# Estimating frontier cost models using extremiles

Abdelaati Daouia* and Irène Gijbels†

May 30, 2011

**Abstract**

In the econometric literature on the estimation of production technologies, there has been considerable interest in estimating so called cost frontier models that relate closely to models for extreme non-standard conditional quantiles (Aragon *et al.* (2005)) and expected minimum input functions (Cazals *et al.* (2002)). In this paper, we introduce a class of extremile-based cost frontiers which includes the family of expected minimum input frontiers and parallels the class of quantile-type frontiers. The class is motivated via several angles, which reveals its specific merits and strengths. We discuss nonparametric estimation of the extremile-based costs frontiers and establish asymptotic normality and weak convergence of the associated process. Empirical illustrations are provided.

*Key words* : cost function, expected minimum, extremile, probability-weighted moment.

## 1 Introduction

In the analysis of productivity and efficiency, for example of firms, the interest lies in estimating a production frontier or cost function. Among the basic references in economic theory are Koopmans (1951), Debreu (1951) and Shephard (1970). The activity of a production unit (e.g. a firm) is characterized via a set of outputs, $y \in \mathbb{R}_+^q$ that is produced by a set of inputs $x \in \mathbb{R}_+^p$. The set of attainable points can be characterized as

$$\Psi = \{(y, x) \in \mathbb{R}_+^{q+p} |\ y \text{ can be produced by } x\}.$$

This set can be described mathematically by its sections. In the input space one has the input requirement sets $C(y) = \{x \in \mathbb{R}_+^p | (y, x) \in \Psi\}$, defined for all possible outputs $y \in \mathbb{R}_+^q$. The radial (or input-oriented) efficiency boundary is then given by $\partial C(y)$, the boundary of the input requirement set. In the case of univariate inputs $\partial C(y) = \min C(y)$, the input-efficiency function, also called the frontier cost function. From an economic point

*Toulouse School of Economics (GREMAQ), University of Toulouse, France (daouia@cict.fr). This research was supported by the French "Agence Nationale pour la Recherche" under grant ANR-08-BLAN-0106-01/EPI project

†Department of Mathematics and Leuven Statistics Research Center, Katholieke Universiteit Leuven, Belgium (Irene.Gijbels@wis.kuleuven.be). This research was supported by the Research Fund KULeuven (GOA/07/04-project) and by the IAP research network P6/03, Federal Science Policy, Belgium.

of view, a monotonicity assumption on this function is reasonable, meaning that higher outputs go along with a higher minimal cost. Different other assumptions can be made on $\Psi$ such as for example free disposability, i.e. if $(y,x) \in \Psi$ then $(y',x') \in \Psi$ for any $x' \geq x$ and $y' \leq y$ (the inequalities here have to be understood componentwise); or convexity, i.e., every convex combination of feasible production plans is also feasible. See Shephard (1970) for more information and economic background.

In this paper we will focus the presentation on the input orientation[1], where we want to estimate the minimal cost frontier in the case of univariate inputs. To our disposal are observations $\mathcal{X}_n = \{(Y_i, X_i)|\ i = 1, \cdots, n\}$ generated by the production process defined through for example the joint distribution of a random vector $(Y, X)$ on $\mathbb{R}_+^q \times \mathbb{R}_+$, where the $q$-dimensional vector $Y$ represents the outputs and the second variable $X$ is the single input. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probability space on which both $Y$ and $X$ are defined. In the case where the production set $\Psi$ is equal to the support of the joint distribution of $(Y, X)$, a probabilistic way for defining the cost frontier is as follows. The cost function $\partial C(y)$ is characterized for a given set of outputs $y$ by the lower boundary of the support of the conditional distribution of $X$ given $Y \geq y$, i.e.

$$\varphi(y) := \inf\{x \geq 0 | \bar{F}_y(x) < 1\} \equiv \partial C(y), \tag{1}$$

where $\bar{F}_y(x) = 1 - F_y(x)$, with $F_y(x) = \mathbb{P}(X \leq x | Y \geq y)$ being the conditional distribution function of $X$ given $Y \geq y$, for $y$ such that $\mathbb{P}(Y \geq y) > 0$. The frontier function $\varphi$ is monotone nondecreasing, which corresponds to the free disposability property of the support $\Psi$. When the support boundary $\partial C(\cdot)$ is not assumed to be monotone, $\varphi(\cdot)$ is in fact the largest monotone function which is smaller than or equal to the lower boundary $\partial C(\cdot)$. See Cazals *et al.* (2002) for this formulation and a detailed discussion on the concept of frontier cost function.

There is a vast literature on the estimation of frontier functions from a random sample of production units $\mathcal{X}_n$. There have been developments along two main approaches: the deterministic frontier models which suppose that with probability one, all the observations in $\mathcal{X}_n$ belong to $\Psi$, and the stochastic frontier models where random noise allows some observations to be outside of $\Psi$.

In deterministic frontier models, two different nonparametric methods based on envelopment techniques have been around. The free disposal hull (FDH) technique and the data envelopment analysis (DEA) technique. Deprins *et al.* (1984) introduced the FDH estimator that relies only on the free disposability assumption on $\Psi$. The DEA estimator initiated by

[1]The presentation for the output orientation, where we want to estimate the maximal production frontier in the case of univariate outputs, is a straightforward adaptation of what is done here.

Farrell (1957) and popularized as linear programming estimator by Charnes *et al.* (1978), requires stronger assumptions, it relies on the free disposability assumption and the convexity of $\Psi$. Such a convexity assumption is widely used in economics, but it is not always valid. Because of the additional assumption of convexity the FDH estimator is a more general estimator than the DEA estimator. The asymptotic distribution of the FDH estimator was derived by Park *et al.* (2000) in the particular case where the joint density of $(Y, X)$ has a jump at the frontier and by Daouia *et al.* (2010) in the general setting. The asymptotic distribution of the DEA estimator was derived by Gijbels *et al.* (1999). Today, most statistical theory of these estimators is available. See Simar and Wilson (2008), among others.

In stochastic frontier models, where noise is allowed, one often imposes parametric restrictions on the shape of the frontier and on the data generating process to allow identification of the noise from the cost frontier and subsequently estimation of this frontier. These parametric methods may lack robustness if the distributional assumptions made do not hold.

Since nonparametric deterministic frontier models rely on very few assumptions, they are quite appealing. Moreover, the FDH estimator of the frontier cost function can simply be viewed as a plug-in version of (1) obtained by just replacing the conditional distribution function by its empirical analog $\widehat{F}_y(x)$ resulting into

$$\hat{\varphi}(y) = \inf\{x \in \mathbb{R}_+ |\ \hat{F}_y(x) > 0\} = \min_{i:Y_i \geq y} X_i \ ,$$

with $\hat{F}_y(x) = \sum_{i=1}^{n} 1\!\!I(X_i \leq x, Y_i \geq y) / \sum_{i=1}^{n} 1\!\!I(Y_i \geq y)$.

The FDH estimator, as well as the DEA estimator, however are very sensitive to outlying observations. In the literature two robust nonparametric estimators of (partial) cost frontiers have been proposed to deal with this sensitivity. Cazals *et al.* (2002) introduced the concept of expected minimal cost of order $m \in \{1, 2, 3, ...\}$. It is defined as the expected minimal cost among $m$ firms drawn in the population of firms exceeding a certain level of outputs. More precisely, for a given level of outputs $y$, the cost function of order $m$ is given by

$$\varphi_m(y) = E[\min(X_1^y, \cdots, X_m^y)] = \int_0^\infty \{\bar{F}_y(x)\}^m dx,$$

where $(X_1^y, \cdots, X_m^y)$ are $m$ independent identically distributed random variables generated from the distribution of $X$ given $Y \geq y$. Its nonparametric estimator is defined by

$$\hat{\varphi}_{m,n}(y) = \int_0^\infty \{1 - \hat{F}_y(x)\}^m dx.$$

The estimator $\hat{\varphi}_{m,n}(y)$ does not envelop all the data points, and so it is more robust to extreme values than the FDH estimator $\hat{\varphi}(y)$. By choosing $m$ appropriately as a function of the sample size $n$, $\hat{\varphi}_{m,n}(y)$ estimates the cost function $\varphi(y)$ itself while keeping the asymptotic properties of the FDH estimator.

A second approach to deal with the sensitivity to outlying observations was proposed by Aragon *et al.* (2005). They consider extreme quantiles of the conditional distribution of $X$ given $Y \geq y$. Such non-standard conditional quantiles provide another concept of a partial cost frontier as an alternative towards the order-$m$ partial cost frontier introduced by Cazals *et al.* (2002). The duality between expected minimum input frontiers and quantile-type cost frontiers has been investigated by Daouia and Gijbels (2011).

In this paper we introduce a new class of extremile-based cost frontiers which includes the class of order-$m$ expected minimum input frontiers. The class also parallels the class of quantile partial cost frontiers in the sense that it is related to the mean of a random variable rather than the median (or quantile more generally).

The paper is organized as follows. In Section 2 we introduce the class of extremile-based cost frontier functions, and discuss the relation with the order-$m$ partial cost functions and the quantile-type cost functions. Some basic properties of the new class of frontier functions are provided. Section 3 is devoted to nonparametric estimation of an extremile-based cost frontier, and studies the asymptotic properties of the estimators. An empirical study on a simulation model and on a real data example is provided in Section 4. Section 5 concludes.

# 2 The extremile-based cost function

Consider a real $\gamma \in (0,1)$ and let $K_\gamma$ be a measure on $[0,1]$ whose distribution function is

$$K_\gamma(t) = \begin{cases} 1-(1-t)^{s(\gamma)} & \text{if } \ 0<\gamma\leq \frac{1}{2} \\ t^{s(1-\gamma)} & \text{if } \ \frac{1}{2}\leq\gamma<1 \end{cases}$$

where

$$s(\gamma) = \frac{\log(1/2)}{\log(1-\gamma)} \geq 1 \quad \text{for} \quad \gamma\in[0,1/2].$$

Define the score function $J_\gamma(\cdot)$ to be the density of the measure $K_\gamma$ on $(0,1)$.

**Definition 1.** *The extremile-based cost function of order $\gamma$ denoted by $\xi_\gamma(y)$ is the real function defined on $\mathbb{R}^q_+$ as*

$$\xi_\gamma(y) = \mathbb{E}\left[XJ_\gamma\left(F_y(X)\right)|Y\geq y\right]$$

*where we assume the existence of this expectation.*

As a matter of fact, the partial cost function $\xi_\gamma(y)$ coincides with the $\gamma$th extremile (see Daouia and Gijbels 2009) of the conditional distribution of $X$ given $Y \geq y$. The following proposition is a basic property of extremiles.

**Proposition 1.** *If $\mathbb{E}(X|Y \geq y) < \infty$, then $\xi_\gamma(y)$ exists for all $\gamma \in (0,1)$.*

*Proof.* Following Daouia and Gijbels (2009, Proposition 1 (i)), the $\gamma$th extremile exists provided that the underlying distribution has a finite absolute mean which corresponds here to $\mathbb{E}(X|Y \geq y) < \infty$. $\square$

From an economic point of view, the quantity $X$ of inputs-usage is often assumed to be bounded or at least to have a finite mean and so, in this case, the $\gamma$th cost function is well defined for any order $\gamma$ in $(0,1)$ and all $y \in \mathbb{R}^q_+$ such that $\mathbb{P}(Y \geq y) > 0$.

More specifically, the extremile function $\xi_\gamma(y)$ is proportional to conditional probability-weighted moments:

$$\xi_\gamma(y) = \begin{cases} s(\gamma)\mathbb{E}\left[X\left\{\bar{F}_y(X)\right\}^{s(\gamma)-1}|Y \geq y\right] & \text{for} \quad 0 < \gamma \leq \frac{1}{2} \\ s(1-\gamma)\mathbb{E}\left[X\left\{F_y(X)\right\}^{s(1-\gamma)-1}|Y \geq y\right] & \text{for} \quad \frac{1}{2} \leq \gamma < 1. \end{cases}$$

In the special case where $\gamma \leq 1/2$ with $s(\gamma)$ being a positive integer, $\xi_\gamma(y)$ equals the expectation of the minimum of $s(\gamma)$ independent random variables $(X_1^y, \ldots, X_{s(\gamma)}^y)$ generated from the distribution of $X$ given $Y \geq y$. Whence

$$\xi_\gamma(y) = \mathbb{E}\left[\min\left(X_1^y, \ldots, X_{s(\gamma)}^y\right)\right] = \varphi_{s(\gamma)}(y).$$

Thus the class of our conditional extremiles includes the family of expected minimum input functions introduced by Cazals *et al.* (2002). Likewise, if $\gamma \geq 1/2$ with $s(1-\gamma) = 1, 2, \ldots$ we have $\xi_\gamma(y) = \mathbb{E}\left[\max\left(X_1^y, \ldots, X_{s(1-\gamma)}^y\right)\right]$, where $X_1^y, \ldots, X_{s(1-\gamma)}^y$ are independent random variables generated from the distribution of $X$ given $Y \geq y$.

**Proposition 2.** *If the conditional distribution of $X$ given $Y \geq y$ has a finite mean, then it can be characterized by the subclass $\{\xi_\gamma(y) : s(\gamma) = 1, 2, \ldots\}$ or $\{\xi_\gamma(y) : s(1-\gamma) = 1, 2, \ldots\}$.*

*Proof.* This follows from the well known result of Chan (1967) which states that a distribution with finite absolute first moment can be uniquely defined by its expected maxima or expected minima. $\square$

The non-standard conditional distribution of $X$ given $Y \geq y$ whose $\mathbb{E}(X|Y \geq y) < \infty$ is uniquely defined by its discrete extremiles. This means that no two such non-standard distributions with finite means have the same expected minimum input functions.

Of particular interest is the left tail $\gamma \leq 1/2$ where the partial $\gamma$th cost function has the following interpretation

$$\mathbb{E}\left[\min\left(X_1^y, \ldots, X_{[s(\gamma)]+1}^y\right)\right] \leq \xi_\gamma(y) \leq \mathbb{E}\left[\min\left(X_1^y, \ldots, X_{[s(\gamma)]}^y\right)\right]$$

where $[s(\gamma)]$ denotes the integer part of $s(\gamma)$ and $X_1^y, X_2^y, \ldots$ are iid random variables of distribution function $F_y$. In other words, we have $\varphi_{[s(\gamma)]+1}(y) \leq \xi_\gamma(y) \leq \varphi_{[s(\gamma)]}(y)$ for $\gamma \leq 1/2$. Hence $\xi_\gamma(y)$ benefits from a similar "benchmark" interpretation as expected minimum input functions. For the manager of a production unit working at level $(x, y)$, comparing its inputs-usage $x$ with the benchmarked value $\xi_\gamma(y)$, for a sequence of few decreasing values of $\gamma \searrow 0$, could offer a clear indication of how efficient its firm is compared with a fixed number of $(1 + [s(\gamma)])$ potential firms producing more than $y$.

Yet, there is still another way of looking at $\xi_\gamma(y)$. Let $\mathcal{X}_\gamma^y$ be a random variable having cumulative distribution function

$$F_{\mathcal{X}_\gamma^y} = \begin{cases} 1 - \left\{\bar{F}_y\right\}^{s(\gamma)} & \text{if} \quad 0 < \gamma \leq \frac{1}{2} \\ \left\{F_y\right\}^{s(1-\gamma)} & \text{if} \quad \frac{1}{2} \leq \gamma < 1. \end{cases}$$

**Proposition 3.** *We have* $\xi_\gamma(y) = \mathbb{E}(\mathcal{X}_\gamma^y)$ *provided this expectation exists.*

*Proof.* Since $\mathbb{E}|\mathcal{X}_\gamma^y| = \mathbb{E}(\mathcal{X}_\gamma^y) < \infty$, we have $\mathbb{E}(\mathcal{X}_\gamma^y) = \int_0^1 F_{\mathcal{X}_\gamma^y}^{-1}(t)dt$ in view of a general property of expectations (see Shorack 2000, p.117). On the other hand, it is easy to check that $\xi_\gamma(y) = \int_0^1 J_\gamma(t) F_y^{-1}(t)dt = \int_0^1 F_y^{-1}(t)dK_\gamma(t) = \int_0^1 F_{\mathcal{X}_\gamma^y}^{-1}(t)dt$. $\square$

This allows to establish how our class of extremile-based cost functions is related to the family of quantile-based cost functions defined by Aragon *et al.* (2005) as

$$Q_\gamma(y) = F_y^{-1}(\gamma) := \inf\{x \in \mathbb{R}_+ | \; F_y(x) \geq \gamma\} \quad \text{for} \quad 0 < \gamma < 1.$$

Indeed, while $\xi_\gamma(y)$ equals the mean of the random variable $\mathcal{X}_\gamma^y$, it is easy to see that the quantile function $Q_\gamma(y)$ coincides with the median of the same variable $\mathcal{X}_\gamma^y$. Consequently the $\gamma$th extremile-based cost function is clearly more tail sensitive and more efficient than the $\gamma$th quantile-based cost function. The latter means that the (asymptotic) variance for the extremile-based cost function estimator is smaller than the (asymptotic) variance of the $\gamma$th quantile-based cost function. Recall that for many population distributions (such as e.g. a normal distribution) the sample mean has a smaller asymptotic variance than the sample median, when both are estimating the same quantity. See for example Serfling (1980).

One way of defining $\xi_\gamma(y)$, with $0 \leq \gamma \leq 1$, is as the following explicit quantity.

**Proposition 4.** *If* $\mathbb{E}(\mathcal{X}_\gamma^y) < \infty$, *we have*

$$\xi_\gamma(y) = \begin{cases} \varphi(y) + \int_{\varphi(y)}^\infty \left\{\bar{F}_y(x)\right\}^{s(\gamma)} dx & \textit{for} \quad 0 \leq \gamma \leq \frac{1}{2} \\ \varphi(y) + \int_{\varphi(y)}^\infty \left(1 - \left\{F_y(x)\right\}^{s(1-\gamma)}\right) dx & \textit{for} \quad \frac{1}{2} \leq \gamma \leq 1. \end{cases} \tag{2}$$

*Proof.* We have $\xi_\gamma(y) = \mathbb{E}(\mathcal{X}_\gamma^y)$ by Proposition 3 and $\mathbb{E}(\mathcal{X}_\gamma^y) = \int_0^\infty \left\{1 - F_{\mathcal{X}_\gamma^y}(x)\right\} dx = \varphi(y) + \int_{\varphi(y)}^\infty \left\{1 - F_{\mathcal{X}_\gamma^y}(x)\right\} dx$ by a general property of expectations (Shorack 2000, p.117). □

This explicit expression is very useful when it comes to proposing an estimator for $\xi_\gamma(y)$. Obviously, the central extremile-based cost function $\xi_{1/2}(y)$ reduces to the regression function $\mathbb{E}(X|Y \geq y)$. The conditional extremile $\xi_\gamma(y)$ is clearly a continuous and increasing function in $\gamma$ and it maps $(0,1)$ onto the range $\{x \geq 0 | 0 < F_y(x) < 1\}$. The left and right endpoints of the support of the conditional distribution function $F_y(\cdot)$ coincide respectively with the lower and upper extremiles $\xi_0(y)$ and $\xi_1(y)$ since $s(0) = \infty$. Hence the range of $\xi_\gamma(y)$ is the entire range of $X$ given $Y \geq y$.

Of interest is the limiting case $\gamma \downarrow 0$ which leads to access the full cost function $\varphi(y) = \xi_0(y)$. Although the limit frontier function $\varphi(\cdot)$ is monotone nondecreasing, the partial cost function $\xi_\gamma(\cdot)$ itself is not necessarily monotone. To ensure the monotonicity of $\xi_\gamma(y)$ in $y$, it suffices to assume, as it can be easily seen from Proposition 4, that the conditional survival function $\bar{F}_y(x)$ is nondecreasing in $y$. As pointed out by Cazals *et al.* (2002), this assumption is quite reasonable from an economic point of view since the chance of spending more than a cost $x$ does not decrease if a firm produces more.

The next proposition provides an explicit relationship between the $\gamma$th quantile and extremile type cost functions at $\gamma \downarrow 0$. Let $\mathrm{DA}(W_\rho)$ denote the minimum domain of attraction of the Weibull extreme-value distribution

$$W_\rho(x) = 1 - \exp\{-x^\rho\} \quad \text{with support} \quad [0, \infty), \quad \text{for some} \quad \rho > 0,$$

*i.e.*, the set of distribution functions whose asymptotic distributions of minima are of the type of $W_\rho$. According to Daouia *et al.* (2010), if there exists a sequence $\{a_n > 0\}$ such that the normalized minima $a_n^{-1}(\hat{\varphi}(y) - \varphi(y))$ converges to a non-degenerate distribution, then the limit distribution function is of the type of $W_\rho$ for a positive function $\rho = \rho(y)$ in $y$.

**Proposition 5.** *Suppose* $\mathbb{E}(X|Y \geq y) < \infty$ *and* $F_y(\cdot) \in DA(W_{\rho(y)})$. *Then*

$$\frac{\xi_\gamma(y) - \varphi(y)}{Q_\gamma(y) - \varphi(y)} \sim \Gamma(1 + \rho^{-1}(y))\{\log 2\}^{-1/\rho(y)} \quad \text{as} \quad \gamma \downarrow 0,$$

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

*Proof.* This follows immediately by applying Proposition 2 (ii) in Daouia and Gijbels (2009) to the distribution of $-X$ given $Y \geq y$. □

Consequently, as $\gamma \downarrow 0$, the quantile curve $Q_\gamma(\cdot)$ is closer to the true cost frontier $\varphi(\cdot)$ than is the extremile curve $\xi_\gamma(\cdot)$ following the value of the tail index $\rho$. In most situations described so far in the econometric literature on frontier analysis, it is assumed that there is a jump of the joint density of $(Y, X)$ at the frontier: this corresponds to the case where the tail index $\rho(y)$ equals the dimension of data $(1+q)$ according to Daouia *et al.* (2010). It was shown in that paper that $\beta(y) = \rho(y) - (1+q)$, where $\beta(y)$ denotes the algebraic rate with which the joint density decreases to 0 when $x$ approaches the point at the frontier function. Since a jump of the joint density at the frontier implies that $\beta(y) = 0$, it follows that $\rho(y) = 1+q$ in that case. In such situations, $Q_\gamma(\cdot)$ is asymptotically closer to $\varphi(\cdot)$ than is $\xi_\gamma(\cdot)$ when $q \leq 2$, but $\xi_\gamma(\cdot)$ is more spread than $Q_\gamma(\cdot)$ when $q > 2$.

On the other hand, the score function $J_\gamma(\cdot)$ being monotone increasing for $\gamma \geq 1/2$ and decreasing for $\gamma \leq 1/2$, the conditional extremile $\xi_\gamma(y)$ depends by construction on all feasible values of $X$ given $Y \geq y$ putting more weight to the high values for $\gamma \geq 1/2$ and more weight to the low values for $\gamma \leq 1/2$. Therefore $\xi_\gamma(y)$ is sensible to the magnitude of extreme values for any order $\gamma \in (0,1)$. In contrast, the conditional quantile $Q_\gamma(y)$ is determined solely by the tail probability (relative frequency) $\gamma$, and so it may be unaffected by desirable extreme values whatever the shape of tails of the underlying distribution. On the other hand, when $Q_\gamma(y)$ breaks down at $\gamma \downarrow 0$ or $\gamma \uparrow 1$, the $\gamma$th conditional extremile, being an L-functional, is more resistant to extreme values. Hence, $\xi_\gamma(y)$ steers an advantageous middle course between the extreme behaviors (insensitivity and breakdown) of $Q_\gamma(y)$.

# 3 Nonparametric estimation

Instead of estimating the full cost function, an original idea first suggested by Cazals *et al.* (2002) and applied by Aragon *et al.* (2005) to quantiles is rather to estimate a partial cost function lying near $\varphi(y)$. Thus the interest in this section will be in the estimation of the extremile function $\xi_\gamma(y)$ for $\gamma \leq 1/2$. The right tail (*i.e.* $\gamma \geq 1/2$) can be handled in a similar way and so is omitted. Results below are easily derived by means of L-statistics theory applied to the dimensionless transformation $Z^y = X \mathbb{1}(Y \geq y)$ of the random vector $(Y, X) \in \mathbb{R}_+^{q+1}$. Let $\bar{F}_{Z^y} = 1 - F_{Z^y}$ be the survival function of $Z^y$. It is easy to check that $\bar{F}_y(X) = \bar{F}_{Z^y}(Z^y)/\mathbb{P}(Y \geq y)$. Then

$$\begin{aligned} \xi_\gamma(y) &= \mathbb{E}\left[Z^y J_\gamma\left(F_y(X)\right)\right]/\mathbb{P}(Y \geq y) \quad \text{for} \quad 0 < \gamma < 1 \\ &= \frac{s(\gamma)}{\mathbb{P}(Y \geq y)} \mathbb{E}\left[Z^y \left\{\frac{\bar{F}_{Z^y}(Z^y)}{\mathbb{P}(Y \geq y)}\right\}^{s(\gamma)-1}\right] = \frac{\xi_{Z^y}(\gamma)}{\{\mathbb{P}(Y \geq y)\}^{s(\gamma)}} \quad \text{for} \quad 0 < \gamma \leq \frac{1}{2} \end{aligned} \tag{3}$$

where $\xi_{Z^y}(\gamma) = s(\gamma)\mathbb{E}\left[Z^y \left\{\bar{F}_{Z^y}(Z^y)\right\}^{s(\gamma)-1}\right] = \int_0^1 F_{Z^y}^{-1} dK_\gamma$ is by definition the ordinary $\gamma$th extremile of the random variable $Z^y$. Therefore it suffices to replace $\mathbb{P}(Y \geq y)$ by its empirical version $\hat{\mathbb{P}}(Y \geq y) = (1/n)\sum_{i=1}^n \mathbb{1}(Y_i \geq y)$ and $\xi_{Z^y}(\gamma)$ by a consistent estimator to obtain a convergent estimate of the conditional extremile $\xi_\gamma(y)$.

As shown in Daouia and Gijbels (2009), a natural estimator of the ordinary extremile $\xi_{Z^y}(\gamma)$ is given by the L-statistic generated by the measure $K_\gamma$:

$$\hat{\xi}_{Z^y}(\gamma) = \sum_{i=1}^{n} \left\{ K_\gamma\left(\frac{i}{n}\right) - K_\gamma\left(\frac{i-1}{n}\right) \right\} Z^y_{(i)}, \tag{4}$$

where $Z^y_{(1)} \leq Z^y_{(2)} \leq \cdots \leq Z^y_{(n)}$ denote the order statistics generated by the sample $\{Z^y_i = X_i 1\!\!I(Y_i \geq y) : i = 1, \cdots, n\}$. It is easy to see that the resulting estimator of the $\gamma$th cost function $\xi_\gamma(y)$, given by

$$\hat{\xi}_\gamma(y) = \hat{\xi}_{Z^y}(\gamma)/\{\hat{\mathbb{P}}(Y \geq y)\}^{s(\gamma)}$$

coincides with the empirical conditional extremile obtained by replacing $F_y(x)$ in expression (2) with its empirical version $\hat{F}_y(x)$, *i.e*,

$$\hat{\xi}_\gamma(y) = \int_0^\infty \left\{1 - \hat{F}_y(x)\right\}^{s(\gamma)} dx = \hat{\varphi}(y) + \int_{\hat{\varphi}(y)}^\infty \left\{1 - \hat{F}_y(x)\right\}^{s(\gamma)} dx. \tag{5}$$

This estimator converges to the FDH input efficient frontier $\hat{\varphi}(y)$ as $\gamma$ decreases to zero. In particular, when the power $s(\gamma)$ is a positive integer $m = 1, 2, \ldots$ we recover the estimator $\hat{\varphi}_{m,n}(y)$, of the expected minimum input function of order $m$ proposed by Cazals *et al.* (2002). See Section 1. The following theorem summarizes the asymptotic properties of $\hat{\xi}_\gamma(y)$ for a fixed order $\gamma$.

**Theorem 1.** *Assume that the support of $(Y, X)$ is compact and let $\gamma \in (0, 1/2]$.*

*(i) For any point $y \in \mathbb{R}^q_+$ such that $P(Y \geq y) > 0$, $\hat{\xi}_\gamma(y) \overset{a.s.}{\to} \xi_\gamma(y)$ as $n \to \infty$, and $\sqrt{n}\left(\hat{\xi}_\gamma(y) - \xi_\gamma(y)\right)$ has an asymptotic normal distribution with mean zero and variance $\mathbb{E}\left[\mathbb{S}_\gamma(y, Y, X)\right]^2$, where $\mathbb{S}_\gamma(y, Y, X) =$*

$$\frac{s(\gamma)}{\{\mathbb{P}(Y \geq y)\}^{s(\gamma)}} \int_0^\infty \{\mathbb{P}(X > x, Y \geq y)\}^{s(\gamma)-1} 1\!\!I(X > x, Y \geq y) dx - \frac{s(\gamma)\xi_\gamma(y)}{\mathbb{P}(Y \geq y)} 1\!\!I(Y \geq y).$$

*(ii) For any subset $\mathcal{Y} \subset \mathbb{R}^q_+$ such that $\inf_{y \in \mathcal{Y}} P(Y \geq y) > 0$, the process $\sqrt{n}\left(\hat{\xi}_\gamma(\cdot) - \xi_\gamma(\cdot)\right)$ converges in distribution in the space of bounded functions on $\mathcal{Y}$ to a $q$-dimensional zero mean Gaussian process indexed by $y \in \mathcal{Y}$ with covariance function*

$$\Sigma_{k,l} = \mathbb{E}\left[\mathbb{S}_\gamma(y^k, Y, X)\, \mathbb{S}_\gamma(y^l, Y, X)\right].$$

*Proof.* Let $m = s(\gamma)$. When $m = 1, 2, \ldots$ the two results (i)-(ii) are given respectively by Theorem 3.1 and Appendix B in Cazals *et al.* (2002). In fact, it is not hard to verify that the proofs of these results remain valid even when the trimming parameter $m$ is not an integer. $\square$

The conditional distribution function $F_y$ even does not need to be continuous, which is not the case for the empirical conditional quantiles $\hat{Q}_\gamma(y) = \hat{F}_y^{-1}(\gamma)$ whose asymptotic normality requires at least the differentiability of $F_y$ at $Q_\gamma(y)$ with a strictly positive derivative (see Aragon et al (2002) for the pointwise convergence and Daouia *et al.* (2008) for the functional convergence).

Next we show that if the FDH estimator $\hat{\varphi}(y)$ converges in distribution, then for a specific choice of $\gamma$ as a function of $n$, $\hat{\xi}_\gamma(y)$ estimates the true full cost function $\varphi(y)$ itself and converges in distribution as well to the same limit as $\hat{\varphi}(y)$ and with the same scaling.

**Theorem 2.** *Suppose the support of* $(Y, X)$ *is compact. If* $a_n^{-1}(\hat{\varphi}(y) - \varphi(y)) \xrightarrow{d} W_{\rho(y)}$*, then* $a_n^{-1}\left(\hat{\xi}_{\gamma_y(n)}(y) - \varphi(y)\right) \xrightarrow{d} W_{\rho(y)}$ *provided*

$$\gamma_y(n) \leq 1 - \left\{1 - \frac{1}{n\hat{\mathbb{P}}(Y \geq y)}\right\}^{\frac{\log(2)}{(\beta+1)\log(Cn)}} \quad or \quad \gamma_y(n) \leq 1 - \exp\left\{\frac{(1+o(1))\log(1/2)}{(\beta+1)n\log(Cn)\mathbb{P}(Y \geq y)}\right\},$$

*with* $\beta > 0$ *such that* $a_n n^\beta \to \infty$ *as* $n \to \infty$*, and* $C$ *being a positive constant.*

*Proof.* We have $a_n^{-1}\left(\hat{\xi}_\gamma(y) - \varphi(y)\right) = a_n^{-1}(\hat{\varphi}(y) - \varphi(y)) + a_n^{-1}\left(\hat{\xi}_\gamma(y) - \hat{\varphi}(y)\right)$. Let $N_y = \sum_{i=1}^n 1\!\!I(Y_i \geq y) = \sum_{i=1}^n 1\!\!I(Z_i^y > 0)$. It is easily seen from (5) that

$$\left(\hat{\xi}_\gamma(y) - \hat{\varphi}(y)\right) = \sum_{j=1}^{N_y}\left\{\frac{N_y - j}{N_y}\right\}^{s(\gamma)}\left(Z^y_{(n-N_y+j+1)} - Z^y_{(n-N_y+j)}\right).$$

The support of $(Y, X)$ being compact, the range of $Z^y$ is bounded and so $\left(\hat{\xi}_\gamma(y) - \hat{\varphi}(y)\right) = O\left(n\left\{1 - \frac{1}{N_y}\right\}^{s(\gamma)}\right)$. Then, for the term $a_n^{-1}\left(\hat{\xi}_\gamma(y) - \hat{\varphi}(y)\right)$ to be $o_p(1)$ as $n \to \infty$, it is sufficient to choose $\gamma = \gamma_y(n)$ such that $\left\{1 - \frac{1}{N_y}\right\}^{s(\gamma_y(n))} = O\left(n^{-(\beta+1)}\right)$ or equivalently $\left\{1 - \frac{1}{N_y}\right\}^{s(\gamma_y(n))} \leq (Cn)^{-(\beta+1)}$ with $C > 0$ being a constant and $\beta > 0$ is such that $a_n^{-1}n^{-\beta} = o(1)$ as $n \to \infty$. Whence the condition $s(\gamma_y(n)) \geq \frac{(\beta+1)\log(Cn)}{\log\left(1-\frac{1}{N_y}\right)\log(1/2)}$, or equivalently, $\gamma_y(n) \leq 1 - \left\{1 - \frac{1}{N_y}\right\}^{\frac{\log 2}{(\beta+1)\log(Cn)}}$. Since $\log\left(1 - \frac{1}{N_y}\right) \sim -\frac{1}{N_y} \sim -\frac{1}{n\mathbb{P}(Y \geq y)}$ as $n \to \infty$, with probability 1, it suffices to assume that $s(\gamma_y(n)) \geq \frac{(\beta+1)n\log(Cn)\mathbb{P}(Y\geq y)}{\log(2)(1+o(1))}$, or equivalently, $\gamma_y(n) \leq 1 - \exp\left\{\frac{(1+o(1))\log(1/2)}{(\beta+1)n\log(Cn)\mathbb{P}(Y\geq y)}\right\}$. $\square$

Note that the condition of Theorem 2 on the order $\gamma_y(n)$ is also provided in the proof in terms of $s\left(\gamma_y(n)\right)$ and reads as follows:

$$s(\gamma_y(n)) \geq \frac{(\beta+1)\log(Cn)}{\log\left(1 - \frac{1}{n\hat{\mathbb{P}}(Y\geq y)}\right)\log(1/2)} \quad \text{or} \quad s(\gamma_y(n)) \geq \frac{(\beta+1)n\log(Cn)\mathbb{P}(Y \geq y)}{\log(2)\left(1 + o(1)\right)}. \quad (6)$$

Note also that in the particular case considered by Cazals *et al.* (2002) where the joint density of $(Y, X)$ is strictly positive at the upper boundary and the frontier function $\varphi(y)$ is continuously differentiable in $y$, the convergence rate $a_n$ satisfies $a_n^{-1} \sim (n\ell_y)^{1/\rho(y)}$ with $\rho(y) = 1 + q$ and $\ell_y > 0$ being a constant (see Park *et al.* (2000)). In this case, the condition $a_n n^{\beta} \to \infty$ reduces to $\beta > 1/(1+q)$.

It should be clear that the main results of Cazals *et al.* (2002) are corollaries of our Theorems 1 and 2. Indeed, when the real parameter $s(\gamma) \in [1, \infty)$ in our approach is taken to be a positive integer $m = 1, 2, \ldots$, we recover Theorems 3.1 and 3.2 of Cazals *et al.* (2002). However, we hope to have shown that the sufficient condition $m_y(n) = O\left(\beta n \log(n) \mathbb{P}(Y \geq y)\right)$ of Cazals *et al.* (2002, Theorem 3.2) on the trimming parameter $m_y(n) \equiv s(\gamma_y(n))$ is somewhat premature and should be replaced by (6).

Alternative estimators of the conditional extremile $\xi_\gamma(y)$ can be constructed from expression (3). Instead of the sample extremile (4), one may estimate the ordinary extremile $\xi_{Z^y}(\gamma)$ by

$$\tilde{\xi}_{Z^y}(\gamma) = \frac{1}{n} \sum_{i=1}^{n} J_\gamma\left(\frac{i}{n+1}\right) Z^y_{(i)}.$$

This estimator which is in fact first-order equivalent with $\hat{\xi}_{Z^y}(\gamma)$ (see Daouia and Gijbels, 2009) leads to the alternative estimator $\tilde{\xi}_\gamma(y)$ of $\xi_\gamma(y)$ defined as $\tilde{\xi}_\gamma(y) = \tilde{\xi}_{Z^y}(\gamma)/\{\hat{\mathbb{P}}(Y \geq y)\}^{s(\gamma)}$. In the particular case considered by Cazals *et al.* (2002) where $s(\gamma)$ is only a positive integer, the statistic

$$\xi^*_{Z^y}(\gamma) = \frac{s(\gamma)}{n} \sum_{i=1}^{n-s(\gamma)+1} \left( \prod_{j=1}^{s(\gamma)-1} \frac{(n-i+1-j)}{(n-j)} \right) Z^y_{(i)}$$

is an unbiased estimator of the ordinary extremile $\xi_{Z^y}(\gamma)$ with the same asymptotic normal distribution as $\hat{\xi}_{Z^y}(\gamma)$ and $\tilde{\xi}_{Z^y}(\gamma)$ (Daouia and Gijbels, 2009). This provides an attractive estimator $\xi^*_\gamma(y) = \xi^*_{Z^y}(\gamma)/\{\hat{\mathbb{P}}(Y \geq y)\}^{s(\gamma)}$ for the order-$s(\gamma)$ expected minimum input function $\xi_\gamma(y) \equiv \varphi_{s(\gamma)}(y)$.

# 4 Empirical illustration

Let $N_y$ be the number of the $Y_i$ observations greater than or equal to $y$, *i.e.* $N_y = \sum_{i=1}^{n} 1\!\!I(Y_i \geq y)$, and, for $j = 1, \ldots, N_y$, denote by $X^y_{(j)}$ the $j$th order statistic of the $X_i$'s such that $Y_i \geq y$. It is then clear that $X^y_{(j)} = Z^y_{(n-N_y+j)}$ for each $j = 1, \ldots, N_y$, and the estimator $\hat{\xi}_\gamma(y)$ can be easily computed as

$$\hat{\xi}_\gamma(y) = \frac{\hat{\xi}_{Z^y}(\gamma)}{\{n^{-1}N_y\}^{s(\gamma)}} = X^y_{(1)} + \sum_{j=1}^{N_y-1} \left(1 - \frac{j}{N_y}\right)^{s(\gamma)} \left\{X^y_{(j+1)} - X^y_{(j)}\right\}.$$

This estimator always lies above the FDH $\hat{\varphi}(y) = X^y_{(1)}$ and so is more robust to extremes and outliers. Moreover, being a linear function of the data, $\hat{\xi}_\gamma(y)$ suffers less than the empirical $\gamma$th quantile $\hat{Q}_\gamma(y)$ to sampling variability or measurement errors in the extreme values $X^y_{(j)}$. The quantile-based frontier only depends on the frequency of tail costs and not on their values. Consequently, it could be too liberal (insensitive to the magnitude of extreme costs $X^y_{(j)}$) or too conservative following the value of $\gamma$. In contrast, putting more weight to high and low observations in the input-orientation, the extremile-based frontier is always sensible to desirable extreme costs. Nevertheless, being a linear function of all the data points (L-statistic), it remains resistant in the sense that it could be only attracted by outlying observations without envelopping them.

We first apply Theorem 2 in conjunction with these sensitivity and resistance properties to estimate the optimal cost of the delivery activity of the postal services in France. The data set contains information about 9521 post offices $(Y_i, X_i)$ observed in 1994, with $X_i$ being the labor cost (measured by the quantity of labor which represents more than 80% of the total cost of the delivery activity) and the output $Y_i$ is defined as the volume of delivered mail (in number of objects). See Cazals *et al.* (2002) for more details. Here, we only use the $n = 4000$ observations with the smallest inputs $X_i$ to illustrate the extremile-based estimator $\hat{\xi}_{\gamma_y(n)}(y)$ of the efficient frontier $\varphi(y)$. The important question of how to pick out the order $\gamma_y(n)$ in practice can be addressed as follows.

We know that the condition of Theorem 2 provides an upper bound on the value of $\gamma_y(n)$. Remember also that in most situations described so far in the econometric literature on frontier analysis, the joint density of $(Y, X)$ is supposed to be strictly positive at the frontier. In this case, the upper bound for $\gamma_y(n)$ is given by

$$\gamma_{(C)} = 1 - \left\{1 - \frac{1}{N_y}\right\}^{\frac{(1+q)\log(2)}{(2+q)\log(Cn)}},$$

where the number of outputs $q$ equals here 1 and the positive constant $C$ should be selected so that $\log(Cn) \neq 0$, *i.e.*, $C > 1/n$. The practical question now is how to choose $C > 0.00025$ in such a way that $\hat{\xi}_{\gamma(C)}$ provides a reasonable estimate of the frontier function $\varphi$. This can be achieved by looking to Figure 1 which indicates how the percentage of points below the curve $\hat{\xi}_{\gamma(C)}$ decreases with the constant $C$. The idea is to choose values of $C$ for which the frontier estimator $\hat{\xi}_{\gamma(C)}$ is sensible to the magnitude of desirable extreme post offices and, at the same time, is robust to outliers (or at least not being drastically influenced by outliers as is the case for the FDH estimator).

The evolution of the percentage in Figure 1 has clearly an "L" structure. This deviation should appear whatever the analyzed data set due to both sensitivity and resistance properties of extremiles. The percentage falls rapidly until the circle, *i.e.*, for $C \leq 0.000257$.

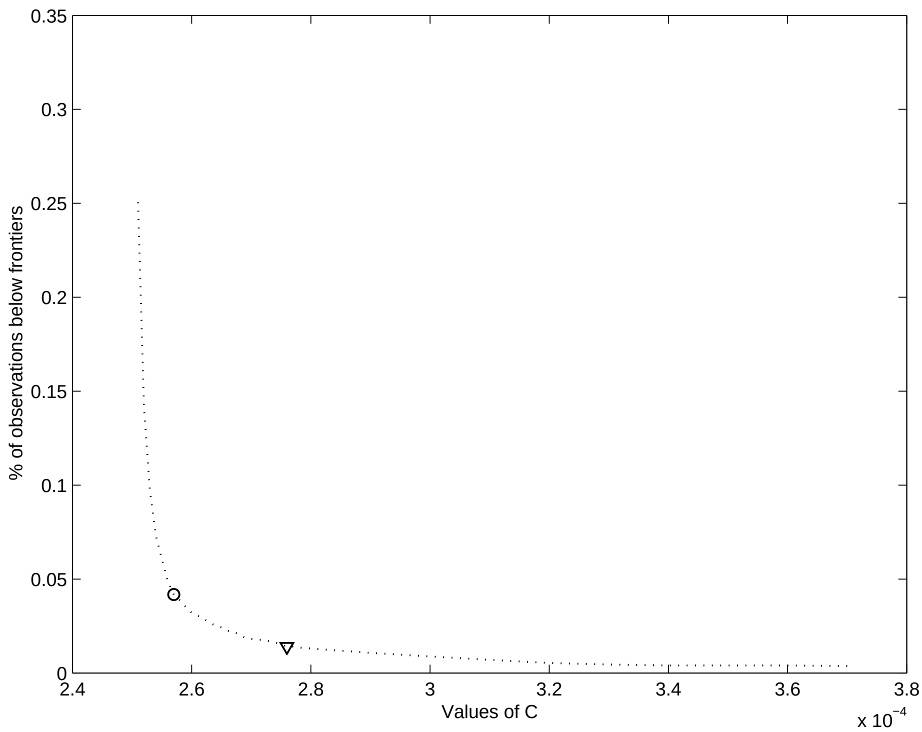


Figure 1: *Evolution of the percentage of observations below the frontier* $\hat{\xi}_{\gamma(C)}$ *with* $C$.

This means that the observations below the frontiers $\{\hat{\xi}_{\gamma(C)} : C < 0.000257\}$ are not really extreme and could be interior observations to the cloud of data points. So it is not judicious to select $C < 0.000257$. In contrast, the percentage becomes very stable from the triangle on (*i.e.* $C \geq 0.000276$), where precisely 1.4% of the 4000 observations are left out. This means that these few 1.4% observations are really very extreme in the input-direction and could be outlying or perturbed by noise. Although the frontier $\hat{\xi}_{\gamma(C)}$, for $C \geq 0.000276$, is resistant to these suspicious extremes, it can be severely attracted by them due to its sensitivity. This suggests to choose $C < 0.000276$. Thus, our strategy leads to the choice of a constant $C$ ranging over the interval $[0.000257, 0.000276)$ where the decrease of the percentage is rather moderate.

The two extreme (lower and upper) choices of the frontier estimator $\hat{\xi}_{\gamma(C)}$ are graphed in Figure 2, where the solid line corresponds to the lower bound $C_\ell = 0.000257$ and the dotted line corresponds to the upper bound $C_u = 0.000276$. The frontier estimator $\hat{\xi}_{\gamma(C)}$ in dashed line corresponds to the medium value $C_m = (C_\ell + C_u)/2$. The obtained curves are quite satisfactory.

Let us now test this strategy on a data set of 100 observations simulated following the

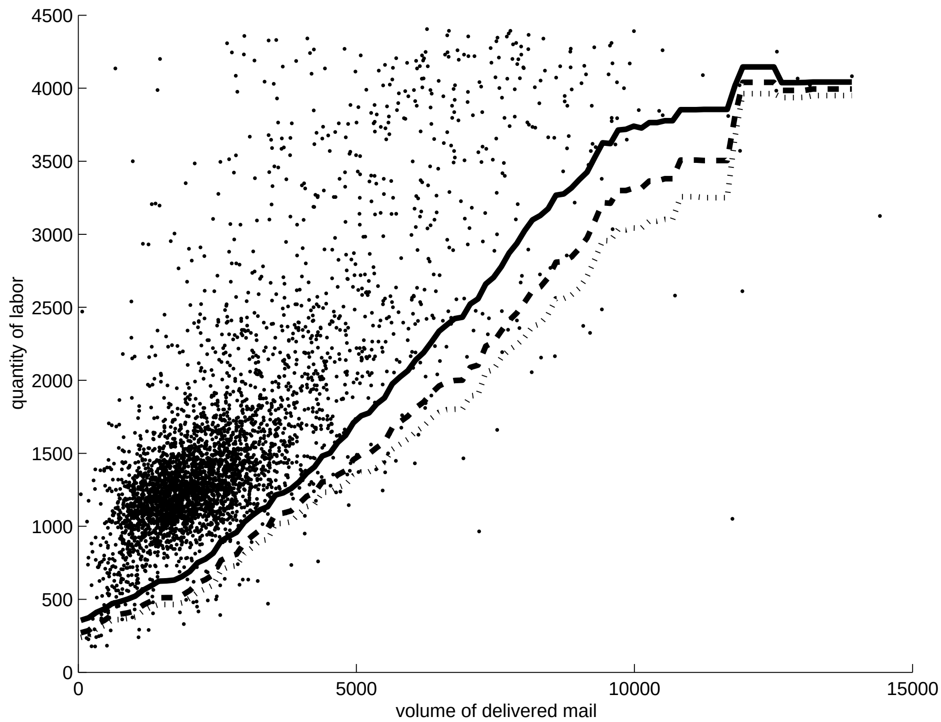


Figure 2: $\hat{\xi}_{\gamma(C_\ell)}$ *in solid line,* $\hat{\xi}_{\gamma(C_u)}$ *in dotted line and* $\hat{\xi}_{\gamma(C_m)}$ *in dashed line.*

model

$$Y = \exp(-5 + 10X)/(1 + \exp(-5 + 10X)) \exp(-U),$$

where $X$ is uniform on $(0, 1)$ and $U$ is exponential with mean $1/3$. Five outliers indicated as "*" in Figure 3, right-hand side, are added to the cloud of data points (here $n = 105$). The picture on the left-hand side of Figure 3 provides the evolution of the percentage of observations below the frontier $\hat{\xi}_{\gamma(C)}$ with $C$. This percentage falls rapidly until the circle (*i.e.*, for $C \leq 0.0122$) and then becomes very stable suggesting thus the value 0.0122 for the constant $C$. The resulting estimator $\hat{\xi}_{\gamma(.0122)}$ and the true frontier $\varphi$ are superimposed in Figure 3, right-hand side. The frontier estimator $\hat{\xi}_{\gamma(.0122)}$ (in solid line) has a nice behavior: it is somewhat affected by the five outliers, but remains very resistant.

We did the same exercise without the five outliers. The results are displayed in Figure 4. The percentage of observations below the extremile-based frontiers becomes stable from the circle on (*i.e.*, for $C \geq 0.0167$) and so it is enough to choose the value 0.0167 for the constant $C$. One can also select $C$ in the interval $[C_\ell = 0.0167, C_u = 0.0244)$ which corresponds to the range of points between the circle and the triangle. As expected, in absence of outliers, both estimators $\hat{\xi}_{\gamma(C_\ell)}$ (solid line) and $\hat{\xi}_{\gamma(C_u)}$ (dashed line) are very close from the FDH frontier (*i.e.*, the largest step and nondecreasing curve envelopping below all observations). However,

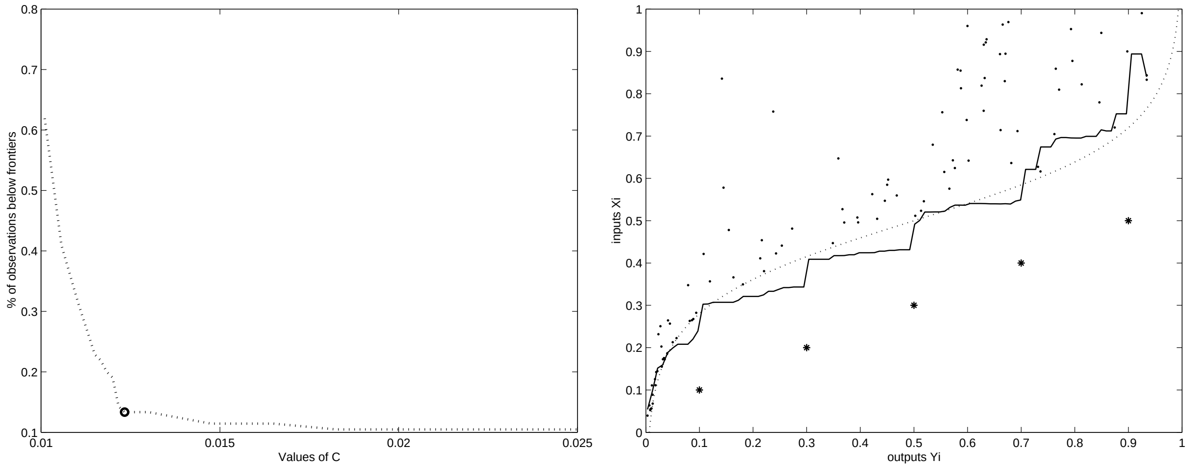


Figure 3: *Left-hand side, the percentage curve. Right-hand side, the frontiers* $\varphi$ *and* $\hat{\xi}_{\gamma(.0122)}$ *superimposed (in dotted and solid lines respectively). Five outliers included as "*".*

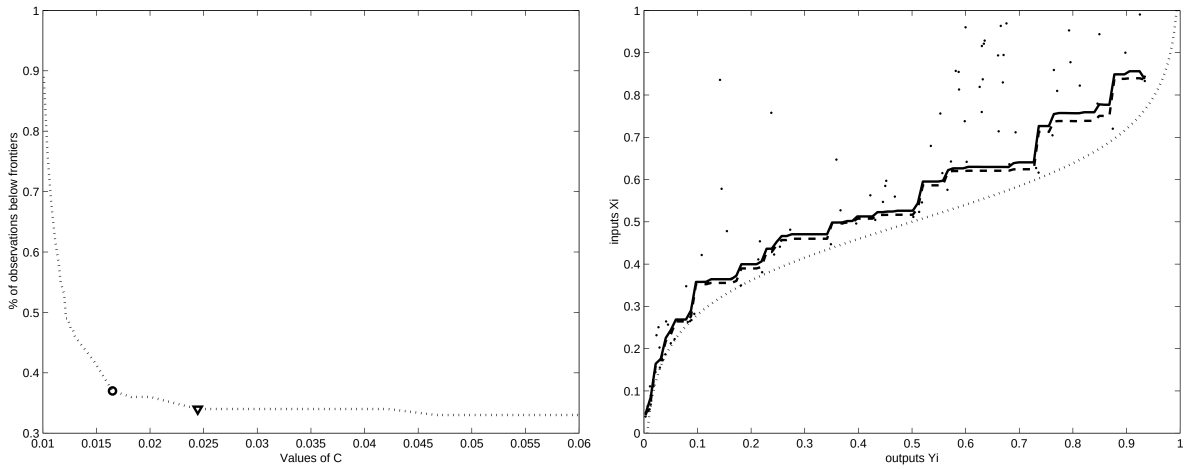


Figure 4: *As above without outliers. Here* $\hat{\xi}_{\gamma(C_\ell)}$ *in solid line and* $\hat{\xi}_{\gamma(C_u)}$ *in dashed line.*

as desired, here also $\hat{\xi}_{\gamma(C_\ell)}$ and $\hat{\xi}_{\gamma(C_u)}$ capture the shape of the efficient boundary of the cloud of data points without envelopping the most extreme observations.

# 5 Conclusions

Instead of estimating the full cost frontier we rather propose in this paper to estimate a boundary well inside the production set $\Psi$ but near its optimal frontier by using extremiles of the same non-standard conditional distribution considered by Cazals *et al.* (2002) and Aragon *et al.* (2005). The extremile cost function of order $\gamma \in (0,1)$ is proportional to a specific conditional probability-weighted moment. It defines a natural concept of a partial cost frontier instead of the $m$-trimmed frontier suggested by Cazals *et al.* (2002). The concept is attractive because the "trimming" is continuous in terms of the transformed index $s(\gamma)$, where $s(\gamma) \in [1, \infty)$, whereas $m \in \{1, 2, \ldots\}$. In the particular case where $s(\gamma)$ is discrete (*i.e.* $s(\gamma) = 1, 2, \ldots$), the corresponding $\gamma$th extremile-based function coincides with the expected minimum input function of order $m = s(\gamma)$. So the family of order-$m$ frontiers of Cazals *et al.* (2002) is a subclass of the order-$\gamma$ extremile frontiers. As a matter of fact, the discrete order $m$ being replaced with the continuous index $s(\gamma)$, the general class of extremile-based cost functions can be viewed as a *fractional* variant of expected minimum input functions. This new class benefits from a similar "benchmark" interpretation as in the discrete case. Moreover, the continuous trimming in terms of the new order $\gamma$ allows the partial $\gamma$th extremile boundaries to cover the attainable set $\Psi$ entirely giving thus a clear information of the production performance, which is not the case for the discrete order-$m$ frontiers.

The class of extremile-type cost functions characterizes the production process in much the same way the quantile-type cost functions introduced by Aragon *et al.* (2005) do. Moreover, while the $\gamma$th quantile-type function can be expressed as the median of a specific power of the underlying conditional distribution, the $\gamma$th extremile-type function is given by its expectation. Being determined solely by the tail probability $\gamma$, the $\gamma$th quantile-based cost frontier may be unaffected by desirable extreme observations, whereas the $\gamma$th extremile-based cost frontier is always sensible to the magnitude of extremes for any order $\gamma$. In contrast, when the $\gamma$-quantile frontier becomes very non-robust (breaks down) at $\gamma \downarrow 0$, the $\gamma$-extremile frontier being an L-functional, is more resistant to outliers. So the class of extremile-based cost frontiers steers an advantageous middle course between the extreme behaviors of the quantile-based cost frontiers. We also show in the standard situation in econometrics where the joint density of $(Y, X)$ has a jump at the frontier that the $\gamma$th quantile frontier is asymptotically closer (as $\gamma \downarrow 0$) to the true full cost frontier than is the $\gamma$th extremile frontier when $q \leq 2$, but the latter is more spread than the former when $q > 2$.

The new concept of a $\gamma$th extremile-based cost frontier is motivated via several angles, which reveals its specific merits and strength. Its various equivalent explicit formulations result in several estimators which satisfy similar asymptotic properties as the nonparametric expected minimum input and quantile-type frontiers. Nevertheless, the underlying conditional distribution function even does not need to be continuous, which is not the case for the empirical conditional quantiles whose asymptotic normality requires at least the differentiability of this distribution function with a strictly positive derivative at the conditional quantile. On the other hand, by choosing the order $\gamma$ as an appropriate function of the sample size $n$, we derive an estimator of the true full cost frontier having the same limit distribution as the conventional FDH estimator. Combining the sensitivity and resistance properties of this frontier estimator with the theoretical conditions on the order $\gamma = \gamma(n)$, we show how to pick out in practice reasonable values of $\gamma(n)$. Our empirical rule is illustrated through a simulated and a real data set providing remarkable results. It should be clear that, unlike the approaches of Cazals *et al.* (2002) and Daouia and Simar (2007), the conditional extremile approach is not extended here to the full multivariate case (multi-inputs and multi-outputs). This problem is worth investigating.

# 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] Chan, L.K. (1967), On a characterization of distributions by expected values of extreme order statistics, *Am. Math. Mthly*, 74, 950-951.

[4] Charnes, A., Cooper, W.W. and Rhodes, E. (1978), Measuring the inefficiency of decision making units, *European Journal of Operational Research*, 2, 429-444.

[5] Daouia, A., J-P. Florens and L. Simar (2008), Functional Convergence of Quantile-type Frontiers with Application to Parametric Approximations, *Journal of Statistical Planning and Inference*, 138, 708–725.

[6] Daouia, A., J-P. Florens and L. Simar (2010), Frontier estimation and extreme value theory, *Bernoulli*, 16 (4), 1039-1063.

[7] Daouia, A. and I. Gijbels (2011), Robustness and inference in nonparametric partial-frontier modeling, *Journal of Econometrics*, 161, 147-165.

[8] Daouia, A. and I. Gijbels (2009), Extremiles, manuscript.

[9] Daouia, A. and L. Simar (2007), Nonparametric Efficiency Analysis: A Multivariate Conditional Quantile Approach, *Journal of Econometrics*, 140(2), 375–400.

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

[11] 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.

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

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

[14] Koopmans, T.C. (1951), An analysis of production as an efficient combination of activities. In T.C. Koopmans (ed.) *Activitity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph 13, Wiley.

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

[16] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons Inc.

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

[18] Shorack, G.R. (2000). *Probability for Statisticians*. Springer, New York.

[19] Simar, L. and P. Wilson (2008). Statistical Inference in Nonparametric Frontier Models: recent Developments and Perspectives, in The Measurement of Productive Efficiency, 2nd Edition, Harold Fried, C.A.Knox Lovell and Shelton Schmidt, editors, Oxford University Press.