N° TSE-411

May 2013

"A Γmoment approach to monotonic boundaries estimation"

A. Daoui, S. Girard and A. Guillou

A Γ-moment approach to monotonic boundaries estimation

A. Daouia*, S. Girard†, A. Guillou‡

* Toulouse School of Economics, University of Toulouse, France (abdelaati.daouia@tse-fr.eu)
† MISTIS (INRIA Grenoble Rhône-Alpes/LJK Laboratoire Jean Kuntzmann), INRIA - Laboratoire Jean Kuntzmann, France (Stephane.Girard@inria.fr)
‡ Université de Strasbourg et CNRS, Institut de Recherche Mathématique Avancée, UMR 7501, France (armelle.guillou@math.unistra.fr)

Abstract

The estimation of optimal support boundaries under the monotonicity constraint is relatively unexplored and still in full development. This article examines a new extreme-value based model which provides a valid alternative for completely envelopment frontier models that often suffer from lack of precision, and for purely stochastic ones that are known to be sensitive to model misspecification. We provide different motivating applications including the estimation of the minimal cost in production activity and the assessment of the reliability of nuclear reactors.

JEL classification: C13, C14, D20

Key words : cost function, edge data, extreme-value index, free disposal hull, moment frontier.

1 Introduction

1.1. *Setting and objectives.* Given a data set $\{(Y_1, X_1), \ldots, (Y_n, X_n)\}$ from a support $\mathcal{P} = \big\{(y, x) \in \mathbb{R}_+^q \times \mathbb{R}_+ | \, \varphi(y) \leq x\big\}$, we wish to estimate the unknown function φ, called boundary or frontier. This problem arised in various contexts, such as for example edge estimation in image reconstruction, where the boundary is typically the interface of areas of different intensities or different color tones (see, *e.g.*, Park (2001) for the literature therein). The major task of this paper is to estimate the lower boundary $\varphi(\cdot)$ of $\mathcal{P}$ under the monotonicity constraint. Our first motivating application concerns frontier estimation in econometrics : the data typically consist of input factors $X_i \in \mathbb{R}_+$ used to produce multiple outputs $Y_i \in \mathbb{R}_+^q$ in a certain firm i, and the aim is to investigate the performance of the firms by looking at the minimal amount of input-usage $\varphi(y)$ needed to produce a given level of outputs y. Econometric considerations lead to the natural assumption that the cost function is isotonic nondecreasing with respect to the partial order in the sense that $y \leq y'$ componentwise implies $\varphi(y) \leq \varphi(y')$, for two vectors $y, y' \in \mathbb{R}_+^q$. Our second motivating field of application is the reliability of nuclear reactors where the objective is to analyze the fracture toughness X_i of the reactor pressure vessel material i as a function of the temperature Y_i. The goal is to estimate the so-called master curve prediction $x = \varphi(y)$ of the lowest fracture toughness as a function of the temperature.

1.2. *Previous work on frontier modeling.* Most of the works on boundary estimation in the statistical literature deal with the output-oriented case, where the problem is rather to estimate the

upper curve $\phi(\cdot)$ of the joint support of (X, Y), characterized by $\mathcal{P} = \big\{(x, y) \in \mathbb{R}_+^p \times \mathbb{R}_+ | \, y \leq \phi(x)\big\}$. It is often assumed that $(X_1, Y_1), \ldots, (X_n, Y_n)$ are independent and identically distributed with a density $f(x, y)$ defined as an algebraic function of the distance from the upper frontier, $\{\phi(x) - y\}$, with a power $\beta_x > -1$. The quantity $\beta_x = 0$ corresponds to a jump of the density at the boundary $\phi(x)$. When $\beta_x \neq 0$, it describes the rate at which the density decays to zero smoothly (in case $\beta_x > 0$) or rises up to infinity (in case $\beta_x < 0$) as it approaches to the boundary. The case $\beta_x > 1$ has been considered in Hall *et al.* (1997), where the estimation of $\phi(x)$ is based on an increasing number of upper order statistics generated by the Y_i's falling into a strip around x. The case of general β_x can be found in Gijbels and Peng (2000), where the maximum of all Y_i observations falling into a strip around x and another frontier estimator based on three large order statistics of the Y_i's in the strip are considered. Piecewise polynomial estimators for $\phi(x)$ have been studied in Härdle *et al.* (1995) in the case of non-sharp boundaries, and polynomial estimators have been used in Hall *et al.* (1998). All of these elegant approaches do not rely, however, on the monotonicity constraint.

There is a vast econometrics literature on monotone frontier analysis, but there are mainly two popular nonparametric methods based on envelopment techniques: the free disposal hull (FDH) estimator introduced by Deprins *et al.* (1984) and defined as the optimal step and monotone surface which envelops all the data points, and the data envelopment analysis (DEA) estimator initiated by Farrell (1957) and which can be defined as the optimal piecewise and concave function covering the FDH estimator. Although their simplicity, the statistical aspects of these envelopment estimators have been explored only during the last decade. See, *e.g.*, Jeong and Park (2006), Kneip *et al.* (2008), Daouia *et al.* (2010) and Park *et al.* (2010) for recent asymptotic developments in the output-orientation. Note that the convexity assumption on $\mathcal{P}$ is not always valid, although it is widely used in economics. Hence, the FDH is a more general estimator than the DEA. Note also that the output-orientation measures the maximal quantity of outputs which can be produced with a given level of inputs, while the input-orientation searches for the minimal cost needed to produce a given amount of outputs. The optimal support boundary, $\mathcal{P}^{\partial}$, is unique and the graphs of the "minimum input" function $\varphi(y)$ and the "maximum output" function $\phi(x)$ are two different ways of describing it.

1.3. *Benchmark free disposal hull model.* In the input-orientation, which is of genuine interest from the economic viewpoint, the estimation of the lower frontier $\varphi(\cdot)$ under the monotonicity constraint is relatively unexplored and still in full development. A closed form expression of $\varphi(y)$ has been suggested by Cazals *et al.* (2002) in terms of the non-standard conditional distribution of X given $Y \geq y$. If $(\Omega, \mathcal{A}, \mathbb{P})$ denotes the probability space on which the random vector $(Y, X) \in \mathbb{R}_+^q \times \mathbb{R}_+$ is defined and $S(x|y) = \mathbb{P}(X > x|Y \geq y)$ is the survival function of X conditioned by $Y \geq y$ assuming $\mathbb{P}(Y \geq y) > 0$, then $\varphi(y)$ can be characterized as the lower endpoint of the

non-standard conditional distribution, *i.e.*,

$$\varphi(y) = \inf\{x \geq 0 \,|\, S(x|y) < 1\}. \tag{1}$$

Generally speaking, $\varphi(y)$ is not the lower boundary $\psi(y)$ of the support of (Y, X) at $Y = y$, but equals $\inf_{y' \geq y} \psi(y')$. Therefore, it is isotonic nondecreasing and envelops the lower support boundary. In the case where the frontier function ψ is nondecreasing, which is the main shape constraint in the present paper, φ coincides with ψ. Then, consideration of φ is advantageous since it is expected to afford estimation at a faster rate than ψ. Because of the local nature of ψ, one can use only the observations in a local strip around y to estimate it, while it is not the case with estimation of φ. Replacing the survival function $S(x|y)$ in (1) with its empirical counterpart $\widehat{S}_n(x|y) = \sum_{i=1}^n \mathbb{1}_{\{X_i > x, Y_i \geq y\}} / \sum_{i=1}^n \mathbb{1}_{\{Y_i \geq y\}}$, Cazals *et al.* (2002) recover the intuitive FDH estimator

$$\widehat{\varphi}_n(y) = \inf\{x \geq 0 \,|\, \widehat{S}_n(x|y) < 1\} \equiv \min\{X_i : 1 \leq i \leq n,\, Y_i \geq y\}, \tag{2}$$

which defines the largest step and monotone function lying below the sample points (Y_i, X_i). Only Park *et al.* (2000) have determined its limit distribution under the restrictive condition that the density of data is strictly positive at the boundary.

1.4. *Contributions and organization of the paper.* Starting from the formulations (1) and (2) of $\varphi(y)$ and $\widehat{\varphi}_n(y)$ as conditional endpoints, the problem of convergence in distribution of suitably normalized FDH is elucidated in a general setting in Section 2.1 by using simple arguments from extreme value theory. An intuitive interpretation of the so-called extreme-value index γ_y, which is involved in the necessary and sufficient condition of convergence as well as in the limit distribution, is provided in terms of the data dimension $(q+1)$ and the shape of the joint density near its support boundary. This allows in particular to recover the results of Park *et al.* (2000) in the special case $\gamma_y = -(q+1)^{-1}$. The convergence of moments is also easily derived in the general setting.

In absence of information on whether the data are measured accurately, it would look awkward for practitioners to assume that only the FDH boundary points $(Y_i, X_i \equiv \widehat{\varphi}_n(Y_i))$ contain valuable information about the lower distribution tail, especially as FDH observations may look so isolated from the cloud points that they seem hardly related to the sample. Different estimation techniques have been developed for so-called stochastic frontier models, where observational errors or random noise allow some observations to be outside of the frontier. The advantages of such (semi-)parametric models come at the cost of explicit assumptions on the functional form of the frontier and/or the distribution of noise; see, *e.g.*, Kumbhakar and Lovell (2000) for a nice survey. In Section 2.2, we show how other top observations, well inside the sample, could help the practitioners to achieve their objective by using ideas from de Haan and Ferreira (2006) which in turn are based on the popular moments device of Dekkers *et al.* (1989).

The resulting moment frontier estimator is quite appealing for huge samples of the order of several thousands, but unfortunately disappoints by its rather large volatility for small and moderate samples. This motivated the quest for alternative estimators in Section 2.3: we have been able to come up with a sensible correction of the FDH boundary for its inherent bias via the convergence of its moments, which then inspired our main methodological innovation in this article, Γ-moment estimation called. Practical guidelines to effect the necessary computations of the estimators are described in Section 2.4: they are first based on the prescription of Ferreira *et al.* (2003) about optimizing the estimation of endpoints, and then on a monotonization technique of the unconstrained estimators. Evidence is given in Section 3 to demonstrate the superiority of the Γ-moment frontiers over the usual moment and FDH estimators. Section 4 returns to our motivating applications and explores boundary estimation for a large dataset on the delivery activity of postal services and two small datasets on the reliability of nuclear reactors and the productivity of electric utility companies.

2 Main results

Let $(y, \varphi(y))$ be the point we want to estimate at the lower boundary of the support of (Y, X) such that $\mathbb{P}(Y \geq y) > 0$ and $\varphi(y) > 0$. For the sake of conciseness, we focus in this section on the monotone nondecreasing case. Similar considerations evidently apply to the nonincreasing case.

2.1 Data envelopment estimation

To estimate $\varphi(y)$ we look at observations X_i having $Y_i \geq y$, and estimate $\varphi(y)$ via extreme order statistics of these X_i's. More precisely, we propose to transform the random vectors (Y_i, X_i) into the dimensionless variables

$$Z_i^y = X_i^{-1} \mathbb{1}_{\{Y_i \geq y\}},\ i = 1, \ldots, n.$$

It is not hard to verify that the common distribution function of the Z_i^y's satisfies

$$F_{Z^y}(z) = \begin{cases} 1 - \mathbb{P}(X < \frac{1}{z}, Y \geq y) & \text{for} \quad z > 0 \\ 1 - \mathbb{P}(Y \geq y) & \text{for} \quad z = 0. \end{cases}$$

Writing $F_{Z^y}^{\leftarrow}(\alpha) := \inf\{z \geq 0 | F_{Z^y}(z) \geq \alpha\}$ for the quantile of order $\alpha \in (0, 1]$ of F_{Z^y}, it is then easily seen from (1) that

$$\varphi(y) \equiv \{F_{Z^y}^{\leftarrow}(1)\}^{-1}.$$

Likewise, it is immediate from (2) that $\widehat{\varphi}_n(y)$ is identical to $\{\max_{1 \leq i \leq n} Z_i^y\}^{-1}$. Therefore, the asymptotic properties of the maximum carry over automatically to the FDH estimator.

In all the sequel, denote by $U_y(t)$ the tail quantile function defined as $U_y(t) := F_{Z^y}^{\leftarrow}(1 - 1/t)$. Assume the classical first order condition

$$\lim_{t \uparrow \infty} \frac{U_y(tx) - U_y(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma} \quad \text{for} \quad x > 0 \tag{3}$$

where $a(.)$ is a positive auxiliary function, regularly varying with index γ, denoted as RV_γ (see Bingham *et al.*, 1987).

Theorem 1.

(i) There exists $\{a_n > 0\}$ such that $a_n^{-1}(\varphi(y) - \widehat{\varphi}_n(y))$ converges to a non-degenerate distribution if and only if, for some $\gamma < 0$ and for all $z > 0$, (3) holds. The limit distribution function is then given by $G_\gamma(x) = \exp\{-(-x)^{-1/\gamma}\}$ with support $(-\infty, 0)$, and we can set $a_n = \varphi^2(y)\left\{\frac{1}{\varphi(y)} - U_y(n)\right\}$.

(ii) Assume that the upper endpoint of the support of X is finite and that $a_n^{-1}\{\varphi(y) - \widehat{\varphi}_n(y)\} \xrightarrow{d} G_\gamma$ with $a_n = \varphi^2(y)\left\{\frac{1}{\varphi(y)} - U_y(n)\right\}$. Then, for any integer $k \geq 1$,

$$\lim_{n\to\infty} \mathbb{E}\left[a_n^{-1}\{\varphi(y) - \widehat{\varphi}_n(y)\}\right]^k = (-1)^k\,\Gamma(1-k\gamma) \tag{4}$$

where $\Gamma(\cdot)$ stands for the gamma function.

Proof. Part (i). According to Resnick (1987, Propositions 0.3 and 1.13), if $b_n = \varphi^{-1}(y) - U_y(n)$, then $b_n^{-1}(\widehat{\varphi}_n^{-1}(y) - \varphi^{-1}(y))$ converges to a non degenerate distribution function if and only if $1 - F_{Z^y}\left(\varphi^{-1}(y) - x^{-1}\right) \in RV_{1/\gamma}$ as $x \to \infty$. Now, using (2.11) in Beirlant *et al.* (2004), this condition is equivalent to $U_y(x) = \varphi^{-1}(y) - x^\gamma \ell_U(x)$ as $x \to \infty$ where ℓ_U is a slowly varying function at infinity, i.e. equivalent to condition (3) (see Corollary 1.2.10, de Haan and Ferreira, 2006). Finally, note that by Slutsky's lemma, saying that $b_n^{-1}(\widehat{\varphi}_n^{-1}(y) - \varphi^{-1}(y))$ converges to a non degenerate distribution function is equivalent to the same convergence for $a_n^{-1}(\varphi(y) - \widehat{\varphi}_n(y))$ with $a_n = b_n\varphi^2(y)$.
Part (ii). Using directly Z^y does not give the desired convergence of moments. Here we shall use the alternative transformation $\widetilde{Z}^y := -X\mathbb{1}_{\{Y\geq y\}} - x_*\mathbb{1}_{\{Y\not\geq y\}}$ and its independent and identically distributed copies $\widetilde{Z}_1^y, \ldots, \widetilde{Z}_n^y$, with x_* being the upper endpoint of the support of X. Clearly, the upper endpoint of the support of $\widetilde{Z}^y$ coincides with $-\varphi(y)$ and the sample maximum $\widetilde{Z}_{(n)}^y = \max_i \widetilde{Z}_i^y$ is identical to $-\widehat{\varphi}_n(y)$. Moreover, it is easy to see that the distribution function of $\widetilde{Z}^y$ is given by $F_{\widetilde{Z}^y}(\tilde{z}) = \{1 - \mathbb{P}[X < -\tilde{z}, Y \geq y]\}\mathbb{1}_{\{-x_*\leq\tilde{z}\leq-\varphi(y)\}}$. Then $F_{\widetilde{Z}^y}(\tilde{z}) = F_{Z^y}(-1/\tilde{z})$ for every $\tilde{z} \in [-x_*, -\varphi(y)]$. Whence, for all $t > \mathbb{P}^{-1}(Y \geq y)$,

$$\begin{aligned}
F_{\widetilde{Z}^y}^{\leftarrow}(1-1/t) &= \inf\left\{\tilde{z} \in [-x_*, -\varphi(y)] : F_{\widetilde{Z}^y}(\tilde{z}) \geq 1 - 1/t\right\} \\
&= -1/\inf\left\{-\tilde{z}^{-1} \in [x_*^{-1}, \varphi^{-1}(y)] : F_{Z^y}(-\tilde{z}^{-1}) \geq 1 - 1/t\right\} \\
&= -1/F_{Z^y}^{\leftarrow}(1-1/t) = -1/U_y(t).
\end{aligned}$$

Therefore, the sequence $c_n := F_{\widetilde{Z}^y}^{\leftarrow}(1) - F_{\widetilde{Z}^y}^{\leftarrow}(1-1/n)$ satisfies $c_n = a_n/[\varphi(y)U_y(n)] \sim a_n$ as $n \to \infty$. Thus, $a_n^{-1}\{\varphi(y) - \widehat{\varphi}_n(y)\} \xrightarrow{d} G_\gamma$ implies $c_n^{-1}(\widetilde{Z}_{(n)}^y - F_{\widetilde{Z}^y}^{\leftarrow}(1)) \xrightarrow{d} G_\gamma$. On the other hand, $\mathbb{E}|\widetilde{Z}^y|^k \leq x_*^k < \infty$ for all $k > 0$. Hence, following *e.g.* Resnick (1987, Proposition 2.1(ii), p.77), we have

$$\lim_{n\to\infty} \mathbb{E}\left[c_n^{-1}\{\widetilde{Z}_{(n)}^y - F_{\widetilde{Z}^y}^{\leftarrow}(1)\}\right]^k = (-1)^k\,\Gamma(1-k\gamma),$$

which ends the proof since $c_n \sim a_n$, $\widetilde{Z}^y_{(n)} = -\widehat{\varphi}_n(y)$ and $F^{\leftarrow}_{\widetilde{Z}^y}(1) = -\varphi(y)$. □

Note that the normalizing sequence $a_n > 0$ and the extreme-value index $\gamma < 0$ of the transformed distribution $F_{Z^y}(\cdot)$ depend on y. For simplicity of notation we do not mention this dependence.

Next we shall provide more motivation for the assumption (3) under which the FDH estimator converges in distribution, and give an intuitive meaning for the parameter γ. The necessary and sufficient condition (3) has the following equivalent representation in terms of the joint distribution of (Y, X) :

$$\mathbb{P}(Y \geq y, X < x) = \left(\frac{x - \varphi(y)}{x}\right)^{-1/\gamma} L_y\left(\frac{x}{x - \varphi(y)}\right) \quad \text{for} \quad x > \varphi(y),$$

where $L_y(\cdot)$ is a slowly varying function. In order to recover the usual assumption in frontier analysis that the joint density of data is an algebraic function of the distance $(x - \varphi(y))$ from the lower support boundary, it is enough to consider the simple class of functions $L_y(\cdot)$ satisfying $L_y(t) = \ell_y > 0$ for all t sufficiently large. Indeed, this leads to the sufficient condition

$$\mathbb{P}(Y \geq y, X < x) = \ell_y \, x^{1/\gamma} \, (x - \varphi(y))^{-1/\gamma} \quad \text{as} \quad x \downarrow \varphi(y). \tag{5}$$

Then, by differentiating both sides of (5) with respect to x and y (assuming that the functions ℓ_y, $\gamma = \gamma_y$ and $\varphi(y)$ are differentiable in y), it is not hard to verify that the density $f(y, x)$ of (Y, X) exhibits algebraic tails, *i.e.*,

$$f(y, x) = c_y \, \{x - \varphi(y)\}^{\beta_y} + o\left(\{x - \varphi(y)\}^{\beta_y}\right) \quad \text{as} \quad x \downarrow \varphi(y), \tag{6}$$

where the shape parameter β_y of the joint density has the explicit expression

$$\beta_y = -\frac{1}{\gamma} - (q + 1), \tag{7}$$

and c_y is a strictly positive constant (provided that $\gamma > -\frac{1}{q}$ and the first partial derivatives of $\varphi(y)$ are strictly positive). As such, the regular-variation exponent γ in (3) turns into a parameter with an intuitive interpretation: When $\gamma > -\{q + 1\}^{-1}$, the density of data decays to zero smoothly as it approaches to the lower boundary; When $\gamma = -\{q + 1\}^{-1}$, the joint density has sudden jumps at the frontier; Finally, the range $\gamma < -\{q + 1\}^{-1}$ corresponds to a density rising up to infinity as it approaches to the frontier. In the sequel, we focus on the more realistic range $\gamma \geq -\{q + 1\}^{-1} \geq -1/2$.

2.2 Moment frontier estimation

Let $Z^y_{(1)} \leq \cdots \leq Z^y_{(n)}$ be the order statistics of the dimensionless sample $Z^y_1, \ldots, Z^y_n$. Instead of the maximum estimator $Z^y_{(n)}$ which is very simple in nature, a prominent way of estimating the

endpoint $F^{\leftarrow}_{Z^y}(1)$ based on other extreme observations has been proposed by Dekkers *et al.* (1989). The underlying idea is to estimate an *anchor* quantile well inside the transformed sample but near the frontier point, and then to shift it to the right place. Our second estimator of the frontier function $\varphi(y) = 1/F^{\leftarrow}_{Z^y}(1)$ follows the prescription of de Haan and Ferreira (2006) which is based on the moment estimator (Dekkers *et al.*, 1989) of the extreme-value index γ. For our problem the tail condition (3) means

$$\varphi(y) = \{U_y(\infty)\}^{-1} \approx \left\{ U_y\left(\frac{n}{k}\right) - \frac{a(\frac{n}{k})}{\gamma} \right\}^{-1} \tag{8}$$

for a sequence $k := k(n) = o(n)$ as $n \to \infty$, that is, the frontier point $\varphi(y)$ is linked to γ, $a(n/k)$ and the quantile $U_y(n/k) = F^{\leftarrow}_{Z^y}(1 - k/n)$. This latter quantity can be estimated simply by its empirical counterpart $Z^y_{(n-k)}$, whereas the quantities γ and $a(n/k)$ are estimated by the moment estimators given by

$$\begin{aligned} \widehat{\gamma}_M &:= M_n^{(1)} + 1 - \frac{1}{2}\left(1 - \frac{(M_n^{(1)})^2}{M_n^{(2)}}\right)^{-1} \\ \widehat{a}\left(\frac{n}{k}\right) &:= Z^y_{(n-k)}\, M_n^{(1)}\, (1 - \widehat{\gamma}_M + M_n^{(1)}), \end{aligned}$$

where

$$M_n^{(r)} = \frac{1}{k}\sum_{i=0}^{k-1} \left(\log Z^y_{(n-i)} - \log Z^y_{(n-k)}\right)^r, \quad r \geq 1.$$

A detailed description of their asymptotic properties can be found in de Haan and Ferreira (2006, Sections 3.5 and 4.2). Hence the approximation (8) motivates the following frontier point estimator

$$\widehat{\varphi}^*_M(y) = \left\{ Z^y_{(n-k)} - \frac{\widehat{a}\left(\frac{n}{k}\right)}{\widehat{\gamma}_M} \right\}^{-1}. \tag{9}$$

In order to be able to derive its asymptotic normality, as usual in extreme value theory, we shall need the following second-order refinement of the relation (3) which specifies the rate of convergence:

$$\lim_{t\uparrow\infty} \frac{\frac{U_y(tx)-U_y(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right) \tag{10}$$

holds for $x > 0$ with $\rho \leq 0$ and A not changing sign and such that $A(t) \to 0$, as $t \to \infty$. Then for $\gamma \neq \rho$, a second-order condition also holds for $\log U_y(t)$ (see *e.g.* de Haan and Ferreira, 2006, p.130), namely

$$\lim_{t\uparrow\infty} \frac{\frac{\log U_y(tx)-\log U_y(t)}{a(t)/U_y(t)} - \frac{x^\gamma - 1}{\gamma}}{Q(t)} = \frac{1}{\rho'}\left(\frac{x^{\gamma+\rho'}-1}{\gamma+\rho'} - \frac{x^\gamma - 1}{\gamma}\right)$$

with $\rho' = \max(\gamma, \rho)$ and Q not changing sign and such that $Q(t) \to 0$ as $t \to \infty$. Finally, assume

$$k = k(n) \to \infty, \quad n/k \to \infty \quad \text{and} \quad \sqrt{k}\, Q(n/k) \to \lambda \in \mathbb{R} \quad \text{as} \quad n \to \infty. \tag{11}$$

The asymptotic distribution of the moment frontier estimator $\widehat{\varphi}_M^*(y)$ of $\varphi(y)$ is given in the next proposition.

Theorem 2. *Given* (10) *and* (11),

$$\frac{\sqrt{k}\,\widehat{\gamma}_M^2}{\widehat{a}\left(\frac{n}{k}\right)\{\widehat{\varphi}_M^*(y)\}^2}\left\{\varphi(y)-\widehat{\varphi}_M^*(y)\right\} \xrightarrow{d} N(\mu_y,\sigma_y^2)$$

where

$$\mu_y=\begin{cases}\frac{\lambda\rho(1-\gamma)}{(\gamma+\rho)(1-\gamma-\rho)(1-2\gamma-\rho)}, & \gamma<\rho\leq 0\\ \frac{\lambda\gamma(1-3\gamma^2)}{(1-\gamma)(1-2\gamma)(1-3\gamma)}, & \rho<\gamma<0\end{cases} \quad \textit{and} \quad \sigma_y^2=\frac{(1-\gamma)^2(1-3\gamma+4\gamma^2)}{(1-2\gamma)(1-3\gamma)(1-4\gamma)}.$$

Proof. On the first hand, $\varphi(y)^{-1}$ coincides with the endpoint $F_{Z^y}^{\leftarrow}(1)$ of our univariate transformed distribution. On the other hand, it is easy to see that $\widehat{\varphi}_M^*(y)^{-1}$ is identical to the moment estimator (Dekkers *et al.*, 1989) of the upper endpoint $F_{Z^y}^{\leftarrow}(1)$. Then, one can simply use the same considerations applied by de Haan and Ferreira (2006, Paragraph 4.3.2, p.140) to obtain the asymptotic normality

$$\frac{\sqrt{k}\,\widehat{\gamma}_M^2}{\widehat{a}\left(\frac{n}{k}\right)}\left\{\frac{1}{\widehat{\varphi}_M^*(y)}-\frac{1}{\varphi(y)}\right\} \xrightarrow{d} N(\mu_y,\sigma_y^2).$$

The stated result follows immediately. □

An alternative option would be to use in (9) the estimator

$$\widehat{\gamma}:=1-\frac{1}{2}\left(1-\frac{(M_n^{(1)})^2}{M_n^{(2)}}\right)^{-1}=\widehat{\gamma}_M-M_n^{(1)}$$

of the negative extreme-value index γ in place of the general moment estimator $\widehat{\gamma}_M$ to get the frontier estimator

$$\widehat{\varphi}_n^*(y):=\left\{Z_{(n-k)}^y-\frac{\widehat{a}\left(\frac{n}{k}\right)}{\widehat{\gamma}}\right\}^{-1}, \tag{12}$$

with the same definition of $\widehat{a}\left(\frac{n}{k}\right)$ as above. Then, following *e.g.* de Haan and Ferreira (2006, p.148), we have under the conditions of Theorem 2,

$$\frac{\sqrt{k}\,\widehat{\gamma}^2}{\widehat{a}\left(\frac{n}{k}\right)\{\widehat{\varphi}_n^*(y)\}^2}\left\{\varphi(y)-\widehat{\varphi}_n^*(y)\right\} \xrightarrow{d} N(\widetilde{\mu}_y,\sigma_y^2)$$

where

$$\widetilde{\mu}_y=\begin{cases}\mu_y & , \quad \gamma<\rho\leq 0\\ \frac{\lambda\gamma(1-3\gamma+3\gamma^2)}{(1-2\gamma)(1-3\gamma)} & , \quad \rho<\gamma<0.\end{cases}$$

From a theoretical point of view, there is no advantage of using the estimator (9) rather than (12) for a fixed y, but in our context of curve estimation, it appears in practice that the latter version provides more sensible and stable results as y varies than the former.

2.3 Γ-moment frontier estimation

According to the analysis of Aarssen and de Haan (1994, Lemma A.3) on endpoint estimation (see also de Haan and Ferreira, 2006, Remark 4.5.5), it turns out that it is most efficient to use the moment estimator $\widehat{\varphi}_n^*(y)$ for estimating $\varphi(y) = \{F_{Z^y}^{\leftarrow}(1)\}^{-1}$ in the range $\gamma > -\frac{1}{2}$. Good estimates may, however, require a large sample size of the order of several thousands. For small samples, as demonstrated in our simulation study, the naive FDH estimator appears to be superior to the moment estimator in terms of mean-squared error, whereas the latter is the winner in terms of bias. On the other hand, when $\gamma < -\frac{1}{2}$, the FDH $\widehat{\varphi}_n(y) = \{Z_{(n)}^y\}^{-1}$ converges faster than $\widehat{\varphi}_n^*(y)$. A similar result has been established by Girard *et al.* (2012) for a high order moments estimator of the endpoint. Nevertheless, even in this latter advantageous setting, $\widehat{\varphi}_n(y)$ overestimates $\varphi(y)$ uniformly in y, with probability 1. In this section, we shall first build intuitive bias-reduced versions of this envelopment estimator.

If $a_n^{-1}\{\varphi(y) - \widehat{\varphi}_n(y)\} \xrightarrow{d} G_\gamma$, then the key element is that $a_n^{-1}\{\varphi(y) - \mathbb{E}[\widehat{\varphi}_n(y)]\} \to \mathbb{E}(G_\gamma) = -\Gamma(1-\gamma)$ as $n \to \infty$, in view of (4). This moment's convergence naturally suggests to use the alternative estimator $\widetilde{\varphi}(y) := \widehat{\varphi}_n(y) - a_n\Gamma(1-\gamma)$ since then $a_n^{-1}\{\varphi(y) - \mathbb{E}[\widetilde{\varphi}(y)]\} \to 0$ and $a_n^{-1}\{\varphi(y) - \widetilde{\varphi}(y)\}$ converges to the centered Weibull extreme-value distribution $G_\gamma + \Gamma(1-\gamma)$, as $n \to \infty$. In applications, the correction term $a_n\Gamma(1-\gamma)$ must, however, be estimated. This can be achieved by substituting the estimated values $\widehat{\gamma}$ and

$$\widetilde{a}_n := -\widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\left\{\widehat{\varphi}_n(y)\right\}^2$$

in place of the extreme-value index γ and the scaling a_n, respectively. The asymptotic distribution of the resulting estimator

$$\begin{aligned}\widetilde{\varphi}_n(y) &:= \widehat{\varphi}_n(y) - \widetilde{a}_n\Gamma(1-\widehat{\gamma}) \\ &= \left\{Z_{(n)}^y + \widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\,\Gamma(1-\widehat{\gamma})\right\}\left(Z_{(n)}^y\right)^{-2}\end{aligned}$$

of $\varphi(y)$ is established in the next theorem. The scale $a(\cdot)$ in the tail conditions above being regularly varying with index γ, it can be written equivalently as $a(x) = x^\gamma\ell(x)$, where $\ell(\cdot)$ is a slowly varying function.

Theorem 3. *Given* (10)-(11) *and assuming that* $\frac{\ell(n)}{\ell(n/k)} \to 1$, *we have*

$$\widetilde{a}_n^{-1}\left\{\varphi(y) - \widetilde{\varphi}_n(y)\right\} \xrightarrow{d} G_\gamma + \Gamma(1-\gamma).$$

Proof. Using the fact that

$$\widetilde{a}_n^{-1}\left\{\varphi(y) - \widetilde{\varphi}_n(y)\right\} = a_n^{-1}\left\{\varphi(y) - \widehat{\varphi}_n(y)\right\}\frac{a_n}{\widetilde{a}_n} + \Gamma(1-\widehat{\gamma}),$$

it is sufficient to prove that $\frac{a_n}{\widetilde{a}_n} \xrightarrow{\mathbb{P}} 1$. To this aim, remark that

$$\begin{aligned}\frac{a_n}{\widetilde{a}_n} &= -\frac{\varphi^2(y)\left\{\frac{1}{\varphi(y)} - U_y(n)\right\}}{\widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\{\widehat{\varphi}_n(y)\}^2} \\ &= \frac{\widehat{\gamma}}{\gamma}\frac{a\left(\frac{n}{k}\right)}{\widehat{a}\left(\frac{n}{k}\right)}k^{\gamma-\widehat{\gamma}}\frac{a(n)}{k^\gamma\, a\left(\frac{n}{k}\right)}\frac{-\gamma[U_y(\infty)-U_y(n)]}{a(n)}\left(\frac{\varphi(y)}{\widehat{\varphi}_n(y)}\right)^2.\end{aligned}$$

According to *e.g.* de Haan and Ferreira (2006, Theorem 4.2.1), we have $\widehat{a}\left(\frac{n}{k}\right)/a\left(\frac{n}{k}\right) \xrightarrow{\mathbb{P}} 1$ and using their Lemma 4.5.4 under (10), we have

$$\lim_{t\to\infty}\frac{\frac{U_y(\infty)-U_y(t)}{a(t)}+\frac{1}{\gamma}}{A(t)} = \frac{1}{\gamma(\rho+\gamma)}.$$

This convergence combined with the assumption that $A(t)\to 0$ lead to $\frac{-\gamma[U_y(\infty)-U_y(t)]}{a(t)} \to 1$. On the other hand, it follows from the representation $a(x) = x^\gamma \ell(x)$ that

$$\frac{a(n)}{a\left(\frac{n}{k}\right)k^\gamma} = \frac{n^\gamma\ell(n)}{k^\gamma\left(\frac{n}{k}\right)^\gamma\ell\left(\frac{n}{k}\right)} = \frac{\ell(n)}{\ell\left(\frac{n}{k}\right)} \to 1, \quad n\to\infty$$

by assumption. Finally, since $\sqrt{k}(\widehat{\gamma}-\gamma)$ converges to a normal distribution by Corollary 3.5.6 in de Haan and Ferreira (2006, p.108), we easily deduce that $k^{\widehat{\gamma}-\gamma} \xrightarrow{\mathbb{P}} 1$ as $n\to\infty$. This concludes the proof of Theorem 3. □

Note that the assumption $\frac{\ell(n)}{\ell(n/k)} \to 1$ in Theorem 3 is not very restrictive. It is in particular satisfied in case where ℓ is asymptotically a constant or a logarithm function under the additional condition that $\frac{\log k}{\log n} \to 0$.

Another option would be to use the estimator

$$\widetilde{\varphi}^*_n(y) := \left\{Z^y_{(n)} - \widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\Gamma(1-\widehat{\gamma})\right\}^{-1},$$

for which it is not hard to verify that Theorem 3 remains still valid when substituting $\widetilde{\varphi}^*_n(y)$ in place of $\widetilde{\varphi}_n(y)$. The motivation of this estimator is to correct the sample maximum for its inherent bias by considering first the hybrid estimator $\widetilde{F}^{\leftarrow}_{Z^y}(1) := Z^y_{(n)} + b_n\Gamma(1-\gamma)$ (see the proof of Theorem 1 for the limit distribution of $Z^y_{(n)}$ and its moment's convergence). Then we replace the unknown quantities γ and $b_n = a_n/\varphi^2(y)$ by $\widehat{\gamma}$ and $\widetilde{a}_n/\widehat{\varphi}^2_n(y)$ to get the endpoint estimator $1/\widetilde{\varphi}^*_n(y)$. This does not seem to have been appreciated in the literature of extreme values before. Even more generally, one can use the following extended versions

$$\begin{aligned}\widetilde{\varphi}^*_{n,m}(y) &:= \left\{Z^y_{(n-m)} - \widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\Gamma(1-\widehat{\gamma})\right\}^{-1}, \\ \widetilde{a}_{n,m} &:= -\widehat{a}\left(\frac{n}{k}\right)\frac{k^{\widehat{\gamma}}}{\widehat{\gamma}}\left\{Z^y_{(n-m)}\right\}^{-2}\end{aligned}$$

for some fixed integer $m\geq 0$. When $m=0$, we recover the estimators $\widetilde{\varphi}^*_n(y)$ and $\widetilde{a}_n$. Next, we show that $\widetilde{\varphi}^*_{n,m}(y)$ converges in distribution as well, with the scaling $\widetilde{a}_{n,m}$ to a different limit distribution.

Theorem 4. *Under the conditions of Theorem 3, we have for any $m \geq 0$,*

$$\widetilde{a}_{n,m}^{-1}\left\{\varphi(y)-\widetilde{\varphi}_{n,m}^{*}(y)\right\} \xrightarrow{d} G_{\gamma,m}+\Gamma(1-\gamma)$$

for the distribution function $G_{\gamma,m}(x)=G_{\gamma}(x)\sum_{i=0}^{m}\left(-\log G_{\gamma}(x)\right)^{i}/i!$.

Proof. To check this result, observe that $\widetilde{a}_{n,m}/\widetilde{a}_{n}=\left\{Z_{(n)}^{y}/Z_{(n-m)}^{y}\right\}^{2} \xrightarrow{\mathbb{P}} 1$, and that

$$\widetilde{a}_{n}^{-1}\left\{\widetilde{\varphi}_{n,m}^{*}(y)^{-1}-\varphi(y)^{-1}\right\}=\frac{\widetilde{a}_{n}^{-1}}{a_{n}^{-1}}\left\{a_{n}^{-1}\left(Z_{(n-m)}^{y}-\varphi(y)^{-1}\right)\right\}+\frac{\Gamma(1-\widehat{\gamma})}{\{\widehat{\varphi}_{n}(y)\}^{2}}.$$

By the proof of Theorem 3, we have $\widetilde{a}_{n}^{-1}/a_{n}^{-1} \xrightarrow{\mathbb{P}} 1$ and $\Gamma(1-\widehat{\gamma})/\{\widehat{\varphi}_{n}(y)\}^{2} \xrightarrow{\mathbb{P}} \Gamma(1-\gamma)/\varphi^{2}(y)$. On the other hand, we have $b_{n}^{-1}\left(\widehat{\varphi}_{n}(y)^{-1}-\varphi(y)^{-1}\right) \equiv b_{n}^{-1}\left(Z_{(n)}^{y}-\varphi(y)^{-1}\right) \xrightarrow{d} G_{\gamma}$ by the proof of Theorem 1 (i), with $b_{n}=a_{n}/\varphi^{2}(y)$. Then, by applying Theorem 21.18 in van der Vaart (1998, p.313), we obtain $b_{n}^{-1}\left(Z_{(n-m)}^{y}-\varphi(y)^{-1}\right) \xrightarrow{d} G_{\gamma,m}$, or equivalently $a_{n}^{-1}\left(Z_{(n-m)}^{y}-\varphi(y)^{-1}\right) \xrightarrow{d} G_{\gamma,m}/\varphi^{2}(y)$. Therefore $\widetilde{a}_{n}^{-1}\left\{\widetilde{\varphi}_{n,m}^{*}(y)^{-1}-\varphi(y)^{-1}\right\} \xrightarrow{d} [G_{\gamma,m}+\Gamma(1-\gamma)]/\varphi^{2}(y)$, which leads to the desired result. □

The motivation via the convergence of moments of the FDH estimator, the interpretation in terms of the involved gamma function, and the impact of the popular moment estimators $\widehat{\gamma}$ and $\widehat{a}(n/k)$ inspired the name Γ-*moment frontiers* for the class of estimators $\widetilde{\varphi}_{n,m}^{*}(y)$.

2.4 Optimizing the frontier estimation

In this section, we provide practical guidelines on how to pick out the intermediate sequences in $\widehat{\varphi}_{n}^{*}(y)$ and $\widetilde{\varphi}_{n,m}^{*}(y)$, and on how to ensure their monotonicity as functions of y.

2.4.1 Optimal choice of $k(n)$ in $\widehat{\varphi}_{n}^{*}(y)$ and $\widetilde{\varphi}_{n,m}^{*}(y)$

The accuracy of the moment estimator $\widehat{\varphi}_{n}^{*}(y) \equiv \widehat{\varphi}_{n}^{*}(y,k)$ depends on the choice of the sequence $k=k(n)$ in (11) for which the approximation (8) is believed to be valid. A promising bootstrap technique by Ferreira *et al.* (2003) to achieve the optimal value of k, for the endpoint estimator $\widehat{\varphi}_{n}^{*}(y)^{-1}$, is a result of balancing variance and bias components. We can employ this bootstrap-based procedure to solve our optimality problem adaptively. Note that what is important in the definition of the frontier function $\widehat{\varphi}_{n}^{*}(y,k)$ is not the sample size n itself, as it is the case in the set-up of Ferreira *et al.* (2003) for endpoint estimation, but only the number $N_{y}:=\sum_{i=1}^{n}\mathbb{1}_{\{Y_{i}\geq y\}}$ of observations (Y_{i},X_{i}) for which $Y_{i}\geq y$ or equivalently $Z_{i}^{y}>0$. Note also that, by construction, the number k ranges from 1 to $N_{y}-1$.

We shall determine the value of k that minimizes the asymptotic mean-squared error:

$$k_{y}(n):=\underset{k}{\operatorname{argmin}}\ \text{asympt. } \mathbb{E}\left(\widehat{\varphi}_{n}^{*}(y,k)-\varphi(y)\right)^{2}. \tag{13}$$

The adaptive method for optimization consists then in replacing the unknown theoretical quantities in (13) with their empirical analogues leading thus to the objective function

$$\mathbb{E}_n \left(\widehat{\varphi}^*_{n,1}(y,k) - \widehat{\varphi}^*_{n,2}(y,k) \right)^2,$$

where $\mathbb{E}_n$ stands for averaging with respect to the empirical distribution function, and

$$\widehat{\varphi}^*_{n,i}(y,k) := \left\{ Z^y_{(n-k)} - \widehat{a}_i \left(\frac{n}{k} \right) / \widehat{\gamma}_i \right\}^{-1} \quad \text{for} \quad i = 1, 2, \tag{14}$$

with $\widehat{\gamma}_1 = \widehat{\gamma}$, $\widehat{a}_1 \left(\frac{n}{k} \right) = \widehat{a} \left(\frac{n}{k} \right)$, $\widehat{\gamma}_2$ and $\widehat{a}_2 \left(\frac{n}{k} \right)$ being the following alternative estimators

$$\widehat{\gamma}_2 := \left(M^{(2)}_n / 2 \right)^{1/2} + 1 - \frac{2}{3} \left(1 - M^{(1)}_n M^{(2)}_n / M^{(3)}_n \right)^{-1},$$
$$\widehat{a}_2 \left(n/k \right) := Z^y_{(n-k)} \, M^{(1)}_n \, (1 - \widehat{\gamma}_2 + M^{(1)}_n).$$

For a fixed y such that $N_y > 2$ is large enough, the guidelines on how to estimate the optimal value $k_y(n)$ in (13) are as follows:

Step 1 Form the transformed set of dimensionless observations $\{Z^y_1, \ldots, Z^y_n\}$ from the multivariate sample $\{(Y_1, X_1), \ldots, (Y_n, X_n)\}$, and extract the subset $\{Z^y_{(n-N_y+1)}, \ldots, Z^y_{(n)}\}$ of non-null Z^y_i's.

Step 2 Select randomly and independently n_y times a member from $\{Z^y_{(n-N_y+1)}, \ldots, Z^y_{(n)}\}$, where $n_y = O(N_y^{1-\varepsilon})$ for some $\varepsilon \in (0, 1/2)$. Indicate by $Z^{y*}_{(1)} \leq \cdots \leq Z^{y*}_{(n_y)}$ the ordered selected members and calculate the corresponding estimates $\widehat{\gamma}^*_1$, $\widehat{\gamma}^*_2$, $\widehat{a}^*_1 \left(\frac{n_y}{k} \right)$, $\widehat{a}^*_2 \left(\frac{n_y}{k} \right)$, $\widehat{\varphi}^*_{n_y,1}(y,k)$ and $\widehat{\varphi}^*_{n_y,2}(y,k)$, for $k = 1, \ldots, n_y - 1$. Then form the quantities

$$Q^*_{n_y}(k) := \left(\widehat{\varphi}^*_{n_y,1}(y,k) - \widehat{\varphi}^*_{n_y,2}(y,k) \right)^2 \quad \text{for} \quad k = 1, \ldots, n_y - 1. \tag{15}$$

Step 3 Repeat step 2 r_y times independently. The number r_y can be taken as big as necessary. Indicate the result by $Q^*_{n_y,1}(k), \cdots, Q^*_{n_y,r_y}(k)$ and compute the average values $\bar{Q}^*_{n_y}(k) = (1/r_y) \sum_{j=1}^{r_y} Q^*_{n_y,j}(k)$, for $k = 1, \ldots, n_y - 1$.

Step 4 Determine the minimizer $\hat{k}^*_n(y)$ of $\bar{Q}^*_{n_y}(k)$ with respect to k over the range of intermediate sequences, say, from $\log n_y$ to $N_y / \log n_y$ (this restriction allows to reject too small values or those very near to n_y, assuming n_y sufficiently large).

Remark 1. In the range $\gamma > -1/2$, which is most frequent in applications, one can get a more refined estimate than $\hat{k}^*_n(y)$, which is asymptotically as good as the theoretical number $k_y(n)$ in (13) thanks to a second bootstrap. We refer to Ferreira *et al.* (2003) for a much more thorough discussion of the rationale for this approach including proofs of its asymptotic optimality.

Remark 2. In what concerns the bootstrap parameters, we used in all our simulations and applications the same considerations as in Ferreira *et al.* (2003) on endpoint estimation, for each fixed y.

Evidence has been given in their Monte Carlo experiments to support that the bootstrap moment estimates are quite stable along n_y. Then they always considered $n_y = N_y^{1-\varepsilon}$ where $\varepsilon = 0.1$ (as mentioned above, the sample size in the set-up of Ferreira *et al.* (2003) corresponds to N_y in our context). They also have recommanded to reject values of $\hat{k}_n^*(y)$ which are smaller than 10 and larger than $0.8n_y$.

Remark 3. It seems that the bootstrap-based procedure affords satisfactory results only for values of y such that N_y exceeds, approximately, 2000 (see the conclusions of Ferreira *et al.*, 2003, Section 2.3.1). For small and moderate samples, as it is the case in our applications, the optimal choice of $k = k_y$ via the bootstrap method is still possible, but is hard to manage for certain values of y following the slope and the curvature of the frontier. This happens when the tail assumption (3) or the approximation (8) is too optimistic. Stated differently, it is in just those values of y that $\widehat{\gamma} \geq 0$ can appear, and this failure leads to a severe bias. To reduce this vexing defect, we used in our numerical illustrations the variant of the moment estimator described in the next proposition instead of (14).

Proposition 1. *If* (3) *holds and* $k = k(n)$ *satisfies* $k/n \to 0$ *with* $k/(\log n)^\delta \to \infty$ *for some* $\delta > 0$, *then we have with probability* 1, *as* $n \to \infty$,

$$\widehat{\varphi}_n^*(y) \equiv \left\{ Z^y_{(n-k)} - \frac{\widehat{a}\left(\frac{n}{k}\right)}{\widehat{\gamma}} \mathbb{1}_{\{\widehat{\gamma}<0\}} \right\}^{-1}, \quad \widetilde{\varphi}_{n,m}^*(y) \equiv \left\{ Z^y_{(n-m)} - \widehat{a}\left(\frac{n}{k}\right) \frac{k^{\widehat{\gamma}}}{\widehat{\gamma}} \Gamma(1-\widehat{\gamma}) \mathbb{1}_{\{\widehat{\gamma}<0\}} \right\}^{-1}.$$

Proof. By Theorem 2.1 of Dekkers *et al.* (1989, p.1834), we have $\widehat{\gamma}_M \xrightarrow{a.s.} \gamma$ as $n \to \infty$. It is also shown in the proof of that theorem (see Equation (2.13), p.1840) that $M_n^{(1)} \xrightarrow{a.s.} \max(0, \gamma) = 0$ as $n \to \infty$. Then $\widehat{\gamma} = \widehat{\gamma}_M - M_n^{(1)} \xrightarrow{a.s.} \gamma$ as $n \to \infty$. Since $\gamma < 0$, we get $\mathbb{1}_{\{\widehat{\gamma}<0\}} = 1$ as $n \to \infty$, with probability 1. □

For computing the number k in the Γ-moment estimator $\widetilde{\varphi}_{n,m}^*(y) = \widetilde{\varphi}_{n,m}^*(y,k)$, we just apply the same scheme as above (step 1 up to 4) by proceeding to step 2 with

$$\begin{aligned} Q_{n_y}^*(k) &:= \left(\widetilde{\varphi}_{n_y,m,1}^*(y,k) - \widetilde{\varphi}_{n_y,m,2}^*(y,k) \right)^2, \\ \widetilde{\varphi}_{n_y,m,i}^*(y,k) &:= \left\{ Z^{y*}_{(n_y-m)} - \widehat{a}_i^* \left(\frac{n_y}{k}\right) \frac{k^{\widehat{\gamma}_i^*}}{\widehat{\gamma}_i^*} \Gamma(1-\widehat{\gamma}_i^*) \mathbb{1}_{\{\widehat{\gamma}_i^*<0\}} \right\}^{-1} \end{aligned}$$

for $i = 1, 2$. As demonstrated in Section 3, this bootstrap procedure provided quite admirable estimates $\widetilde{\varphi}_{n,m}^*(y)$ in terms of both bias and mean-squared error. Even more strongly, the results for $\widetilde{\varphi}_{n,m}^*(y)$ were appreciably better than those for the moment estimator $\widehat{\varphi}_n^*(y)$. So we do not enter here into further theoretic validation of the method.

2.4.2 Isotonized frontier estimators

Yet, there is still another difficulty with our 'pointwise' selection method of $k = k_y$ as y varies. A lack of smoothness of the resulting estimators $\widehat{\varphi}^*_n(y)$ and $\widetilde{\varphi}^*_{n,m}(y)$, as functions of y, can occur for small and moderate samples. Then, these estimators may not automatically inherit the monotonicity property of the true frontier φ. One way to monotonize and stabilize each of these unconstrained estimators is by using the isotonic version $\bar{\varphi}^{\#}_n = [\bar{\varphi}^*_{1n} + \bar{\varphi}^*_{2n}]/2$, with

$$\bar{\varphi}^*_{1n}(y) = \sup_{y' \leq y} \bar{\varphi}^*_n(y') \quad \text{and} \quad \bar{\varphi}^*_{2n}(y) = \inf_{y \leq y'} \bar{\varphi}^*_n(y'), \tag{16}$$

where y and y' run over some domain $\mathbb{D}$ interior to the support of Y and $\bar{\varphi}^*_n$ is either $\widehat{\varphi}^*_n$ or $\widetilde{\varphi}^*_{n,m}$.

Both $\bar{\varphi}^*_{1n}$ and $\bar{\varphi}^*_{2n}$ are monotone nondecreasing on $\mathbb{D}$ with respect to the partial order: $\bar{\varphi}^*_{1n}$ is the smallest monotone function that lies above the unconstrained estimator $\bar{\varphi}^*_n$, and $\bar{\varphi}^*_{2n}$ is the largest monotone function that lies below $\bar{\varphi}^*_n$. As a matter of fact, any convex combination of these envelope estimators would have sufficed as a definition of $\bar{\varphi}^{\#}_n$, but we do not see any reason to bias the restricted estimator one way or the other. Next, we show that the hybrid estimators $\widehat{\varphi}^{\#}_n$ and $\widetilde{\varphi}^{\#}_{n,m}$ are better than the original versions $\widehat{\varphi}^*_n$ and $\widetilde{\varphi}^*_{n,m}$ in the following sense:

$$\sup_{y \in \mathbb{D}} |\widehat{\varphi}^{\#}_n(y) - \varphi(y)| \leq \sup_{y \in \mathbb{D}} |\widehat{\varphi}^*_n(y) - \varphi(y)|, \quad \sup_{y \in \mathbb{D}} |\widetilde{\varphi}^{\#}_{n,m}(y) - \varphi(y)| \leq \sup_{y \in \mathbb{D}} |\widetilde{\varphi}^*_{n,m}(y) - \varphi(y)|.$$

Indeed, using the triangle inequality of the sup-norm, it is easily seen that the # operator is sup-norm contracting in the sense that $\sup_{y \in \mathbb{D}} |r^{\#}(y) - s^{\#}(y)| \leq \sup_{y \in \mathbb{D}} |r(y) - s(y)|$, for any functions $r(\cdot)$ and $s(\cdot)$ defined on $\mathbb{D}$ (see, *e.g.*, Lemma 3.1 in Daouia and Simar, 2005).

In order to compute the restricted estimators $\widehat{\varphi}^{\#}_n(y)$ and $\widetilde{\varphi}^{\#}_{n,m}(y)$, one can use in practice a discrete grid $\mathbb{D}_n$ in place of the domain $\mathbb{D}$ in the definition (16) of the envelope estimators. In the general multivariate case where $y \in \mathbb{R}^q_+$, the idea is to first consider the minimal rectangular set with edges parallel to the coordinate axes covering all the observations Y_i, and then to choose a discrete grid $\mathbb{D}_n$ in this rectangular set including the minimal and maximal points (with respect to the partial order induced by "$\leq$") of the rectangular set. In our simulation study we confine ourselves to a bi-dimensional support of $(Y, X) \in \mathbb{R}^2_+$, where we used n grid points evenly distributed across the entire sample space of Y_i's.

3 Some simulation evidence

We have undertaken some Monte Carlo experiments to evaluate finite-sample performance of the FDH, moment and Γ-moment frontier estimators. The experiments employ the two different frontier functions

$$\varphi^1(y) = y \quad \text{and} \quad \varphi^2(y) = \exp(-5 + 10y)/(1 + \exp(-5 + 10y)).$$

Design points, Y_i, are generated as $U[0,1]$, and responses as

$$X_i = \varphi(Y_i) + \sigma(Y_i)\, V_i, \quad i = 1, \ldots, n,$$

where the lower support boundary φ is either φ^1 or φ^2, the local scale factor $\sigma(y) = (1+y)/2$ is linearly increasing in y, and the V_i's given $Y_i = y$ are independent Beta$(b(y), 1)$, with the Beta parameter being either constant, $b(y) = 1$, or varying above 1 as follows

$$b(y) = 2\left\{\left(\frac{1}{10} + \sin(\pi y)\right)\left(\frac{11}{10} - \frac{1}{2}\exp\left[-64\left(y - \frac{1}{2}\right)^2\right]\right)\right\}^{-1} - 1.$$

In this model, the shape parameter of the joint density β_y and the extreme-value index (EVI) γ_y are given by

$$\beta_y = b(y) - 1 \quad \text{and} \quad \gamma_y = -1/(b(y)+1).$$

All the experiments were performed over 1000 independent samples of size $n = 200$. For each simulated sample, two measures of performance were considered

$$\mathrm{MSE}(\varphi_n) = \frac{1}{L}\sum_{\ell=1}^{L}\{\varphi_n(y_\ell) - \varphi(y_\ell)\}^2, \;\; \mathrm{Bias}(\varphi_n) = \frac{1}{L}\sum_{\ell=1}^{L}\{\varphi_n(y_\ell) - \varphi(y_\ell)\}$$

for $\varphi_n = \widehat{\varphi}_n$, $\widehat{\varphi}_n^*$, $\widehat{\varphi}_n^\#$, $\widetilde{\varphi}_{n,m}^*$, $\widetilde{\varphi}_{n,m}^\#$, where only the most extreme values $m \in \{0,1\}$ were considered to save place, with $y_1 \ldots, y_L$ being 100 points evenly distributed between $Y_{(1)} = \min_{1\le i\le n} Y_i$ and the order statistic $Y_{(n-17)}$ (this choice of the reference points y_ℓ ensures that $N_{y_\ell} \ge 18$ for each ℓ, which in turn allows to use the value 10 (respectively, $0.8n_y$) as a lower (respectively, upper) bound of k in the bootstrap-based method, as mentioned in Remark 2).

In what concerns the number of bootstrap resamples (denoted by r_y in step 3), 100 replications seem fairly enough given that $n = 200$.

To guarantee a fair comparison among the different estimators, our Monte Carlo experiments were first devoted to their accuracy when the oracle quantity

$$Q_{n_y}^*(k) := \big(\varphi_{n_y}(y,k) - \varphi(y)\big)^2, \quad \text{for} \quad \varphi_{n_y} = \widehat{\varphi}_{n_y}^*, \widetilde{\varphi}_{n_y,m}^*, \tag{17}$$

is used in step 2 instead of (15). In Table 1 (respectively, Table 2) we report the Monte Carlo averages of the MSE and the bias, computed over the 1000 replications of the experiment, for the linear (respectively, non-linear) boundary φ in the Beta error model. These results give an overall impression of the precision of the different estimators:

When $\gamma = -1/2$ or equivalently $b(\cdot) = 1$, it may be seen that the FDH estimator $\widehat{\varphi}_n$ outperforms overall both the moment estimator $\widehat{\varphi}_n^*$ and its monotonized version $\widehat{\varphi}_n^\#$ in terms of MSE, whereas there is no winner in terms of bias in all cases. In contrast, the Γ-moment estimators $\widetilde{\varphi}_{n,0}^*$ and $\widetilde{\varphi}_{n,0}^\#$ perform overall clearly better than the FDH and moment estimators in terms of both bias and MSE.

Table 1: *Results for* $\varphi = \varphi^1$ *using the oracle quantity* (17) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 200$.

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.0074	0.0181	0.0106	0.0062	0.0047	0.0099	0.0067
$\gamma > -\frac{1}{2}$	0.0421	0.0706	0.0393	0.0315	0.0244	0.0447	0.0278

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.0757	0.0656	0.0409	0.0237	0.0051	0.0520	0.0255
$\gamma > -\frac{1}{2}$	0.1639	0.1392	0.0610	0.0687	-0.0201	0.1061	0.0060

Table 2: *Results for* $\varphi = \varphi^2$ *using the oracle quantity* (17) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 200$.

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.0081	0.0260	0.0156	0.0071	0.0050	0.0114	0.0066
$\gamma > -\frac{1}{2}$	0.0397	0.0788	0.0474	0.0315	0.0262	0.0453	0.0296

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.0734	0.0774	0.0738	0.0238	0.0033	0.0489	0.0223
$\gamma > -\frac{1}{2}$	0.1587	0.1519	0.0945	0.0695	-0.0157	0.1117	0.0168

Table 3: *Results for* $\varphi = \varphi^1$ *using the oracle quantity* (17) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 201$ *(one outlier included).*

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.1046	0.1572	0.0733	0.0990	0.0987	0.0418	0.0273
$\gamma > -\frac{1}{2}$	0.1046	0.1572	0.0733	0.0990	0.0987	0.0092	0.0071

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	-0.1251	0.2846	0.1944	-0.1703	-0.1966	0.1015	0.0093
$\gamma > -\frac{1}{2}$	-0.1251	0.2846	0.1944	-0.1703	-0.1966	0.0565	0.0340

Table 4: *Results for* $\varphi = \varphi^2$ *using the oracle quantity* (17) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 201$ *(one outlier included).*

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	0.1240	0.0604	0.0474	0.1250	0.1247	0.0091	0.0073
$\gamma > -\frac{1}{2}$	0.1459	0.1612	0.0811	0.1410	0.1399	0.0411	0.0294

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}_n^*$	$\widehat{\varphi}_n^\#$	$\widetilde{\varphi}_{n,0}^*$	$\widetilde{\varphi}_{n,0}^\#$	$\widetilde{\varphi}_{n,1}^*$	$\widetilde{\varphi}_{n,1}^\#$
$\gamma = -\frac{1}{2}$	-0.1896	0.1699	0.1511	-0.2085	-0.2183	0.0560	0.0336
$\gamma > -\frac{1}{2}$	-0.1348	0.3010	0.1905	-0.1745	-0.1982	0.1247	0.0356

When $\gamma > -1/2$ or equivalently $b(\cdot) > 1$, the FDH estimator $\widehat{\varphi}_n$ seems to outperform the moment estimator $\widehat{\varphi}_n^*$ in terms of MSE in all cases, which is not the case for the monotonized version $\widehat{\varphi}_n^\#$. However, both moment estimators $\widehat{\varphi}_n^*$ and $\widehat{\varphi}_n^\#$ seem to be overall superior to the FDH estimator in terms of bias. In contrast, the Γ-moment estimators $\widetilde{\varphi}_{n,0}^*$ and $\widetilde{\varphi}_{n,0}^\#$ have uniformly smaller MSE and bias than the FDH estimator. Moreover, we can see that the monotonized

version $\widetilde{\varphi}^{\#}_{n,0}$ is clearly the winner in all respects and all cases compared to the standard FDH and moment estimators.

Table 5: *Results for* $\varphi = \varphi^1$ *using the bootstrap quantity* (15) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 200$.

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.0075	0.0215	0.0199	0.0055	0.0046	0.0091	0.0074
$\gamma > -\frac{1}{2}$	0.0428	0.0836	0.0471	0.0332	0.0266	0.0475	0.0320

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.0760	0.1108	0.1058	0.0299	0.0170	0.0614	0.0475
$\gamma > -\frac{1}{2}$	0.1657	0.1897	0.0918	0.0817	-0.0033	0.1247	0.0282

Table 6: *Results for* $\varphi = \varphi^2$ *using the bootstrap quantity* (15) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 200$.

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.0080	0.0328	0.0324	0.0063	0.0053	0.0099	0.0075
$\gamma > -\frac{1}{2}$	0.0401	0.0927	0.0623	0.0330	0.0278	0.0466	0.0319

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.0730	0.1113	0.1276	0.0270	0.0107	0.0557	0.0376
$\gamma > -\frac{1}{2}$	0.1587	0.1951	0.1305	0.0738	-0.0081	0.1195	0.0274

Table 7: *Results for* $\varphi = \varphi^1$ *using the bootstrap quantity* (15) *in step 2 and* 1000 *Monte-Carlo simulations with* $n = 201$ *(one outlier included).*

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.0808	0.0970	0.0678	0.0850	0.0867	0.0092	0.0079
$\gamma > -\frac{1}{2}$	0.1043	0.2314	0.0934	0.1013	0.1005	0.0417	0.0277

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	-0.1753	0.2258	0.2003	-0.2010	-0.2162	0.0452	0.0238
$\gamma > -\frac{1}{2}$	-0.1262	0.3741	0.2315	-0.1681	-0.1993	0.1197	0.0307

Table 8: *Results for* $\varphi = \varphi^2$ *using the bootstrap quantity* (15) *in step 2 and* 1000 *Monte-Carlo simulations* $n = 201$ *(one outlier included).*

MSE

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	0.1238	0.1294	0.0892	0.1251	0.1247	0.0088	0.0072
$\gamma > -\frac{1}{2}$	0.1462	0.2610	0.1162	0.1417	0.1392	0.0418	0.0307

Bias

EVI	$\widehat{\varphi}_n$	$\widehat{\varphi}^{*}_n$	$\widehat{\varphi}^{\#}_n$	$\widetilde{\varphi}^{*}_{n,0}$	$\widetilde{\varphi}^{\#}_{n,0}$	$\widetilde{\varphi}^{*}_{n,1}$	$\widetilde{\varphi}^{\#}_{n,1}$
$\gamma = -\frac{1}{2}$	-0.1897	0.2600	0.2359	-0.2092	-0.2179	0.0578	0.0388
$\gamma > -\frac{1}{2}$	-0.1346	0.4079	0.2679	-0.1746	-0.1972	0.1333	0.0447

It may be also seen that the behavior of the Γ-moment frontier $\widetilde{\varphi}^{\#}_{n,1}$ is quite respectable, but $\widetilde{\varphi}^{\#}_{n,0}$ appears to behave better in almost all cases. However, when the data are contaminated by adding outliers sufficiently far from the true frontier, the bias-corrected estimator $\widetilde{\varphi}^{\#}_{n,0}$ shares a

serious defect with the FDH estimator: even a single outlying observation (chosen at the point $(0.8, 0.2)$ and indicated by '*' in Figures 1-4) causes the estimators to take values arbitrarily far from their values at the initial samples, as can be seen from Tables 3 and 4. Only the Γ-moment estimators with $m = 1$, especially the monotonized version $\widetilde{\varphi}^{\#}_{n,1}$, do appear to be more resistant and appreciably better than all the other extreme-value estimators.

Turning to the comparison of the performance of the frontier estimators when the quantity $Q^*_{n_y}(k)$ in (15) itself is used in step 2 of the bootstrap algorithm, the Monte Carlo estimates shown in Tables 5-8 exhibit qualitatively similar results to those that we observed upon earlier for the oracle quantity (17).

A typical realization of the experiment in each scenario is shown in Figures 1-4. The top panel (three pictures) corresponds to the linear boundary φ^1, while the bottom panel corresponds to φ^2. In each panel, we see the unconstrained estimators $\widehat{\varphi}^*_n$ (left), $\widetilde{\varphi}^*_{n,0}$ (middle) and $\widetilde{\varphi}^*_{n,1}$ (right) as the lighter blue curves, while the darker blue curves are the monotonized versions $\widehat{\varphi}^{\#}_n$, $\widetilde{\varphi}^{\#}_{n,0}$ and $\widetilde{\varphi}^{\#}_{n,1}$. In each picture, we see the sample observations as points, the true lower boundary as the green curve and the FDH estimator as the red curve. As is to be expected, Figures 1 and 2 show that

- the unconstrained estimators in the lighter blue curves exhibit some unstability due to their pointwise construction and to the small sample size;
- the restricted estimators in the darker blue curves reduce considerably the unsmoothness of the original versions;
- the standard moment estimator $\widehat{\varphi}^*_n$ (left pictures) disappoints by its severe volatility: good results may require a large sample size of the order of several thousands (see also Remark 3). The monotonized version $\widehat{\varphi}^{\#}_n$ provides more stable estimates, but may offset by its rather large bias;
- the variability and the bias of the Γ-moment estimators $\widetilde{\varphi}^{\#}_{n,m}$ are quite respectable with an added advantage for $\widetilde{\varphi}^{\#}_{n,0}$ (middle pictures) in terms of bias.

The comparison of the different estimators graphed in Figures 3 and 4 illustrates the resistance of the Γ-moment frontiers with $m = 1$ and the fragility of the other frontier estimators in withstanding the influence of the outlying observation indicated by '*'. We repeated the same exercise with different and more outliers and obtained similar results.

Our tentative conclusion is to favor in practice the use of the Γ-moment estimator $\widetilde{\varphi}^{\#}_{n,m}$ with $m = 0$ when the model is nearly correct, and with $m > 0$ otherwise. However, with real data the question: *"Are isolated extreme observations anomalous data or is the density very flat near the boundary?"* is a tedious matter. Figure 6 with the data from the French post offices illustrates exactly the problem. A diagnostic tool which allows to select adequate values for m in this case is proposed below.

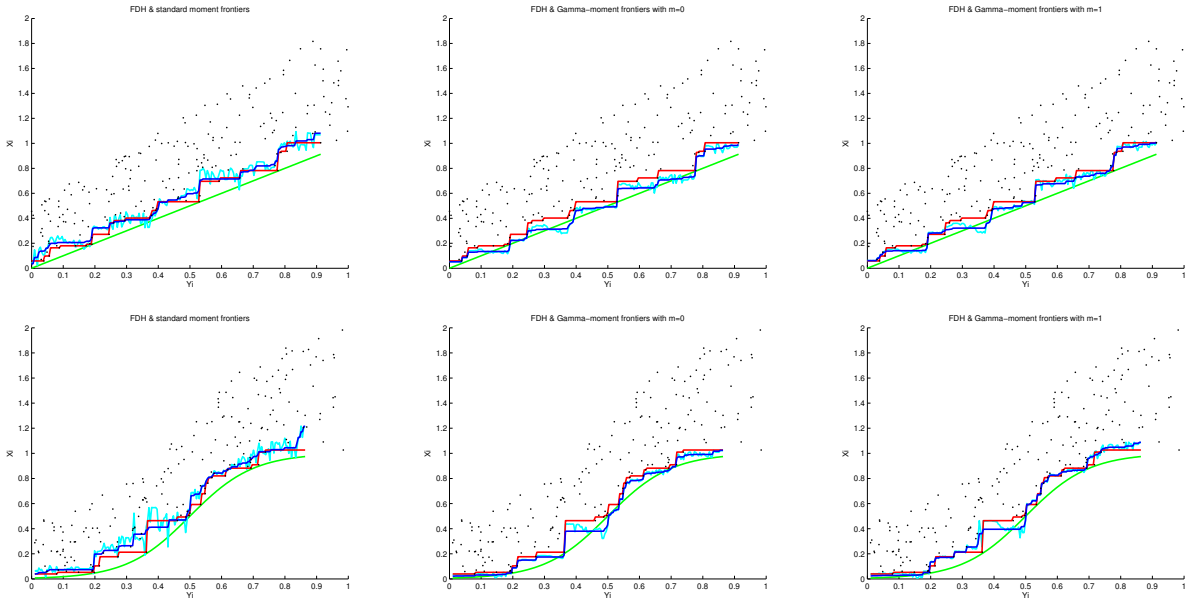

Figure 1: *Results for* $\gamma_y = -1/2$ *and* $n = 200$ *using the bootstrap quantity* (15) *in step 2.*

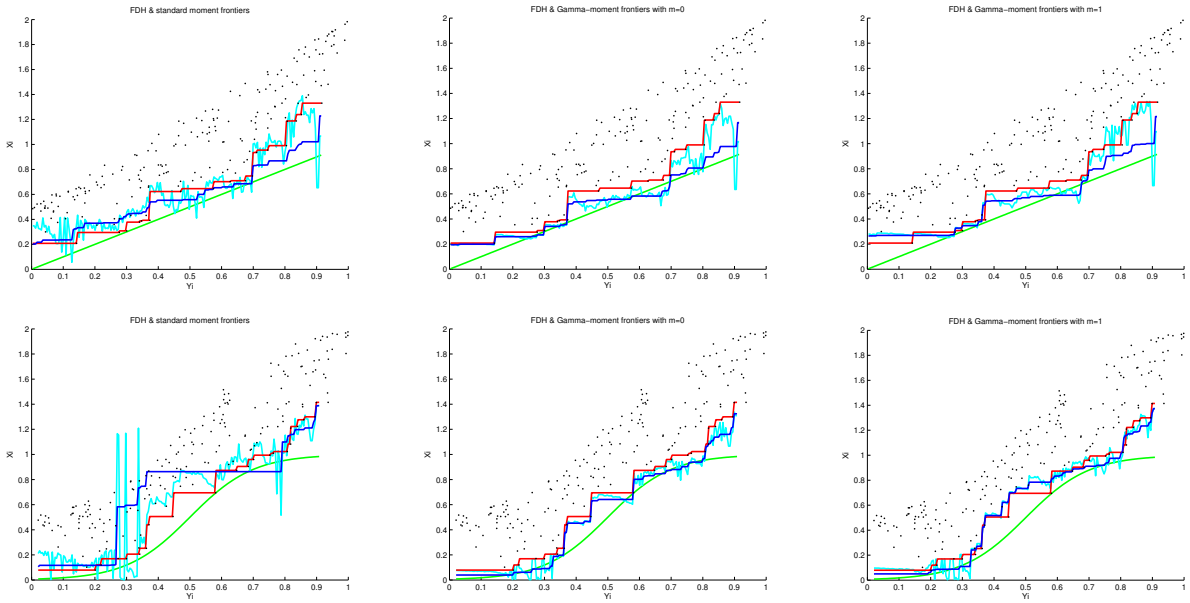

Figure 2: *Results for* $\gamma_y > -1/2$ *and* $n = 200$ *using the bootstrap quantity* (15) *in step 2.*

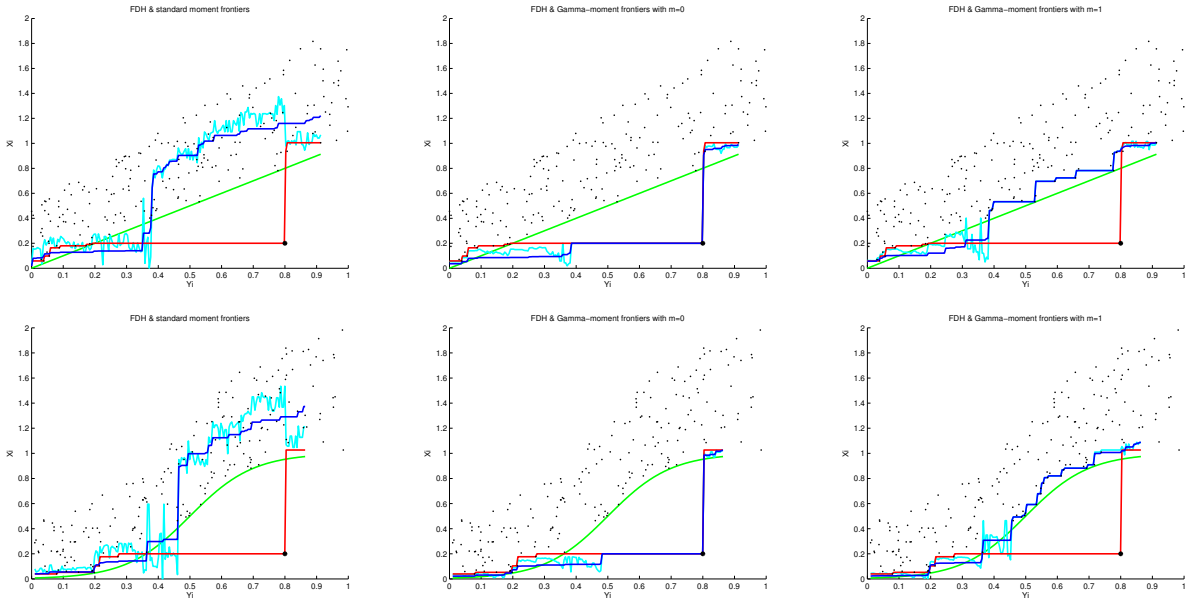

Figure 3: *Results for* $\gamma_y = -1/2$ *and* $n = 201$ *(as in Figure 1 with one outlier included)*

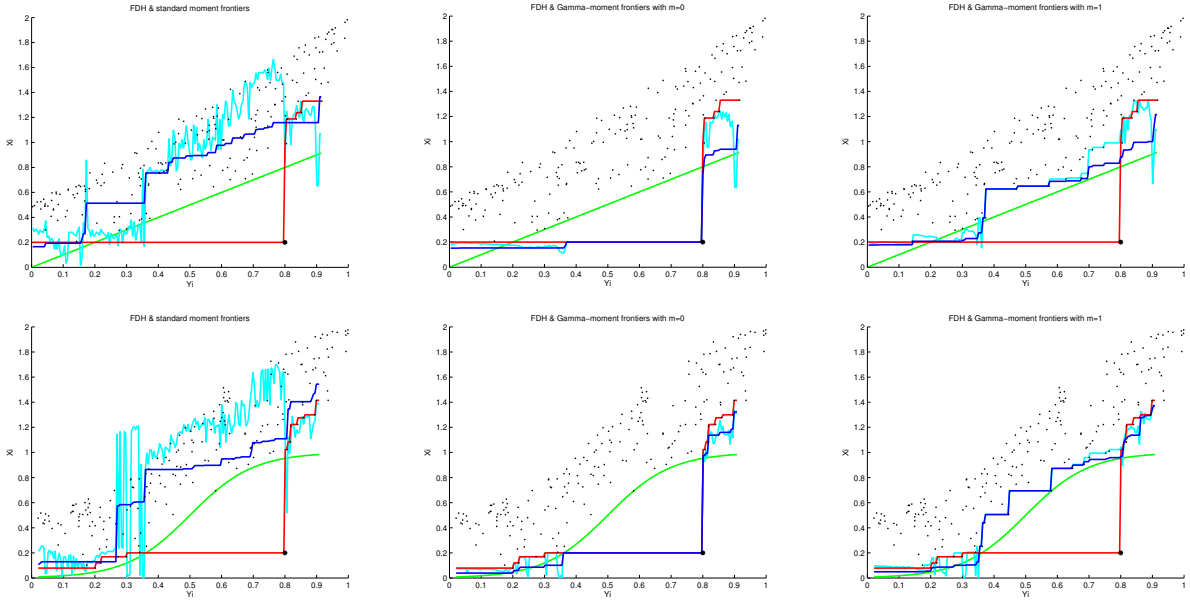

Figure 4: *Results for* $\gamma_y > -1/2$ *and* $n = 201$ *(as in Figure 2 with one outlier included).*

Trade-offs Bias-Robustness: If the density of data near the support boundary is low, then the few observations one is likely to observe near or at the sample boundary are quite valuable and dispensing with them (*i.e.,* letting them be outside the frontier estimator) could be costly. In this case, it would be more reasonable to use the bias-corrected estimator which corresponds to $m = 0$. In contrast, if the practitioner thinks there is a small probability of any observation being outlying or mis-recorded, it would be unrealistic to use the frontier estimator $\widetilde{\varphi}^{\#}_{n,0}$ which envelops all the data points. In this case, robustification via the choice of a 'trimming' order $m \neq 0$ is very important.

In absence of information on whether the data are measured accurately, one way to choose an appropriate number m is by looking to the evolution of the 'Euclidean' distance $D(m)$ between $\widetilde{\varphi}^{\#}_{n,m}$ and the non-robust FDH frontier $\widehat{\varphi}_n$ as m varies, where

$$D(m) := \sqrt{\sum_{\ell} \left(\widehat{\varphi}_n(y_\ell) - \widetilde{\varphi}^{\#}_{n,m}(y_\ell)\right)^2},$$

with the y_ℓ's being evenly distributed between the order statistics $Y_{(1)}$ and $Y_{(n-17)}$. One may distinguish between two possible scenarios:

- In presence of outliers sufficiently isolated from the sample, the curve should show a severe increasing jump from some value m_1, followed by a stable evolution from a larger value m_2. In other words, the frontiers $\widetilde{\varphi}^{\#}_{n,m}$ with $m \leq m_1$ are expected to lie very near to the FDH frontier so that they are drastically influenced by the outlying points, whereas those with $m > m_1$ would be more resistant, especially those with $m \geq m_2$. In this case, it suffices to pick up a value $m \in (m_1, m_2]$ to avoid systematic underestimation when $m \leq m_1$ and possible overestimation when $m > m_2$.
- In absence of influential outliers, no severe increasing jump followed by a stable evolution would appear. In this case, it is most efficient to use $\widetilde{\varphi}^{\#}_{n,0}$.

For the samples used in Figures 1-2 (respectively, Figures 3-4), where each row of 3 pictures corresponds to the same sample, we obtain the top (respectively, bottom) evolution curves displayed in Figure 5. As expected the top panels, which correspond to the 4 uncontaminated samples, suggest to select the value $m = 0$ since there is no severe increasing jump followed by stable oscillations. The first 3 bottom panels (from left to right) indicate an indisputable sharp positive slope of the graph of $D(m)$ at $m = 0$, followed by smooth oscillations from $m = 1$, which favors the use of $\widetilde{\varphi}^{\#}_{n,1}$ to estimate the boundary in the corresponding contaminated samples. Looking to the last bottom panel, the evolution of $D(m)$ becomes stable from $m = 2$, and so it is preferable in this case to use either $\widetilde{\varphi}^{\#}_{n,1}$ or $\widetilde{\varphi}^{\#}_{n,2}$.

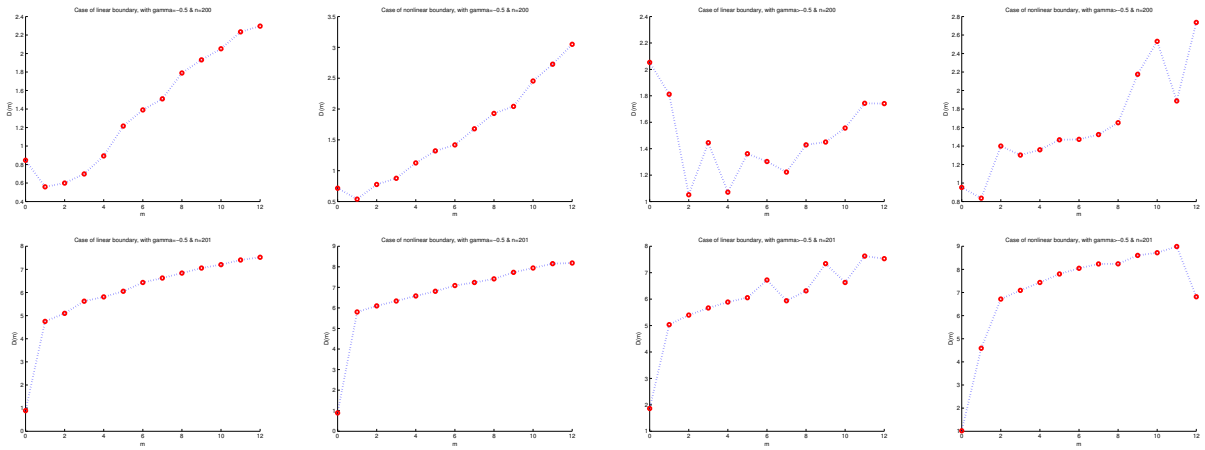

Figure 5: *Evolution of $D(m)$ with m: the top (bottom) panels correspond to $n = 200$ ($n = 201$). From left to right:* $(\varphi = \varphi^1, \gamma = -1/2)$, $(\varphi = \varphi^2, \gamma = -1/2)$, $(\varphi = \varphi^1, \gamma > -1/2)$ *and* $(\varphi = \varphi^2, \gamma > -1/2)$.

4 Applications

This section provides results from production and nuclear data examples where one would expect to find monotonic optimal boundaries.

4.1 Estimation of the minimal cost in production activity

To illustrate our methodology on a large dataset, we analyzed the cost of the delivery activity of the postal services in France (Cazals *et al.*, 2002). For each post office $i = 1, \ldots, 5138$, we have the quantity of labor X_i which represents more than 80% of the total cost of the delivery activity. The volume of the delivered mail defines the output Y_i. Figure 6 (top) plots the observed data, the cost X_i (vertical axis) against the output Y_i (horizontal axis), along with the monotonic estimators $\widehat{\varphi}_n$ in red curve and $\widehat{\varphi}_n^{\#}$ in green curve. The diagnostic graph in Figure 7 (l-h.s) shows a severe increasing jump from the value $m_1 = 2$, followed by stable oscillations from the value $m_2 = 6$. A sensible practice would be then to select the order m in $\widetilde{\varphi}_{n,m}^{\#}$ between the values 3 and 6. Both $\widetilde{\varphi}_{n,3}^{\#}$ and $\widetilde{\varphi}_{n,6}^{\#}$ were superimposed in Figure 6 (top) in darker and lighter blue curves, respectively. As in the simulations discussed previously, we used $n_y = [N_y^{0.9}]$ (see also Remark 2). For the number of bootstrap resamples, $r_y = 500$ seems fairly enough in this case. The range taken to look for the optimal $\hat{k}_n^*(y)$ is the same as in simulations, namely 10 for the lower bound and $[0.8n_y]$ for the upper bound.

It appears that the standard moment estimator $\widehat{\varphi}_n^{\#}$ is too 'conservative' because of its severe robustness to extreme observations, whereas the Γ-moment estimators $\widetilde{\varphi}_{n,m}^{\#}$ are more 'liberal' in the sense that they are sensitive to the magnitude of valuable extreme post offices but, in the same time, they remain resistant to the influence of some suspicious isolated observations: The points

left outside the frontier $\widetilde{\varphi}^{\#}_{n,6}$ or $\widetilde{\varphi}^{\#}_{n,3}$ look so extreme that they seem hardly related to the sample and should be then analyzed carefully because they could be outlying or perturbed by noise. On the other hand, it might also be seen that both the Γ-moment frontiers look like the largest convex minorant of the standard moment estimator $\widehat{\varphi}^{\#}_{n}$, exhibiting thus non-increasing returns to scale; *i.e.*, they do not allow the volume of delivered mail to increase faster than the quantity of labor everywhere. Moreover, they indicate that the joint density of the production process decays to zero smoothly as it approaches its efficient support boundary.

We also considered for our illustration purposes a small dataset which consists of 123 American electric utility companies. As in the set-up of Gijbels *et al.* (1999), we used the measurements of the variables $Y_i = \log(Q_i)$ and $X_i = \log(C_i)$, where Q_i is the production output of the company i and C_i is the total cost involved in the production. Given that the original inputs are constituted of negative values, we shifted the X_i's so that all inputs become strictly positive. Figure 6 (bottom) shows the observations together with the minimum cost function estimators $\widehat{\varphi}_n$, $\widehat{\varphi}^{\#}_{n}$ and $\widetilde{\varphi}^{\#}_{n,m}$. Here, we used $m = 0$ as suggested by the evolution of $D(m)$ in Figure 7 (middle). In what concerns the bootstrap parameters, we always considered $n_y = [N_y^{0.9}]$ with the same bounds for the choice of $\hat{k}^*_n(y)$ as before. For the number of bootstrap resamples, we set $r_y = 100$ as in simulations.

As in the previous example, the Γ-moment frontier estimator does not indicate an ideal production activity. Even the standard economic situation hoped for by producers, where the density of data should be strictly positive at the efficient boundary, does not occur in this sector of productivity since most of the companies operate on the interior of the joint support of (Y, X) rather than near or at its optimal boundary estimator $\widetilde{\varphi}^{\#}_{n,0}$. It may be noticed that the electric utility data do not contain any potential outlier, which is not the case for the postal data. Note also that the tail of the data generating process is obviously heavier in regions where there are less points. It is then not surprising that the difference between the FDH frontier $\widehat{\varphi}_n$ and the bias-corrected estimator $\widetilde{\varphi}^{\#}_{n,0}$ is wider at places where the sparsity of data is greater.

4.2 Assessment of the reliability of nuclear reactors

The knowledge of the behaviour of the pressure vessel is of prime importance in a nuclear power plant lifetime program. The structural integrity relies upon accurate knowledge of the change in fracture toughness of the reactor pressure vessel materials over the time of operation. Fracture toughness is very dependent on material temperature as illustrated on Figure 8. The dataset from the US Electric Power Research Institute (EPRI) consists of 254 toughness results obtained from non-irradiated representative steels. For each steel i, fracture toughness X_i and temperature Y_i were measured. The goal is to estimate the so-called master curve prediction $x = \varphi(y)$ of the lowest fracture toughness as a function of the temperature. Physical considerations permit to establish that φ is nondecreasing.

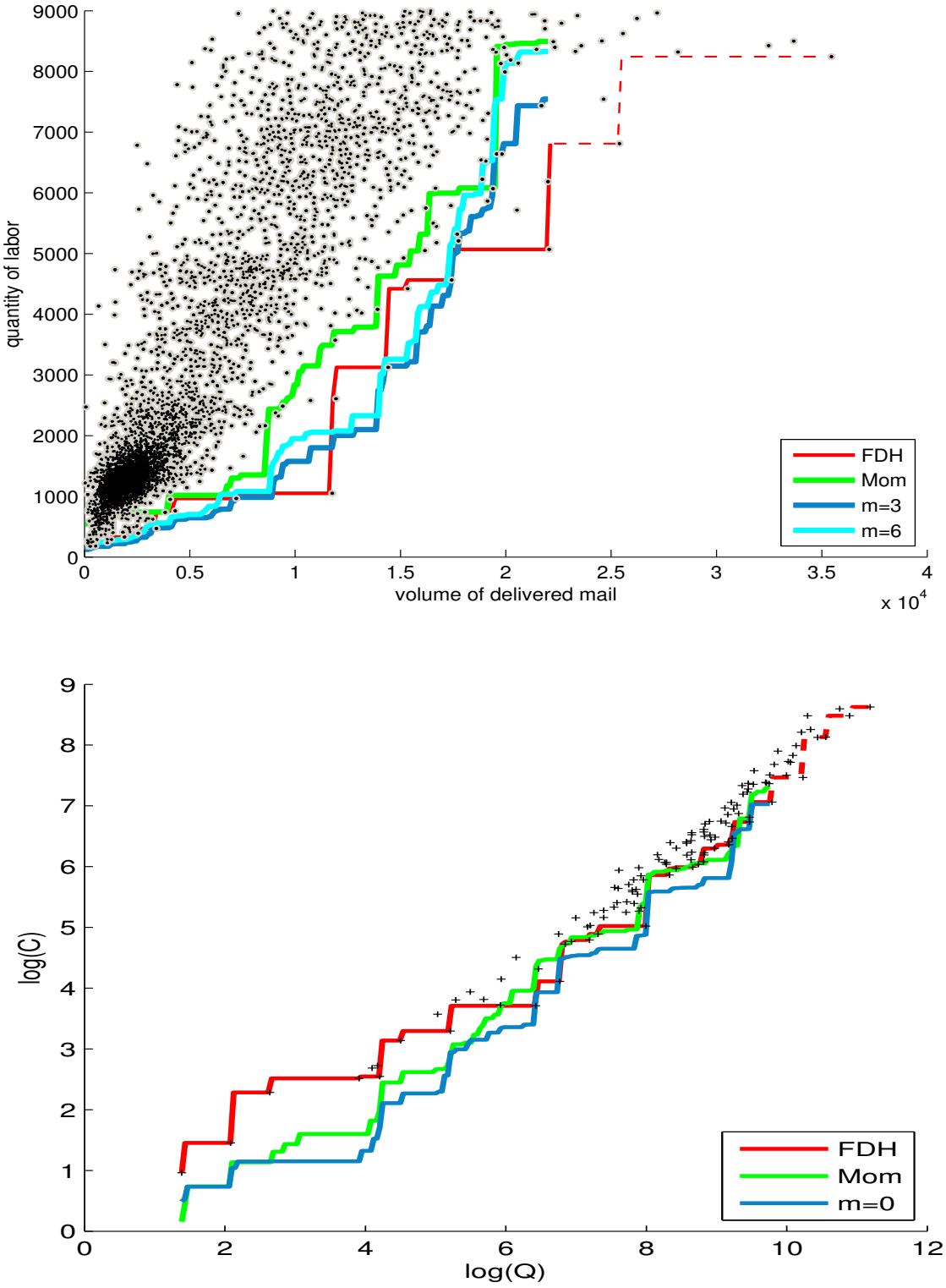

Figure 6: *Top— Scatterplot of the* 5138 *post offices, with the minimum cost function estimators* $\widehat{\varphi}_n$ *(red),* $\widehat{\varphi}_n^{\#}$ *(green),* $\widetilde{\varphi}_{n,3}^{\#}$ *(darker blue) and* $\widetilde{\varphi}_{n,6}^{\#}$ *(lighter blue). Bottom— The* 123 *electric utility data, with the frontier estimators* $\widehat{\varphi}_n$ *(red),* $\widehat{\varphi}_n^{\#}$ *(green) and* $\widetilde{\varphi}_{n,0}^{\#}$ *(blue).*

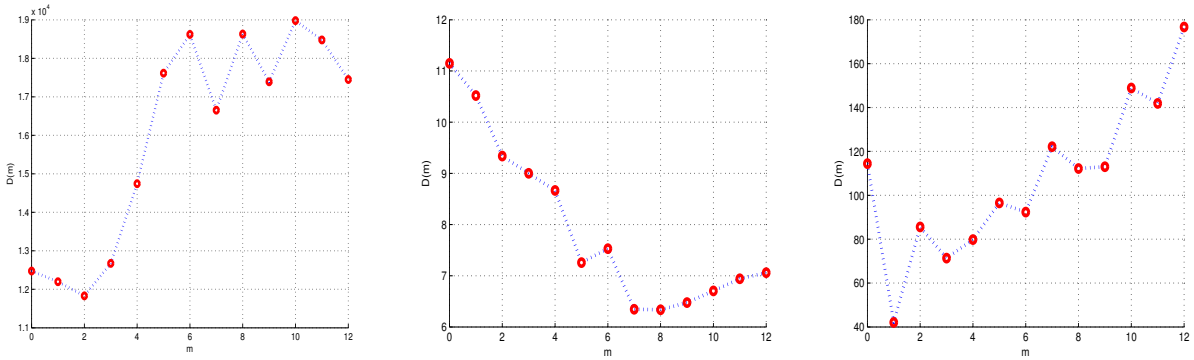

Figure 7: *Evolution of $D(m)$ with m, for the post offices, electric utility companies and nuclear reactors, respectively, from left to right.*

Figure 8 depicts the minimum fracture toughness function estimators $\widehat{\varphi}_n$, $\widehat{\varphi}_n^{\#}$ and $\widetilde{\varphi}_{n,0}^{\#}$. Here also, the choice $m = 0$ in the Γ-moment estimator appears to be more appropriate as can be seen from the graph of $D(m)$ in Figure 7 (r-h.s). Notice that the same bootstrap parameters as in simulations and the electric utility data example were used. Surprisingly, it may be seen that the standard moment estimator $\widehat{\varphi}_n^{\#}$ has a quite remarkable behavior although the small number of data. Nevertheless, the Γ-moment frontier $\widetilde{\varphi}_{n,0}^{\#}$ is stable as well, with the added advantage of improving better the FDH boundary $\widehat{\varphi}_n$ for its inherent bias.

5 Conclusion

Frontier modeling is clearly a problem belonging to extreme value theory. Reliable estimation of boundaries from this perspective involves, however, many delicate issues when the sample size is not sufficiently large. The Γ-moment method seems to offer a viable approach under the monotonicity constraint. Simulation evidence suggests that the Γ-moment frontiers are appreciably more efficient than the popular moment and FDH estimators. Doubtless, further work on the optimal selection of the intermediate sequence will yield new refinements. Codes for all of the procedures described in this paper are available upon request, so we hope that this will encourage others to explore the Γ-moment device.

References

[1] Aarssen, K. and de Haan, L. (1994). On the maximal life span of humans, *Mathematical Population Studies*, **4**, 259–281.

[2] Beirlant, J., Goegebeur, Y, Segers, J. and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley.

[3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge, U.K.: Cambridge University Press.

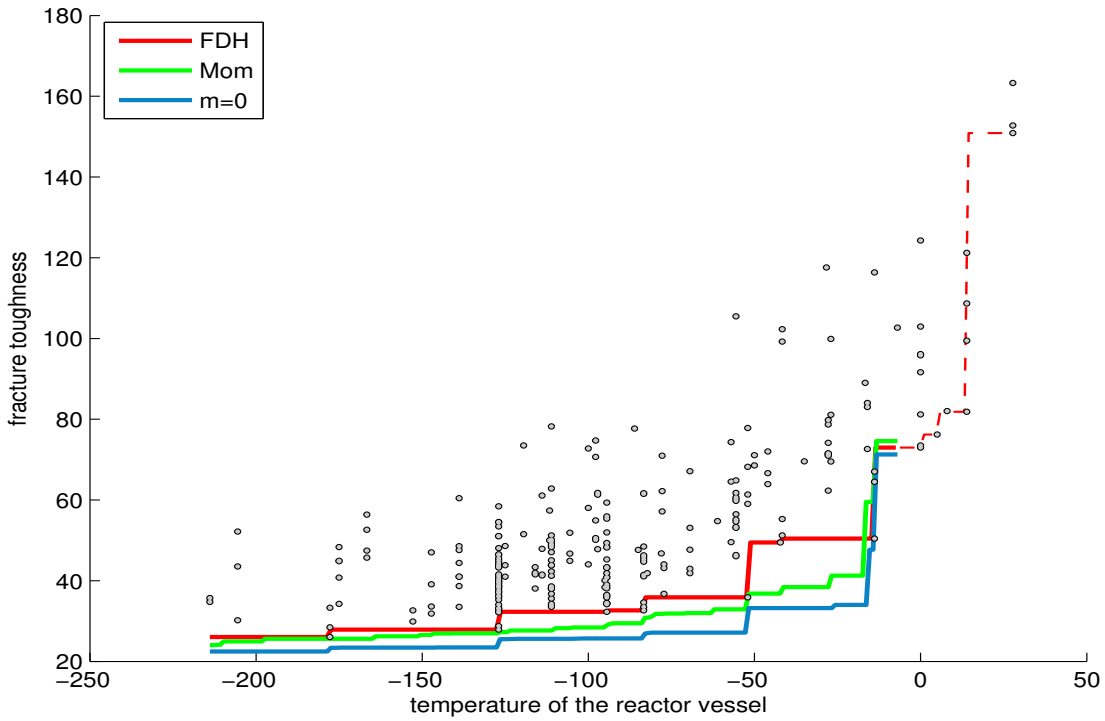

Figure 8: *Scatterplot of the* 254 *nuclear reactors' data, with the minimum fracture toughness function estimators* $\widehat{\varphi}_n$ *(red),* $\widehat{\varphi}_n^{\#}$ *(green) and* $\widetilde{\varphi}_{n,0}^{\#}$ *(blue).*

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

[5] Daouia, A., Florens, J.P. and Simar, L. (2010). Frontier Estimation and Extreme Value Theory, *Bernoulli*, **16**, 1039–1063.

[6] Daouia, A. and Simar, L. (2005). Robust nonparametric estimators of monotone boundaries, *Journal of Multivariate Analysis*, **96**, 311–331.

[7] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer-Verlag, New York.

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

[9] Deprins, D., Simar, L. and Tulkens, H. (1984). Measuring labor efficiency in post offices. *In: M. Marchand, P. Pestieau and H. Tulkens (eds), The performance of Public Enterprises: Concepts and Measurements*, 243–267. Amsterdam: North-Holland.

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

[11] Ferreira, A., de Haan, L. and Peng, L. (2003). On optimizing the estimation of high quantiles of a probability distribution, *Statistics*, **37**, 401–434.

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

[13] Gijbels, I. and Peng, L. (2000). Estimation of a support curve via order statistics, *Extremes*, **3**, 251–277.

[14] Girard, S., Guillou, A. and Stupfler, G. (2012). Estimating an endpoint with high order moments in the Weibull domain of attraction, *Statistics and Probability Letters*, **82**, 2136–2144.

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

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

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

[18] Jeong, S.-O. and Park, B.U. (2006). Large sample approximation of the distribution for convex-hull estimators of boundaries, *Scandinavian Journal of Statistics*, **33**, 139–151.

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

[20] Kumbhakar, S.C. and Lovell, C.A. (2000). *Stochastic Frontier Analysis*, Cambridge University Press.

[21] Park, B.U. (2001). On nonparametric estimation of data edges, *Journal of the Korean Statistical Society*, **30**, 265–280.

[22] Park, B.U., Jeong, S.-O. and Simar, L. (2010). Asymptotic distribution of conical-hull estimators of directional edges, *The Annals of Statistics*, **38**, 1320–1340.

[23] Park, B.U., Simar, L. and Weiner, C. (2000). The FDH estimator for productivity efficiency scores: Asymptotic properties, *Econometric Theory*, **16**, 855-877.

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

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