ converges in distribution, the conditional quantile-based estimator $\hat{\varphi}_{\alpha_n}(x)$, for $\alpha_n := 1 - k/n\hat{F}_X(x) < 1$ (*i.e.* $k = 1, 2, \ldots$ in Theorem 2.2 (i)), estimates the frontier function $\varphi(x)$ itself and converges in distribution as well, with the same scaling but a different limit distribution (here $nb_n^{-1}(1-\alpha_n) \overset{a.s.}{\to} \infty$). 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 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 [3] for an analysis of its quantitative and qualitative robustness properties.

2.3 Conditional tail index estimation

An important question in the setup of the obtained 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 estimation of the so-called extreme value index based rather on a sample of *univariate* random variables (see *e.g.* [10] and the references therein). An attractive estimation method has been proposed by Pickands [28]. 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:

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

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 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 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

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

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.* Resnick [30] 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 [8]) 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

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

The next theorem gives quite general conditions under which $\tilde{\rho}_x$ is consistent and asymptotically normal.

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 $g(t) = t^{1+\frac{2}{\rho_x}}[\{\log\varphi(x) - \log U(t)\}/B(t)]^2$.

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 [7] 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 [29], 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

Another question of particular interest is how to derive asymptotically normal estimates and/or asymptotic confidence intervals for high partial frontiers $\varphi_\alpha(x)$, when $\alpha = \alpha(n) \uparrow 1$, and for the true full frontier $\varphi(x)$ itself by making use of large frontiers $\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$ in the finite case where k is fixed and in the limiting situation where $k = k_n$ grow without bound.

2.4.1 Using differences of large empirical partial frontiers

The following theorem enables one to construct confidence intervals for $\varphi(x)$ and for high quantile-type frontiers $\varphi_{1-\frac{p_n}{F_X(x)}}(x)$ 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 $F'(\varphi(x) - \frac{1}{t}|x) \in RV_{1-\rho_x}$. Then*

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

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

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

Then

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

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

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

Then

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

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

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

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

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

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

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

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. In view of the formulas of $V_2(\rho_x)$ and $V_3(\rho_x)$, 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 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 $L_x\left(\frac{1}{\varphi(x)-y}\right) = \ell_x$ in (2.3), we have $U'(\frac{n}{2k_n}) = \frac{1}{\rho_x}(\frac{1}{\ell_x})^{1/\rho_x}(\frac{2k_n}{n})^{1+1/\rho_x}$.

Remark 2.9. In the particular case where $L_x\left(\frac{1}{\varphi(x)-y}\right) = \ell_x$ in (2.3), 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 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 $\gamma_x = -1/\rho_x$ of our proofs), so that the left-hand side of Equation (2.3) in Dekkers and de Haan [7] 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 2.5 *(ii)*. The next theorem gives another variant of this result.

Theorem 2.7. *Suppose the condition of Corollary 2.1 holds. Then*

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

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 2.7,*

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

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

Remark 2.10. It is easy to see that the value of k_n which minimizes the asymptotic mean squared error, MSE $= \mathcal{B}_n^2 + \mathcal{B}_n^2/(\rho_x^2 k_n)$, does not even depend on n. So the optimization of the MSE of $\hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x)$, or equivalently $\hat{\varphi}_{\alpha_n}(x)$, is not the appropriate theoretical criteria for selecting the optimal sequence k_n.

2.4.2 Using sums of large empirical partial frontiers

We shall now construct asymptotic confidence intervals for both $\varphi(x)$ and $\varphi_{1-\frac{p_n}{F_X(x)}}(x)$ using the sums $M_n^{(1)}$ and $M_n^{(2)}$. We first get the following results in the infinite case ($k = k_n$ unbounded) similarly to Theorem 2.5.

Theorem 2.8. *(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 a moment estimator for $\varphi(x)$ as*

$$\hat{\varphi}(x) = \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)\left\{1 + M_n^{(1)}(1+\tilde{\rho}_x)\right\}.$$

Then

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

where

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

Next we consider the estimation of $\varphi(x)$ and $\varphi_{1-\frac{p_n}{F_X(x)}}(x)$ when the k occurring in $\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$ and $M_n^{(1)}$ is fixed. In this situation we propose to estimate $\varphi_{1-\frac{p_n}{F_X(x)}}(x)$ as follows:

$$\hat{\phi}_{p_n}(x) = \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)\left\{1 + M_n^{(1)}(1+\rho_{x,n})\left(1-(k/np_n)^{-1/\rho_{x,n}}\right)\right\}$$

with $\rho_{x,n}$ being any consistent estimate of ρ_x.

Theorem 2.9. (i) *Suppose $p_n \to 0$ and $np_n \to c \in]0,\infty[$ as $n\to\infty$. Let the number k occurring in the definition of $\hat{\phi}_{p_n}(x)$ be fixed with $k > c$. Then, provided* (2.2) *holds,*

$$\frac{\hat{\phi}_{p_n}(x) - \varphi_{1-\frac{p_n}{F_X(x)}}(x)}{M_n^{(1)}\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)}$$
$$\xrightarrow{d} (1+\rho_x)\{1-(k/c)^{-1/\rho_x}\} + \rho_x\{(Q_k/c)^{-1/\rho_x} - 1\}\Big/\left\{\frac{1}{k}\sum_{i=0}^{k-1}\rho_x\left[1-\exp\left(-\sum_{j=i}^{k-1}\frac{E_j}{j\rho_x}\right)\right]\right\},$$

with $Q_k, E_0, \ldots, E_{k-1}$ independent, Q_k gamma with k degrees of freedom, and E_i, $i = 0,\ldots,k-1$, i.i.d. exponential.

(ii) *Suppose* (2.2) *holds and let k in the definition of $\hat{\varphi}(x)$ be fixed. Then*

$$\frac{\hat{\varphi}(x)-\varphi(x)}{M_n^{(1)}\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)} \xrightarrow{d} (1+\rho_x) + \left\{\frac{1}{k}\sum_{i=0}^{k-1}\exp\left(-\sum_{j=i}^{k-1}\frac{E_j}{j\rho_x}\right) - 1\right\}^{-1}.$$

Remark 2.11. In Theorem 2.9, it should be clear that the number k used in the definition of $M_n^{(1)}$ remains bounded whereas, if for $\rho_{x,n}$ [occurring in the definition of $\hat{\phi}_{p_n}(x)$] one uses Pickands estimate $\hat{\rho}_x$ or the moment estimate $\tilde{\rho}_x$, one needs to use an unbounded number k' in their definitions.

Remark 2.12. Theorems 2.1-2.6 and Theorems 2.8-2.9 follow easily by applying the elegant devices of Dekkers and de Haan [7] and Dekkers et al [8], among others, in conjunction with the simple and clever idea that $\varphi(x)$ coincides with the right endpoint of the common distribution of the *univariate* random variables $Z_i^x := Y_i 1\!\!1(X_i \le x)$, $i = 1,\ldots,n$. It is also clear from Lemma .1(i) in Appendix that $\hat{\rho}_x$ and $\tilde{\rho}_x$ as well as the estimates of the high partial frontiers and of the full frontier can be easily computed in practice by employing the order statistics $Z^x_{(n-k)} = \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$ for each $k = 0,1,\ldots,n\hat{F}_X(x)-1$. 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 [17]) could be adapted to our problem. Of course, the selected number k_n of extreme observations Z^x_i involved in the definition of the estimators $\hat{\rho}_x$, $\tilde{\rho}_x$, $\hat{\varphi}^*_1(x)$, $\tilde{\varphi}^*_1(x)$, $\hat{\varphi}(x)$,... should depend on the fixed level $x \in \mathbb{R}^p$ of inputs-usage. We do not enter in this paper into the question of how to choose theoretically $k_n(x)$ in an optimal way. 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.

2.5 Examples

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 $D = \{(x,y)\,|\,0 \le x \le 1,\, 0 \le y \le x\}$. In this case (see Daouia and Ruiz-Gazen [3], Daouia and Simar [4], 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 2.1-2.7 hold (see Remark 2.9).

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 [13]). 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 $\rho_x = 2$ 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.

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 $k_n(x)$.

The tables below (from Table 1 to 6) 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.

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 $\tilde{\varphi}_1^*(x)$ and $k_n(x)$ for $\tilde{\rho}_x$ and $\hat{\varphi}(x)$. Then, as it can be expected, if $k_n(x)$ or N_x is too small, the variance of the estimator of ρ_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 the estimator. For instance, 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 (2.6) may be negative or too large. In particular, the use of small input values 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). The same is true for the other estimators.

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 $\rho_x = 2$).

The tables show also the results for larger values of x and, as expected, the estimators $\hat{\rho}_x$ 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%.

Table 1: *Pickands and known* ρ_x *cases. Bias (B) and Mean Squared Error (MSE) of the estimates over 2000 Monte-Carlo simulations, sample size* $n = 1000$

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

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}^*_1(x)}$	$MSE_{\widehat{\varphi}^*_1(x)}$	$B_{\widetilde{\varphi}^*_1(x)}$	$MSE_{\widetilde{\varphi}^*_1(x)}$
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

$x = 0.50$ $\quad \bar{N}_x = 250$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.027821$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.000984$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}^*_1(x)}$	$MSE_{\widehat{\varphi}^*_1(x)}$	$B_{\widetilde{\varphi}^*_1(x)}$	$MSE_{\widetilde{\varphi}^*_1(x)}$
62.1	0.86492	1022.60093	0.16285	35.07153	0.00067	0.00154
58.3	0.77734	118.17269	0.13897	3.84698	0.00045	0.00154
54.6	2.06630	3042.01785	0.35339	85.80974	0.00052	0.00149
50.8	1.58387	750.31800	0.26137	21.46568	0.00049	0.00142
47.0	0.17297	326.01713	0.02348	8.78341	0.00023	0.00139
43.2	1.03623	969.17487	0.15767	22.84893	-0.00025	0.00142
39.4	8.17365	138392.15086	1.09759	2490.23917	0.00008	0.00139
35.6	0.60146	664.65620	0.09650	15.59017	-0.00062	0.00146
31.8	0.23675	407.42921	0.04079	5.64263	-0.00071	0.00139
28.1	1.10798	3006.45644	0.11228	34.33301	-0.00045	0.00137

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

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}^*_1(x)}$	$MSE_{\widehat{\varphi}^*_1(x)}$	$B_{\widetilde{\varphi}^*_1(x)}$	$MSE_{\widetilde{\varphi}^*_1(x)}$
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

$x = 1.00$ $\quad \bar{N}_x = 1000$ $\quad$ FDH: $B_{\hat{\varphi}_1(x)} = -0.027473$ $\quad$ and $MSE_{\hat{\varphi}_1(x)} = 0.000953$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}^*_1(x)}$	$MSE_{\widehat{\varphi}^*_1(x)}$	$B_{\widetilde{\varphi}^*_1(x)}$	$MSE_{\widetilde{\varphi}^*_1(x)}$
250.0	0.12371	1.07603	0.04634	0.14997	0.00060	0.00143
234.0	3.26046	19628.31961	1.20848	2701.85640	0.00143	0.00143
218.0	0.15243	0.65257	0.05318	0.07912	0.00131	0.00143
202.0	0.17401	0.72369	0.05785	0.07620	0.00147	0.00143
186.0	0.24197	7.54008	0.07658	0.70279	0.00141	0.00143
170.0	0.27017	12.07866	0.08205	1.06230	0.00110	0.00140
154.0	0.27529	6.77502	0.07947	0.51830	0.00109	0.00141
138.0	0.27553	1.75793	0.07667	0.13168	0.00145	0.00146
122.0	0.32059	2.08038	0.08371	0.13564	0.00143	0.00150
106.0	0.41294	5.18075	0.10076	0.29355	0.00108	0.00152

Table 2: *Pickands and known* ρ_x *cases. Average Lengths (avl) and Coverages (cov) of the 95% confidence intervals, over 2000 Monte-Carlo simulations, sample size* $n = 1000$

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

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
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

$x = 0.50 \quad \bar{N}_x = 250$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
62.1	929.5066	0.8870	172.1322	0.8815	0.1497	0.9315
58.3	115.6243	0.8810	20.8087	0.8725	0.1496	0.9375
54.6	2869.5863	0.8860	481.9853	0.8745	0.1496	0.9390
50.8	753.0965	0.8850	127.2271	0.8850	0.1496	0.9475
47.0	338.1331	0.8840	55.3762	0.8825	0.1494	0.9445
43.2	1062.7489	0.8755	163.1122	0.8675	0.1491	0.9295
39.4	156622.5426	0.8635	21009.3710	0.8610	0.1494	0.9400
35.6	784.5760	0.8540	119.6340	0.8430	0.1489	0.9415
31.8	531.2606	0.8665	62.1563	0.8560	0.1488	0.9395
28.1	4235.2917	0.8575	451.2239	0.8540	0.1490	0.9460

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

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
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

$x = 1.00 \quad \bar{N}_x = 1000$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\hat{\varphi}_1^*(x)}$	$cov_{\hat{\varphi}_1^*(x)}$	$avl_{\tilde{\varphi}_1^*(x)}$	$cov_{\tilde{\varphi}_1^*(x)}$
250.0	2.4226	0.9310	0.8867	0.9285	0.1496	0.9495
234.0	9074.1624	0.9320	3366.6221	0.9275	0.1499	0.9500
218.0	2.4451	0.9255	0.8347	0.9260	0.1498	0.9510
202.0	2.6171	0.9355	0.8564	0.9330	0.1499	0.9590
186.0	6.3956	0.9270	1.9769	0.9275	0.1499	0.9570
170.0	9.2088	0.9200	2.7443	0.9155	0.1498	0.9565
154.0	6.6668	0.9180	1.8731	0.9200	0.1498	0.9460
138.0	4.0256	0.9180	1.0966	0.9125	0.1499	0.9495
122.0	4.5997	0.9115	1.1767	0.9010	0.1499	0.9465
106.0	7.3086	0.8985	1.7440	0.8965	0.1497	0.9445

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 (B) and Mean Squared Error (MSE) of the estimates over 2000 Monte-Carlo simulations, sample size* $n = 5000$

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

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}_1^*(x)}$	$MSE_{\widehat{\varphi}_1^*(x)}$	$B_{\widetilde{\varphi}_1^*(x)}$	$MSE_{\widetilde{\varphi}_1^*(x)}$
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

$x = 0.50$ $\bar{N}_x = 1250$ FDH: $B_{\hat{\varphi}_1(x)} = -0.012563$ and $MSE_{\hat{\varphi}_1(x)} = 0.000200$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}_1^*(x)}$	$MSE_{\widehat{\varphi}_1^*(x)}$	$B_{\widetilde{\varphi}_1^*(x)}$	$MSE_{\widetilde{\varphi}_1^*(x)}$
312.1	0.09248	0.22503	0.01696	0.00735	0.00026	0.00029
297.0	0.09311	0.24340	0.01668	0.00759	0.00012	0.00029
281.9	0.09124	0.24958	0.01595	0.00742	-0.00001	0.00029
266.8	0.09201	0.27538	0.01579	0.00780	-0.00009	0.00029
251.7	0.08954	0.29784	0.01490	0.00797	-0.00042	0.00030
236.6	0.09840	0.33195	0.01584	0.00831	-0.00049	0.00030
221.5	0.11387	0.38048	0.01768	0.00893	-0.00043	0.00030
206.3	0.12297	0.47557	0.01840	0.01038	-0.00060	0.00030
191.2	0.12060	0.43562	0.01720	0.00881	-0.00081	0.00030
176.1	0.14573	0.72946	0.01989	0.01371	-0.00080	0.00029

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

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}_1^*(x)}$	$MSE_{\widehat{\varphi}_1^*(x)}$	$B_{\widetilde{\varphi}_1^*(x)}$	$MSE_{\widetilde{\varphi}_1^*(x)}$
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

$x = 1.00$ $\bar{N}_x = 5000$ FDH: $B_{\hat{\varphi}_1(x)} = -0.012663$ and $MSE_{\hat{\varphi}_1(x)} = 0.000202$

$k_n(x)$	$B_{\hat{\rho}_x}$	$MSE_{\hat{\rho}_x}$	$B_{\widehat{\varphi}_1^*(x)}$	$MSE_{\widehat{\varphi}_1^*(x)}$	$B_{\widetilde{\varphi}_1^*(x)}$	$MSE_{\widetilde{\varphi}_1^*(x)}$
1250.0	0.02755	0.04085	0.01025	0.00540	0.00078	0.00028
1188.0	0.02863	0.04254	0.01047	0.00537	0.00085	0.00028
1126.0	0.02780	0.04643	0.00991	0.00557	0.00065	0.00029
1064.0	0.02689	0.05068	0.00953	0.00575	0.00064	0.00030
1002.0	0.02890	0.05241	0.00981	0.00559	0.00061	0.00029
940.0	0.02670	0.05545	0.00875	0.00552	0.00032	0.00029
878.0	0.02738	0.06064	0.00872	0.00564	0.00029	0.00029
816.0	0.02877	0.06738	0.00882	0.00577	0.00024	0.00028
754.0	0.03001	0.07071	0.00899	0.00562	0.00037	0.00028
692.0	0.03686	0.07869	0.01065	0.00583	0.00065	0.00029

Table 4: *Pickands and known* ρ_x *cases. Average Lengths (avl) and Coverages (cov) of the 95% confidence intervals, over 2000 Monte-Carlo simulations, sample size* $n = 5000$

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

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\widehat{\varphi}_1^*(x)}$	$cov_{\widehat{\varphi}_1^*(x)}$	$avl_{\widetilde{\varphi}_1^*(x)}$	$cov_{\widetilde{\varphi}_1^*(x)}$
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

$x = 0.50$ $\quad \bar{N}_x = 1250$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\widehat{\varphi}_1^*(x)}$	$cov_{\widehat{\varphi}_1^*(x)}$	$avl_{\widetilde{\varphi}_1^*(x)}$	$cov_{\widetilde{\varphi}_1^*(x)}$
312.1	1.7798	0.9295	0.3232	0.9195	0.0670	0.9485
297.0	1.8330	0.9255	0.3248	0.9245	0.0669	0.9490
281.9	1.8810	0.9250	0.3247	0.9240	0.0669	0.9475
266.8	1.9457	0.9220	0.3269	0.9240	0.0669	0.9460
251.7	2.0095	0.9200	0.3279	0.9145	0.0668	0.9505
236.6	2.1038	0.9195	0.3329	0.9165	0.0668	0.9420
221.5	2.2256	0.9150	0.3409	0.9100	0.0668	0.9390
206.3	2.3707	0.9115	0.3506	0.9075	0.0668	0.9440
191.2	2.4375	0.9105	0.3468	0.9085	0.0667	0.9455
176.1	2.7460	0.9155	0.3754	0.9080	0.0667	0.9440

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

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\widehat{\varphi}_1^*(x)}$	$cov_{\widehat{\varphi}_1^*(x)}$	$avl_{\widetilde{\varphi}_1^*(x)}$	$cov_{\widetilde{\varphi}_1^*(x)}$
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

$x = 1.00$ $\quad \bar{N}_x = 5000$

$k_n(x)$	$avl_{\hat{\rho}_x}$	$cov_{\hat{\rho}_x}$	$avl_{\widehat{\varphi}_1^*(x)}$	$cov_{\widehat{\varphi}_1^*(x)}$	$avl_{\widetilde{\varphi}_1^*(x)}$	$cov_{\widetilde{\varphi}_1^*(x)}$
1250.0	0.8019	0.9645	0.2909	0.9605	0.0670	0.9540
1188.0	0.8238	0.9625	0.2914	0.9595	0.0670	0.9555
1126.0	0.8463	0.9535	0.2914	0.9495	0.0670	0.9425
1064.0	0.8707	0.9510	0.2915	0.9445	0.0670	0.9435
1002.0	0.8994	0.9530	0.2922	0.9455	0.0670	0.9475
940.0	0.9273	0.9445	0.2918	0.9420	0.0669	0.9460
878.0	0.9614	0.9420	0.2923	0.9450	0.0669	0.9420
816.0	1.0002	0.9450	0.2932	0.9440	0.0669	0.9500
754.0	1.0426	0.9475	0.2939	0.9460	0.0669	0.9550
692.0	1.0976	0.9455	0.2966	0.9430	0.0670	0.9455

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 moment 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.

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 2.8. 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.

So 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 ($N_x = 5000$), the Pickands estimator becomes more accurate than the moment estimator.

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. However for inference purpose on the frontier function itself, the estimate of the asymptotic variance of $\hat{\varphi}(x)$ does not provide reliable confidence intervals even for relatively large values of N_x. It would be better to use in the latter case estimates obtained by the bootstrap (but at a computational cost).

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 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 very large) and second use the estimator $\widetilde{\varphi}_1^*(x)$, as if ρ_x was known, by plugging the estimated values $\tilde{\rho}_x$ at the place of ρ_x. In the next section, we suggest some *ad hoc* procedure for determining appropriate values of $k_n(x)$ with a real data set.

Table 5: *Moment Estimators. Bias (B), Mean Squared Error (MSE) and Average Lengths (avl) and Coverages (cov) of the 95% confidence intervals, over 2000 Monte-Carlo simulations, sample size* $n = 1000$

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

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
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

$x = 0.50$ $\bar{N}_x = 250$

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
138.0	0.39692	2.60024	-0.10693	0.02356	3.1218	0.8590	0.2101	0.2850
125.3	0.38007	1.72622	-0.08996	0.01569	2.9481	0.8695	0.1974	0.3410
113.0	0.37306	1.66457	-0.07481	0.01268	3.0650	0.8695	0.2031	0.4060
100.4	0.38515	2.11384	-0.05954	0.01194	3.4478	0.8795	0.2240	0.4655
88.0	0.37768	3.32329	-0.04718	0.01508	4.2040	0.8945	0.2657	0.5315
75.5	0.45592	8.06492	-0.03055	0.03275	6.8652	0.8900	0.4282	0.5960
62.9	0.52777	22.24659	-0.01650	0.07686	14.6330	0.8940	0.8597	0.6450
50.4	0.44658	32.58286	-0.01254	0.09383	21.6793	0.8870	1.1609	0.6895
38.0	0.87725	397.77359	0.01816	1.06690	266.8216	0.8870	13.8012	0.7215
25.5	0.07688	171.77509	-0.01379	0.24232	138.6520	0.8650	5.1985	0.7535
6.8	-1.35243	5682.59492	-0.04653	1.83434	8816.0056	0.7310	148.0389	0.6485

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

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
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

$x = 1.00$ $\bar{N}_x = 1000$

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
500.0	0.20400	0.24793	-0.20017	0.04425	1.1180	0.8300	0.1473	0.0825
450.0	0.19013	0.23874	-0.16962	0.03277	1.1669	0.8610	0.1525	0.1280
400.0	0.17333	0.22090	-0.14244	0.02384	1.2216	0.8910	0.1566	0.1895
350.0	0.16584	0.23578	-0.11662	0.01721	1.3024	0.9075	0.1621	0.2775
300.0	0.15620	0.22091	-0.09339	0.01182	1.3952	0.9265	0.1663	0.3885
250.0	0.14678	0.25498	-0.07248	0.00847	1.5263	0.9295	0.1716	0.4915
200.0	0.15006	0.31032	-0.05223	0.00596	1.7232	0.9310	0.1789	0.5955
150.0	0.15083	0.39042	-0.03440	0.00443	2.0127	0.9275	0.1867	0.6935
100.0	0.17579	0.64048	-0.01732	0.00395	2.5892	0.9120	0.2024	0.7815
50.0	0.22896	5.70734	-0.00257	0.02059	6.5478	0.9025	0.3844	0.8170
25.0	0.99667	2025.55489	0.02869	3.47448	1594.6012	0.8555	65.7858	0.7970

Table 6: *Moment Estimators. Bias (B), Mean Squared Error (MSE) and Average Lengths (avl) and Coverages (cov) of the 95% confidence intervals, over 2000 Monte-Carlo simulations, sample size* $n = 5000$

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

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
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

$x = 0.50$ $\bar{N}_x = 1250$

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
600.5	0.16644	0.16966	-0.09657	0.01004	0.9860	0.8375	0.0645	0.0575
550.5	0.16412	0.16874	-0.08407	0.00776	1.0281	0.8590	0.0667	0.0890
500.4	0.16750	0.17596	-0.07212	0.00588	1.0818	0.8735	0.0691	0.1360
450.5	0.17133	0.18419	-0.06106	0.00440	1.1442	0.8970	0.0715	0.2155
400.5	0.16370	0.19777	-0.05158	0.00334	1.2099	0.9085	0.0733	0.2945
350.5	0.15716	0.20738	-0.04270	0.00250	1.2897	0.9225	0.0751	0.3815
300.5	0.16437	0.23740	-0.03370	0.00182	1.4051	0.9335	0.0778	0.4775
250.4	0.15151	0.25663	-0.02649	0.00137	1.5307	0.9430	0.0794	0.5650
200.5	0.13915	0.28167	-0.01987	0.00101	1.7031	0.9415	0.0811	0.6475
150.5	0.12971	0.36589	-0.01373	0.00082	1.9765	0.9305	0.0836	0.7180
50.5	0.29865	6.19391	0.00098	0.00356	6.8895	0.8895	0.1734	0.8000
13.0	-0.58590	9410.59672	-0.01445	1.57034	10243.4270	0.8150	131.6029	0.7550

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

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
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

$x = 1.00$ $\bar{N}_x = 5000$

$k_n(x)$	$B_{\tilde{\rho}_x}$	$MSE_{\tilde{\rho}_x}$	$B_{\widehat{\varphi}(x)}$	$MSE_{\widehat{\varphi}(x)}$	$avl_{\tilde{\rho}_x}$	$cov_{\tilde{\rho}_x}$	$avl_{\widehat{\varphi}(x)}$	$cov_{\widehat{\varphi}(x)}$
2000.0	0.13502	0.05141	-0.14729	0.02230	0.5207	0.7685	0.0664	0.0000
1800.0	0.13019	0.05132	-0.12609	0.01649	0.5471	0.8140	0.0682	0.0025
1600.0	0.12099	0.04935	-0.10701	0.01202	0.5765	0.8455	0.0697	0.0145
1400.0	0.11212	0.05190	-0.08930	0.00855	0.6129	0.8595	0.0712	0.0455
1200.0	0.10555	0.05445	-0.07261	0.00584	0.6593	0.8965	0.0727	0.1055
1000.0	0.09393	0.05677	-0.05771	0.00388	0.7168	0.9180	0.0740	0.2325
800.0	0.07446	0.05965	-0.04469	0.00251	0.7911	0.9245	0.0748	0.3680
600.0	0.07713	0.07992	-0.03069	0.00148	0.9179	0.9310	0.0771	0.5615
400.0	0.06905	0.10581	-0.01877	0.00087	1.1221	0.9415	0.0790	0.7255
200.0	0.07559	0.20770	-0.00744	0.00059	1.6176	0.9365	0.0830	0.8375
100.0	0.09821	0.49803	-0.00225	0.00067	2.4204	0.9095	0.0896	0.8465
50.0	0.15884	1.20953	0.00051	0.00083	3.9082	0.8920	0.1034	0.8420

3.2 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.

First we have observed in our Monte-Carlo exercise that the optimal value for selecting $k_n(x)$ when estimating the tail index ρ_x is not necessarily the same than the value for estimating the frontier function $\varphi(x)$.

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 function $\varphi(x)$ 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 $k_n(x)$).

For one sample generated with $n = 1000$ in the uniform case of our Monte-Carlo exercise above, we obtain the results shown in Figures 1 and 2. In Figure 1 the estimator $\tilde{\varphi}_1^*(x)$ is first computed with the true value $\rho_x = 2$ (top 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 $\tilde{\rho}_x$ (bottom panel). The pointwise confidence intervals are also displayed. The three right 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 top panel results and according our suggestion at the end of the preceding section, if ρ_x is unknown, we should use the bottom panel results, where we replace ρ_x by its moment estimator $\tilde{\rho}_x$ (here $N_x \le 1000$) and continue as if ρ_x was known. It is quite admirable that both panels are very similar.

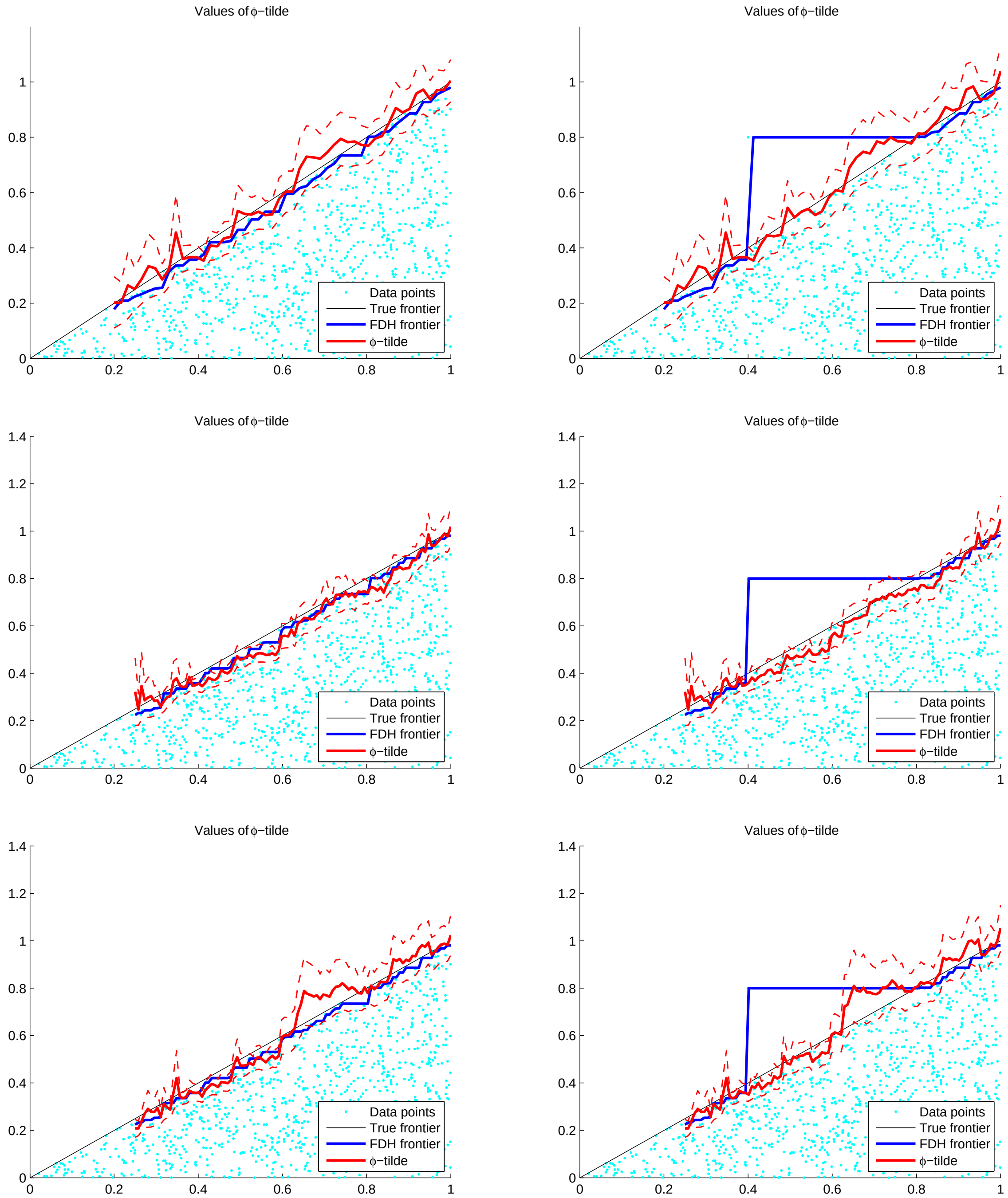

Figure 1: *Resulting estimator* $\tilde{\varphi}_1^*(x)$ *for a uniform data set of size* $n = 1000$ *(plus one outlier for the right panels), from top to bottom, we have the cases* $\rho_x = 2$*, plugging* $\hat{\rho}_x$*, plugging* $\tilde{\rho}_x$*.*

We know that with small values of N_x (when $x = 0.25$, N_x is of the order 60) we cannot hope to have remarkable results when using the Pickands and the moment estimator of $\varphi(x)$. This is confirmed in Figure 2 where we illustrate the behavior of a plug-in version of the Pickands estimator $\hat{\varphi}_1^*(x)$ (top panel) and of the moment estimator $\hat{\varphi}(x)$ (bottom panel). It is a plug-in version because we used here too a two-step estimator: first step, estimation of ρ_x then plugging the obtained values of $\hat{\rho}_x$ (resp. $\tilde{\rho}_x$) for computing the estimators $\hat{\varphi}_1^*(x)$ (resp. $\hat{\varphi}(x)$) and their variances. Again we know by our Monte-Carlo experiment that for small values of x (and so of N_x) the confidence intervals provided by the moment estimators are too narrow and we see that those obtained by the Pickands estimator are quite unstable. However, as expected, we observe a reasonably nice behavior of the estimators themselves. We did the same exercise with a sample of size $n = 5000$, showing even better results but we do not reproduced them for saving space.

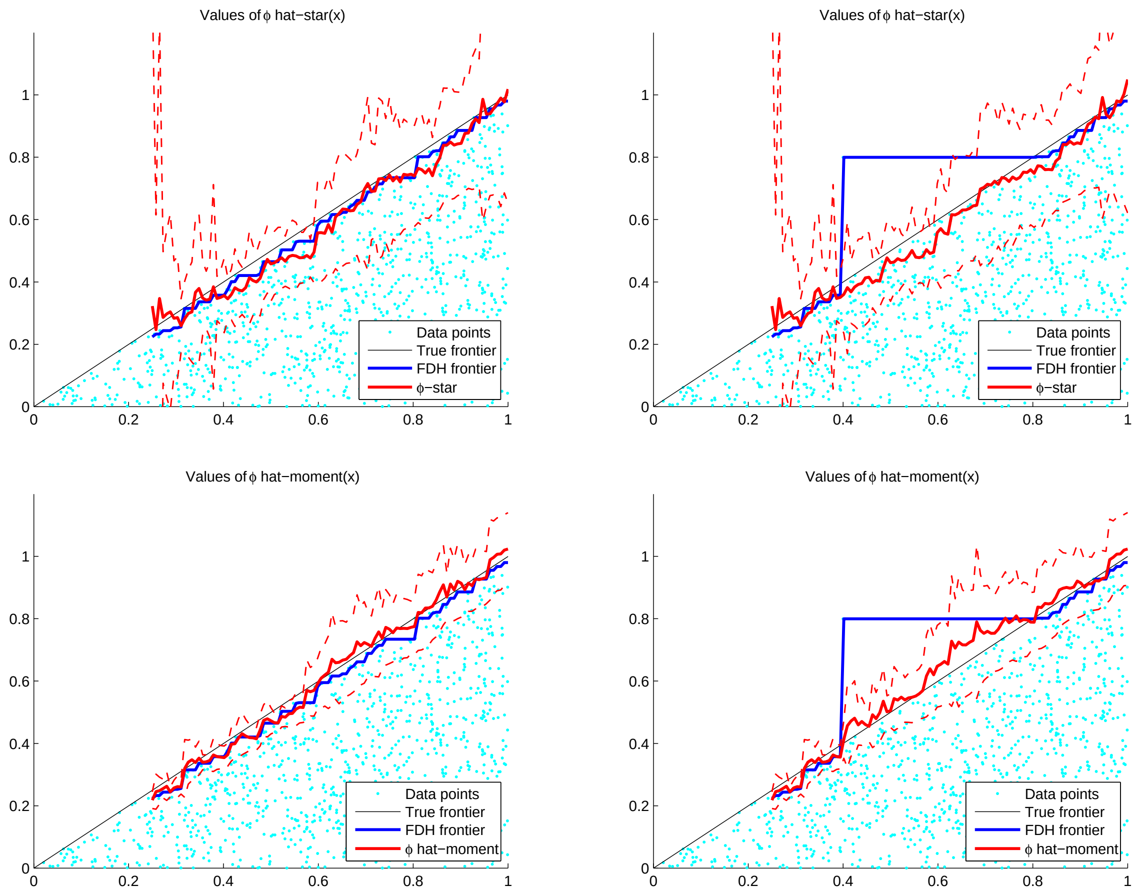

Figure 2: *The uniform data set of size $n = 1000$ (plus one outlier for the right panels), from top to bottom, the plug-in versions of the Pickands $\hat{\varphi}_1^*(x)$ and of the moment estimator $\hat{\varphi}(x)$.*

3.3 An application

We use the same real data example as in Cazals et al [2] and Daouia and Simar [4] 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 $k_n(x)$. We illustrate the same estimators described in the preceding section.

The cloud of points and the resulting estimates are provided in Figures 3 and 4. 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 figures), we obtain the pictures of the left panels: the FDH estimator changes drastically, whereas the extreme-values based estimators are 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.

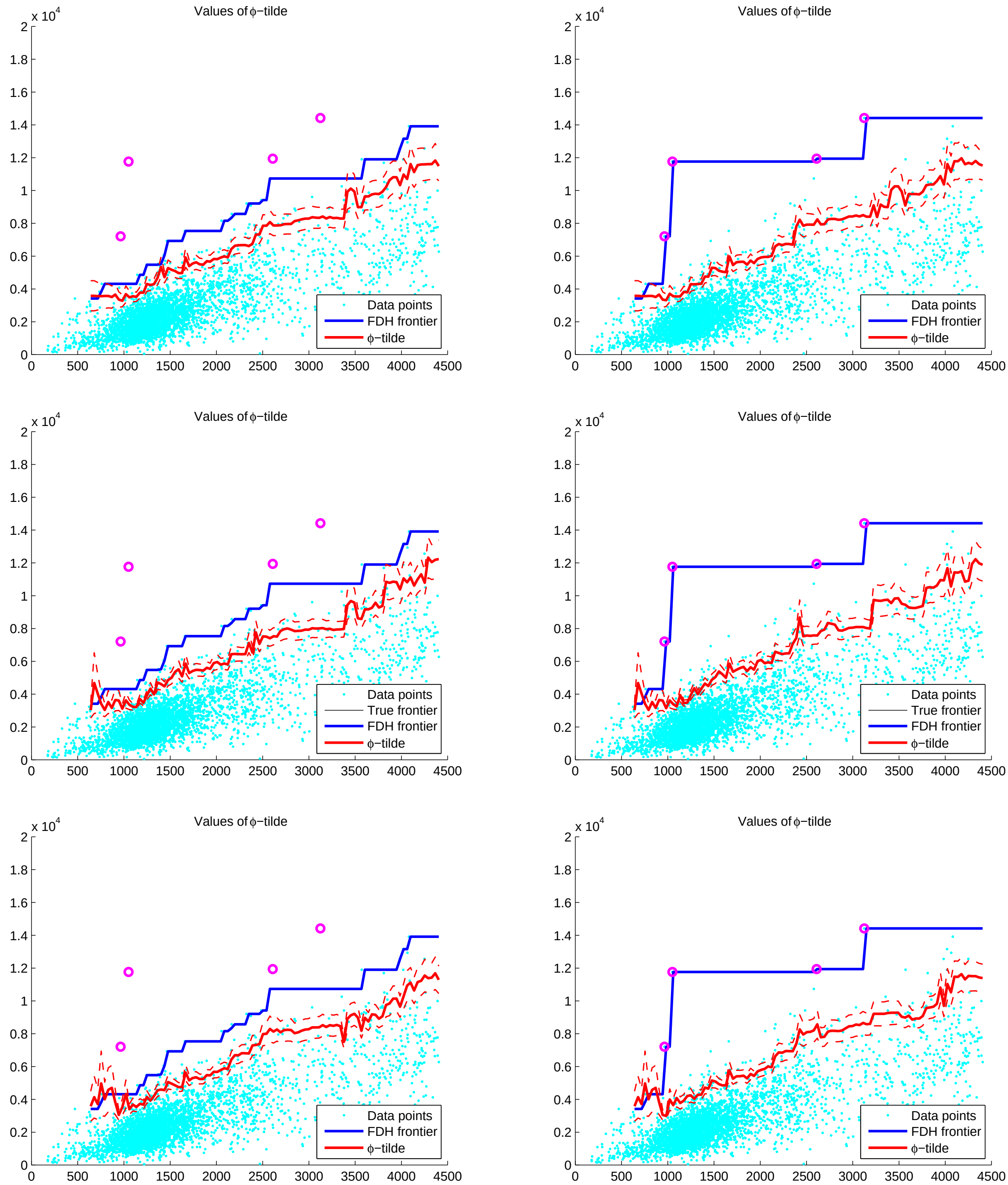

Figure 3: *Resulting estimator* $\tilde{\varphi}_1^*(x)$ *for the French post offices. We include 4 extreme data points (circles) for the right panels. From top to bottom, we have the cases* $\rho_x = 2$*, plugging* $\hat{\rho}_x$*, plugging* $\tilde{\rho}_x$*.*

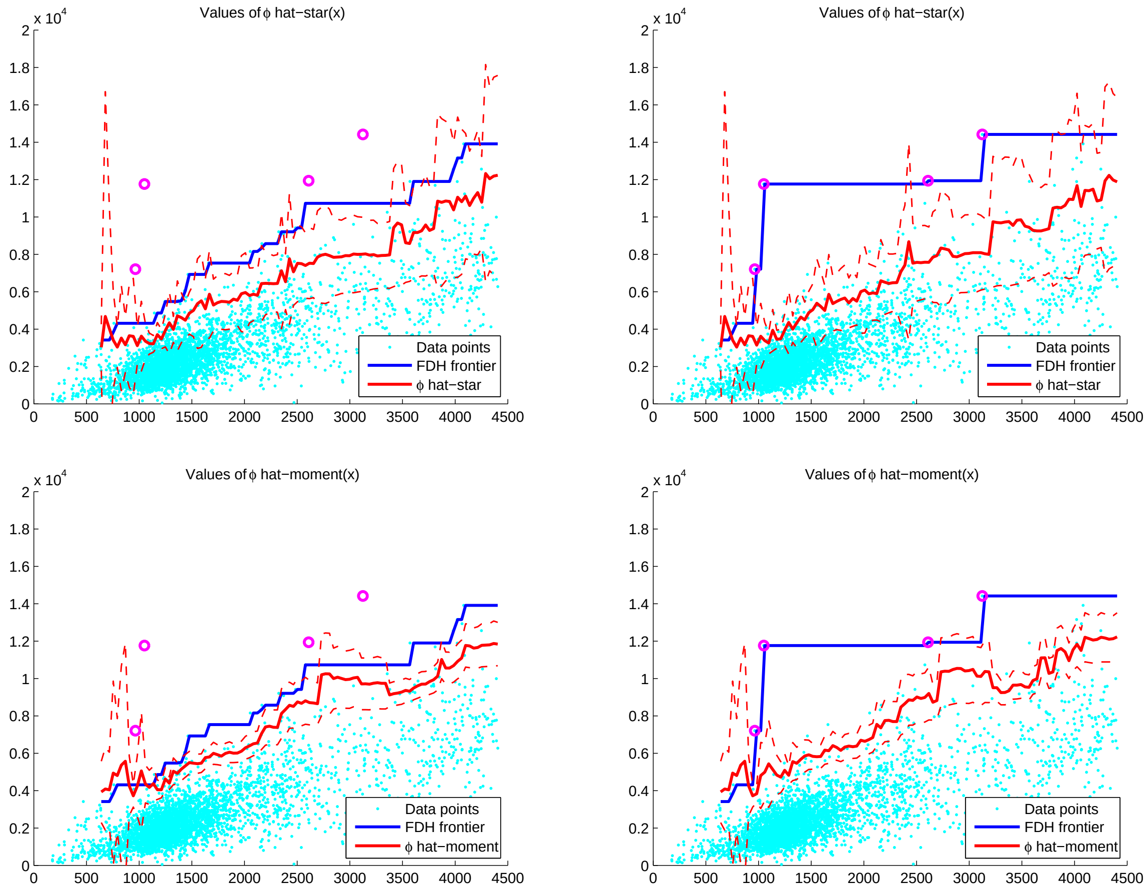

Figure 4: *The French post offices, where we include 4 extreme data points (circles) for the right panels. From top to bottom, the plug-in versions of the Pickands* $\hat{\varphi}_1^*(x)$ *and of the moment estimator* $\hat{\varphi}(x)$.

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 Aragon et al [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 dimension $p+1$ of the data 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 Park et al [27] and Hwang et al [22] to the general setting where $p \geq 1$ and $\beta = \beta_x$ may depend on x.

There is a vast literature on nonparametric estimation of the boundary of the joint support $\mathbb{T}$

of (X,Y). From a practical point of view, compared with the extreme-value based estimators (see *e.g.* [12], [15], [16]), projection estimators (see *e.g.* [23]) or piecewise polynomial estimators (see *e.g.* [26], [21]), our frontier estimators $\hat{\varphi}_1(x)$, $\hat{\varphi}_{\alpha_n}(x)$, $\hat{\varphi}_1^*(x)$, $\tilde{\varphi}_1^*(x)$ and $\hat{\varphi}(x)$ have the advantage to not be limited to a bi-dimensional support since they do not require a partition of $\mathbb{T}$. Moreover our estimators 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.* [19], [20], [24]). From a theoretical point of view, the limit laws of our frontier estimators are 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. Moreover, the new estimators $\hat{\varphi}_1^*(x)$, $\tilde{\varphi}_1^*(x)$ and $\hat{\varphi}(x)$ are asymptotically normally distributed and provide useful asymptotic confidence bands for the monotone frontier function $\varphi(x)$. 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 ([7], [8],...).

A closely related work in boundary estimation via extreme-values theory includes [18] 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 [14] 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 [18] 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 [14] 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 Hall et al [18], 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 [14] 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 [14] and [18].

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 the frontier $\varphi(x)$ and the conditional tail index 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 the frontier and/or the tail index 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. As shown in Remark 2.8 the estimator $\tilde{\varphi}_1^*(x)$ of $\varphi(x)$ based on the true value of ρ_x is theoretically more efficient than the estimator $\hat{\varphi}_1^*(x)$ based on the Pickands estimate of ρ_x.

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 $\varphi(x)$ provide more accurate estimations than the Pickands estimates when considering bias and MSE. When the sample size becomes very large (say of the order of several thousands), the Pickands estimators improve a lot and even seem to outperform the moment estimators. As far as the inference on ρ_x is concerned, the moment estimator provides also quite reliable confidence intervals. However, when inference about the frontier function itself is concerned, the moment estimator provides very poor results. The problem seems to come from an underestimation of the sampling standard deviation of the estimator. This might advocate for the use of alternative bootstrap methods.

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 huge), then we use the asymptotic results for $\tilde{\varphi}_1^*(x)$, as if ρ_x was known.

Appendix: Proofs

Proof of Theorem 2.1 (i) 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 $\{Z_i^x = Y_i\mathbb{1}(X_i \le x), i = 1,\ldots,n\}$ 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 $\max_{i=1,\ldots,n} Z_i^x$ 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 non-degenerate distribution function G, then $G(y) = e^{-(-y)^{\rho}}$ with support $]-\infty, 0]$ and $\rho > 0$ (see *e.g.* Embrechts et al [10], Theorem 3.2.3, p. 121).

(ii) On the other hand (see *e.g.* [10], Theorem 3.3.12, p. 135), there exist norming constants b_n such that $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G$ (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 $\varphi(x) - \inf\{y \ge 0 | F_x(y) \ge 1 - \frac{1}{n}\}$ which coincides with $\varphi(x) - \inf\{y \ge 0 | F(y|x) \ge 1 - \frac{1}{nF_X(x)}\}$.

(iii)-(iv) Under the given regularity conditions, we know that (A.1) holds and it is easy to see that $\mathbb{E}[|Z^x|^k] = F_X(x)\mathbb{E}(Y^k | X \le x) < \infty$. Then it is immediate (see *e.g.* Resnick [29], 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 [29] (see Corollary 2.3, p.83). $\square$

Proof of Corollary 2.1 Notation: a random variable W follows the distribution Weibull$(1, \rho_x)$ if W^{ρ_x} is Exponential with parameter 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. □

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 $x \mapsto \varphi(x)$. □

For the next proofs we need the following lemma.

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 for each $k \in \{0,1,\ldots,n\hat{F}_X(x)-1\}$,

$$\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k)}.$$

(ii) For any fixed integer $k \geq 0$,

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

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

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

The lemma is quite easy to prove but for the sake of completeness we include a proof.

Proof of Lemma .1 (i) First note that since $n\hat{F}_X(x)$ is an integer and $\hat{F}_X(x) > 0$, then $n\hat{F}_X(x) \geq 1$. Note also that for each $k \in \{0,1,\ldots,n\hat{F}_X(x)-1\}$, we have $0 < 1 - \frac{k}{n\hat{F}_X(x)} \leq 1$. Then the empirical conditional quantile $\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$ is well defined and we have by its definition

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

(ii) Since $F_X(x) > 0$ and $\hat{F}_X(x) \xrightarrow{a.s.} F_X(x)$ as $n\to\infty$, we have the event $\{\hat{F}_X(x) > 0$ and $0 \leq \frac{k}{n\hat{F}_X(x)} < 1$ as $n\to\infty\}$ almost surely. This event is equivalent to $\{\hat{F}_X(x) > 0$ and $k \in \{0,1,\ldots,n\hat{F}_X(x)-1\}$ as $n\to\infty\}$ which is contained in the event $\{\hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k)}$ as $n\to\infty\}$ in view of Lemma .1 (i). Thus the later event has a probability 1.

(iii) Here also, since $F_X(x) > 0$ and $k_n/n \to 0$, the event $\{\hat{F}_X(x)\to F_X(x),\ n\to\infty\}$ implies the event $\{\hat{F}_X(x) > 0$ and $0\leq\frac{k_n}{n\hat{F}_X(x)} < 1,\ n\to\infty\} \equiv \{\hat{F}_X(x) > 0$ and $k_n \in \{0,1,\ldots,n\hat{F}_X(x)-1\},\ n\to\infty\}$ which is itself contained in the event $\{\hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x) = Z^x_{(n-k_n)},\ n\to\infty\}$ in view of Lemma .1 (i). Therefore the last event has a probability 1 since $\hat{F}_X(x)\to F_X(x)$ with probability 1. □

Proof of Theorem 2.2 (i) Let $Z_{(i)}^x$ be the i^{th} order statistic generated by the random variables $Z_1^x,\ldots,Z_n^x$. Since $\varphi(x) = F_x^{-1}(1)$ (the endpoint of the distribution F_x of the Z_i^x's) and $\hat{\varphi}_1(x) = Z_{(n)}^x$ for all $n \geq 1$, we have $(\hat{\varphi}_1(x) - \varphi(x)) = (Z_{(n)}^x - F_x^{-1}(1))$. Hence, if $b_n^{-1}(\hat{\varphi}_1(x) - \varphi(x)) \xrightarrow{d} G_x$, then $b_n^{-1}(Z_{(n)}^x - F_x^{-1}(1))$ converges as well to the same non-degenerate distribution G_x. Therefore, following the standard extreme-value theory (see *e.g.* [32], Theorem 21.18, p. 313), $b_n^{-1}(Z_{(n-k)}^x - F_x^{-1}(1)) \xrightarrow{d} H_x$ for any fixed integer $k \geq 0$, where $H_x(y) = G_x(y)\sum_{i=0}^k(-\log G(y))^i/i!$. Finally since $Z_{(n-k)}^x \stackrel{a.s.}{=} \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)$, 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 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 $nb_n^{-1}(1-\alpha_n) \to 0$. □

Proof of Theorem 2.3 (i) Let us consider again 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 [28] estimate of the exponent of variation $\gamma_x < 0$ is given by:

$$\hat{\gamma}_x := (\log 2)^{-1} \log \frac{Z_{(n-k+1)}^x - Z_{(n-2k+1)}^x}{Z_{(n-2k+1)}^x - Z_{(n-4k+1)}^x}.$$

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

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

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

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

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 [7] 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 [7]. This gives the weak consistency of $\hat{\rho}_x$ since $\hat{\gamma}_x \stackrel{a.s.}{=} -1/\hat{\rho}_x$, 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 [7] 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 [7] (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 $\sqrt{k_n}(\hat{\rho}_x - \rho_x) \stackrel{a.s.}{=} \sqrt{k_n}(-\frac{1}{\hat{\gamma}_x} + \frac{1}{\gamma_x})$, 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 $\pm\left\{t^{-1-\frac{1}{\gamma_x}}F_x'(\varphi(x) - \frac{1}{t}) - \delta F_X(x)\right\} \in \mathrm{RV}_{-\kappa}$. The desired conclusion follows then immediately from Theorem 2.5 of Dekkers and de Haan [7] in conjunction with Lemma .1 (iii). □

Proof of Theorem 2.4 (i)-(ii) 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 [8]) for the index γ_x 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 Dekkers et al [8].

(iii) We know by Corollary 3.2 in [8] that $\sqrt{k_n}\{\tilde{\gamma}_x-\gamma_x\}$ has asymptotically a normal distribution with mean 0 and variance

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

provided that $k_n\to\infty$ and $k_n=o(n/g^{-1}(n))$, where $g(t)=t^{1-2\gamma_x}[\{\log\varphi(x)-\log U(t)\}/B(t)]^2$. The desired conclusion follows by using $\rho_x=-1/\gamma_x$, $\tilde{\rho}_x\overset{a.s.}{=}-1/\tilde{\gamma}_x$ as $n\to\infty$, and applying the delta method. $\square$

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 [7] (see Theorem 3.1), $\sqrt{2k_n}\,\frac{Z^x_{(n-k_n+1)}-F_x^{-1}(1-p_n)}{Z^x_{(n-k_n+1)}-Z^x_{(n-2k_n+1)}}$ 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 $\hat{\varphi}_1^*(x)\overset{a.s.}{=}\frac{Z^x_{(n-k_n+1)}-Z^x_{(n-2k_n+1)}}{2^{-\hat{\gamma}_x}-1}+Z^x_{(n-k_n+1)}$ as $n\to\infty$. Then following Theorem 3.2 in [7], $\sqrt{2k_n}\,\frac{\hat{\varphi}_1^*(x)-\varphi(x)}{Z^x_{(n-k_n+1)}-Z^x_{(n-2k_n+1)}}$ is asymptotically normal with mean 0 and variance $3\gamma_x^2 2^{2\gamma_x-1}/(2^{\gamma_x}-1)^6$. This completes the proof. $\square$

Proof of Theorem 2.6 (i) 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

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

The first term at the right hand side tends to zero as established by Dekkers and de Haan ([7], 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 [7]. 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) in conjunction with the fact that

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

(ii) As indicated in the proof of Theorem 2.3 (i), the extreme-value condition (1.1) in Dekkers and de Haan [7] holds under (2.2). Therefore Theorem 3.4 of [7] implies that $\frac{\hat{\varphi}_1^*(x)-\varphi(x)}{Z^x_{(n-k+1)}-Z^x_{(n-2k+1)}}$ 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 [7]. This completes the proof of (ii) in conjunction with the fact that

$$\frac{\overline{\varphi}_1^*(x)-\varphi(x)}{\hat{\varphi}_{1-\frac{k-1}{n\hat{F}_X(x)}}(x)-\hat{\varphi}_{1-\frac{2k-1}{n\hat{F}_X(x)}}(x)} \overset{a.s.}{=} \frac{\hat{\varphi}_1^*(x)-\varphi(x)}{Z^x_{(n-k+1)}-Z^x_{(n-2k+1)}} \quad \text{as} \quad n\to\infty$$

in view of Lemma .1 (ii). □

Proof of Theorem 2.7 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

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

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,

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

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

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

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

$$\frac{\rho_x k_n^{1/2}}{(k_n/n\ell_x)^{1/\rho_x}}\left[Z^x_{(n-k_n+1)} + (k_n/n\ell_x)^{1/\rho_x} - \varphi(x)\right] \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as} \quad n \to \infty. \tag{A.3}$$

We conclude by using $Z^x_{(n-k_n+1)} \overset{a.s.}{=} \hat{\varphi}_{1-\frac{k_n-1}{n\hat{F}_X(x)}}(x)$ as $n \to \infty$. □

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.3) holds. This ends the proof. □

Proof of Theorem 2.8 (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 [8] in conjunction with (A.2), we get

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

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

(ii) Since $Z^x_{(n-k_n)} \overset{a.s.}{=} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)$ and $\tilde{\gamma}_x \overset{a.s.}{=} -1/\tilde{\rho}_x$ as $n \to \infty$ (as shown in the proof of Theorem 2.4 (i)), we have

$$\hat{\varphi}(x) \overset{a.s.}{=} Z^x_{(n-k_n)} M_n^{(1)}(1 - 1/\tilde{\gamma}_x) + Z^x_{(n-k_n)}, \quad n \to \infty. \tag{A.4}$$

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 Dekkeres et al [8] 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 [8] with $\gamma_x = -1/\rho_x < 0$. Then according to Theorem 5.2 in [8], we have

$$\sqrt{k_n}\,\frac{\hat{x}_n^* - F_x^{-1}(1)}{M_n^{(1)} Z^x_{(n-k_n)}(1-\tilde{\gamma}_x)} \xrightarrow{d} \mathcal{N}(0, V_5(-1/\gamma_x))$$

which gives to the desired convergence in distribution of Theorem 2.8 (ii) since $F_x^{-1}(1) = \varphi(x)$, $\hat{x}_n^* \overset{a.s.}{=} \hat{\varphi}(x)$, $\tilde{\gamma}_x \overset{a.s.}{=} -1/\tilde{\rho}_x$ and $Z^x_{(n-k_n)} \overset{a.s.}{=} \hat{\varphi}_{1-\frac{k_n}{n\hat{F}_X(x)}}(x)$ as $n \to \infty$. □

Proof of Theorem 2.9 (i) Putting $\gamma_{x,n} := -1/\rho_{x,n}$ as any consistent estimate of $\gamma_x = -1/\rho_x < 0$, we have by Lemma .1 (ii),

$$\hat{\phi}_{p_n}(x) \overset{a.s.}{=} \frac{(k/np_n)^{\gamma_{x,n}} - 1}{\gamma_{x,n}}\,\frac{Z^x_{(n-k)} M_n^{(1)}}{1/(1-\gamma_{x,n})} + Z^x_{(n-k)}, \quad n \to \infty.$$

We also have $M_n^{(1)} \overset{a.s.}{=} (1/k)\sum_{i=0}^{k-1} \log Z^x_{(n-i)} - \log Z^x_{(n-k)}$, as $n \to \infty$, by Lemma .1 (ii). Then it is easy to check that $\hat{\phi}_{p_n}(x)$ coincides almost surely, for all n large enough, with the estimator $\hat{x}_{p,n}$ of the unconditional quantile $F_x^{-1}(1-p_n)$ introduced by Dekkeres et al [8] in Equation (4.3). Therefore

$$\frac{\hat{x}_{p,n} - F_x^{-1}(1-p_n)}{M_n^{(1)} Z^x_{(n-k)}} \overset{a.s.}{=} \frac{\hat{\phi}_{p_n}(x) - \varphi_{1-\frac{p_n}{F_X(x)}}(x)}{M_n^{(1)} \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)}, \quad n \to \infty.$$

Hence, the desired convergence in distribution follows immediately from Theorem 4.1 in [8].

(ii) By Lemma .1 (ii), $\hat{\varphi}(x)$ coincides here also almost surely, for all n large enough, with the endpoint estimator $\hat{x}_n^*$ introduced by Dekkeres et al [8] in Equation (4.8) and we have

$$\frac{\hat{x}_n^* - F_x^{-1}(1)}{Z^x_{(n-k)} M_n^{(1)}} \overset{a.s.}{=} \frac{\hat{\varphi}(x) - \varphi(x)}{M_n^{(1)} \hat{\varphi}_{1-\frac{k}{n\hat{F}_X(x)}}(x)}, \quad n \to \infty.$$

Moreover, we have by Theorem 4.3 in [8],

$$\frac{\hat{x}_n^* - F_x^{-1}(1)}{Z^x_{(n-k)} M_n^{(1)}} \xrightarrow{d} \left(1 - \frac{1}{\gamma_x}\right) + \left\{ \frac{1}{k} \sum_{i=0}^{k-1} \exp\left(\gamma_x \sum_{j=i}^{k-1} \frac{E_j}{j} \right) - 1 \right\}^{-1}, \quad n \to \infty.$$

This completes the proof by replacing the index γ_x with $-1/\rho_x$. $\square$

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] 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] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

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

[15] Girard, S. and Jacob, P. (2003), Extreme values and Haar series estimates of point process boundaries. *Scandinavian Journal of Statistics*, 30(2), 369–384.

[16] Girard, S. and Jacob, P. (2004), Extreme values and kernel estimates of point processes boundaries. *ESAIM: Probability and Statistics*, 8, 150–168.

[17] 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.

[18] 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.

[19] Hall, P. and Park, B.U. (2004), Bandwidth choice for local polynomial estimation of smooth boundaries, *Journal of Multivariate Analysis*, 91(2), 240–261.

[20] 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.

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

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

[23] Jacob, P. and Suquet, P. (1995), Estimating the edge of a Poisson process by orthogonal series. *Journal of Statistical Planning and Inference*, 46, 215–234.

[24] Knight, K. (2001), Limiting distributions of linear programming estimators. *Extremes*, 4(2), 87–103.

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

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

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

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

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

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

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

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