Nonparametric Frontier Estimation : A Conditional Quantile-based Approach

Y. ARAGON, A. DAOUIA, C. THOMAS-AGNAN[*]

Groupe de Recherche en Economie Mathématique et Quantitative
et Laboratoire de Statistique et Probabilités, Université de Toulouse
33(0)561128548

Septembre 10, 2003

Abstract

In frontier analysis, most of the nonparametric approaches (FDH,DEA) are based on envelopment ideas and their statistical theory is now mostly available. However, by construction, they are very sensitive to outliers. Recently, a robust nonparametric estimator has been suggested by Cazals, Florens and Simar (2002). In place of estimating the full frontier, they propose rather to estimate an expected frontier of order m. Similarly, we construct a new nonparametric estimator of the efficient frontier. It is based on conditional quantiles of an appropriate distribution associated with the production process. We show how these quantiles are interesting in efficiency analysis. We provide the statistical theory of the obtained estimators. We illustrate with some simulated examples and a frontier analysis of French post offices, showing the advantage of our estimators compared with the estimators of the expected maximal output frontiers of order m.

[*]Key words : conditional quantile, frontier, production function, nonparametric estimation, resistance to outliers. E-mail addresses : aragon@cict.fr, daouia@cict.fr, cthomas@cict.fr.

1 Introduction

An important problem in productivity and efficiency analysis is to characterize and to estimate the production frontier, i.e., the set of the most efficient production process. The idea is to analyze how firms combine their inputs to produce in an efficient way the output. We are then interested in the production frontier because it represents a reasonable benchmark value or reference frontier. Let us introduce the basic concepts and notation.

According to economic theory (Koopmans, 1951, Debreu, 1951, Shephard, 1970), the production set, where the activity is described through a set of p inputs $x \in \mathbb{R}_+^p$ used to produce an univariate output $y \in \mathbb{R}_+$, is defined as the set of physically attainable points (x, y)

$$\Psi = \{(x, y) \in \mathbb{R}_+^{p+1} | \ x \text{ can produce } y\}.$$

This set can be described mathematically by its sections $Y(x) = \{y \in \mathbb{R}_+ | \ (x, y) \in \Psi\}$ where, for any level of inputs x, the requirement set $Y(x)$ represents the set of all outputs that a firm can produce using x as inputs. Assuming that Ψ is compact, the maximal achievable level of output for a given level of inputs x defines the output-efficient function $\partial Y(x) = \max Y(x)$. From an economic point of view, this function is supposed monotone nondecreasing and it is then called the production function and its graph, which represents the efficient boundary of Ψ, is called the production frontier. Different other assumptions can be assumed on Ψ like free disposability, i.e., if $(x, y) \in \Psi$ then $(x', y') \in \Psi$ for any $x' \geq x$ and $y' \leq y$; or convexity, i.e., every convex combination of feasible production plans is also feasible; or no free lunch, i.e., for all $y > 0$ we have $y \notin Y(0)$; ... (see, e.g., Shephard, 1970).

The production process, which generates observations $\mathcal{X}_n = \{(X_i, Y_i) | \ i = 1, \cdots, n\}$ is defined, e.g., through the joint distribution of a random vector (X, Y) on $\mathbb{R}_+^p \times \mathbb{R}_+$, where X represents the inputs and Y is the output. In the case where Ψ is equal to the support of the distribution of (X, Y), another way for defining the production frontier is given as follows. The production function, which we denote from now on by φ, is characterized for a given level of inputs x by the upper boundary of the support of the conditional distribution of Y given $X \leq x$, i.e.,

$$\varphi(x) = \sup\{y \in \mathbb{R}_+ | \ F(y/x) < 1\}, \tag{1}$$

where $F(\cdot/x) = F(x, \cdot)/F_X(x)$ is the conditional distribution function of Y given $X \leq x$, with F being the joint distribution function of (X, Y) and F_X the marginal distribution function of X. It is supposed here that $F_X(x) > 0$ or that x is an interior point of the support of the distribution of X. The inequality $X \leq x$ has to be understood componentwise. As a matter of fact, the function φ is the smallest monotone nondecreasing function which is larger than or equal to the output-efficient function $\partial Y(\cdot)$. Its graph defines the production frontier. If the efficient boundary of Ψ is monotone nondecreasing (a quite reasonable assumption in practice), it coincides with the production frontier. So, we have, in some sense, just reparametrized the definition of the efficient frontier of Ψ. This new formulation of the production frontier is due to Cazals, Florens and Simar (2002).

A large amount of literature is devoted to the estimation of the production frontier from a random sample of production units $\mathcal{X}_n$. Two different approaches have been mainly developed: the deterministic frontier models which suppose that with probability one, all

the observations in $\mathcal{X}_n$ belong to Ψ, and the stochastic frontier models where random noise allows some observations to be outside of Ψ.

In deterministic frontier models, there are mainly two nonparametric methods based on envelopment techniques: the free disposal hull (FDH) and the data envelopment analysis (DEA). The FDH estimator was introduced by Deprins, Simar and Tulkens (1984) and relies only on the free disposability assumption on Ψ. The DEA estimator, which was initiated by Farrell (1957) and popularized as linear programming estimator by Charnes, Cooper and Rhodes (1978), requires stronger assumptions: it relies on the free disposability assumption and the convexity of Ψ. Note that the convexity assumption is widely used in economics, but it is not always valid. The production set might admit increasing returns to scale, i.e., the output increases faster than the inputs, or there might be lumpy goods, i.e., fractional values of inputs or outputs do not exist. Hence, the FDH is a more general estimator than the DEA. The asymptotic distribution of the FDH estimator was derived by Park, Simar and Weiner (2000) in the case of multivariate input and output, and the asymptotic distribution of the DEA estimator was derived by Gijbels, Mammen, Park and Simar (1999) in the univariate case. The statistical theory of these estimators is now mostly available. See Simar and Wilson (2000) for a recent survey of the available results.

In stochastic frontier models, where noise is allowed, only parametric restrictions on the shape of the frontier and on the data generating process allow identification of the noise from the efficiency frontier and estimation of this frontier. Aigner, Lovell and Schmidt (1977), Meeusen and van den Broek (1977), Olsen, Schmidt and Waldman (1980), Stevenson (1980) and Battese and Coelli (1988) specified a model for the production function and a specific distributional form for the error and then used maximum likelihood methods to estimate the parameters of the production function. These methods may lack robustness if the assumed distributional form does not hold. In particular, outliers in the data may unduly affect the estimate of the frontier function, or, it may be biased if the error structure is not correctly specified. Furthermore, as illustrated by Caudill, Ford and Groper (1995), heteroscedasticity in the error term, if not properly accounted for, can lead to significant biases when estimating the production frontier.

Nonparametric deterministic frontier models are very appealing because they rely on very few assumptions but, by construction, they are very sensitive to extreme values and to outliers. Recently, a robust nonparametric envelopment estimator of the production frontier has been suggested by Cazals, Florens and Simar (2002). They introduce the concept of expected maximal output frontier of order $m \in \mathbb{N}^*$, where $\mathbb{N}^*$ denotes the set of all integers $m \geq 1$. It is defined as the expected maximum achievable level of output among m firms drawn in the population of firms using less than a given level of inputs. Formally, for a fixed integer $m \in \mathbb{N}^*$ and a given level of inputs x, the frontier function of order m is defined as

$$\varphi_m(x) = E[\max(Y^1, \cdots, Y^m)] = \int_0^\infty (1 - [F(y/x)]^m) dy,$$

where $(Y^1, \cdots, Y^m)$ are m independent identically distributed random variables generated by the distribution of Y given $X \leq x$. Its nonparametric estimator is defined by

$$\widehat{\varphi}_{m,n}(x) = \int_0^\infty (1 - [\widehat{F}(y/x)]^m) dy,$$

3

where $\widehat{F}(y/x) = \widehat{F}(x,y)/\widehat{F}_X(x)$ is the empirical version of $F(y/x)$, with

$$\widehat{F}(x,y) = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x, Y_i \le y) \quad \text{and} \quad \widehat{F}_X(x) = \frac{1}{n}\sum_{i=1}^{n} 1(X_i \le x).$$

As pointed out in Cazals et al. (2002), the FDH estimator of the production function can be viewed as a plug-in estimator of $\varphi(x)$, where the unknown $F(y/x)$ in the formulae (1) has been replaced by its empirical analog $\widehat{F}(y/x)$. It is given by

$$\widehat{\varphi}_n(x) = \sup\{y \ge 0|\ \widehat{F}(y/x) < 1\} = \max_{i|X_i \le x} Y_i.$$

Due to the trimming nature of the order-m frontier, the estimator $\widehat{\varphi}_{m,n}(x)$ does not envelop all the data points, and so it is more robust to extreme values than the FDH estimator $\widehat{\varphi}_n(x)$. By choosing m appropriately as a function of the sample size n, $\widehat{\varphi}_{m,n}(x)$ estimates the production function $\varphi(x)$ itself while keeping the asymptotic properties of the FDH estimator.

Hendricks and Koenker (1992, p. 58) stated, "In the econometric literature on the estimation of production technologies, there has been considerable interest in estimating so called frontier production models that correspond closely to models for extreme quantiles of a stochastic production surface". The present paper can be viewed as the first work to actually implement the idea of Hendricks and Koenker: we construct a new nonparametric estimator of the production frontier which is more robust to extreme values than the standard DEA/FDH estimators and than the nonparametric estimator of Cazals et al. It is based on extreme quantiles of the conditional distribution of Y given $X \le x$. These non standard conditional quantiles define a natural concept of a partial production frontier in place of the m-trimmed frontier. The idea is nice and attractive, because here the "trimming" is continuous in terms of the order-α quantile where $\alpha \in [0,1]$. Quantile methods are known for their robustness. More precisely, conditional quantiles are not very sensitive to large observations in the output direction. We show that our new partial frontier and its resulting estimator share most of the properties of the order-m frontier and its estimator.

The paper follows the structure of Cazals et al. (2002) initially very closely, adapting their technique to the output oriented case and extending their basic ideas, thus sharing similar comments. It is organized as follows. Section 2 motivates our concept of quantile-frontier of order α and investigates its properties and its relation to the order-m frontier and to the true production frontier. In section 3, we define a nonparametric estimator of our order-α frontier, which is very easy to derive, very fast to compute and does not envelop all the observed data points. In section 4, we show that this estimator converges at the rate $\sqrt{n}$ and is asymptotically normally distributed. We also derive a nonparametric estimator of the efficient production frontier and analyze its asymptotic distribution. In section 5, a numerical illustration is proposed with some simulated examples and a data set on labor (as input) and mail volumes (as output) about 10.000 French post offices. We show how resistant to outliers our estimators are compared with the estimators of the expected maximal output frontiers of order m. Section 6 concludes the paper.

2   A New Concept of Production Frontier

Let $(\Omega, \mathcal{A}, P)$ be the probability space on which the vector of inputs X and the output variable Y are defined. In this approach, we define the attainable set Ψ to be the support of the joint distribution of (X, Y) and we will concentrate on the set $\Psi^* = \{(x, y) \in \Psi | F_X(x) > 0\}$ which contains the interior of Ψ.

From its definition, $\varphi(x)$, the value of the production function coincides with the order one quantile of the law of Y given $X \leq x$,

$$q_1(x) = \inf\{y \geq 0 | F(y/x) = 1\}.$$

This suggests to introduce a concept of production function of continuous order $\alpha \in [0, 1]$, as the quantile function of order α of the law of Y given that X does not exceed a given level of inputs. This function takes, for a given level of inputs x, the value

$$q_\alpha(x) := F^{-1}(\alpha/x) = \inf\{y \geq 0 | F(y/x) \geq \alpha\}.$$

This conditional quantile is the production threshold exceeded by $100(1 - \alpha)\%$ of firms that use less than the level x as inputs. The function $F^{-1}(./x)$ is the so called generalized inverse of $F(\cdot/x)$. If the distribution function $F(\cdot/x)$ is strictly increasing, its inverse coincides with the generalized inverse $F^{-1}(./x)$. Using this property, we easily obtain the following result.

Proposition 2.1. *Assume that for every x such that $F_X(x) > 0$, the conditional distribution function $F(\cdot/x)$ is strictly increasing on the support $[0, \varphi(x)]$. Then,*

$$\forall (x, y) \in \Psi^* \text{ we have } y = q_\alpha(x) \text{ with } \alpha = F(y/x). \tag{2}$$

From property (2), we see that any production unit (x, y) in Ψ^* belongs to some α-order quantile curve. Then unit (x, y) produces more than $100\alpha\%$ of all production units using inputs smaller than or equal to x and produces less than the $100(1 - \alpha)\%$ remaining units. Thus the quantile function $q_\alpha(x)$ quantifies the production efficiency of unit (x,y) by comparing it with all units which use the same level of inputs x as well as with those which use strictly less than x. This motivates our interest in the distribution of Y given $X \leq x$.

But the most attractive property of this quantile function is that it can be easily nonparametrically estimated without the drawbacks of the methods trying to estimate the frontier function itself: it will be less sensitive to noise, extreme values or outliers. This is developed in the next section.

As it is shown by property (2), the quantile curves $\{(x, q_\alpha(x)) | F_X(x) > 0\}$ cover the whole production set Ψ^*. As can be seen in the next proposition, this does not hold for expected order-m frontiers of Cazals et al. $\{(x, \varphi_m(x)) | F_X(x) > 0\}$.

Proposition 2.2. *Under the assumption of Proposition 2.1 and if we assume furthermore the free disposability of outputs, i.e.,*

$$y \in Y(x) \text{ and } y' \leq y \implies y' \in Y(x),$$

then the functions φ_m do not satisfy the following property

$$\forall (x, y) \in \Psi^*, \ \exists m \in \mathbb{N}^* \ s.t. \ y = \varphi_m(x).$$

Let us compare how the expected maximal production function and the quantile function can be useful in terms of practical efficiency analysis. Suppose a production unit uses a quantity of inputs x_0 and produces an output y_0, $\varphi_m(x_0)$ gives the expected maximum production among a fixed number of m firms using less than x_0 as inputs. This value indicates how efficient the unit (x_0, y_0) is, compared with these m units. This is achieved by comparing its level y_0 with the value of $\varphi_m(x_0)$. For this particular unit, we know that it belongs to a quantile frontier. The order of this frontier, which is known, gives the proportion of units that produce less than y_0 among all firms using less than x_0. Hence the quantile function gives a clearer indication on the production performance and it can be viewed as a reasonable benchmark value.

We can however establish an asymptotic relationship between the two families of production functions φ_m and q_α. Namely,

Proposition 2.3. *For every x such that the conditional distribution function $F(\cdot/x)$ is twice differentiable with first derivative $f(\cdot/x)$ strictly positive on the support $[0, \varphi(x)]$, we have as $m \to \infty$ and $\alpha \to 1$,*

$$\varphi_m(x) - q_\alpha(x) = \left\{ \frac{1}{f(\varphi(x)/x)} + (\alpha - 1)\left[\psi_x'(\alpha) + o(\alpha)\right] \right\}$$
$$\times O\left(m^{-3/4}(\log m)^{1/2}(\log \log m)^{1/4}\right) + o(\alpha) \qquad (3)$$

where $\psi_x'(\alpha) = -F''(q_\alpha(x)/x)/f^3(q_\alpha(x)/x)$.

From its definition, it is clear that for any fixed x such that $F_X(x) > 0$, $q_\alpha(x)$ is a monotone nondecreasing function of α. The limiting case when $\alpha \to 1$ is of particular interest. It converges to the efficient frontier: by letting m tend to infinity in (3) and using $\lim_{m\to\infty} \varphi_m(x) = \varphi(x)$, we obtain $\varphi(x) - q_\alpha(x) = o(\alpha)$ when $\alpha \to 1$. We can prove this property directly by using the monotonicity of quantiles $q_\alpha(x)$ with respect to α as indicated by the next proposition. Even more strongly it is shown, under some regularity conditions, that the order-α production function q_α converges uniformly to the true production function φ.

Proposition 2.4.

1. *For any fixed value of x such that $F_X(x) > 0$, we have $\lim_{\alpha\to 1} \nearrow q_\alpha(x) = \varphi(x)$.*

2. *Assume that for every $\alpha \in [0, 1]$, the quantile function $q_\alpha(.)$ is continuous on the interior of the support of X. Then for any compact K interior to the support of X,*

$$\sup_{x \in K} |q_\alpha(x) - \varphi(x)| \longrightarrow 0 \quad as \quad \alpha \nearrow 1.$$

The function q_α converges to a monotone nondecreasing function φ as $\alpha \to 1$, but it is not monotone nondecreasing itself unless we add the following assumption

$$\forall y \geq 0, \ \forall x_1 \leq x_2 \text{ s.t. } F_X(x_1) > 0, \text{ we have } F(y/x_1) \geq F(y/x_2). \qquad (4)$$

This assumption is not needed for all the results of this paper except for the next Proposition, but it appears to be quite reasonable: it says that the chance of producing less than a value y decreases if a firm uses more inputs. This assumption is necessary as well as sufficient.

Proposition 2.5. *The quantile function $x \mapsto q_\alpha(x)$ is monotone nondecreasing on the set $\{x \in \mathbb{R}_+^p | F_X(x) > 0\}$ for every order $\alpha \in [0, 1]$ if and only if the function $x \mapsto F(y/x)$ is monotone nonincreasing on the set $\{x \in \mathbb{R}_+^p | F_X(x) > 0\}$ for any output $y \in \mathbb{R}_+$.*

Note that the results established in Proposition 2.4 are very similar to those obtained for the order-m frontier. Indeed $\varphi_m(x)$ converges simply and uniformly to $\varphi(x)$ as $m \to \infty$. However for Proposition 2.5, Cazals et al. (2002, Theorem A.3) only prove that if assumption (4) holds then $\varphi_m(x)$ is monotone nondecreasing in x.

3 Nonparametric Estimation

To estimate the conditional quantile $q_\alpha(x)$, it is natural to use the conditional empirical quantile obtained by inverting the conditional empirical distribution function $\widehat{F}(\cdot/x)$,

$$\widehat{q}_{\alpha,n}(x) := \widehat{F}^{-1}(\alpha/x) = \inf\{y | \widehat{F}(y/x) \geq \alpha\}.$$

This estimator may be computed explicitly as follows. Let N_x be the number of observations X_i smaller than or equal to x, i.e., $N_x = \sum_{i=1}^n 1(X_i \leq x)$, and, for $j = 1, \cdots, N_x$, denote by $Y_{(i_j)}$ the j-th order statistic of the observations Y_i such that $X_i \leq x : Y_{(i_1)} \leq Y_{(i_2)} \leq \ldots \leq Y_{(i_{N_x})}$. We have, for x such that $N_x \neq 0$,

$$\widehat{F}(y/x) = \frac{\sum_{i|X_i \leq x} 1(Y_i \leq y)}{N_x} = \frac{\sum_{j=1}^{N_x} 1(Y_{(i_j)} \leq y)}{N_x}.$$

Hence,

$$\widehat{F}(y/x) = \begin{cases} 0 & \text{if} \quad y < Y_{(i_1)} \\ k/N_x & \text{if} \quad Y_{(i_k)} \leq y < Y_{(i_{\{k+1\}})}, \ 1 \leq k \leq N_x - 1 \\ 1 & \text{if} \quad y \geq Y_{(i_{N_x})}. \end{cases}$$

Therefore, we obtain for every $\alpha > 0$,

$$\widehat{\varphi}_n(x) = Y_{(i_{N_x})} = \max_{i|X_i \leq x} Y_i, \quad \widehat{q}_{\alpha,n}(x) = \begin{cases} Y_{(i_{\{\alpha N_x\}})} & \text{if} \quad \alpha N_x \in \mathbb{N}^* \\ Y_{(i_{\{[\alpha N_x]+1\}})} & \text{otherwise}, \end{cases} \tag{5}$$

where $[\alpha N_x]$ denotes the integral part of αN_x: the largest integer less than or equal to αN_x. The conditional empirical quantile $\widehat{q}_{\alpha,n}(x)$ is thus computed very easily as being the simple empirical quantile of observations y_i such that $x_i \leq x$.

For comparison, note that an exact formula is available in order to compute $\widehat{\varphi}_{m,n}(x)$. It is as simple as the formula (5) but restricted to the case of no ties among the inputs. The nonparametric estimator $\widehat{\varphi}_{m,n}(x)$ can also be approximated in practice by using a Monte-Carlo algorithm, even in the full multivariate case (several inputs and several outputs), which we do not treat in our paper. For instance, in the univariate output case, the Monte-Carlo method can be described as follows. For a given x, draw a random sample of size m with replacement among these y_i such that $x_i \leq x$ and denote this sample by $(y_b^1, \cdots, y_b^m)$. Then

compute $\widetilde{\varphi}_{b,m}(x) = \max_{i=1,\cdots,m}(y_b^i)$. Redo this for $b = 1, \cdots, B$ where B is large. Finally, we have

$$\widetilde{\varphi}_{m,n}(x) \approx \frac{1}{B} \sum_{i=1}^{B} \widetilde{\varphi}_{b,m}(x),$$

where the quality of the approximation can be tuned by the choice of B.

Note also that the relation between the order-m frontier and the true frontier remains valid with their estimators $\widehat{\varphi}_{m,n}(x)$ and $\widehat{\varphi}_n(x)$, i.e., $\lim_{m\to\infty} \nearrow \widehat{\varphi}_{m,n}(x) = \widehat{\varphi}_n(x)$. Similarly it is easily seen, for any fixed value of inputs x for which the estimator $\widehat{q}_{\alpha,n}(x)$ is well defined for every order $\alpha \in [0,1]$, that $\widehat{q}_{\alpha,n}(x)$ is a monotone nondecreasing function of α, and thus

$$\lim_{\alpha\to 1} \nearrow \widehat{q}_{\alpha,n}(x) = \widehat{\varphi}_n(x).$$

Note that even for large values of $\alpha < 1$, the estimator $\widehat{q}_{\alpha,n}(x)$ is less sensitive to extreme values than the FDH estimator $\widehat{\varphi}_n(x)$ which by construction, envelopes all the observations. The asymptotic theory is discussed in the next section. Note also that $\widehat{q}_{\alpha,n}(x)$ is not necessarily monotone nondecreasing with respect to x. Indeed, even if assumption (4) is assumed for the true conditional distribution function, it could happen that its empirical counterpart does not satisfy it. Of course we know that for large sample size n, it will mostly be the case.

Another property that $\widehat{q}_{\alpha,n}(x)$ shares with $\widehat{\varphi}_{m,n}(x)$ lies in the fact that both the nonparametric partial frontiers underestimate the full frontier $\varphi(x)$, for every order. In our case, for any value of inputs x for which $\varphi(x)$ and $\widehat{q}_{\alpha,n}(x)$ are well defined for any order $\alpha \in [0,1]$, we have

$$\widehat{q}_{\alpha,n}(x) \leq \widehat{\varphi}_n(x) \leq \varphi(x) \ a.s, \ \forall \alpha \in [0,1].$$

Indeed, since the production function $\varphi(\cdot)$ is monotone nondecreasing and greater than or equal to the efficient-output function $\partial Y(.)$, for each i such that $X_i \leq x$ we have almost surely $Y_i \leq \partial Y(X_i) \leq \varphi(X_i) \leq \varphi(x)$. Therefore $\widehat{\varphi}_n(x) = \max\{Y_i \mid X_i \leq x\} \leq \varphi(x) \ a.s.$ On the other hand we have $\widehat{q}_{\alpha,n}(x) \leq \widehat{q}_1(x) = \widehat{\varphi}_n(x)$ for every $\alpha \in [0,1]$.

4 Asymptotic Properties

For the unconditional case where ξ_α denotes the order-α quantile of a distribution function F_Z of a random variable Z, and $\widehat{\xi}_\alpha$ denotes the empirical quantile of a sample $(Z_1, \cdots, Z_n)$ of Z, if F_Z is differentiable in ξ_α and such that $F_Z'(\xi_\alpha) > 0$, the Bahadur representation theorem gives,

$$\sqrt{n}\left(\widehat{\xi}_\alpha - \xi_\alpha\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\alpha(1-\alpha)}{[F_Z'(\xi_\alpha)]^2}\right) \quad \text{as} \quad n \to \infty.$$

The direct application of this result to the distribution function $F_Z(\cdot) = F(\cdot/x)$ does not serve our purpose because our data do not yield a sample from this distribution. However, as for unconditional quantiles ξ_α, we focus here on pairs (x, α) which satisfy the following property

$$F(\cdot/x) \text{ is differentiable at } q_\alpha(x) \text{ s.t. } F'(q_\alpha(x)/x) > 0. \tag{6}$$

As a consequence of this property, $F(\cdot/x)$ is a bijective transformation from a neighborhood of $q_\alpha(x)$ onto a neighborhood of α. In particular the generalized inverse $F^{-1}(\cdot/x)$ is equal to the inverse of $F(\cdot/x)$ in the neighborhood of α. This property will be used in the proof of the following theorem which summarizes the asymptotic properties of our estimator $\widehat{q}_{\alpha,n}(x)$.

Theorem 4.1. *Let $\alpha \in (0,1)$ be a fixed order and let x be a fixed value such that $F_X(x) > 0$. Assume that the conditional distribution function $F(\cdot/x)$ is differentiable at $q_\alpha(x)$ with derivative $f(q_\alpha(x)/x) > 0$. Then,*

1. $\widehat{q}_{\alpha,n}(x) \overset{P}{\longrightarrow} q_\alpha(x) \quad as \quad n \to \infty;$

2. $\sqrt{n}\left(\widehat{q}_{\alpha,n}(x) - q_\alpha(x)\right) \overset{\mathcal{L}}{\longrightarrow} N\left(0, \sigma^2(x,\alpha)\right) \quad as \quad n \to \infty,$

where

$$\sigma^2(x,\alpha) = \alpha(1-\alpha)/\left(f^2(q_\alpha(x)/x)F_X(x)\right).$$

It is important to note that here also, the equivalent properties hold with the nonparametric estimator of the order-m frontier. Indeed it is easy to see that $\widehat{\varphi}_{m,n}(x)$ converges at the rate $\sqrt{n}$, is asymptotically unbiased and normally distributed: $\mathcal{L}\left(\sqrt{n}(\widehat{\varphi}_{m,n}(x) - \varphi_m(x))\right) \to N\left(0, \sigma^2(x,m)\right)$, where $\sigma^2(x,m) = E[\Gamma_m^2(x,X,Y)]$, with

$$\Gamma_m(x,X,Y) = \frac{m}{F_X(x)}1(X_i \leq x)\int_0^\infty F^{m-1}(y/x)\left[F(y/x) - 1(Y_i \leq y)\right]dy.$$

Moreover for a vector $(\widehat{\varphi}_{m,n}(x^1), \cdots, \widehat{\varphi}_{m,n}(x^r))$, the asymptotic r-variate normal distribution is obtained with asymptotic covariances given by $\Sigma_m(x^k, x^l) = E[\Gamma_m(x^k, X, Y)\Gamma_m(x^l, X, Y)]$. Similarly we have the following more general result for the estimator of the conditional quantile frontier function.

Theorem 4.2. *Let $x^1, \cdots, x^r$ be r levels of the input X which satisfy the assumption of Theorem 4.1 for a given order $\alpha \in (0,1)$. Then,*

$$\sqrt{n}\left(\widehat{q}_{\alpha,n}(x^1) - q_\alpha(x^1), \cdots, \widehat{q}_{\alpha,n}(x^r) - q_\alpha(x^r)\right) \overset{\mathcal{L}}{\longrightarrow} N_r(0, \Sigma_\alpha) \quad as \quad n \to \infty,$$

where

$$\Sigma_\alpha(x^k, x^l) = E\left[h_\alpha(x^k, X, Y)h_\alpha(x^l, X, Y)\right],$$

with

$$h_\alpha(x, X, Y) = \frac{\alpha 1(X \leq x) - 1(X \leq x, Y \leq q_\alpha(x))}{f(q_\alpha(x)/x)F_X(x)}.$$

In applied work, the variance factors $\sigma^2(x,\alpha)$ and $\Sigma_\alpha(x^k, x^l)$ must be estimated. For instance, consistent estimators for these factors can be obtained by plugging nonparametric estimators for the conditional density $f(\cdot/x)$ and the marginal distribution function $F_X(x)$ and taking the empirical mean for the expectation. Note that, as for unconditional quantiles, quantiles in the tail of the conditional distribution where the conditional density is low are inherently more difficult to estimate.

Note also that Cazals et al. (2002) obtained an asymptotic representation for their nonparametric estimator $\widehat{\varphi}_{m,n}(x)$ where the error term is uniform in x, whereas the error term involved in our approach depends on x (see the proof of Theorem 4.1). This can be explained as follows. Both $\widehat{\varphi}_{m,n}(x)$ and $\widehat{q}_{\alpha,n}(x)$ are representable as functionals of the empirical distribution function $\widehat{F}$. The corresponding functional for $\widehat{\varphi}_{m,n}(x)$ is differentiable in the Frechet sense w.r.t. the sup-norm, whereas that corresponding to $\widehat{q}_{\alpha,n}(x)$ is only differentiable in the Gâteaux sense. The uniformity of the error term allowed Cazals et al. (2002, Appendix B) to improve the convergence results of $\widehat{\varphi}_{m,n}(x)$ by a functional limit theorem, which is not the case in our approach.

It is also interesting to compare $\widehat{q}_{\alpha,n}(x)$ with the estimator of the standard conditional quantile of the distribution of Y given $X = x$. First note that this latter estimate requires a smoothing procedure which is not the case when the distribution of Y is conditioned by $X \leq x$. To compare their asymptotic variance, let us recall that the smooth estimators of the quantiles $\xi_\alpha(x)$ of the distribution function F_x of Y given $X = x$, are obtained by inverting a kernel estimator of F_x and satisfy the following result

$$\sqrt{nh_n} \left(\widehat{\xi}_\alpha(x) - \xi_\alpha(x) \right) \xrightarrow{\mathcal{L}} N \left(0, \mu^2(x, \alpha) \right),$$

where $\mu^2(x, \alpha) = \alpha(1 - \alpha)R(K)/f_x^2(\xi_\alpha(x))$, with $f_x(y) = \frac{\partial}{\partial y}F_x(y)$, $R(K) = \int K^2(u)du$, and K and h_n are respectively a kernel and a bandwidth satisfying some specific constraints (see, e.g., Berlinet et al., 2001, Ducharme et al., 1995).

Let us now turn to the convergence to the full frontier function $\varphi(x)$. We know that the estimator $\widehat{q}_{\alpha,n}(x)$ converges to the FDH estimator $\widehat{\varphi}_n(x)$ as $\alpha \to 1$. We also know from Park, Simar and Weiner (2000), that under regularity conditions, as $n \to \infty$, the FDH estimator $\widehat{\varphi}_n(x)$ converges to the true unknown frontier $\varphi(x)$. The idea is then to define α as a function of n such that $\alpha(n) \to 1$ and $\widehat{q}_{\alpha(n),n}(x) \to \varphi(x)$ as $n \to \infty$. We thus derive an estimator of the true production frontier $\varphi(x)$ and show in the next theorem that it converges to the same asymptotic distribution as the FDH estimator and as the nonparametric envelopment estimator of Cazals et al. (2002). The rate of convergence of the order $\alpha(n)$ to 1 is provided.

Theorem 4.3. *Assume that the joint probability measure of (X, Y) on the compact support Ψ provides a strictly positive density on the frontier $\{(x, \varphi(x)) | F_X(x) > 0\}$ and that the function φ is continuously differentiable. Then, for any x interior to the support of X we have as $n \to \infty$,*

$$n^{1/(p+1)} \left(\varphi(x) - \widehat{q}_{\alpha(n),n}(x) \right) \xrightarrow{\mathcal{L}} Weibull(\mu_x^{p+1}, p + 1)$$

where μ_x is a constant and the order $\alpha(n)$ is such that

$$n^{(p+2)/(p+1)} \left(1 - \alpha(n) \right) \longrightarrow 0 \quad as \quad n \to \infty.$$

The constant μ_x appearing in the limiting Weibull depends on the slope of the frontier and the value of the density near the frontier point $(x, \varphi(x))$. A consistent nonparametric estimator of this unknown constant has been proposed in Park et al. (2000).

Like the approach of Cazals et al. (2002), here also we loose the $\sqrt{n}$-consistency because we use $\widehat{q}_{\alpha,n}(x)$ to estimate the full frontier $\varphi(x)$ and not the partial frontier $q_\alpha(x)$.

5 Numerical Illustrations

In this section, we illustrate our procedure through some numerical examples with simulated and real data. In the simulation study, the observations are simulated according the same data generating process used in Simar (2003).

5.1 Example 1

We first consider a situation where the attainable set is convex. We simulate a sample of $n = 500$ data points (x_i, y_i) according to the Cobb-Douglas log-linear frontier model given by $Y = X^{0.5} * \exp(-U)$, where X is uniform on $(0, 1)$ and U is exponential with mean $1/3$. The true frontier function is $\varphi(x) = x^{0.5}$.

Figure 1 illustrates the simulated data and the quantile curves $\widehat{q}_{\alpha,n}$ and the expected maximal frontiers $\widehat{\varphi}_{m,n}$ ($B = 1000$) for several different values of α and m. In solid lines, the estimates $\widehat{q}_{\alpha,n}$ on the left with $\alpha = .7, .97, .98, .99, 1$, are compared with the estimates $\widehat{\varphi}_{m,n}$ on the right with $m = 2, 25, 50, 75, \infty$. The true frontier φ is in dash-dotted lines. The frontiers are monotone nondecreasing with respect to the order. For Figure 2, we add in the data set three outliers and we plot the same frontiers $\widehat{q}_{\alpha,n}$ and $\widehat{\varphi}_{m,n}$.
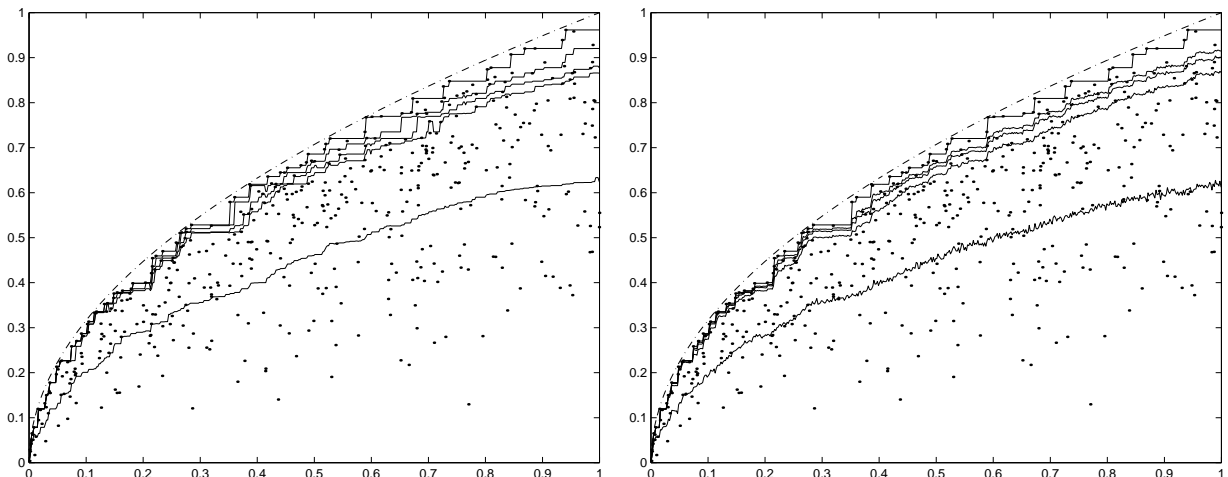

Figure 1: n=500: Comparison between $\widehat{q}_{\alpha,n}$ (left) and $\widehat{\varphi}_{m,n}$ (right), output vs input.

From Figures 1 and 2, it is clear that the frontiers $\widehat{\varphi}_{m,n}$ are more resistant to the three outliers than the FDH frontier, but they are less resistant to the outliers than the quantile frontiers of orders $\alpha < 1$. Indeed, the quantile frontier $\widehat{q}_{.99,n}$ is influenced by only one outlier and it comes back down immediately, whereas the frontiers $\widehat{\varphi}_{m,n}$ with $m = 25, 50, 75$ are attracted by all the outliers and moreover continue to grow after each jump. So in this particular example the frontier $\widehat{q}_{.99,n}$ is more robust to the outliers than the three frontiers $\widehat{\varphi}_{m,n}$, whereas it envelopes all these frontiers in absence of the three outliers.
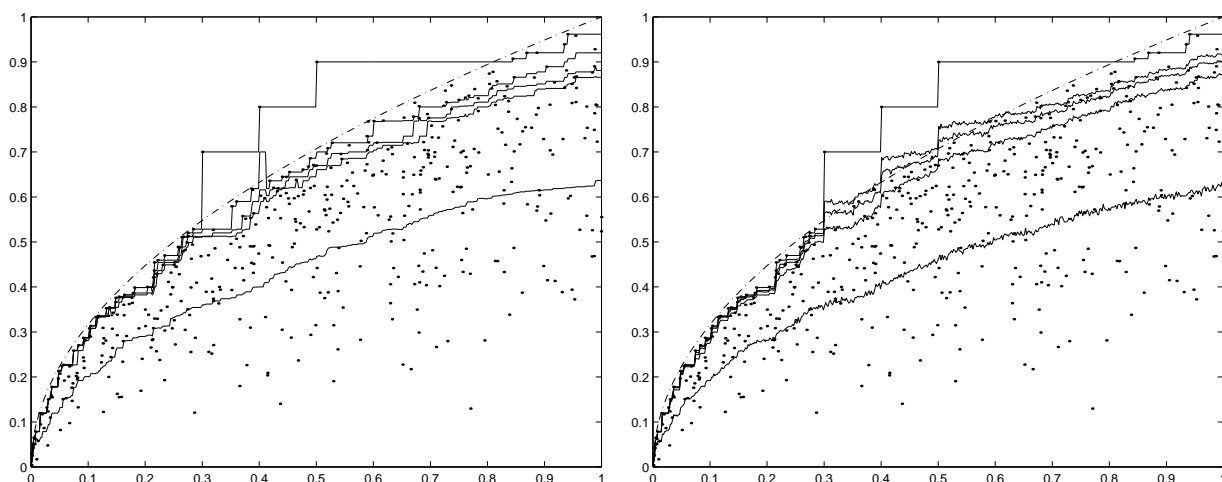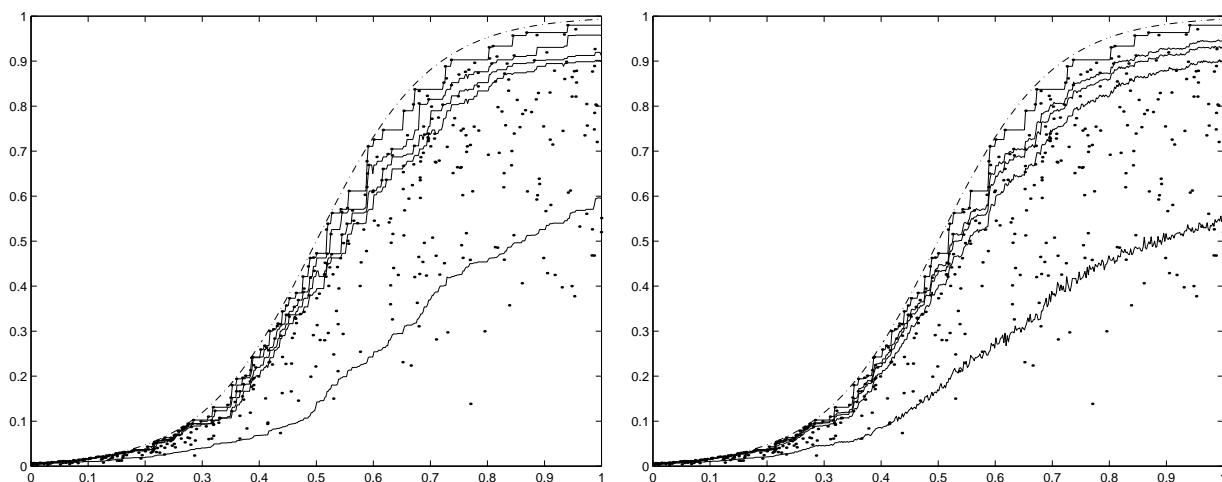
11

Figure 2: n=503: Same as above with 3 outliers included, output vs input.

5.2 Example 2

We now simulate a sample of $n = 500$ data points (x_i, y_i) with a non-convex production set. We choose here the model $Y = \exp\left(-5 + 10X\right)/\left(1 + \exp\left(-5 + 10X\right)\right)\exp\left(-U\right)$, where X is uniform on $(0, 1)$ and U is exponential with mean $1/3$.

Figure 3: n=500: Comparison between $\widehat{q}_{\alpha,n}$ (left) and $\widehat{\varphi}_{m,n}$ (right), output vs input.

Figure 3 plots the simulated data and, in solid lines, the frontiers $\widehat{q}_{\alpha,n}$ and $\widehat{\varphi}_{m,n}$ ($B = 1000$) with the same orders as in the preceding example and, in dash-dotted lines, the true frontier φ. Note that, here also, the frontier $\widehat{q}_{.99,n}$ is above all the frontiers $\widehat{\varphi}_{m,n}$. We again add in the data set three outliers, as shown in Figure 4, and we plot the frontiers $\widehat{q}_{\alpha,n}$ and $\widehat{\varphi}_{m,n}$ for the same orders. It is clear that the quantile curves of orders $\alpha < 1$ are more resistant to the three outliers than the expected maximal output frontiers $\widehat{\varphi}_{m,n}$ and the FDH frontier $\widehat{q}_{1,n}$.
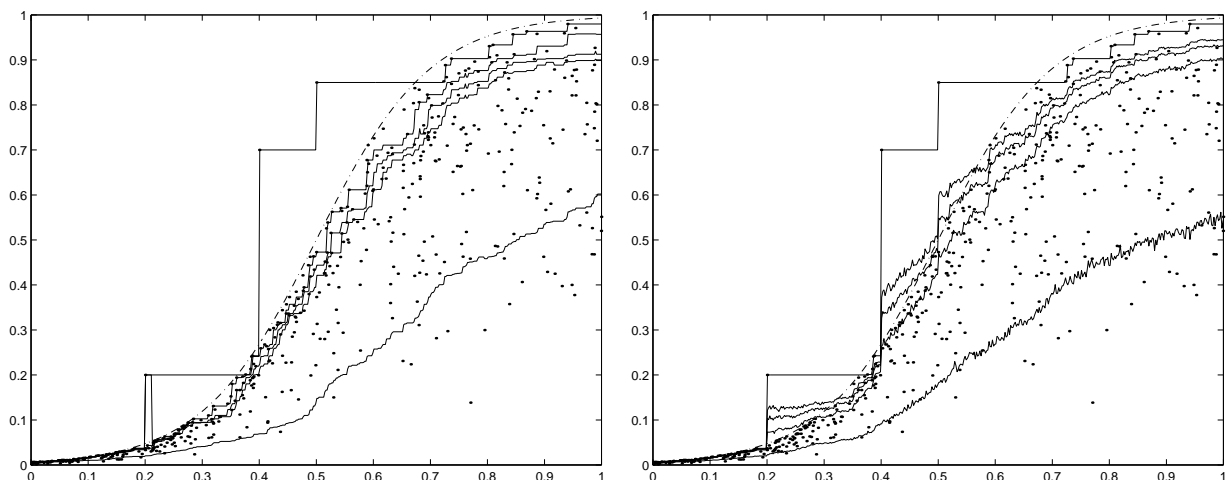
Figure 4: n=503: Same as above with 3 outliers included, output vs input.

5.3 Example 3

We now test the robustness of both estimators $\widehat{\varphi}_{m,n}$ and $\widehat{q}_{\alpha,n}$ for a small sample size n= 100. In each side of Figure 5 we plot, in dotted line, the quantile frontier of order $\alpha = .93$, and in solid lines, the frontiers $\widehat{\varphi}_{m,n}$ ($B = 1000$) of orders $m = 5,\ 7,\ 50,\ 75$. On the left-hand side, the data points are simulated according to the same model used in Example 1, and on the right-hand side, they are simulated according to the same model used in Example 2. Observe that the quantile frontier $\widehat{q}_{.93,n}$ is below the frontiers $\widehat{\varphi}_{m,n}$ of orders $m = 50,\ 75$, and is above those of orders $m = 5,\ 7$.
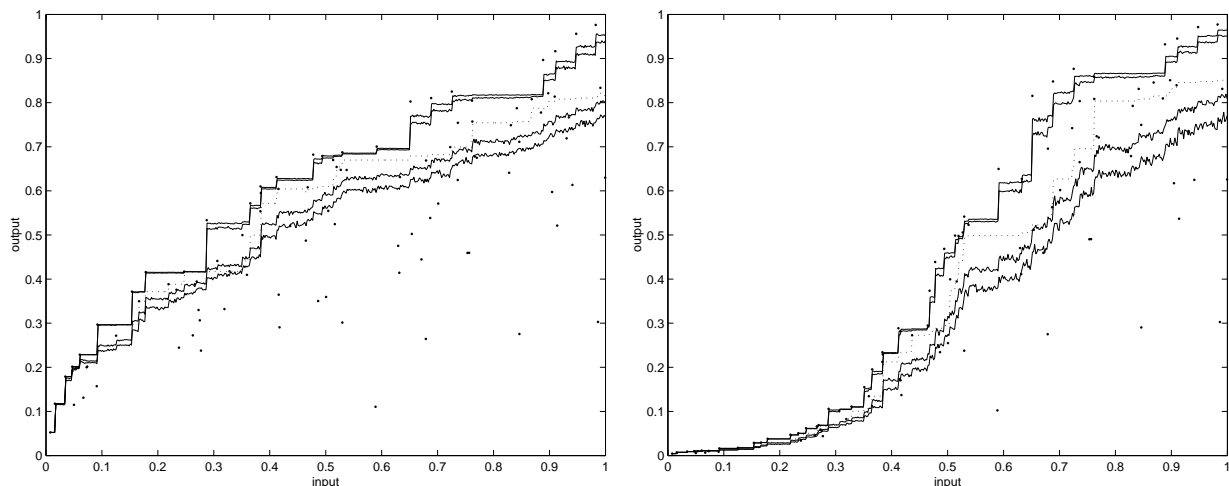

Figure 5: n=100: In solid lines, the order-m frontiers $\widehat{\varphi}_{5,n}$, $\widehat{\varphi}_{7,n}$, $\widehat{\varphi}_{50,n}$, $\widehat{\varphi}_{75,n}$, and in dotted line, the quantile frontier $\widehat{q}_{.93,n}$.

In Figure 6 we add to the above two data sets the same three outliers used in Examples 1 and 2 and we plot the same frontiers. We remark that the frontiers $\widehat{\varphi}_{m,n}$ of orders $m = 50,\ 75$
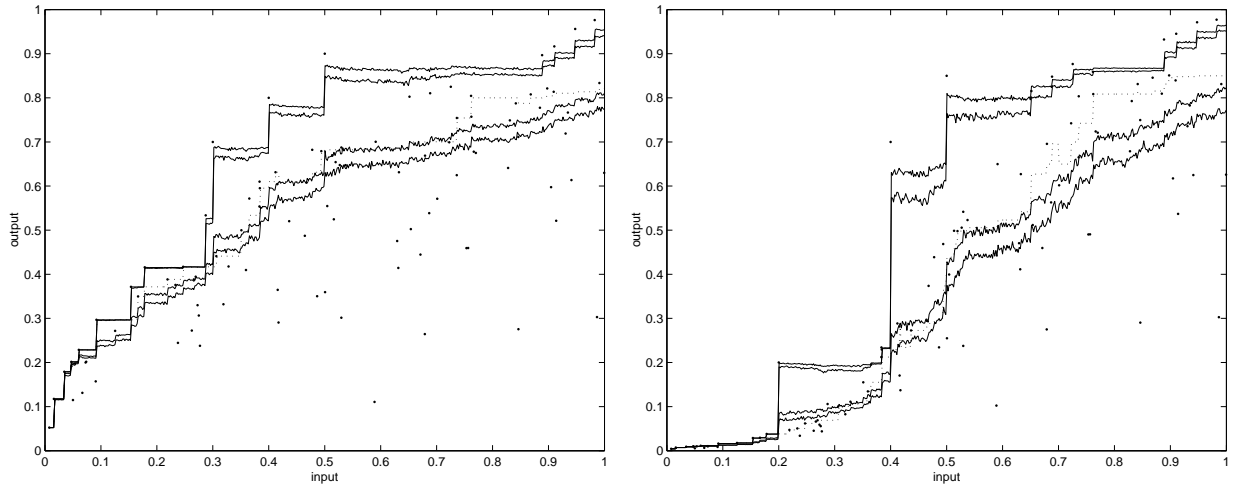
13

Figure 6: n=103: Same as above with 3 outliers included.

are highly influenced, even those of very low orders $m = 5$, 7 are attracted by the three outliers, while the quantile frontier is slightly perturbed.

We repeated the same exercise with many other simulated data sets leading to the same kind of results.

5.4 Frontier analysis of French post offices

We examine here a real data set in an univariate situation: this data set about the cost of the delivery activity of the postal services in France is analyzed by Cazals, Florens and Simar (2002). There are $n = 9521$ post offices observed in 1994. For each post office i, the input x_i is the labor cost measured by the quantity of labor, which represents more than 80% of the total cost of the delivery activity. The output y_i is defined as the volume of delivered mail (in number of objects).

The 4000 observed post offices with the smallest input levels are plotted in Figure 7, along with the estimates of quantile frontiers on the left and of expected maximal output frontiers on the right, for several different orders α and m. Here we obtain the frontiers $\widehat{\varphi}_{m,n}$ with $B = 2000$ bootstrap loops.

By using (5), it is very easy to check that every post office i belongs to the quantile curve of order $\alpha_i = \widehat{F}(y_i/x_i)$. On the other hand, the frontiers $\widehat{\varphi}_{m,n}$ do not cover the observations below the first frontier $\widehat{\varphi}_{1,n}$ (12% of the observed data) and the observations between the frontiers of successive orders $\widehat{\varphi}_{m,n}$ and $\widehat{\varphi}_{m+1,n}$. This disadvantage of frontiers $\widehat{\varphi}_{m,n}$ with respect to frontiers $\widehat{q}_{\alpha,n}$ is due to the fact that the order m is discrete.

Note that the order α_i of the quantile frontier which passes through the post office (x_i, y_i) is equal to the percentage of post offices that produce less than y_i among all the post offices using inputs smaller than or equal to x_i. In other words, this order indicates that the ith post office produces more than $100\alpha\%$ of all post offices using inputs smaller or equal to x_i and produces less than the $100(1 - \alpha)\%$ remaining post offices. This is why one sees in the left side of Figure 7 that, if α_i is close to one, then the post office (x_i, y_i) can be seen to

14

be performing relatively efficiently, and likewise, if α_i is close to zero, then the post office would be performing relatively inefficiently. Thus the order of the empirical quantile frontier $\widehat{q}_{\alpha,n}$ defines a reasonable benchmark value. Note also that the nonparametric estimation of the expected frontier $\widehat{\varphi}_{m,n}$ can be viewed as a mark of good practice for post offices when studying their performance. However, this benchmark is less clear than the empirical quantile frontier because it is less easy to interpret and does not cover the whole data set.
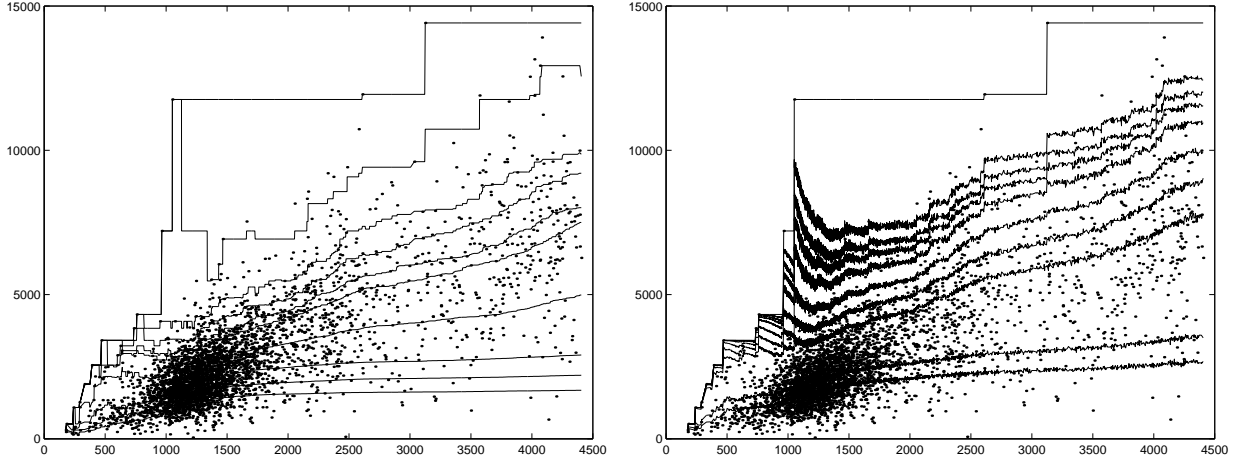

Figure 7: The 4000 observations with the smallest input levels. On the left, the frontiers $\widehat{q}_{\alpha,n}$ of orders $\alpha = .3, .5, .7, .9, .97, .98, .99, .995, .999$, and the FDH frontier ($\alpha = 1$). On the right, the frontiers $\widehat{\varphi}_{m,n}$ of orders $m = 1, 2, 25, 50, 100, 200, 300, 400, 600$, and the FDH frontier ($m = \infty$). Output vs input.

We also remark that the frontiers $\widehat{\varphi}_{m,n}$ (right side of Figure 7) are perturbed by the extreme observations from the order $m = 25$, whereas the frontiers $\widehat{q}_{\alpha,n}$ are not influenced except for those having orders almost equal to one ($\alpha \geq .999$).

The left side (resp. the right side) of Figure 8 indicates how the percentage $p(\alpha)$ (resp. $p(m)$) of observations above the quantile estimates $\widehat{q}_{\alpha,n}$ (resp. the expected maximum cost estimates $\widehat{\varphi}_{m,n}$) decreases with α (resp. m). We remark that the percentage $p(\alpha)$ decreases very slowly until the order $\alpha = .8$ of approximately 24% of observations. It means that the quantile frontiers of orders $0 \leq \alpha \leq .8$ are very tight. The 24% observations below the frontier $\widehat{q}_{.8,n}$ have an intermediate production performance and can be relatively inefficient. However, the percentage $p(\alpha)$ falls dramatically from the order $\alpha = .8$, which means that the quantile frontiers of extreme orders $.8 \leq \alpha \leq 1$ are very spaced and are spread out over 76% of the observations. In particular, 10% of the observations are above the frontier $\widehat{q}_{.995,n}$ and 3% of the observations are above the frontier $\widehat{q}_{.999,n}$. It is what explains notably the fact that only quantile frontiers of orders very close to one are influenced by super-efficient units.

On the right-hand side of Figure 8, we observe an opposite phenomenon: first the percentage $p(m)$ falls severely until the order $m = 50$ of approximately 80% of observations and then it continues to decrease but very slowly. Consequently the frontiers $\widehat{\varphi}_{m,n}$ of orders $m \geq 50$ are very tight. In particular we just have 9% of observations between the two frontiers $\widehat{\varphi}_{100,n}$ and $\widehat{\varphi}_{600,n}$, and only 3% of observations are above the frontier $\widehat{\varphi}_{600,n}$. The 20% observations above the frontier $\widehat{\varphi}_{50,n}$ are extreme and could be outliers or noisy observations.
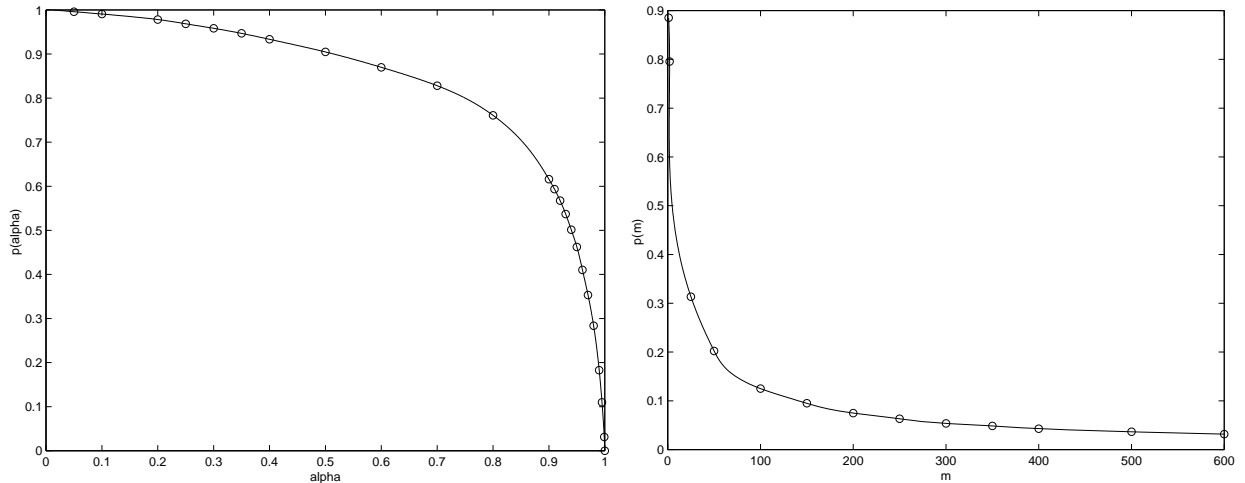
Figure 8: Evolution of the percentage of observations above the frontiers $\widehat{q}_{\alpha,n}$ (left) and $\widehat{\varphi}_{m,n}$ (right).

In summary the frontiers $\widehat{\varphi}_{m,n}$ are very tight from the order $m = 50$ and are spread out over extreme observations, it is then natural that these frontiers be more sensitive to extreme values than the quantile frontiers.

We can illustrate this result more clearly by considering the following inverse problem: for a given percentage p_0, denote by $\alpha(p_0)$ (resp. $m(p_0)$) the order of the frontier $\widehat{q}_{\alpha,n}$ (resp. $\widehat{\varphi}_{m,n}$) above which the percentage of observations is equal to p_0, we have $p(\alpha(p_0)) = p(m(p_0)) = p_0$. Inverting the relationship between α and $p(\alpha)$ and between m and $p(m)$ in Figure 8, we get the evolution of α and m as functions of p. When the percentage p varies between 0 and 10%, we remark that the order $\alpha(p)$ is almost constant ($\alpha(p) \approx 1$) while the order $m(p)$ falls rapidly from $m = 600$ to $m \approx 100$. This means that the 10% extreme observations influence all the frontiers $\widehat{\varphi}_{m,n}$ with orders $100 \leq m \leq \infty$, whereas only the frontiers $\widehat{q}_{\alpha,n}$ with orders almost equal to 1 are influenced by these extreme observations. This can be understood since the FDH frontier $\widehat{q}_1$ envelopes all the observed data.

This result is also illustrated in Figure 9 where the curve of evolution of $\alpha(p)$ with respect to $m(p)$ is nearly flat from the point $(100, .995)$ which corresponds to the percentage $p \approx 10\%$. This plot establishes an empirical relationship between the two families of frontiers $\{\widehat{q}_{\alpha,n}\}$ and $\{\widehat{\varphi}_{m,n}\}$. Given a frontier $\widehat{q}_{\alpha,n}$, we can determine the frontier $\widehat{\varphi}_{m,n}$ above which we have the same percentage of observations, and vice versa.

6   Conclusions

In this paper, we propose a new statistical concept of a production frontier which allows a more subtle tuning than the expected maximal output frontier of order $m \in \mathbb{N}^*$ (Cazals, Florens and Simar, 2002). We define a frontier of continuous order $\alpha \in [0,1]$ of the production set Ψ, for a given level of inputs x, by the conditional α-quantile of the distribution of Y given $X \leq x$.

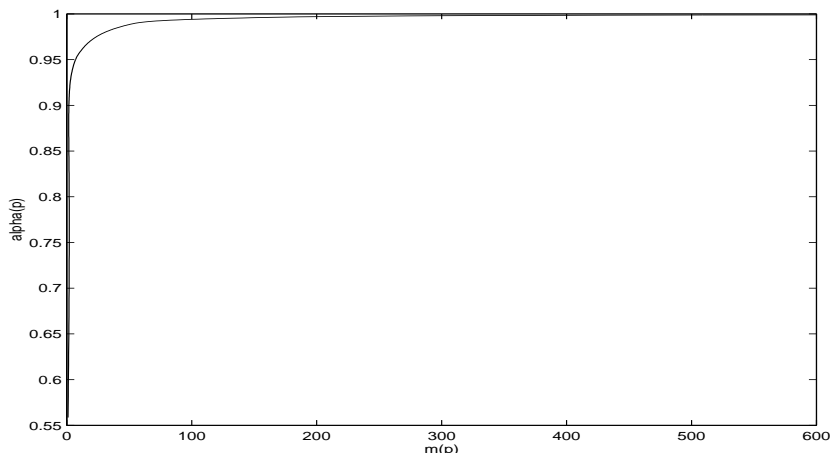Our quantile frontiers satisfy at least the same statistical properties as the expected

16

Figure 9: An empirical relationship between $\widehat{q}_{\alpha,n}$ and $\widehat{\varphi}_{m,n}$.

maximal output frontiers of order m. Moreover they have the advantage, from an economic point of view, to cover the interior of the attainable set entirely thus given a clearer indication on the production efficiency. This benefit is due to the continuity of the index α of our conditional quantiles.

A nonparametric estimator of the quantile function of order $\alpha < 1$ is very easy to derive by inverting the empirical version of the conditional distribution function. It does not envelop all the observed data points, and so it is more robust to extreme values than the standard DEA/FDH nonparametric envelopment estimators. Also it is easier to interpret than the nonparametric estimator of the expected function of order m. Moreover our estimator achieves the $\sqrt{n}$-consistency, is asymptotically unbiased and normally distributed, which is reasonable since the conditioning set $X \leq x$ has a positive probability measure. By choosing α as an appropriate function of n, it estimates the true frontier function and satisfies the asymptotic properties of the FDH estimator.

The method is illustrated using simulated and real data. It shows that the nonparametric quantile frontiers are more resistant to large observations in the output direction than the nonparametric estimates of expected maximal output frontiers, and that the continuous order α represents a good benchmark value. The robustness revealed by the numerical illustrations needs to be confirmed by some theoretical properties. This question is currently investigated.

It should be clear that, unlike the approach of Cazals et al. (2002), the conditional quantile approach is not extended here to multivariate Y. Serfling (2001) stated, "Despite the absence of a natural ordering of Euclidean space for dimension greater than one, effort to define vector-valued quantile functions for multivariate distributions has generated several approaches". The methods based on depth functions recommended by Serfling might be adapted to generalize in a reasonable way our univariate conditional quantiles. This problem is worth investigating.

17

References

Aigner, D.J., C.A.K. Lovell, & P. Schmidt (1977) Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6, pp. 21-37.

Battese, G.E. & T.J. Coelli (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. Journal of Econometrics 3, pp. 387-399.

Berlinet, A., B. Cadre, & A. Gannoun (2001) On the conditional L1-median and its estimation. Journal of Nonparametric Statistics 13, pp. 631-645.

Caudill, S.B., J.M. Ford, & D.M. Groper (1995) Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroscedasticity. Journal of Business & Economic Statistics 13, pp. 105-111.

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

Charnes, A., W.W. Cooper, & E. Rhodes (1978) Measuring the efficiency of decision making units. European Journal of Operations Research 2, pp. 429-444.

Debreu, G. (1951) The Coefficient of Resource Utilization. Econometrica 19(3), pp. 273-292.

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

Ducharme, G.R., A. Gannoun, M.C. Guertin, & J.C. Jequier (1995) Reference values obtained by kernel-based estimation of quantile regression. Biometrics 51, pp. 1105-1116.

Farrell, M.J. (1957) The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A 120, pp. 253-281.

Ghosh, J.K. (1971) A new proof of the Bahadur representation of quantiles and an application. The Annals of Mathematical Statistics 42, pp. 1957-1961.

Gijbels, I., E. Mammen, B.U. Park, & L. Simar (1999) On estimation of monotone and concave frontier functions. Journal of the American Statistical Association 94, pp. 220-228.

Hendricks, W. & R. Koenker (1992) Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity. Journal of the American Statistical Association 87, pp. 58-68.

Koopmans, T.C. (1951) An Analysis of Production as an Efficient Combination of Activities. In T. C. Koopmans (ed), Activity Analysis of Production and Allocation. Cowles Commis-

sion for Research in Economics, Monograph 13. New York: John Wiley & Sons Inc.

Meeusen, W., & van den Broek (1977) Efficiency estimation from Cobb-Douglas production function with composed error. International Economic Review 8, pp. 435-444.

Olsen, J.A., P. Schmidt, & D.M. Waldman (1980) A Monte Carlo study of Estimators of Stochastic Frontier Production Functions. Journal of Econometrics 13, pp. 67-82.

Park, B.U., L. Simar, & C. Weiner (2000) The FDH estimator for productivity efficiency scores. Econometric Theory 16, pp. 855-877.

Schwartz, L. (1991) Analyse. I. Hermann, éditeurs des sciences et des arts, Paris.

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

Serfling, R.J. (2002) Quantile functions for multivariate analysis: approaches and applications. Statistica Neerlandica 56(2), pp. 214-232.

Shephard , R.W. (1970) Theory of Cost and Production Function. Princeton, New Jersey: Princeton University Press.

Shorack, G.R., & J.A. Wellner (1986) Empirical Processes with Applications to Statistics. New York: John Wiley & Sons Inc.

Simar, L. (2003) Detecting Outliers in Frontier Models: A Simple Approach. Discussion paper #0146, UCL (http://www.stat.ucl.ac.be), to appear in Journal of Productivity Analysis.

Simar, L., & P. Wilson (2000) Statistical inference in nonparametric frontier models: The state of the art. Journal of Productivity Analysis 13, pp. 49-78.

Stevenson, R.E. (1980) Likelihood Functions for Generalised Stochastic Frontier Estimation. Journal of Econometrics 13, pp. 57-66.

APPENDIX: PROOFS

Proof of Proposition 2.1. Let $(x, y) \in \Psi^*$ and set $\alpha = F(y/x)$. It is an immediate consequence of the strict monotonicity of $F(\cdot/x)$ that $q_\alpha(x) = F^{-1}(\alpha/x) = y$. $\qquad\square$

Proof of Proposition 2.2. Let us assume the contrary. Then we obtain,

$$\forall x \in \widehat{\overset{\circ}{Supp}(X)}, \ \forall y \in Y(x), \ \exists m_{x,y} \in \mathbb{N}^* : \ y = \varphi_{m_{x,y}}(x) \tag{7}$$

19

where $Supp(X)$ is the support of the distribution of X and $\overset{\overset{\circ}{\frown}}{Supp(X)}$ denotes its interior. Let $x \in \overset{\overset{\circ}{\frown}}{Supp(X)}$ be fixed such that $\partial Y(x) > 0$. Since the production function φ is greater than or equal to the output-efficient function $\partial Y(.)$, we have $\varphi(x) = \partial Y(x)$ or $\varphi(x) > \partial Y(x)$.

If $\varphi(x) = \partial Y(x)$, we know from Cazals, Florens and Simar (2002, Appendix A) that $\varphi(x) = \lim \nearrow_{m \to \infty} \varphi_m(x)$, so there exists an integer $m_x \in \mathbb{N}^*$ such that $\varphi_{m_x}(x) < \varphi_{m_x+1}(x) \leq \partial Y(x)$ (else, we would have $\varphi_m(x) = \varphi_{m+1}(x)$ for every $m \in \mathbb{N}^*$, so that $\varphi(x) = \varphi_1(x)$, consequently we would obtain $\int_0^{\varphi(x)} F(y/x)dy = 0$ which is impossible because the function $F(\cdot/x)$ is strictly increasing on $[0, \varphi(x)]$). Let y be a real number such that $\varphi_{m_x}(x) < y < \varphi_{m_x+1}(x)$. Using the free disposability assumption of outputs, it is easily seen that $Y(x) \equiv [0, \partial Y(x)]$, so that $y \in Y(x)$. Then by (7), there exists an integer $m_{x,y} \in \mathbb{N}^*$ such that $y = \varphi_{m_{x,y}}(x)$. It follows that $\varphi_{m_x}(x) < \varphi_{m_{x,y}}(x) < \varphi_{m_x+1}(x)$, and thus $m_x < m_{x,y} < m_x + 1$ since $\varphi_m(x)$ is a monotone nondecreasing function of m. This contradicts the fact that $m_{x,y}$ is an integer.

Now if $\varphi(x) > \partial Y(x)$, first note that $\partial Y(x) \in Y(x)$ yields by (7) that $\partial Y(x) = \varphi_{m_{x,\partial Y(x)}}(x)$ where $m_{x,\partial Y(x)} \in \mathbb{N}^*$. Due to $\lim \nearrow_{m \to \infty} \varphi_m(x) = \varphi(x)$, there exists an order $m_x > m_{x,\partial Y(x)}$ such that $\varphi(x) \geq \varphi_m(x) > \partial Y(x)$ when $m \geq m_x$, and $\varphi_m(x) \leq \partial Y(x)$ when $m < m_x$. For any $y \in Y(x)$ we have $y \leq \partial Y(x)$ so that $y < \varphi_{m_x}(x)$. We also have by (7) $y = \varphi_{m_{x,y}}(x)$ where $m_{x,y} \in \mathbb{N}^*$. Hence $\varphi_{m_{x,y}}(x) < \varphi_{m_x}(x)$. Therefore, using again the monotonicity of $\varphi_m(x)$ with respect to m, we get $m_{x,y} < m_x$. In summary,

$$\forall y \in Y(x), \ \exists m_{x,y} \in \{1, \cdots, m_x - 1\} : \ y = \varphi_{m_{x,y}}(x).$$

Since $\varphi_m(x) \in [0, \partial Y(x)] = Y(x)$ for every $m < m_x$, the map $m \mapsto \varphi_m(x)$ is well defined and is onto from $\{1, \cdots, m_x - 1\}$ to $Y(x)$. As a consequence, the finite set $\{\varphi_1(x), \cdots, \varphi_{m_x-1}(x)\}$ coincides with the interval $[0, \partial Y(x)]$, which implies the contradiction. $\square$

Proof of Proposition 2.3. Let x be an input which satisfies the condition of Proposition 2.3. We have by definition $\varphi_m(x) = E[\max(Y^1, \cdots, Y^m)]$, where $Y^1, \cdots, Y^m$ are m independent identically distributed random variables generated by the distribution function $F(\cdot/x)$. Let $F_{m,x}(y) = \frac{1}{m} \sum_{i=1}^m 1(Y^i \leq y)$ be the empirical distribution function of $(Y^1, \cdots, Y^m)$. The empirical quantile of order $\alpha \in (0, 1]$ of this sample is then defined by

$$q_\alpha^m(x) := F_{m,x}^{-1}(\alpha) = \inf\{y| \ F_{m,x}(y) \geq \alpha\}.$$

We know that $q_\alpha^m(x)$ is equal to $Y^{(\alpha m)}$ if αm is an integer and to $Y^{([\alpha m]+1)}$ otherwise. Then $q_1^m(x) = Y^{(m)}$. Since the family $(q_\alpha^m(x))_{0<\alpha<1}$ increases to $q_1^m(x)$ when $\alpha \to 1$, the dominated convergence theorem yields $\varphi_m(x) = \lim_{\alpha \to 1} E[q_\alpha^m(x)]$, and thus we can write $\varphi_m(x) = E[q_\alpha^m(x)] + \varepsilon_1(\alpha)$, where $\varepsilon_1(\alpha) = o(\alpha)$ when $\alpha \to 1$. On the other hand, according to the representation theorem of Bahadur (see, e.g., Serfling, 1980, Theorem 2.5.1, p. 91), we have for every $\alpha \in (0, 1)$,

$$q_\alpha^m(x) - q_\alpha(x) = \frac{\alpha - F_{m,x}(q_\alpha(x))}{f(q_\alpha(x)/x)} + R_{m,x}(\alpha),$$

where, with probability 1, $R_{m,x}(\alpha) = O\left(m^{-3/4}(\log m)^{3/4}\right)$ as $m \to \infty$. By using Kiefer's Theorem (see, e.g., Serfling, 1980, Theorem D, p. 101), it can easily be seen that a more

20

precise expression of the remainder is given by

$$R_{m,x}(\alpha) = \frac{1}{f(q_\alpha(x)/x)} R_{m,x}$$

where, almost surely and uniformly in α, we have

$$R_{m,x} = O\left(m^{-3/4}(\log m)^{1/2}(\log\log m)^{1/4}\right), \quad m \to \infty.$$

It follows that,

$$\varphi_m(x) - q_\alpha(x) = \frac{1}{f(q_\alpha(x)/x)} O\left(m^{-3/4}(\log m)^{1/2}(\log\log m)^{1/4}\right) + \varepsilon_1(\alpha), \quad m \to \infty.$$

Now consider the function $\psi_x(p) = 1/f(q_p(x)/x)$, $p \in [0,1]$. We have as $m \to \infty$,

$$\varphi_m(x) - q_\alpha(x) = [\psi_x(1) - (\psi_x(1) - \psi_x(\alpha))] O\left(m^{-3/4}(\log m)^{1/2}(\log\log m)^{1/4}\right) + \varepsilon_1(\alpha).$$

Since $F(\cdot/x)$ has a positive continuous density $f(./x)$ in the neighborhood $(0, \varphi(x))$ of $q_\alpha(x)$, for any $\alpha \in (0,1)$, we obtain according to Shorack and Wellner (1986, Proposition 6, p. 9) that the partial derivative $(\partial/\partial\alpha)q_\alpha(x)$ exists and equals $1/f(q_\alpha(x)/x)$. Then, for every $\alpha \in (0,1)$, the derivative of $\psi_x(\alpha)$ with respect to α is given by

$$\psi_x'(\alpha) = -\frac{\partial}{\partial\alpha} q_\alpha(x) F''\left(q_\alpha(x)/x\right) / f^2(q_\alpha(x)/x).$$

Using the fact that $\psi_x(1) - \psi_x(\alpha) = (1-\alpha)\psi_x'(\alpha) + (1-\alpha)\varepsilon_2(\alpha)$, where $\varepsilon_2(\alpha) = o(\alpha)$ when $\alpha \to 1$, we obtain as $m \to \infty$ and $\alpha \to 1$,

$$\varphi_m(x) - q_\alpha(x) = \{\psi_x(1) - (1-\alpha)[\psi_x'(\alpha) + o(\alpha)]\} O\left(m^{-3/4}(\log m)^{1/2}(\log\log m)^{1/4}\right) + o(\alpha),$$

which proves (3). $\qquad\square$

Proof of Proposition 2.4. 1. Since the family $\{q_\alpha(x)\}_{0\le\alpha\le1}$ is monotone nondecreasing and bounded by $q_1(x)$, $q_\alpha(x)$ converges pointwise to $q_1(x)$ when $\alpha \to 1$.

2. Let K be a compact subset of $\widehat{Supp(X)}$. $\{q_\alpha(.)\}_{0<\alpha<1}$ is a nondecreasing sequence of real valued functions which are continuous on K. Moreover it converges pointwise to the continuous function $q_1(.)$ as α tends to one. Then by Dini's theorem (Schwartz, 1991, p. 325) the convergence is uniform on K. $\qquad\square$

Proof of Proposition 2.5. Suppose that for every $y \ge 0$, the function $x \mapsto F(y/x)$ is monotone nonincreasing on $\{x \in \mathbb{R}^p_+ | F_X(x) > 0\}$. Let $\alpha \in [0,1]$ and $x_1 \le x_2$ such that $F_X(x_1) > 0$. Then $F(q_\alpha(x_2)/x_1) \ge F(q_\alpha(x_2)/x_2) \ge \alpha$. It follows that $q_\alpha(x_2) \ge \inf\{y| F(y/x_1) \ge \alpha\} = q_\alpha(x_1)$.

Conversely, suppose that the quantile function is monotone nondecreasing for every order on $\{x \in \mathbb{R}^p_+ | F_X(x) > 0\}$. Let $y \in \mathbb{R}+$ and $x_1 \le x_2$ such that $F_X(x_1) > 0$. Set $\alpha = F(y/x_2)$. We have $q_\alpha(x_2) = \inf\{u| F(u/x_2) \ge \alpha\}$, so that $y \ge q_\alpha(x_2)$. Since $q_\alpha(x_1) \le q_\alpha(x_2)$, $y \ge q_\alpha(x_1)$, and thus $F(y/x_1) \ge F(q_\alpha(x_1)/x_1) \ge \alpha = F(y/x_2)$. $\qquad\square$

The following lemma will be useful in the proof of Theorem 4.1.

Lemma 6.1. *Let $\{V_n\}$, $\{W_n\}$ be two sequences of random variables satisfying the following conditions*

i) For all $\delta > 0$, there exists a λ (depending on δ) s.t. $P(|W_n| > \lambda) < \delta$.

ii) For all k and all $\varepsilon > 0$

$$\lim_{n \to \infty} P(V_n \leq k, W_n \geq k + \varepsilon) = 0$$
$$\lim_{n \to \infty} P(V_n \geq k + \varepsilon, W_n \leq k) = 0.$$

Then $V_n - W_n \xrightarrow{P} 0$ as $n \to \infty$.

The proof of this lemma can be found in Ghosh (1971, Lemma 1, p. 1958). Now let us demonstrate Theorem 4.1.

Proof of Theorem 4.1. Consider the statistical functional $T^{\alpha,x}$ which associates to a distribution function G on $\mathbb{R}^2$ the real value

$$T^{\alpha,x}(G) = \inf\left\{y \mid \frac{G(x,y)}{G(x,\infty)} \geq \alpha\right\}.$$

The conditional quantile $q_\alpha(x)$ and its estimator $\widehat{q}_{\alpha,n}(x)$ are then given by $q_\alpha(x) = T^{\alpha,x}(F)$ and $\widehat{q}_{\alpha,n}(x) = T^{\alpha,x}(\widehat{F})$. Let

$$R_n^{\alpha,x} = (T^{\alpha,x}(\widehat{F}) - T^{\alpha,x}(F)) - \frac{1}{n}\sum_{i=1}^{n} d_1 T^{\alpha,x}\left(F; 1(X_i \leq ., Y_i \leq .) - F\right) \tag{8}$$

where

$$d_1 T^{\alpha,x}\left(F; 1(X_i \leq ., Y_i \leq .) - F\right) = \frac{d}{d\lambda} T^{\alpha,x}(F + \lambda(1(X_i \leq ., Y_i \leq .) - F))|_{\lambda=0^+}$$

is the first Gâteaux differential of $T^{\alpha,x}$ at F in the direction of $1(X_i \leq ., Y_i \leq .)$. Using property (6), we obtain by a straightforward computation,

$$T^{\alpha,x}(F + \lambda(1(X_i \leq ., Y_i \leq .) - F)) = F^{-1}\left(\alpha + \frac{\alpha\lambda 1(X_i \leq x)}{(1-\lambda)F_X(x)}/x\right)1(Y_i > q_\alpha(x))$$

$$+ F^{-1}\left(\alpha - \frac{\lambda(1-\alpha)1(X_i \leq x)}{(1-\lambda)F_X(x)}/x\right)1(Y_i < q_\alpha(x)) + q_\alpha(x)1(Y_i = q_\alpha(x)).$$

Hence,

$$d_1 T^{\alpha,x}\left(F; 1(X_i \leq ., Y_i \leq .) - F\right)$$
$$= \frac{\alpha 1(X_i \leq x)1(Y_i > q_\alpha(x)) + (\alpha - 1)1(X_i \leq x)1(Y_i < q_\alpha(x))}{f(q_\alpha(x)/x)F_X(x)}.$$

Therefore,

$$\frac{1}{n}\sum_{i=1}^{n} d_1 T^{\alpha,x}\left(F; 1(X_i \leq ., Y_i \leq .) - F\right) = \frac{1}{n}\sum_{i=1}^{n} h_\alpha(x, X_i, Y_i) \tag{9}$$

22

where

$$h_\alpha(x, X, Y) = \frac{\alpha 1(X \le x) - 1(X \le x, Y \le q_\alpha(x))}{f(q_\alpha(x)/x)F_X(x)}.$$

We have

$$Var[h_\alpha(x, X, Y)] = E[(h_\alpha(x, X, Y))^2] = \sigma^2(x, \alpha),$$

so that, by the central limit theorem,

$$W_n^{\alpha,x} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} d_1 T^{\alpha,x} (F; 1(X_i \le ., Y_i \le .) - F) \xrightarrow{\mathcal{L}} N\left(0, \sigma^2(x, \alpha)\right) \quad \text{as} \quad n \to \infty, \quad (10)$$

and by the law of large numbers,

$$\frac{1}{\sqrt{n}} W_n^{\alpha,x} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \quad (11)$$

Let $V_n^{\alpha,x} = \sqrt{n} \left(\widehat{q}_{\alpha,n}(x) - q_\alpha(x)\right)$. Then we obtain via (8)

$$\sqrt{n} R_n^{\alpha,x} = V_n^{\alpha,x} - W_n^{\alpha,x}.$$

Using Lemma 6.1, we will show that $\sqrt{n} R_n^{\alpha,x}$ converges in probability to zero. We have for every real number t,

$$\{V_n^{\alpha,x} \le t\} = \left\{\widehat{q}_{\alpha,n}(x) \le q_\alpha(x) + \frac{t}{\sqrt{n}}\right\} = \left\{\alpha \le \widehat{F}(q_\alpha(x) + \frac{t}{\sqrt{n}}/x)\right\} = \{Z_{t,n} \le T_n\}, \quad (12)$$

where

$$Z_{t,n} = \frac{\sqrt{n}\widehat{F}_X(x)}{f(q_\alpha(x)/x)F_X(x)}[F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) - \widehat{F}(q_\alpha(x) + \frac{t}{\sqrt{n}}/x)]$$

and

$$T_n = \frac{\sqrt{n}\widehat{F}_X(x)}{f(q_\alpha(x)/x)F_X(x)}[F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) - \alpha].$$

Since $F(\cdot/x)$ is differentiable at $q_\alpha(x)$ with derivative $f(q_\alpha(x)/x)$, $F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) - \alpha = \frac{t}{\sqrt{n}}f(q_\alpha(x)/x) + \frac{t}{\sqrt{n}}o(1)$ as $n \to \infty$, which implies $T_n = \frac{\widehat{F}_X(x)}{F_X(x)}t + \frac{t\widehat{F}_X(x)}{f(q_\alpha(x)/x)F_X(x)}o(1)$ as $n \to \infty$. We know from the law of large numbers that $\widehat{F}_X(x) \xrightarrow{P} F_X(x)$, and thus,

$$T_n \xrightarrow{P} t \quad \text{as} \quad n \to \infty. \quad (13)$$

On the other hand we have

$$Z_{t,n} - W_n^{\alpha,x} = \frac{\sqrt{n}\widehat{F}_X(x)}{f(q_\alpha(x)/x)F_X(x)}$$
$$\times \left[\left(F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) - \widehat{F}(q_\alpha(x) + \frac{t}{\sqrt{n}}/x)\right) - \left(\alpha - \widehat{F}(q_\alpha(x)/x)\right)\right].$$

23

By a simple computation we find that,

$$\left\{ f^2(q_\alpha(x)/x) F_X^2(x) \right\} E[(Z_{t,n} - W_n^{\alpha,x})^2] = \alpha F_X(x) - \alpha^2 F_X(x) - F_X(x) F^2(q_\alpha(x) + \frac{t}{\sqrt{n}}/x)$$

$$+ F_X(x) F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) + 2\alpha F_X(x) F(q_\alpha(x) + \frac{t}{\sqrt{n}}/x) - 2F_X(x) F(q_\alpha(x) + (\frac{t}{\sqrt{n}} \wedge 0)/x).$$

Using the continuity of $F(\cdot/x)$ in $q_\alpha(x)$, we get $E[(Z_{t,n} - W_n^{\alpha,x})^2] \to 0$ as $n \to \infty$, and thus,

$$Z_{t,n} - W_n^{\alpha,x} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty. \tag{14}$$

Now, using the results (9) and (12)-(14), we will show that $V_n^{\alpha,x}$ and $W_n^{\alpha,x}$ satisfy the two conditions of Lemma 6.1. As $E[(W_n^{\alpha,x})^2] = \sigma^2(x,\alpha)$, the first condition follows from a trivial application of the Markov inequality. For any k and any $\varepsilon > 0$, setting $t = k$, we have by (12),

$$\{V_n^{\alpha,x} \le k, W_n^{\alpha,x} \ge k + \varepsilon\} = \{Z_{t,n} \le T_n, W_n^{\alpha,x} \ge t + \varepsilon\} \subset \{(W_n^{\alpha,x} - Z_{t,n}) \ge \varepsilon - (T_n - t)\}.$$

Hence,

$$P(V_n^{\alpha,x} \le k, W_n^{\alpha,x} \ge k + \varepsilon) \le P\left(|(W_n^{\alpha,x} - Z_{t,n}) + (T_n - t)| \ge \varepsilon\right).$$

Therefore, combined with (13) and (14), $\lim_{n\to\infty} P(V_n^{\alpha,x} \le k, W_n^{\alpha,x} \ge k + \varepsilon) = 0$. Now applying (12) to $t = k + \varepsilon$, we get

$$P(V_n^{\alpha,x} \ge k + \varepsilon, W_n^{\alpha,x} \le k) \le P\left(|(Z_{t,n} - W_n^{\alpha,x}) - (T_n - t)| \ge \varepsilon\right).$$

Then (13) and (14) yield $\lim_{n\to\infty} P(V_n^{\alpha,x} \ge k + \varepsilon, W_n^{\alpha,x} \le k) = 0$. Hence, the second condition of Lemma 6.1 is also satisfied. Therefore $V_n^{\alpha,x} - W_n^{\alpha,x} \xrightarrow{P} 0$, as $n \to \infty$, that is,

$$\sqrt{n} R_n^{\alpha,x} \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty.$$

In particular $R_n^{\alpha,x}$ converges in probability to zero as $n \to \infty$. Thus,

$$\widehat{q}_{\alpha,n}(x) - q_\alpha(x) = \frac{1}{\sqrt{n}} W_n^{\alpha,x} + R_n^{\alpha,x} = \frac{1}{\sqrt{n}} W_n^{\alpha,x} + o_p(1) \quad \text{as} \quad n \to \infty, \tag{15}$$

$$\sqrt{n}(\widehat{q}_{\alpha,n}(x) - q_\alpha(x)) = W_n^{\alpha,x} + \sqrt{n} R_n^{\alpha,x} = W_n^{\alpha,x} + o_p(1) \quad \text{as} \quad n \to \infty. \tag{16}$$

The consistency follows from results (11) and (15), and the asymptotic normality is obtained by (10) and (16). $\qquad\square$

Proof of Theorem 4.2. It follows from (16), as $n \to \infty$,

$$\sqrt{n} \left(\widehat{q}_{\alpha,n}(x^1) - q_\alpha(x^1), \cdots, \widehat{q}_{\alpha,n}(x^r) - q_\alpha(x^r) \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(h_\alpha(x^1, X_i, Y_i), \cdots, h_\alpha(x^r, X_i, Y_i) \right) + o_p(1).$$

24

Hence, the multivariate central limit theorem yields,

$$\sqrt{n}\left(\widehat{q}_{\alpha,n}(x^1) - q_\alpha(x^1), \cdots, \widehat{q}_{\alpha,n}(x^r) - q_\alpha(x^r)\right) \xrightarrow{\mathcal{L}} N_r(0, \Sigma_\alpha) \quad \text{as} \quad n \to \infty,$$

where

$$\Sigma_\alpha(x^k, x^l) = Cov\left(h_\alpha(x^k, X_1, Y_1), h_\alpha(x^l, X_1, Y_1)\right).$$

This ends the proof. □

Proof of Theorem 4.3. From Park, Simar and Weiner (2000) and Cazals, Florens and Simar (2002) we know that

$$n^{1/(p+1)}(\varphi(x) - \widehat{\varphi}_n(x)) \xrightarrow{\mathcal{L}} Weibull(\mu_x^{p+1}, p+1) \quad \text{as} \quad n \to \infty.$$

So by using the following decomposition

$$n^{1/(p+1)}(\varphi(x) - \widehat{q}_{\alpha,n}(x)) = n^{1/(p+1)}(\varphi(x) - \widehat{\varphi}_n(x)) + n^{1/(p+1)}(\widehat{\varphi}_n(x) - \widehat{q}_{\alpha,n}(x))$$

we want to find a function α of n such that,

$$\alpha(n) \to 1 \quad \text{and} \quad n^{1/(p+1)}(\widehat{\varphi}_n(x) - \widehat{q}_{\alpha(n),n}(x)) \xrightarrow{\mathcal{L}} 0, \quad \text{as} \quad n \to \infty.$$

From (5) we have for any $\alpha > 0$,

$$\begin{aligned}
\widehat{\varphi}_n(x) - \widehat{q}_{\alpha,n}(x) &= \left(Y_{(i_{N_x})} - Y_{(i_{\{\alpha N_x\}})}\right) 1(\alpha N_x \in \mathbb{N}^*) \\
&+ \left(Y_{(i_{N_x})} - Y_{(i_{\{[\alpha N_x]+1\}})}\right) 1(\alpha N_x \notin \mathbb{N}^*).
\end{aligned} \tag{17}$$

Set for every $k \in \{1, \cdots, N_x - 1\}$,

$$C_{x,k}(n) = \frac{Y_{(i_{N_x})} - Y_{(i_k)}}{1 - \frac{k}{N_x}},$$

and let $C_x(n) = \max\{C_{x,k}(n) \mid 1 \le k \le N_x - 1\}$. Then we have

$$\begin{aligned}
\left(Y_{(i_{N_x})} - Y_{(i_{\{\alpha N_x\}})}\right) 1(\alpha N_x \in \mathbb{N}^*) &= \left(Y_{(i_{N_x})} - Y_{(i_{\{\alpha N_x\}})}\right) \sum_{k=1}^{N_x} 1(\alpha N_x = k) \\
&= \sum_{k=1}^{N_x - 1} \left(Y_{(i_{N_x})} - Y_{(i_k)}\right) 1(\alpha N_x = k) \\
&= \sum_{k=1}^{N_x - 1} C_{x,k}(n)(1 - \frac{k}{N_x})1(\alpha = \frac{k}{N_x}) \\
&\le C_x(n) \sum_{k=1}^{N_x} (1 - \alpha)1(\alpha N_x = k) \\
&= C_x(n)(1 - \alpha)1(\alpha N_x \in \mathbb{N}^*) \tag{18}
\end{aligned}$$

25

since $1(\alpha N_x \in \mathbb{N}^*) = \sum_{k=1}^{N_x} 1(\alpha N_x = k)$.

Now, using that $1(\alpha N_x \notin \mathbb{N}^*) = \sum_{k=0}^{N_x-1} 1(\alpha N_x \in (k, k+1))$, we get

$$
\begin{aligned}
\left(Y_{(i_{N_x})} - Y_{(i_{\{[\alpha N_x]+1\}})}\right) 1(\alpha N_x \notin \mathbb{N}^*) &= \sum_{k=0}^{N_x-1} \left(Y_{(i_{N_x})} - Y_{(i_{\{k+1\}})}\right) 1(\alpha N_x \in (k, k+1)) \\
&= \sum_{k=0}^{N_x-2} \left(Y_{(i_{N_x})} - Y_{(i_{\{k+1\}})}\right) 1(\alpha N_x \in (k, k+1)) \\
&= \sum_{k=0}^{N_x-2} C_{x,k+1}(n)(1 - \frac{k+1}{N_x}) 1(\alpha N_x \in (k, k+1)) \\
&\leq C_x(n)(1-\alpha) \sum_{k=0}^{N_x-2} 1(\alpha N_x \in (k, k+1)) \\
&\leq C_x(n)(1-\alpha) 1(\alpha N_x \notin \mathbb{N}^*). \qquad (19)
\end{aligned}
$$

It follows from (17)-(19) that,

$$
\begin{aligned}
\widehat{\varphi}_n(x) - \widehat{q}_{\alpha,n}(x) &\leq C_x(n)(1-\alpha) 1(\alpha N_x \in \mathbb{N}^*) + C_x(n)(1-\alpha) 1(\alpha N_x \notin \mathbb{N}^*) \\
&= C_x(n)(1-\alpha),
\end{aligned}
$$

so that

$$
n^{1/(p+1)}(\widehat{\varphi}_n(x) - \widehat{q}_{\alpha,n}(x)) \leq n^{1/(p+1)} C_x(n)(1-\alpha). \qquad (20)
$$

Since the support Ψ of (X, Y) is compact, the support of Y is bounded. Let $M > 0$ be its upper bound. Then for any $k = 1, \cdots, N_x - 1$,

$$
Y_{(i_{N_x})} - Y_{(i_k)} \leq Y_{(i_{N_x})} \leq M \ a.s, \quad \text{and} \quad \frac{1}{1 - \frac{k}{N_x}} \leq N_x.
$$

Hence,

$$
\forall k = 1, \cdots, N_x - 1 : \quad C_{x,k}(n) = \frac{Y_{(i_{N_x})} - Y_{(i_k)}}{1 - \frac{k}{N_x}} \leq M N_x \ a.s.
$$

Therefore,

$$
C_x(n) = \max_{1 \leq k \leq N_x - 1} C_{x,k}(n) \leq M N_x \ a.s.
$$

We deduce from (20),

$$
n^{1/(p+1)}(\widehat{\varphi}_n(x) - \widehat{q}_{\alpha,n}(x)) \leq M N_x n^{1/(p+1)}(1-\alpha) = M \widehat{F}_X(x) n^{(p+2)/(p+1)}(1-\alpha) \ a.s.
$$

We know from the strong law of large numbers that $\widehat{F}_X(x) \xrightarrow{a.s} F_X(x)$. So to achieve our goal, it is sufficient to choose $\alpha(n)$ such that,

$$
n^{(p+2)/(p+1)}(1 - \alpha(n)) \to 0 \quad \text{as} \quad n \to \infty.
$$

Indeed we find,

$$
\alpha(n) \to 1 \quad \text{and} \quad n^{1/(p+1)}(\widehat{\varphi}_n(x) - \widehat{q}_{\alpha(n),n}(x)) \xrightarrow{a.s} 0 \quad \text{as} \quad n \to \infty.
$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$