A Smooth Nonparametric Conditional Quantile Frontier Estimator

Carlos Martins-Filho
Department of Economics
Oregon State University
Ballard Hall 303
Corvallis, OR 97331-3612 USA
email: carlos.martins@oregonstate.edu
Voice: + 1 541 737 1476

and

Feng Yao
Department of Economics
University of North Dakota
P.O. Box 8369
Grand Forks, ND 58203 USA
email: feng.yao@mail.business.und.edu
Voice: + 1 701 777 3360

March, 2006

Abstract. Traditional estimators for nonparametric frontier models (DEA, FDH) are very sensitive to extreme values/outliers. Recently, Aragon, Daouia, and Thomas-Agnan (2005) proposed a nonparametric α-frontier model and estimator based on a suitably defined conditional quantile which is more robust to extreme values/outliers. Their proposed estimator is simple to construct but produces a nonsmooth estimated α-frontiers even when the underlying technology induces smooth frontiers. In this paper, we propose a new smooth nonparametric conditional quantile estimator for the α-frontier model. Our estimator is a kernel based conditional quantile estimator that builds on early work of Azzalini (1981). It is computationally simple, resistant to outliers and extreme values, and smooth. In addition, the estimator is also shown to be consistent and $\sqrt{n}$ asymptotically normal under mild regularity conditions. We also show that our estimator's variance is of smaller order than that of the estimator proposed by Aragon et al. A simulation study confirms the asymptotic theory predictions and contrasts our estimator with that of Aragon et al.

Keywords and Phrases. conditional quantile estimation; nonparametric frontier; production function.

JEL Classifications. C14, C22

1 Introduction

The specification and estimation of output set boundaries or production frontiers, and the measurement of the associated efficiency level of production units has been the subject of a vast and expanding literature since the seminal work of Farrell (1957). The main objective of this literature can be stated simply. Consider $(X, Y) \in \Re_+^d \times \Re_+$ where Y describes the output of a production unit and X describes the d inputs used in production. The output set is given by $\Psi = \{(x, y) \in \Re_+^d \times \Re_+ : x \text{ can produce } y\}$ and the production function or frontier associated with Ψ is $g(x) = sup\{y \in \Re_+ : (x, y) \in \Psi\}$ for all $x \in \Re_+^d$. Let $(x_0, y_0) \in \Psi$ characterize the performance of a production unit and define $0 \leq R_0 \equiv \frac{y_0}{g(x_0)} \leq 1$ to be this unit's (inverse) Farrell output efficiency measure. The main objective in production and efficiency analysis is, given a random sample of production units $\chi_n \equiv \{(X_i, Y_i)\}_{i=1}^n$ that share the set Ψ, to obtain estimates of $g(\cdot)$ and by extension $R_i = \frac{Y_i}{g(X_i)}$ for $i = 1, \cdots, n$.

Deterministic frontier models and estimators have gained popularity among applied researchers because their construction relies on very mild assumptions on Ψ.[1] These models, represented largely in econometrics by Charnes et al. (1978) data envelopment analysis (DEA) and Deprins et al. (1984) free disposal hull (FDH), are based on the assumption that χ_n lie in Ψ, i.e., $P((X, Y) \in \Psi) = 1$, where P is the probability measure associated with the random vector (X, Y). The most appealing characteristic of such models is that there is no need to assume any restrictive parametric structure on $g(\cdot)$ or the probability measure P to perform estimation. In addition to accommodating a flexible nonparametric structure, the appeal of DEA and FDH estimators has increased since Gijbels et al. (1999) and Park et al. (2000) obtained their asymptotic distributions under some fairly reasonable assumptions.[2]

Although desirable from various perspectives, DEA and FDH type estimators have two serious deficiencies. First, since they are based on the idea of enveloping the observed data, these estimators are very sensitive to outliers or extreme observations and are inherently biased. Second, even in cases where the production technology induces a smooth production frontier, estimated frontiers based on FDH and DEA are

[1] See Simar and Wilson (2006) for a review of deterministic frontiers and illustrations of their widespread use.
[2] See the earlier work of Banker (1993) and Korostelev et al. (1995) for some preliminary asymptotic results.

discontinuous or piecewise linear function, respectively. Efforts to remedy such deficiencies have appeared in different nonparametric frontier modeling contexts. Knight (2001) proposed a local polynomial frontier estimator that envelops the data as a smooth function of input usage. Unfortunately, Knight's estimator has no explicit asymptotic distribution making it difficult to conduct inference and to correct its inherent bias. Similarly, the piecewise polynomial estimator of Hall et al. (1998) does not have an explicit asymptotic distribution rendering difficult its practical use. Girard and Jacob (2004) proposed a smooth kernel estimator of the frontier based on a convolution representation that extends the estimator of Geffroy (1964). The asymptotic properties of their estimator, however, are obtained for the case where the data χ_n is a Poisson process, rather than under the more conventional setting where χ_n is a random sample. Martins-Filho and Yao (2006) propose a nonparametric frontier estimator that is smooth, more robust to outliers, and asymptotically normal. However, since their estimator is based on local linear regression, convergence is slowed by the number of inputs (curse of dimensionality).

Prominent among these recent developments is the contributions of Aragon et al. (2005). They propose an alternative definition for the production function,

$$h(x) = \sup\{y \in \Re_+ : F(y/x) < 1\} \equiv \inf\{y \in \Re_+ : F(y/x) = 1\} \tag{1}$$

where $F(y/x) = \frac{F(x,y)}{F_X(x)}$, $F(x,y) = P(\{(X,Y) : X \leq x, Y \leq y\})$ and $F_X(x)$ is the associated marginal distribution of X. Since $F_X(x) > 0$, they restrict attention to $\Psi^* = \{(x,y) \in \Psi : F_X(x) > 0\}$. If the frontier $g(x)$ is monotone nondecreasing, a typical assumption in economic theory, then $h(x) = g(x)$ for all x such that $(x,y) \in \Psi^*$. Note that the assumption that $g(x)$ is monotone nondecreasing is equivalent to $F(y/x)$ being monotone nonincreasing on the set $\{x \in \Re_+^d : F_X(x) > 0\}$.[3] Aragon et al. observed that $g(x)$ is the order one quantile for the conditional distribution of Y given that $X \leq x$, where the inequality should be understood componentwise, and therefore $g(x) \equiv q_1(x) = \inf\{y \in \Re_+ : F(y/x) = 1\}$. As natural extension, they suggest the concept of a production function of continuous order $\alpha \in [0,1]$ given by

$$q_\alpha(x) = \inf\{y \in \Re_+ : F(y/x) \geq \alpha\}. \tag{2}$$

[3]See proposition 2.5 in Aragon et al. (2005)

The usefulness of this concept rests in the fact that if $F(\cdot/x)$ is strictly increasing on the support $[0, g(x)]$, then $q_\alpha(x) = F^{-1}(\alpha/x)$ where $F^{-1}(\cdot/x)$ is the inverse of $F(\cdot/x)$. In this context, any production plan $(x, y) \in \Psi^*$ belongs to some α-order conditional quantile curve, and is such that y represents an output level that is greater than 100α percent of the output of all production plans using inputs X such that $X \leq x$. Thus, rather than relying on $g(X_i)$ to define production efficiency of firm i, the conditional quantile function $q_\alpha(X_i)$ compares the production plan (X_i, Y_i) of firm i to all other $\{(X_j, Y_j)\}_{j \neq i}$ such that $X_j \leq X_i$.

Aragon et al. propose an estimator for $q_\alpha(x)$ that is based on conditional empirical quantile, obtained from inverting the empirical conditional distribution function $\hat{F}(y/x)$. Although their estimator has desirable properties of consistency and $\sqrt{n}$ asymptotic normality, it is well known from the unconditional distribution and quantile estimation literature (Azzalini, 1981; Falk, 1985; Yang, 1985) that smoothing beyond that given by the empirical distribution can produce significant gains in finite samples. In this paper, we propose a smooth nonparametric kernel estimator for the α-frontier ($q_\alpha(x)$). Our estimator is an extension of the seminal idea of Nadaraya (1964) and is based on a smooth estimator of the conditional distribution $F(y/x)$. Besides having the properties of consistency and $\sqrt{n}$-asymptotic normality, the variance of our estimator has an order of magnitude that is smaller than that of the estimator proposed by Aragon et al., confirming that the gains first identified by Azzalini in unconditional quantile estimation extend to conditional quantile estimation. Our simulations also confirm the superior performance of our proposed estimator.

Besides this introduction, this paper has five additional sections. Section 2 describes the stochastic model in detail, contrasts its assumptions with those in the past literature and describes the estimation procedure. Section 3 provides supporting lemmata and the main theorems establishing the asymptotic behavior of our estimator. Section 4 contains a Monte Carlo study that implements the estimator, sheds some light on its finite sample properties and compares its performance with that of the estimator proposed by Aragon et al. Section 5 provides an empirical illustration of our estimation procedure using data on electric utilities from the United States. Lastly, section 6 provides a summary and some directions for future work.

2 Stochastic Model and Estimation

2.1 α Frontier Estimator

We start by considering $\chi_n = \{(X_i, Y_i)\}_{i=1}^n$ a sequence of independent random vectors taking values in Ψ^* and having the same distribution F as the vector (X, Y). Throughout the paper, X will represent a d-vector of inputs used in the production process and Y will represent a scalar measure of output. F is taken to be absolutely continuous with associated density function given by f. The marginal distribution and density functions of X are denoted by F_X and f_X respectively. Given that our interest is on the estimation of the α-frontier, which coincides with conditional quantile $q_\alpha(x)$ for $\alpha \in [0, 1]$, we start by providing a definition of an estimator $\hat{F}(y/x)$ for $F(y/x)$. We write

$$\hat{F}(y/x) = \begin{cases} 0 & \text{if } y = 0, \\ \frac{\hat{F}(x,y)}{\hat{F}(x)} & \text{if } y > 0. \end{cases} \tag{3}$$

where $\hat{F}(y, x) = (nh_n)^{-1} \sum_{i=1}^n \int_0^y K\left(\frac{Y_i - \gamma}{h_n}\right) d\gamma I(\{X_i : X_i \leq x\})$ and $\hat{F}(x) = n^{-1} \sum_{i=1}^n I(\{X_i : X_i \leq x\})$, $I(A)$ is the indicator function for the set A, $K(\cdot)$ is a suitably defined kernel function and h_n is a nonstochastic sequence of bandwidths such that $0 < h_n \to 0$ as $n \to \infty$. The estimator is different from that proposed by Aragon et al. in that their estimator for $F(x, y)$ is given by $F_n(x, y) = n^{-1} \sum_{i=1} I(\{(X_i, Y_i) : X_i \leq x, Y_i \leq Y\})$, i.e., the joint empirical distribution function of (X, Y). In essence, rather than estimating $F(y/x)$ by the empirical distribution of the data such that $X_i \leq x$ for $i = 1, \cdots, n$, we estimate $F(y/x)$ by integrating a smooth Rosenblatt density estimator constructed using the observations $\{(X_i, Y_i)\}_{i \in \{i : X_i \leq x\}}$. It is easy to demonstrate that $\hat{F}(y/x)$ is asymptotically a distribution function in that the following properties hold for suitably defined kernels: (a) $\hat{F}(y/x)$ is nondecreasing in y; (b) $\hat{F}(y/x)$ is right continuous in $\Re_+$; (c) $lim_{y \to 0} \hat{F}(y/x) = 0$; and (d) there exists $N(x)$ such that for all $n > N(x)$ we have $lim_{y \to \infty} \hat{F}(y/x) = 1$.

Assuming that $q_\alpha(x)$ is the unique α order quantile for the conditional distribution $F(y/x)$, i.e., the unique root for $F(q_\alpha(x)/x) = \alpha$, we define the estimator $q_{\alpha,n}(x)$ as the root of

$$\hat{F}(q_{\alpha,n}(x)/x) = \alpha \text{ for } \alpha \in (0, 1] \text{ and } x \in \Re_+^d. \tag{4}$$

Using the mean value theorem, absolute continuity of F and smoothness of the kernel function we can write,

$$q_{\alpha,n}(x) - q_\alpha(x) = \frac{F(q_\alpha(x)/x) - \hat{F}(q_\alpha(x)/x)}{\hat{f}(\bar{q}_{\alpha,n}(x)/x)} \tag{5}$$

where $\hat{f}(y/x) = \frac{\partial \hat{F}(y/x)}{\partial y} = \frac{(nh_n)^{-1}\sum_{i=1}^n K\left(\frac{Y_i - y}{h_n}\right) I(\{X_i : X_i \leq x\})}{\hat{F}_X(x)}$ for $y \geq 0$ ($\hat{f}(y/x) = 0$ for $y < 0$) and $\bar{q}_{\alpha,n}(x) = \lambda q_{\alpha,n}(x) + (1-\lambda) q_\alpha(x)$ for $\lambda \in (0,1)$.

2.2 Assumptions

The stochastic properties of the estimator (4) are obtained under the following regularity conditions:

Assumption A1. a. $\chi_n = \{(X_i, Y_i)\}_{i=1}^n$ is a sequence of independent random vectors taking values in Ψ^* and having the same distribution F as the vector (X, Y), with support in Ψ^*. We assume that F is absolutely continuous with associated density function given by f. The marginal distribution and density functions of X are denoted by F_X and f_X respectively; b. Ψ^* is compact and $0 < f(x,y) < B_f$ for all $(x,y) \in \Psi^*$.

The assumption that χ_n is an independent and identically distributed sequence, and the existence of the density f as a bounded function in Ψ is standard in the deterministic frontier literature (Aragon et al., 2005, Cazals et al., 2002; Gijbels et al., 1999; Martins-Filho and Yao, 2006; Park et al., 2000). The following assumption A2 is standard in nonparametric estimation and involves only the kernel $K(\cdot)$. We observe that A2 is satisfied by commonly used kernels such as Biweight, Epanechnikov and others.

Assumption A2. a. $K(\gamma) : S_K \to \Re$ is a symmetric bounded function with compact support $S_K = [-B_K, B_K]$ such that: b. $\int_{-B_K}^{B_K} K(\gamma)d\gamma = 1$; c. $\int_{-B_K}^{B_K} \gamma K(\gamma)d\gamma = 0$, $\int_{-B_K}^{B_K} \gamma^2 K(\gamma)d\gamma = \sigma_K^2$; d. for all $\gamma, \gamma' \in S_K$ we have $|K(\gamma) - K(\gamma')| \leq m_K|\gamma - \gamma'|$ for some $0 < m_K < \infty$; e. for all $\gamma, \gamma' \in \Re$ we have $|\kappa(\gamma) - \kappa(\gamma')| \leq m_\kappa|\gamma - \gamma'|$ for some $0 < m_\kappa < \infty$, where $\kappa(\lambda) = \int_{-B_K}^{\lambda} K(\gamma)d\gamma$.

Assumption A3. a. f is continuous in Ψ^*; b. for all x such that $F_X(x) > 0$ and for all $\alpha \in (0,1]$, $f(q_\alpha(x)/x) > 0$, where $f(\cdot/x)$ is the derivative of $F(\cdot/x)$; c. for all $(x,y), (x,y') \in \Psi^*$, $|f(x,y') - f(x,y)| \leq m_f|y - y'|$ for some $0 < m_f < \infty$; d. F is twice continuously differentiable in the interior of Ψ^*.

A3.b is assumed by Aragon et al. (2005), and the Lipschitz condition in A3.c. is also assumed by Park et al. (2000).

Assumption A4. For all $y, y' \in G$, where G is a compact subset of $(0, \infty)$, we have $\left|\int_{g^{-1}([y,y'])} dX\right| \leq m_{g^{-1}}|y - y'|$ for some $0 < m_{g^{-1}} < \infty$. Here, let $x = (x_1, \cdots, x_d)'$, then for any two sets $A \subseteq C_x = \times_{i=1}^d [0, x_i]$ and $B \subseteq [0, g(x)]$, $g(A) = \{g(x) : x \in A\}$ and $g^{-1}(B) = \{x : x \in C_x,\ g(x) \in B\}$.

Assumption A4 imposes a Lipschitz type condition on the inverse image g^{-1} of g. Note, for example, that if $g : \Re_+ \to \Re_+$ is bijective with inverse g^{-1}, assumption A4 is equivalent to $\left|g^{-1}(y) - g^{-1}(y')\right| \leq m_{g^{-1}}|y - y'|$ for some $0 < m_{g^{-1}} < \infty$.

3 Asymptotic Characterization of the Estimator

We first establish the following auxiliary Lemma which provides the order for the bias and variance of the estimator $\hat{F}(x, y)$ for the joint distribution function $F(x, y)$ at $(x, y) \in \Psi^*$. The proofs for all lemmata and theorems are collected in Appendix 1. We observe that the proofs of lemmata and theorems that follow depend on repeated use of Lebesgue's Dominated Convergence (LDC) theorem, and as will be observed the technical difficulties involve verifying whether or not the conditions necessary to use the LDC theorem are met.

Lemma 1 *For all $(x, y) \in \Psi^*$ and under assumptions A1, A2.a, A2.b, A2.c, and A3, we have the following asymptotic representations for $E(\hat{F}(x, y))$ and $V(\hat{F}(x, y))$,*

(a)

$$E(\hat{F}(x,y)) = \begin{cases} F(x,y) + \frac{1}{2}h_n^2\sigma_K^2 \int_{g^{-1}([y,g(x)])} f^{(1)}(X,y)dX + o(h_n^2) & \text{if } 0 < y < g(x), \\ F(x,y) + o(h_n^2) & \text{if } y > g(x), \\ F(x,y) + o(h_n) & \text{if } y = g(x). \end{cases}$$

(b)

$$V(\hat{F}(x,y)) = \begin{cases} n^{-1}F(x,y)(1 - F(x,y)) - 2n^{-1}h_n\sigma_\kappa \int_{g^{-1}([y,g(x)])} f(X,y)dX + o\left(h_n/n\right) & \text{if } 0 < y < g(x), \\ n^{-1}F(x,y)(1 - F(x,y)) + o\left(h_n/n\right) & \text{if } y \geq g(x). \end{cases}$$

where $\kappa(x) = \int_{-B_K}^{x} K(\gamma)d\gamma$, $\sigma_\kappa = \int_{-B_K}^{B_K} \gamma\kappa(\gamma)K(\gamma)d\gamma$, $f^{(1)}(X, y)$ denotes the first derivative of f with respect to Y, and $0 < h_n \to 0$ is a nonstochastic sequence of bandwidths.

Lemma 1 can be viewed as an extension to the multivariate case of the second order results of Azzalini (1981), where the nonparametric distribution function estimator for $F(x, y)$ is given by $\hat{F}(x, y)$. We observe that $E(\hat{F}(x, y)) \to F(x, y)$ and $V(\hat{F}(x, y)) - n^{-1}F(x, y)(1 - F(x, y)) = o(1)$ as $n \to \infty$ demonstrating that the estimator for F that we propose has the same asymptotic bias and variance as the multivariate empirical distribution function estimator, the difference between the estimators is manifested in the order at which the bias and variance converge to zero. To obtain consistence of $q_{\alpha,n}(x)$ in theorem 1, we establish the following uniform convergence results on $\hat{F}(x,y)$.

Lemma 2 *Let $0 < h_n \to 0$ be a nonstochastic sequence of bandwidths with $nh_n \to \infty$ as $n \to \infty$. Assume that for a given $x \in \Re_+$ and some n we have that $min_{\{i:X_i \leq x\}} Y_i \geq h_n B_K$ and A1, A2, A3 and A4. Then, we have that,*

(a) $sup_{y\in[0,g(x)]}|\hat{F}(x,y) - E(\hat{F}(x,y))| = o_p(1)$

(b) $sup_{y\in[0,g(x)]}|E(\hat{F}(x,y)) - F(x,y))| = o(1)$.

Since, $h_n \to 0$ as $n \to \infty$, the assumption that $min_{\{i:X_i \leq x\}} Y_i \geq h_n B_K$ implies that even as the number of observations that satisfy $\{i : X_i \leq x\}$ grows to infinity, the associated output levels Y_i are bounded away from zero. Although reasonable in most contexts, it is certainly an assumption that could be violated by certain data generating processes. We now state,

Theorem 1 *Let $0 < h_n \to 0$ be a nonstochastic sequence of bandwidths with $nh_n \to \infty$ as $n \to \infty$. Assume that for a given $x \in \Re_+$ and some n we have that $min_{\{i:X_i \leq x\}} Y_i \geq h_n B_K$ and A1, A2, A3 and A4. Then, we have that,*

$$q_{\alpha,n}(x) - q_\alpha(x) = o_p(1). \tag{6}$$

Asymptotic normality of $q_{\alpha,n}(x)$ under suitable normalization is obtained in theorem 2, which relies on the following lemma.

Lemma 3 *Let $0 < h_n \to 0$ be a nonstochastic sequence of bandwidths with $nh_n^2 \to \infty$ as $n \to \infty$. Assume A1, A2, A3 and A4, then for a compact subset G of $(0, g(x))$ we have,*

$$sup_{y\in G} \left| \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right) I(\{X_i : X_i \leq x\}) - \int_{g^{-1}([y,g(x)])} f(X,y)dX \right| = o_p(1).$$

We now state,

Theorem 2 *Let $0 < h_n \to 0$ be a nonstochastic sequence of bandwidths with $nh_n^2 \to \infty$ and $nh_n^4 = O(1)$ as $n \to \infty$. Assume that for a given $x \in \Re_+$ and some n we have that $min_{\{i:X_i \leq x\}} Y_i \geq h_n B_K$ and A1, A2, A3* *and A4. Then, for all $\alpha \in (0,1)$ we have*

$$v_n(x)^{-1}\sqrt{n}(q_{\alpha,n}(x) - q_\alpha(x) - B_n(x)) \xrightarrow{d} N(0,1) \tag{7}$$

where $B_n(x) = -\frac{1}{2}h_n^2\sigma_K^2 \frac{\int_{g^{-1}([q_\alpha(x),g(x)])} f^{(1)}(\gamma, q_\alpha(x))d\gamma}{F_X(x)f(q_\alpha(x)/x)} + o(h_n^2)$ *and*

$$v_n^2(x) = \frac{1}{(F_X(x)f(q_\alpha(x)/x))^2}\left(F(x,q_\alpha(x)) - \frac{F^2(x,q_\alpha(x))}{F_X(x)} - 2h_n\sigma_\kappa \int_{g^{-1}([q_\alpha(x),g(x)])} f(\gamma, q_\alpha(x))d\gamma\right) + o(h_n).$$

It is important to note that the conditional quantile estimator proposed by Aragon at al. (2005) is also consistent and $\sqrt{n}$ asymptotically normal under similar assumptions, however there are some important differences between the estimators. First, we observe that although our estimator depends on kernel smoothing, and therefore a bandwidth h_n is necessary in constructing the estimator, there is no asymptotic cost as the rate of convergence to normality occurs at the parametric rate $\sqrt{n}$. Hence, as one would expect given the definition of the estimator, the number of inputs d has no impact on the convergence rate of the estimator, but most importantly even though there is smoothing in the construction of $\hat{F}(y/x)$ it produces no slowing effect on the convergence in distribution, a result obtained by Falk (1985) in the context of unconditional distribution functions. Second, although the extra smoothing we propose might impose modest computational costs compared to the estimator proposed by Aragon et al., theorem 2 reveals that the extra smoothness produces a variance of smaller order. Note that the variance of the asymptotic distribution of their estimator is given by

$$\frac{\alpha(1-\alpha)}{f^2(q_\alpha(x)/x)F_X(x)} \equiv \frac{1}{(F_X(x)f(q_\alpha(x)/x))^2}\left(F(x,q_\alpha(x)) - \frac{F^2(x,q_\alpha(x))}{F_X(x)}\right),$$

and given that the extra term that appears in v_n^2 is nonnegative, the variance of our estimator is smaller for all n finite. Third, the extra smoothing we propose does introduce a bias term $B_n(x) = O(h_n^2)$, but provided that $nh_n^4 = o(1)$ the bias vanishes asymptotically. We note that this condition is consistent with the conditions on h_n necessary to obtain theorem 2. Finally, we observe that given that $B_n(x) = O(h_n^2)$ and the variance is of order $O(n^{-1} + h_n n^{-1})$ the optimal bandwidth rate for minimization of the asymptotic mean integrated squared error is $h_n \propto n^{-1/3}$.

We now give an expanded asymptotic normality result for the asymptotic distribution of our α-frontier estimator. The result is similar to that in theorem 4.2. in Aragon et al. (2005).

Theorem 3 *Under the assumptions of Theorem 2, let $x^1, x^2, \cdots, x^r$ be r levels of input X such that at x^l for $l \in \{1, 2, \cdots, r\}$ and $\alpha \in (0,1)$. Then,*

$$\sqrt{n}\left(q_{\alpha,n}(x^1) - q_\alpha(x^1) - B(x^1), q_{\alpha,n}(x^2) - q_\alpha(x^2) - B(x^2), \cdots, q_{\alpha,n}(x^r) - q_\alpha(x^r) - B(x^r)\right)' \xrightarrow{d} N(0, Q)$$

where $B(x^l) = -\frac{1}{f(q_\alpha(x^l)/x^l)F_X(x^l)}\sigma_K^2 \frac{h_n^2}{2}\int_{g^{-1}([q_\alpha(x^l), g(x^l)])} f^{(1)}(X, q_\alpha(x^l))dX + o(h_n^2)$ and Q is an $r \times r$ matrix with $(l,m)^{th}$ element $Q_{l,m}$ given by

(1) $Q_{l,l} = \frac{\alpha(1-\alpha)}{f^2(q_\alpha(x^l)/x^l)F_X(x^l)}$ if $l = m$,

(2) $Q_{l,m} = \frac{1}{f(q_\alpha(x^l)/x^l)F_X(x^l)f(q_\alpha(x^m)|x^m)F_X(x^m)}[F(x^{lm}, q_\alpha(x^l))(1-\alpha) - \alpha F(x^{lm}, q_\alpha(x^m)) + \alpha^2 F_X(x^{lm})]$ if $l \neq m$, and $q_\alpha(x^l) \leq q_\alpha(x^m)$,

(3) $Q_{l,m} = \frac{1}{f(q_\alpha(x^l)/x^l)F_X(x^l)f(q_\alpha(x^m)|x^m)F_X(x^m)}[F(x^{lm}, q_\alpha(x^m))(1-\alpha) - \alpha F(x^{lm}, q_\alpha(x^l)) + \alpha^2 F_X(x^{lm})]$ if $l \neq m$, and $q_\alpha(x^l) \geq q_\alpha(x^m)$, where $x^{lm} = \{\min(x_1^l, x_1^m), \min(x_2^l, x_2^m) \cdots, \min(x_d^l, x_d^m)\}'$,

As is typical in applied work, for inference purposes, the unknown higher order components of the variance terms in theorems 2 and 3 must be estimated *via* consistent nonparametric estimators. $f(q_\alpha(x)/x)$ can be estimated by $\hat{f}(q_{\alpha,n}(x)/x)$ the conditional Rosenblatt density estimator, using the *rule-of-thumb* bandwidth of Silverman (1986). Note the consistency of $\hat{f}(q_{\alpha,n}(x)/x)$ has been established in the proof of theorem 2. Furthermore, $F_X(x)$ can be consistently estimated by $\hat{F}(x) = n^{-1}\sum_{i=1}^n I(\{X_i : X_i \leq x\})$, and $F(x, q_\alpha(x'))$ can be consistently by $\hat{F}(x, q_{\alpha,n}(x'))$.

We now turn our attention to the estimation of the true frontier $g(x)$, or alternatively, $q_1(x)$. We first show that if $min_{\{i:X_i\leq x\}}Y_i \geq h_n B_K$, $q_{1,n}(x)$ is asymptotically equivalent to the FDH estimator. We then rely on Park et al. (2000) to obtain the asymptotic distribution of $q_{1,n}(x)$.

Theorem 4 *Assume that $min_{\{i:X_i\leq x\}}Y_i \geq h_n B_K$, and that A1, A2 hold with Ψ^* compact. In addition, assume that the density f is strictly positive on the frontier $\{(x, g(x)) : F_X(x) > 0\}$, and that $g(x)$ is continuously differentiable. Then for any x in the interior of the support of X we have that*

a) There exists $N(x) > 0$ such that for all $n > N(x)$, $q_{1,n}(x) = max_{\{i:X_i\leq x\}}Y_i + h_n B_K$.

b) $n^{1/(d+1)}(q_1(x) - q_{1,n}(x) + h_n B_K) \xrightarrow{d} Weibull(\mu_x^{d+1}, d+1)$.

μ_x is a constant depending on the slope of $g(\cdot)$ and the value of f at the frontier. Park et al. (2000) provide the exact expression for μ_x as well as a consistent estimator for μ_x. We note that by their theorem 3.3, it is a direct consequence of the assumptions in theorem 4 that

$$E(q_1(x) - q_{1,n}(x)) = \Gamma\left(\frac{d+2}{d+1}\right)\mu_x^{-1}n^{-1/(d+1)} - h_n B_K + o(n^{-1/(d+1)})$$

which suggests that the bias associated with the estimation of the true frontier $q_1(x)$ via $q_{1,n}(x)$ could be smaller than that associated with the FDH estimator. We now turn our attention to a small Monte Carlo study designed to compare our smooth estimator with that proposed by Aragon et al. (2005).

4 Monte Carlo Simulation Study

In this section, we perform a Monte Carlo study to implement our smooth α-frontier estimator and illustrate its finite sample performance. For comparison purpose we also include the empirical α-frontier estimator by Aragon et al. (2005) in the simulations. The data are simulated according to the model

$$Y_i = g(X_i)R_i$$

where Y_i represent the output, the univariate input X_i are pseudo random variables generated from a uniform distribution with support given by $[b_l, b_u]$. $R_i = exp(-Z_i)$ and Z_i are independently generated pseudo random variables from an exponential distribution with parameter $\beta = \frac{1}{3}$, therefore the efficiency R_i has support $(0, 1]$ with global average level of efficiency $E(R_i) = 0.75$.

We consider two specifications for $g(\cdot)$:

$$g_1(x) = \sqrt{x} \text{ with } [b_l, b_u] = [4, 25] \text{ and } g_2(x) = x^3 \text{ with } [b_l, b_u] = [1, 2]$$

which are associated with convex and nonconvex production technologies respectively. The DGP used in the simulation has been considered in Park et al. (2000), Aragon et al. (2005), Gijbels et al. (1999), Martins-Filho and Yao (2006) and is regarded as reasonable with respect to many applications found in the econometric literature (Gijbels et al., 1999, p. 224).

For each specification of $g(x)$ we consider three sample sizes $n = 100, 200$ and 400 and perform 1000 repetitions at each experiment design. We estimate the $\alpha-$frontier at different levels of $\alpha - 0.75, 0.97, 0.98, 0.99,$ and 1 since high levels of α's are likely to be of interest in production frontier estimation, at three different points, namely $x_0 = 9, 16,$ and 25 for $g_1(x)$ and $x_0 = 1.33, 1.66,$ and 2 for $g_2(x)$. The values of x_0 correspond to the 33^{rd}, 66^{th} and 100^{th} percentiles of the support of X_i. We also construct 95% asymptotic confidence intervals for the $\alpha-$frontier at $\alpha = 0.75, 0.97, 0.98, 0.99$ with the asymptotic distribution results in previous section.

To implement our estimator, we use the Epanechnikov kernel and consider the following *rule-of-thumb* bandwidth as in Azzalini (1981),

$$\hat{h}_{ROT} = 1.3\hat{\sigma}_Y(x)n^{-\frac{1}{3}}$$

where $\hat{\sigma}_Y(x)$ is the standard deviation of observations Y_i such that $X_i \leq x$, keeping in mind that the conditions on bandwidth in previous sections are satisfied. Given the asymptotic results, we know that the bias of our smooth estimator is of order $O(h_n^2)$ and the variance is of order $O(\frac{1}{n} + \frac{h_n}{n})$, so the *rule-of thumb* bandwidth is also of the order that minimizes the mean squared error of our smooth estimator.

Given the convergence results asymptotic confidence intervals for the $\alpha-$frontier can be constructed. With the *rule-of-thumb* bandwidth $\hat{h}_{ROT}$, we have the asymptotic bias of order $O(\sqrt{n}\hat{h}_{ROT}^2) = O(n^{-\frac{1}{6}}) = o(1)$. Hence, the bias term disappears asymptotically and for 97.5% quantile $Z_{0.975}$ of a standard normal distribution, we obtain

$$\lim_{n\to\infty} P(q_{\alpha,n}(x) - n^{-\frac{1}{2}}(\hat{S}_2^2)^{\frac{1}{2}}Z_{0.975} \leq q_\alpha(x) \leq q_{\alpha,n}(x) + n^{-\frac{1}{2}}(\hat{S}_2^2)^{\frac{1}{2}}Z_{0.975}) = 0.95$$

where $\hat{S}_2^2 = \frac{\alpha(1-\alpha)}{\hat{F}(x)(\hat{f}(q_{\alpha,n}(x)/x))^2}$, $\hat{F}(x)$ is the empirical distribution function, and

$$\hat{f}(q_{\alpha,n}(x)/x) = \frac{\frac{1}{nh_n}\sum_{i=1}^n K(\frac{Y_i - q_{\alpha,n}(x)}{h})I(\{X_i : X_i \leq x\})}{\hat{F}(x)}$$

Since $\hat{f}(q_{\alpha,n}(x)/x)$ is essentially the Rosenblatt density estimator, we utilize the *rule-of-thumb* bandwidth of Silverman (1986). Note the consistency of $\hat{f}(q_{\alpha,n}(x)/x)$ has been established in previous section. The asymptotic confidence interval is constructed similarly with the empirical α-frontier estimator by Aragon et al. (2005).

A typical simulated dataset from the DGP has been plotted in Figures 1 and 2 for $g_1(x)$ and $g_2(x)$ respectively with sample size $n = 100$, together with the true $\alpha-$frontier, estimated smooth and empirical frontier for $\alpha = 1$ and 0.75. The appearance of the two estimators are similar for $\alpha = 0.75$ and they deviate to a larger degree with $\alpha = 1$. For illustration purpose, we also depict the true $\alpha-$frontier with estimated smooth and empirical $\alpha-$frontier for α ranging over $0.02, 0.04 \cdots, 1$ in Figure 3, for a simulated dataset of size $n = 50$ with $g_1(x) = g_1(25)$. As expected, our $\alpha-$frontier estimate is smooth function of α and the empirical $\alpha-$frontier is not.

The simulation results are given in Tables 1-3. Tables 1 and 2 provide the bias and root mean squared error of $\alpha-$frontier estimators at three values of x_0 and five values of α we indicate above, for $g_1(x)$ and $g_2(x)$, respectively. The empirical estimator always shows a negative bias and in most experiment designs has larger bias than the smooth estimator, where exceptions occur only for $\alpha \in (0, 1)$. As can be inferred from the tables but not reported here to save space, the standard deviation of the smooth estimator is always smaller than that of the empirical estimator across different experiments for $\alpha \in (0, 1)$, but it is not the case for $\alpha = 1$. The above results are consistent with the asymptotic results established in previous section for $\alpha \in (0, 1)$, that the variance of the smooth estimator is smaller than that of the empirical estimator by a term of second order magnitude, but the bias could be larger. For $\alpha = 1$ the bias of the smooth estimator is smaller than that of the empirical estimator (FDH). This confirms the comment we made following our theorem 4. In terms of root mean squared error, the smooth estimator outperforms the empirical estimator for all sample sizes, different x_0 points evaluated, α values and the $g(x)$ specifications.

We find that as sample size increases, with some exceptions for the bias, the bias and root mean squared error of both estimators decrease, which confirms the asymptotic results in previous section. We also observe that the bias and root mean squared error for both estimators are larger when evaluating the $\alpha = 1$ frontier than when evaluating frontiers with $\alpha < 1$. This observation is consistent with the asymptotics that both α-frontier estimators converge at a speed of $\sqrt{n}$ for $\alpha \in (0, 1)$, which is in general faster than the convergence rate when evaluating the full frontier with $\alpha = 1$. The fact that it is more difficult to estimate the full frontier in terms of bias and root mean squared error is also intuitively expected, as there are relatively less representative data available in estimating the full frontier.

The empirical coverage probability (the frequency that the estimated confidence interval contains the true $\alpha-$frontier in 1000 repetitions) is given in Table 3. For all experiment designs, we observe that smooth estimator is superior to empirical estimator, i.e., the empirical coverage probability with smooth estimator is closer to the target value 95% than that with empirical estimator. As sample size increases, there is tendency that the empirical coverage probabilities from both estimators get closer to 95% with some exceptions. There is also weak evidence that the as α decreases, the empirical coverage probability gets closer to 95%.

To provide further evidence on the finite sample distribution of the two estimators, we provide in Figure 4 the kernel density estimates for the smooth and empirical α-frontier estimators centered around the true value $q_{0.99}(25)$, the $\alpha = 0.99$ frontier function evaluated at $x_0 = 25$, based on 1000 simulations from $g_1(x)$ of sample sizes $n = 100$ and 400. The kernel density estimates are calculated using the Epanechnikov kernel and the *rule-of-thumb* bandwidth of Silverman (1986). From Figure 4 we notice that the kernel density with smooth estimator is shifted to the right and more tightly distributed around zero, illustrating the improvement over the empirical estimator. Also, the estimated densities have taller and more pronounced spikes as the sample size increases, confirming the asymptotic results. Based on above results, it might be reasonable to conclude that in our simulation experiments, the smooth estimator performs better *vis a vis* the empirical estimator and the extra smoothing pays off at least in finite sample.

5 Empirical Illustration

To illustrate our methodology, we employ data on 123 utility companies from the United States reported in Greene (1990). These data consist of variables on production cost, output, input prices, and has been analyzed by Christensen and Greene (1976), Greene (1990) and Gijbels et al. (1999). Following Gijbels et al. (1999), we utilize only the measurements on the output variable with $Y = Ln(Q)$ and input or cost variable defined as $X = Ln(C)$, where Q is the production output for a firm, and C is the total cost involved in the production. For detailed description of the data set and analysis, see Christensen and Greene (1976) and Greene (1990).

We provide in Figure 5 a scatterplot of the data, together with the estimated smooth and empirical frontiers for $\alpha = 0.99$ and 0.90. For illustration purpose, we restrict the estimation region to be $x \in [0, 6]$, where 116 out of the 123 observations are located. The bandwidth for our smooth estimator is selected according to the *rule-of-thumb* $\hat{h}_{ROT}$ as described in the simulation section. As expect based on the graphs produced in our simulation section, the graphs for both frontier estimators are similar for $\alpha = 0.90$. In the case where $\alpha = 0.99$ and for $x \leq 3$, the smooth estimator gives slightly larger prediction for the $\alpha = 0.99$ frontier, but they are alike for $x \geq 3$. Given the comments in Aragon et al. (2005) regarding the robustness of the empirical frontier to extreme observations, we conjecture that for $\alpha \in (0, 1)$ our smooth estimator should also be reasonably robust to extreme values and outliers. In Figure 6, we construct 95% confidence intervals for the $\alpha = 0.90$ frontiers using the smooth and empirical estimators following the steps outlined in the simulation section. Besides the similarity of both estimates, we note that the confidence bands are wider in regions of the input space where there are a smaller number of observations. This follows from our definition for asymptotic confidence intervals and Theorem 4.1 of Aragon et al. Indeed the width of the confidence interval depends on the density $f(q_\alpha(x)/x)$ and marginal probability $F_X(x)$. In regions of the input space where there are more data, both the density and marginal probability will be larger, and hence it is natural to observe narrower confidence intervals.

6 Summary

In this paper we proposed a nonparametric α-frontier estimator based on a smooth kernel estimator of a conditional quantile of order α. Our estimator is an alternative to the conditional quantile estimator proposed by Aragon et al. (2005), which is based on empirical distribution functions. The estimator is easily implementable and we show that it is consistent and $\sqrt{n}$ asymptotically normal, i.e., the extral smoothing and the need to select a bandwidth produces no harm to the convergence rate. In addition, the extra smoothness pays off in that our estimator's variance is of smaller order than that of the estimator proposed by Aragon et al. (2005). Our simulation study confirms the asymptotic theory predictions and contrasts our estimator with that of Aragon et al. In the simulations, our smooth estimator outperforms the empirical distribution based estimator of Aragon et al. (2005). Future work is needed in the context of α-frontiers, specifically estimators that can produce smooth boundaries for the output set are desirable in the applied economics literature.

Appendix 1- Proofs

Lemma 1. *Proof* (a): Let $C_x = \times_{i=1}^{d}[0, x_i]$ where x_i is the i^{th} component of x. Since $h_n \to 0$ as $n \to \infty$, there exists $N \in \Re_+$ such that for all $n > N$,

$$E(\hat{F}(x, y)) = \int_{C_x}\int_{[0,g(X)]}\int_{-B_K}^{(y-Y)/h_n} K(\gamma)d\gamma f(X, Y)dYdX.$$

Let $F_f(x, y) = \int_{[0,y]} f(x, \gamma)d\gamma$, then $E(\hat{F}(x, y)) = \int_{C_x}\int_{[0,g(X)]}\kappa\left(\frac{y-Y}{h_n}\right)\frac{\partial F_f(X,Y)}{\partial Y}dYdX$. Using integration by parts

$$\int_{[0,g(X)]}\kappa\left(\frac{y-Y}{h_n}\right)dF_f(X, Y) = \kappa\left(\frac{y-g(X)}{h_n}\right)F_f(X, g(X)) + \int_{\frac{y-g(X)}{h_n}}^{y/h_n} F_f(X, y - h_n\gamma)K(\gamma)d\gamma.$$

By A3.d and Taylor's theorem $F_f(X, y - h_n\gamma) = F_f(X, y) - h_n\gamma f(X, y) + \frac{1}{2}h_n^2\gamma^2 f^{(1)}(X, y) + o(h_n^2)$ where $f^{(1)}(X, y) = \frac{\partial f(X,y)}{\partial Y}$. Hence, we write $E(\hat{F}(x, y)) = E_{1n} + E_{2n} - E_{3n} + E_{4n} + o(h_n^2)$, where

$$\begin{aligned}
E_{1n} &= \int_{C_x}\kappa\left(\frac{y-g(X)}{h_n}\right)F_f(X, g(X))dX \\
E_{2n} &= \int_{C_x}F_f(X, y)\int_{(y-g(X))/h_n}^{y/h_n}K(\gamma)d\gamma dX \\
E_{3n} &= h_n\int_{C_x}f(X, y)\int_{(y-g(X))/h_n}^{y/h_n}\gamma K(\gamma)d\gamma dX \\
E_{4n} &= \frac{h_n^2}{2}\int_{C_x}f^{(1)}(X, y)\int_{(y-g(X))/h_n}^{y/h_n}\gamma^2 K(\gamma)d\gamma dX.
\end{aligned}$$

For $(x, y) \in \Psi^*$, if $y \leq 0$ then $\hat{F}(x, y) = 0$. We now consider the limiting behavior of each term when: (1) $0 < y < g(x)$; (2) $y > g(x)$; (3) $y = g(x)$.

(1): For any $A \subseteq C_x$ and $B \subseteq [0, g(x)]$, let $g(A) = \{g(x) : x \in A\}$ and $g^{-1}(B) = \{x : x \in C_x, g(x) \in B\}$. Then, $E_{1n} = \int_{g^{-1}([0,y])}\kappa\left(\frac{y-g(X)}{h_n}\right)F_f(X, g(X))dX + \int_{g^{-1}([y,g(x)])}\kappa\left(\frac{y-g(X)}{h_n}\right)F_f(X, g(X))dX = E_{11,n} + E_{12,n}$. First, observe that $\left|\kappa\left(\frac{y-g(X)}{h_n}\right)\right||F_f(X, g(X)| < B < \infty$ for some $0 < B < \infty$ given A1. Note that in the case of $E_{11,n}$ we have $X \in g^{-1}([0, y])$ and $\kappa\left(\frac{y-g(X)}{h_n}\right) \to 1$. Hence, by Lebesgue's dominated convergence theorem $E_{11,n} \to \int_{g^{-1}([0,y])}\int_{[0,g(X)]} f(X, Y)dXdY$. For $E_{12,n}$, since $X \in g^{-1}([y, g(x)])$ we have that $\kappa\left(\frac{y-g(X)}{h_n}\right) \to 0$, hence by Lebesgue's dominated convergence (LDC) theorem, we have $E_{12,n} \to 0$.

$E_{2n} = \int_{g^{-1}([0,y])}F_f(X, y)\int_{(y-g(X))/h_n}^{y/h_n}K(\gamma)d\gamma dX + \int_{g^{-1}([y,g(x)])}F_f(X, y)\int_{(y-g(X))/h_n}^{y/h_n}K(\gamma)d\gamma dX$. For $X \in$

$g^{-1}([0,y])$ we have that $\int_{(y-g(X))/h_n}^{y/h_n} K(\gamma)d\gamma \to 0$, and for $X \in g^{-1}([y,g(x)])$ we have $\int_{(y-g(X))/h_n}^{y/h_n} K(\gamma)d\gamma \to 1$. Therefore, $E_{2n} \to \int_{g^{-1}([y,g(x)])}\int_{[0,y]} f(X,Y)dYdX$.

$E_{3n} = h_n \int_{g^{-1}([0,y])} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma K(\gamma)d\gamma dX + h_n \int_{g^{-1}([y,g(x)])} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma K(\gamma)d\gamma dX = E_{31,n} + E_{32,n}$. For $X \in g^{-1}([0,y])$ we have that $\int_{(y-g(X))/h_n}^{y/h_n} \gamma K(\gamma)d\gamma \to 0$, hence $h_n^{-1}E_{31,n} \to 0$. By A2.c, for $X \in g^{-1}([y,g(x)])$ we have that $\int_{(y-g(X))/h_n}^{y/h_n} \gamma K(\gamma)d\gamma \to 0$, and therefore $h_n^{-1}E_{3n} \to 0$. Now,

$$\begin{aligned} E_{4n} &= \frac{h_n^2}{2}\int_{g^{-1}([0,y])} f^{(1)}(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma^2 K(\gamma)d\gamma dX \\ &+ \frac{h_n^2}{2}\int_{g^{-1}([y,g(x)])} f^{(1)}(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma^2 K(\gamma)d\gamma dX \\ &= E_{41,n} + E_{42,n} \end{aligned}$$

For $X \in g^{-1}([0,y])$ we have that $\int_{(y-g(X))/h_n}^{y/h_n} \gamma^2 K(\gamma)d\gamma \to 0$, hence $h_n^{-2}E_{41,n} \to 0$. For $X \in g^{-1}([y,g(x)])$ we have that $\int_{(y-g(X))/h_n}^{y/h_n} \gamma^2 K(\gamma)d\gamma \to \sigma_K^2$ by A2.c, and $h_n^{-2}E_{4n} \to \frac{1}{2}\sigma_K^2 \int_{g^{-1}([y,g(X)])} f^{(1)}(X,y)dX$. Hence, for $0 < y < g(x)$ we have $E(\hat{F}(x,y)) = F(x,y) + \frac{h_n^2}{2}\sigma_K^2 \int_{g^{-1}([y,g(x)])} f^{(1)}(X,y)dX + o(h_n^2)$.

(2): If $y > g(x)$, $E_{1n} = \int_{g^{-1}([0,g(x)])} \kappa\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX$ and since $X \in g^{-1}([0,g(x)])$ we have that $g(X) < y$ and $\kappa\left(\frac{y-g(X)}{h_n}\right) \to 1$. Hence, by LDC theorem $E_{1n} \to \int_{C_x} F_f(X,g(X))dX = F(x,y)$. Similarly, $E_{2n} \to 0$, $h_n^{-1}E_{3n} \to 0$, $h_n^{-2}E_{4n} \to 0$, since $\frac{y-g(X)}{h_n} \to \infty$ as $n \to \infty$. Consequently, $E(\hat{F}(x,y)) = F(x,y) + o(h_n^2)$.

(3): If $y = g(x)$, then $f(X,y)$ is not differentiable and $F_f(X,y-h_n\gamma) = F_f(X,y) - h_n\gamma f(X,y) + o(h_n)$. Hence, we write $E(\hat{F}(x,y)) = E_{1n} + E_{2n} - E_{3n} + o(h_n)$. In this case, $E_{1n} = \int_{g^{-1}([0,g(x)))} \kappa\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX$ and since for $X \in g^{-1}([0,g(x)))$ we have that $g(X) < y$, $\kappa\left(\frac{y-g(X)}{h_n}\right) \to 1$ as $n \to \infty$, and by LDC theorem $E_{1n} \to \int_{g^{-1}([0,g(x)))} F_f(X,g(X))dX = F(x,y)$. Similarly, $E_{2n} = \int_{C_x} F_f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} K(\gamma)d\gamma dX \to 0$ and $E_{3n} = h_n \int_{C_x} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma K(\gamma)d\gamma dX \to 0$ as $n \to \infty$. Hence, when $y = g(x)$ we have that $E(\hat{F}(x,y)) = F(x,y) + o(h_n)$.

(b) $V(\hat{F}(x,y)) = \frac{1}{n}(V_{1n} - V_{2n})$ where $V_{1n} = E\left(\left(\frac{1}{h_n}\int_0^y K\left(\frac{\gamma-Y}{h_n}\right)d\gamma\right)^2 I(\{X_i : X_i \leq x\})\right)$ and $V_{2n} = \left(E\left(\frac{1}{h_n}\int_0^y K\left(\frac{\gamma-Y}{h_n}\right)d\gamma I(\{X_i : X_i \leq x\})\right)\right)^2$. We first examine the term V_{1n}. Given that $h_n \to 0$, there exists $N \in \Re_+$such that for all $n > N$ we have that $V_{1n} = \int_{C_x}\int_{[0,g(X)]} \kappa^2\left(\frac{y-Y_i}{h_n}\right)\frac{\partial F_f(X,Y)}{\partial Y}dYdX$. As in part (a), using integration by parts and the fact that $F_f(X, y-h_n\gamma) = F_f(X,y) - h_n\gamma f(X,y) + o(h_n)$, we obtain

$$\begin{aligned} V_{1n} &= \int_{C_x} \kappa^2\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX + 2\int_{C_x} F_f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \kappa(\gamma)K(\gamma)d\gamma dX \\ &- 2h_n\int_{C_x} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma\kappa(\gamma)K(\gamma)d\gamma dX + o(h_n) = V_{11,n} + V_{12,n} + V_{13,n} + o(h_n) \end{aligned}$$

We consider the asymptotic behavior of each term for (1) $0 < y < g(x)$ and (2) $y \geq g(x)$.

(1): $V_{11,n} = \int_{g^{-1}([0,y])} \kappa^2\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX + \int_{g^{-1}([y,g(x)])} \kappa^2\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX = v_{1n} + v_{2n}$. Given that $y < g(x)$, and our assumptions allow for repeated use of LDC theorem to obtain $v_{1n} \to \int_{g^{-1}([0,y))} F_f(X,g(X))dX$ and $v_{2n} \to 0$ as $n \to \infty$. Consequently, $V_{11,n} \to \int_{g^{-1}([0,y))} F_f(X,g(X))dX$. Note that,

$$\begin{aligned} V_{12,n} &= 2\left(\int_{g^{-1}([0,y])} F_f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \kappa(\gamma)K(\gamma)d\gamma dX \right. \\ &+ \left. \int_{g^{-1}([y,g(x)])} F_f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \kappa(\gamma)K(\gamma)d\gamma dX\right), \end{aligned}$$

and by LDC theorem we have

$$V_{12,n} \to 2\int_{g^{-1}([y,g(x)])} F_f(X,y)\int_{-B_K}^{B_K} \kappa(\gamma)K(\gamma)d\gamma dX = \int_{g^{-1}([y,g(x)])} F_f(X,y)dX,$$

since $\int_{-B_K}^{B_K} \kappa(\gamma)K(\gamma)d\gamma = \frac{1}{2}$. Now,

$$\begin{aligned} V_{13,n} &= -2h_n\left(\int_{g^{-1}([0,y))} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma\kappa(\gamma)K(\gamma)d\gamma dX \right. \\ &+ \left. \int_{g^{-1}([y,g(x)])} f(X,y)\int_{(y-g(X))/h_n}^{y/h_n} \gamma\kappa(\gamma)K(\gamma)d\gamma dX\right) \end{aligned}$$

and by LDC theorem $\frac{V_{13,n}}{h_n} \to -2\sigma_\kappa \int_{g^{-1}([y,g(x)])} f(X,y)dX$, where $\sigma_\kappa = \int_{-B_K}^{B_K} \gamma\kappa(\gamma)K(\gamma)d\gamma$. Hence, we conclude that $V_{1n} = F(x,y) - 2h_n\sigma_\kappa \int_{g^{-1}([y,g(x)])} f(X,y)dX + o(h_n)$.

(2): $V_{11,n} = \int_{g^{-1}([0,g(x)])} \kappa^2\left(\frac{y-g(X)}{h_n}\right) F_f(X,g(X))dX$ and by LDC theorem we have

$$V_{11,n} \to \int_{g^{-1}([0,g(x)])} F_f(X,g(X))dX = F(x,y).$$

Similarly, $V_{12,n} \to 0$ and $\frac{V_{13,n}}{h_n} \to 0$ as $n \to \infty$. Consequently, $V_{1n} = F(x,y) + o(h_n)$. Since $V_{2n} = \left(E(\hat{F}(x,y))\right)^2$, we obtain directly from the results on part (a) that

$$V_{2n} = \begin{cases} F^2(x,y) + O(h_n^2) & \text{if } 0 < y < g(x), \\ F^2(x,y) + o(h_n) & \text{if } y = g(x). \\ F^2(x,y) + o(h_n^2) & \text{if } y > g(x). \end{cases} \tag{8}$$

Consequently, combining the results for V_{1n} and V_{2n} we have,

$$V(\hat{F}(x,y)) = \begin{cases} n^{-1}F(x,y)(1-F(x,y)) - 2n^{-1}h_n\sigma_\kappa \int_{g^{-1}([y,g(x)])}^{x} f(X,y)dX + o\left(h_n/n\right) & \text{if } 0 < y < g(x), \\ n^{-1}F(x,y)(1-F(x,y)) + o\left(h_n/n\right) & \text{if } y \geq g(x). \end{cases}$$

Lemma 2 . *Proof* (a) Under the assumption that $min_{\{i:X_i \leq x\}} Y_i \geq h_n B_K$ $\hat{F}(x,y) = \frac{1}{n}\sum_{i=1}^{n} \kappa\left(\frac{y-Y}{h_n}\right) I(\{X_i : X_i \leq x\})$. Let $G(x) = [0, g(x)]$, and note that since $G(x)$ is compact, there exists $y_0 \in G(x)$ such that $G(x) \subseteq B(y_0, r_x)$ where $B(y_0, r_x) = \{y \in \Re : |y - y_0| < r_x\}$ where $r_x \in \Re_+$. Let $0 < h_n \to 0$ as $n \to \infty$, then by the Heine-Borel theorem, for every n there exists a finite collection of sets $\{B(y_k, (n/h_n^a)^{-1/2})\}_{k=1}^{l_n}$, $a > 0$ such that $G(x) \subset \cup_{k=1}^{l_n} B(y_k, (n/h_n^a)^{-1/2})$ for $y_k \in G(x)$ with $l_n < r_x(n/h_n^a)^{1/2}$. For $y \in B(y_k, (n/h_n^a)^{-1/2})$ we have that

$$\begin{aligned} |\hat{F}(x,y) - \hat{F}(x,y_k)| &\leq \frac{1}{n}\sum_{i=1}^{n}\left|\kappa\left(\frac{y-Y}{h_n}\right) - \kappa\left(\frac{y_k-Y}{h_n}\right)\right| I(\{X_i : X_i \leq x\}) \\ &\leq m_\kappa h_n^{-1}|y-y_k| \text{ by A2.e and the fact that } I(\{X_i : X_i \leq x\}) \leq 1 \\ &\leq m_\kappa (nh_n^{2-a})^{-1/2}. \end{aligned}$$

Also,

$$\begin{aligned} |E(\hat{F}(x,y)) - E(\hat{F}(x,y_k))| &\leq \int_{C_x}\int_{[0,g(X)]}\left|\kappa\left(\frac{y-Y}{h_n}\right) - \kappa\left(\frac{y_k-Y}{h_n}\right)\right| f(X,Y)dYdX \\ &\leq m_\kappa F_X(x)(nh_n^{2-a})^{-1/2} \text{ by A2.e.} \end{aligned}$$

Therefore, we can write $|\hat{F}(x,y) - E(\hat{F}(x,y))| \leq |\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| + m_\kappa(nh_n^{2-a})^{-1/2}(1+F_X(x))$ and conclude that $sup_{y\in G(x)}|\hat{F}(x,y) - E(\hat{F}(x,y))| \leq max_{1\leq k\leq l_n}|\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| + m_\kappa(nh_n^{2-a})^{-1/2}(1+F_X(x))$. Taking $a = 1$ and under the assumption that $nh_n \to \infty$ then $m_\kappa(nh_n^{2-a})^{-1/2} \to 0$ and we need only show that for all $\varepsilon_n > 0$, $lim_{n\to\infty} P\left(max_{1\leq k\leq l_n}|\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| \geq \varepsilon_n\right) = 0$. For this limit to hold it suffices to establish that $lim_{n\to\infty}\sum_{l=1}^{l_n} P\left(|\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| \geq \varepsilon_n\right) = 0$. Note that, $|\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| = \left|n^{-1}\sum_{i=1}^{n} W_{in}\right|$ where $W_{in} = \kappa\left(\frac{y_k - Y_i}{h_n}\right) I(\{X_i : X_i \leq x\}) - E\left(\kappa\left(\frac{y_k-Y_i}{h_n}\right) I(\{X_i : X_i \leq x\})\right)$, where $E(W_{in}) = 0$ and $|W_{in}| \leq 2$ given that $I(\cdot), \kappa(\cdot) \leq 1$. Given the measurability of κ and I and the fact that χ_n forms an independent sequence, by Bernstein's inequality we have $P\left(|\hat{F}(x,y_k) - E(\hat{F}(x,y_k))| \geq \varepsilon\right) < 2exp\left(\frac{-n\varepsilon_n^2}{2\bar{\sigma}^2 + \frac{4}{3}\varepsilon_n}\right)$, with $\bar{\sigma}^2 = n^{-1}\sum_{i=1}^{n} V(W_{in}) = V_{1n}(x,y_k) - V_{2n}(x,y_k) \to F(x,y_k)(1-F(x,y_k))$ for $y_k \in$

$G(x)$. Let $c_n = 2\bar{\sigma}^2 + \frac{4}{3}\varepsilon_n$ and $\varepsilon_n = \left(\frac{ln(n)}{n}\right)^{1/2} \Delta$ for some $\Delta > 0$. Then, $P\left(|\hat{F}(x, y_k) - E(\hat{F}(x, y_k))| \geq \varepsilon\right) \leq 2r_x \left(\frac{n}{h_n}\right)^{1/2} n^{-\Delta/c_n} \leq r_x(nh_n)^{-1/2}$, where the last inequality follows for Δ sufficiently large. Hence, provided that $nh_n \to \infty$ we have that $lim_{n\to\infty} P\left(\left(\frac{n}{ln(n)}\right)^{1/2} sup_{y\in G(x)}|\hat{F}(x,y) - E(\hat{F}(x,y))| \geq \Delta\right) = 0$ and consequently $sup_{y\in G(x)}|\hat{F}(x,y) - E(\hat{F}(x,y))| = o_p(1)$.

(b) Note that

$$\begin{aligned} sup_{y\in[0,g(x)]}|E(\hat{F}(x,y)) - F(x,y))| &\leq sup_{y\in[0,g(x)]}\left|E_{1n}(y) - \int_{g^{-1}([0,y])}\int_{[0,g(X)]} f(X,Y)dYdX\right| \\ &+ sup_{y\in[0,g(x)]}\left|E_{2n}(x,y) - \int_{g^{-1}([y,g(x)])}\int_{[0,y]} f(X,Y)dYdX\right| \\ &+ sup_{y\in[0,g(x)]}|E_{3n}(x,y)| \end{aligned}$$

where

$$\begin{aligned} E_{1n}(y) &= \int_{g^{-1}([0,y])}\int_{[0,g(X)]} \kappa\left(\frac{y-Y}{h_n}\right) f(X,Y)dYdX, \\ E_{2n}(x,y) &= \int_{g^{-1}([y,g(x)])}\int_{[0,y]} \kappa\left(\frac{y-Y}{h_n}\right) f(X,Y)dYdX, \end{aligned}$$

and

$$E_{3n}(x,y) = \int_{g^{-1}([y,g(x)])}\int_{[y,g(X)]} \kappa\left(\frac{y-Y}{h_n}\right) f(X,Y)dYdX.$$

Hence, to complete the proof we show that each supremum on the r.h.s. is $o(1)$.

For the first term on the r.h.s. we have: a) $X \in g^{-1}([0,y])$ which implies that $g(X) \leq y$ and as $n \to \infty$ $\frac{y-Y}{h_n} > B_K$ and $\kappa\left(\frac{y-Y}{h_n}\right) \to 1$, hence by LDC theorem $E_{1n}(y) \to \int_{g^{-1}([0,y])}\int_{[0,g(X)]} f(X,Y)dYdX$ for every $y \in [0, g(x)]$; b) since κ is a nondecreasing function and h_n is a decreasing function of n we have that for all $y \in [0, g(x)]$ $E_{1n}(y) \leq E_{1,n+1}(y)$. Now, given that $f(X,Y) < B_f$, κ satisfies a Lipschitz condition, and $\left|\int_{g^{-1}([y,y'])}\right| \leq m_{g^{-1}}|y - y'|$ for all $y, y' \in [0, g(x)]$, we have that $E_{1n}(y)$ is continuous. In addition, since $\int_{g^{-1}([0,y])}\int_{[0,g(X)]} f(X,Y)dYdX$ is continuous in y, by theorem 7.13 in Rudin (1976) we have that $sup_{y\in[0,g(x)]}\left|E_{1n}(y) - \int_{g^{-1}([0,y])}\int_{[0,g(X)]} f(X,Y)dYdX\right| = o(1)$.

For the second term on the r.h.s. we have that $X \in g^{-1}([y, g(x)])$ which implies that $y \leq g(X) \leq g(x)$. Given our assumptions, we can verify that $E_{2n}(x,y) \to \int_{g^{-1}([y,g(x)])}\int_{[0,y)} f(X,Y)dYdX$, $E_{21n}(x,y) \leq E_{21(n+1)}(x,y)$. In addition, given that $f(X,Y) < B_f$, κ satisfies a Lipschitz condition, and $\left|\int_{g^{-1}([y,y'])}\right| \leq$

$m_{g^{-1}}|y-y'|$ for all $y, y' \in [0, g(x)]$, we have that $E_{2n}(x,y)$ is continuous with respect to the argument y. Hence, by theorem 7.13 in Rudin (1976) we have that

$$sup_{y\in[0,g(x)]}\left|E_{2n}(x,y) - \int_{g^{-1}([y,g(x)])}\int_{[0,y]} f(X,Y)dYdX\right| = o(1).$$

Following a similar argument we have $sup_{y\in[0,g(x)]}|E_{3n}(x,y)| = o(1)$. Consequently,

$$sup_{y\in[0,g(x)]}|E(\hat{F}(x,y)) - F(x,y))| = o(1).$$

Lemma 3 . *Proof* Let $s_{0,x}(y) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right) I(\{X_i : X_i \le x\})$, and from Lemma 1 in Martins-Filho and Yao (2006), if $nh_n^2 \to \infty$ we have that $sup_{y\in G}|s_{0,x}(y) - E(s_{0,x}(y))| = O_p\left(\left(\frac{ln(n)}{nh_n}\right)^{1/2}\right)$. Now, $E(s_{0,x}(y)) = \int_{C_x}\int_{[0,g(X)]}\frac{1}{h_n}K\left(\frac{Y-y}{h_n}\right) f(X,Y)dYdX = \int_{C_x}\int_{[-y/h_n,(g(X)-y)/h_n]} K(\gamma) f(X, y + h_n\gamma)d\gamma dX$ and by A3c

$$\begin{aligned}
\left|E(s_{0,x}(y)) - \int_{C_x}\int_{[-y/h_n,(g(X)-y)/h_n]} K(\gamma) f(X,y)d\gamma dX\right| &\le m_f h_n \int_{C_x}\int_{[-y/h_n,(g(X)-y)/h_n]} |\gamma| K(\gamma)\, d\gamma dX \\
&\le m_f h_n \int_{C_x}\int_{[-B_K,B_K]} |\gamma| K(\gamma)\, d\gamma dX = O(h_n).
\end{aligned}$$

Given that $y \in G \subset (0, g(x))$, there exists N_+ such that for all $n > N_+$ we have

$$\begin{aligned}
\int_{C_x}\int_{[-y/h_n,(g(X)-y)/h_n]} K(\gamma) f(X,y)d\gamma dX &= \int_{g^{-1}([0,g(x)])} \kappa\left(\frac{g(X)-y}{h_n}\right) f(X,y)dX \\
&= H_{1n}(x,y) + H_{2n}(x,y)
\end{aligned}$$

where $H_{1n}(x,y) = \int_{g^{-1}([0,y])}\kappa\left(\frac{g(X)-y}{h_n}\right) f(X,y)dX$ and $H_{2n}(x,y) = \int_{g^{-1}([y,g(x)])}\kappa\left(\frac{g(X)-y}{h_n}\right) f(X,y)dX$. For $H_{1n}(x,y)$ we have that $X \in g^{-1}([0,y])$ which implies that $g(X) \le y$ and consequently $\kappa\left(\frac{g(X)-y}{h_n}\right) \to 0$. Since $\kappa(\cdot) \le 1$ and $\int_{g^{-1}([0,y])} f(X,y)dX < \infty$, by LDC theorem $H_{1n}(x,y) \to 0$. In addition, since $\kappa(\cdot)$ is a nondecreasing function we have that $H_{1n}(x,y) \ge H_{1(n+1)}(x,y)$. Now, given A2e, A3c, A4 and the fact that $f(x,y) < B_f$, we have that for fixed x and n , $H_{1n}(x,y)$ is continuous in the argument y. Hence, by theorem 7.13 in Rudin (1976) $sup_{y\in G}|H_{1n}(x,y)| = o(1)$. A similar argument gives $sup_{y\in G}\left|H_{2n}(x,y) - \int_{g^{-1}([y,g(x)])} f(X,y)dX\right| = o(1)$ and consequently,

$$sup_{y\in G}\left|\int_{C_x}\kappa\left(\frac{g(X)-y}{h_n}\right) f(X,y)dX - \int_{g^{-1}([y,g(x))} f(X,y)dX\right| = o(1).$$

Thus, we obtain $sup_{y\in G}\left|\frac{1}{nh_n}\sum_{i=1}^{n}K\left(\frac{Y_i-y}{h_n}\right)I(\{X_i:X_i\leq x\})-\int_{g^{-1}([y,g(x)])}f(X,y)dX\right|=o_p(1)$.

Theorem 1. *Proof* From Nadaraya (1964), given that $q_\alpha(x)$ is unique, for all $\epsilon>0$ there exists $0<\delta(\epsilon,x)$ where $\delta(\epsilon,x)=\ \min\{F(q_\alpha(x)+\epsilon/x)-F(q_\alpha(x)/x),F(q_\alpha(x)/x)-F(q_\alpha(x)-\epsilon/x)\}$ and

$$P\left(|q_{\alpha,n}(x)-q_\alpha(x)|>\epsilon\right)\leq P\left(|F(q_{\alpha,n}(x)/x)-F(q_\alpha(x)/x)|>\delta(\epsilon,x)\right).$$

Since, $|F(q_{\alpha,n}(x)/x)-F(q_\alpha(x)/x)|\leq\ \sup_{y\in\Re_+}|\hat{F}(y/x)-F(y/x)|$ and we can write $sup_{y\in\Re_+}|\hat{F}(y/x)-F(y/x)|\leq\frac{1}{\hat{F}(x)}sup_{y\in\Re_+}|\hat{F}(x,y)-F(x,y)|+\left|\frac{1}{F_X(x)}-\frac{1}{\hat{F}(x)}\right|F_X(x)$. Now, $sup_{y\in\Re_+}|\hat{F}(x,y)-F(x,y)|\leq sup_{y\in[0,g(x)]}|\hat{F}(x,y)-F(x,y)|+sup_{(g(x),\infty)}|\hat{F}(x,y)-F(x,y)|$, and since from Lemma 2 we have that $sup_{y\in[0,g(x)]}|\hat{F}(x,y)-E(\hat{F}(x,y))|=o_p(1)$ and $sup_{y\in[0,g(x)]}|E(\hat{F}(x,y))-F(x,y)|=o(1)$, it follows that $sup_{y\in[0,g(x)]}|\hat{F}(x,y)-F(x,y))|=o_p(1)$. Now, for all $y\in(g(x),\infty)$ we have that $F(x,y)=F(x,g(x))=\int_{C_x}\int_{[0,g(X)]}f(X,Y)dYdX=F_X(x)$. In addition, under the assumption that $min_{\{i:X_i\leq x\}}Y_i\geq h_nB_K$ and given that $0<Y\leq g(x)$, we have that, for all $y\in(g(x),\infty)$, $y-Y>0$. Hence, there exists N_+ such that for all $n>N$ and we have that $\hat{F}(x,y)=n^{-1}\sum_{i=1}^{n}\int_{-B_K}^{B_K}K(\gamma)d\gamma I(\{X_i:X_i\leq x\})=n^{-1}\sum_{i=1}^{n}I(\{X_i:X_i\leq x\})=\hat{F}(x)$. Therefore, $sup_{(g(x),\infty)}|\hat{F}(x,y)-F(x,y)|=sup_{(g(x),\infty)}|\hat{F}(x)-F_X(x)|=o_p(1)$, where the last equality follows from the fact that $E(\hat{F}(x))=F_X(x)$ and $V(\hat{F}(x))=\frac{1}{n}F_X(x)(1-F_X(x))\to 0$, hence by Chebyshev's inequality $\hat{F}(x)-F_X(x)=o_p(1)$. To complete the proof, note that $\hat{F}(x)-F_X(x)=o_p(1)$ implies that $\hat{F}(x)=O_p(1)$, and provided that $F_X(x)>0$ we have by Slutsky theorem $\hat{F}(x)^{-1}-F_X(x)^{-1}=o_p(1)$.

Theorem 2 . *Proof* We write $q_{\alpha,n}(x)-q_\alpha(x)=(A_n+C_n)\left(\frac{1}{f(q_\alpha(x)/x)}+\beta_n\right)$, where $A_n=F(q_\alpha(x)/x)-\frac{E(\hat{F}(x,q_\alpha(x)))}{E(\hat{F}(x))}$, $\beta_n=\hat{f}^{-1}(\bar{q}_{\alpha,n}(x)/x)-f^{-1}(q_\alpha(x)/x)$ and $C_n=\frac{E(\hat{F}(x,q_\alpha(x)))}{E(\hat{F}(x))}-\hat{F}(q_\alpha(x)/x)$. Note that the theorem follows if we prove: a) $\beta_n=o_p(1)$; b) $A_n=-\frac{1}{2}h_n^2\sigma_K^2\frac{\int_{g^{-1}([q_\alpha(x),g(x)])}f^{(1)}(\gamma,q_\alpha(x))d\gamma}{F_X(x)}+o(h_n^2)$; c) $\left(\frac{s_n(x)}{\hat{F}(x)}\right)^{-1}\sqrt{n}C_n\xrightarrow{d}N(0,1)$ where

$$s_n^2(x)=F(x,q_\alpha(x))-\frac{F(x,q_\alpha(x))^2}{F_X(x)}-2h_n\sigma_\kappa\int_{g^{-1}([q_\alpha(x),q(x)])}f(X,q_\alpha(x))dX+o(h_n).$$

a) It suffices to show that $\hat{f}(\bar{q}_{\alpha,n}(x)/x)-f(q_\alpha(x)/x)=o_p(1)$ for all $\alpha\in(0,1)$. Since we have already shown that $q_{\alpha,n}(x)-q_\alpha(x)=o_p(1)$, by theorem 21.6 in Davidson (1994) it suffices to show that $sup_{y\in G}|\hat{f}(y/x)-$

$f(y/x)| = o_p(1)$, where $G \subset (0, g(x))$, G compact. Note that

$$\begin{aligned} sup_{y\in G}|\hat{f}(y/x) - f(y/x)| &\leq \frac{1}{\hat{F}(x)} sup_{y\in G} \left| \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right) I(\{X_i : X_i \leq x\}) - \int_{g^{-1}([y,g(x)])} f(X, y)dX \right| \\ &+ \left| \frac{1}{F_X(x)} - \frac{1}{\hat{F}(x)} \right| sup_{y\in G} \int_{g^{-1}([y,g(x)])} f(X, y)dX. \end{aligned}$$

Since $f(x, y) < B_f$, we have that $sup_{y\in G} \int_{g^{-1}([y,g(x)])} f(X, y)dX \leq B_f \int_{g^{-1}([y,g(x)])} dX = O(1)$ for all finite x. In addition, given that $\hat{F}(x)^{-1} - F_X(x)^{-1} = o_p(1)$ the second term on the r.h.s. of the inequality is $o_p(1)$. By Lemma 3 $sup_{y\in G} \left| \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{Y_i - y}{h_n}\right) I(\{X_i : X_i \leq x\}) - \int_{g^{-1}([y,g(x)])} f(X, y)dX \right| = o_p(1)$, hence $\beta_n = o_p(1)$.

b) $A_n = (E(\hat{F}(x)))^{-1}(A_{1n}(x) + A_{2n}(x))$ where $A_{1n}(x) = F(q_\alpha(x)/x)E(\hat{F}(x)) - F(x, q_\alpha(x))$ and $A_{2n}(x) = F(x, q_\alpha(x)) - E(\hat{F}(x, q_\alpha(x)))$. Since $E(\hat{F}(x)) = F_X(x)$, $A_{1n}(x) = 0$. In addition, given that $0 < \alpha < 1$, we have $0 < q_\alpha(x) < g(x)$ and therefore from Lemma 1 $A_{2n}(x) = -\frac{1}{2}h_n^2\sigma_K^2 \int_{g^{-1}([q_\alpha(x),g(x)])} f^{(1)}(X, q_\alpha(x))dX + o(h_n^2)$. Thus, $A_n = -\frac{1}{F_X(x)}\frac{h_n^2}{2}\sigma_K^2 \int_{g^{-1}([q_\alpha(x),g(x)])} f^{(1)}(X, q_\alpha(x))dX + o(h_n^2)$.

c) We start by noting that,

$$\begin{aligned} C_n &= \frac{E(\hat{F}(x, q_\alpha(x)))}{E(\hat{F}(x))} - \frac{\hat{F}(x, q_\alpha(x))}{\hat{F}(x)} \\ &= -\frac{1}{\hat{F}(x)}\left(\frac{1}{nh_n}\sum_{i=1}^{n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i - \gamma}{h_n}\right) I(\{X_i : X_i \leq x\})d\gamma - \hat{F}(x)\frac{E(\hat{F}(x, q_\alpha(x)))}{E(\hat{F}(x))}\right) \\ &= -\frac{1}{\hat{F}(x)}\frac{1}{n}\sum_{i=1}^{n} c_{in} \end{aligned}$$

where $c_{in} = \frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \leq x\})d\gamma - I(\{X_i : X_i \leq x\})\frac{E(\hat{F}(x,q_\alpha(x)))}{F_X(x)}$. Hence, we write $\sqrt{n}C_n = -\frac{1}{\hat{F}(x)}\sum_{i=1}^{n} Z_{in}$ where $Z_{in} = \frac{1}{\sqrt{n}}c_{in}$, and note that $E(Z_{in}) = 0$, $s_n^2 = \sum_{i=1}^{n} E(Z_{in}^2) = E(c_{in}^2)$ by assumption A1. In addition, we note that $E(c_{in}^2) = s_{1n} + s_{2n} + s_{3n}$, where

$s_{1n} = E\left(\frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \leq x\})d\gamma\right)^2$,

$s_{2n} = E(I(\{X_i : X_i \leq x\}))\frac{(E(\hat{F}(x,q_\alpha(x)))^2}{F_X(x)^2}$ and

$s_{3n} = -2E\left(\frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \leq x\})d\gamma\right)\frac{E(\hat{F}(x,q_\alpha(x))}{F_X(x)}$. From Lemma 1- b) we have hat $s_{1n} = F(x, q_\alpha(x)) - 2h_n\sigma_\kappa \int_{g^{-1}([q_\alpha(x),q(x)])} f(X, q_\alpha(x))dX + o(h_n)$, and from Lemma 1 - a)

$s_{2n} = (F_X(x))^{-1}E(\hat{F}(x, q_\alpha(x)))^2 = F_X(x)^{-1}\left(F(x, q_\alpha(x)) + \sigma_K^2\frac{h_n^2}{2}\int_{g^{-1}([q_\alpha(x),q(x)])} f^{(1)}(X, q_\alpha(x))dX + o(h_n^2)\right)^2$

and therefore we have

$s_{2n} = F_X(x)^{-1}F(x,q_\alpha(x))^2 + o(h_n)$. Note that $s_{3n} = -2s_{2n} = -2\frac{(F(x,q_\alpha(x)))^2}{F_X(x)} + o(h_n)$. Combining these results we have that $s_n^2(x) = F(x,q_\alpha(x)) - \frac{F(x,q_\alpha(x))^2}{F_X(x)} - 2h_n\sigma_\kappa \int_{g^{-1}([q_\alpha(x),g(x)])} f(X,q_\alpha(x))dX + o(h_n)$. By Liapounov's central limit theorem, $\sum_{i=1}^n \frac{Z_{in}}{s_n(x)} \xrightarrow{d} N(0,1)$ provided that $lim_{n\to\infty}\sum_{i=1}^n E\left(\left|\frac{Z_{in}}{s_n(x)}\right|^{2+\delta}\right) = 0$ for some $\delta > 0$. First, note that

$$\begin{aligned}
\sum_{i=1}^n E\left(\left|\frac{Z_{in}}{s_n(x)}\right|^{2+\delta}\right) &= (s_n^2(x))^{-1-\delta/2}\sum_{i=1}^n E(|Z_{in}|^{2+\delta}) \\
&= (s_n(x)^2)^{-1-\delta/2}n^{-\delta/2}E\left(\left|\frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \le x\})d\gamma\right.\right. \\
&- \left.\left. I(\{X_i : X_i \le x\})\frac{E(\hat{F}(x,q_\alpha(x)))}{F_X(x)}\right|^{2+\delta}\right) \\
&\le (s_n^2(x))^{-1-\delta/2}n^{-\delta/2}2^{1+\delta}\left(E\left(\left|\frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \le x\})d\gamma\right|^{2+\delta}\right)\right. \\
&+ \left. E\left(\left|I(\{X_i : X_i \le x\})\frac{E(\hat{F}(x,q_\alpha(x)))}{F_X(x)}\right|\right)\right)
\end{aligned}$$

where the last inequality follows from c_r-inequality. Hence, given

$$s_n^2(x) = F(x,q_\alpha(x)) - \frac{F(x,q_\alpha(x))^2}{F_X(x)} - 2h_n\sigma_\kappa \int_{g^{-1}([q_\alpha(x),q(x)])} f(X,q_\alpha(x))dX + o(h_n)$$

to complete the proof it suffices to show that $a_n = n^{-\delta/2}E\left(\left|\frac{1}{h_n}\int_{[0,q_\alpha(x)]} K\left(\frac{Y_i-\gamma}{h_n}\right) I(\{X_i : X_i \le x\})d\gamma\right|^{2+\delta}\right) = o(1)$, and

$b_n = n^{-\delta/2}E\left(\left|I(\{X_i : X_i \le x\})\frac{E(\hat{F}(x,q_\alpha(x)))}{F_X(x))}\right|\right) = o(1)$. First, note that

$$\begin{aligned}
a_n &= n^{-\delta/2}\int_{C_x}\int_{[0,g(X)]}\left(\int_{-Y/h_n}^{(q_\alpha(x)-Y)/h_n} K(\gamma)\,d\gamma\right)^{2+\delta} f(X,Y)dY\,dX \\
&\le n^{-\delta/2}\int_{C_x}\int_{[0,g(X)]} f(X,Y)dY\,dX \to 0 \text{ as } n\to\infty
\end{aligned}$$

Second, note that $b_n = n^{-\delta/2}E(I(\{X_i : X_i \le x\}))\frac{(E(\hat{F}(x,q_\alpha(x)))^{2+\delta})}{F_X(x)^{2+\delta}} \to 0$ since $E(I(\{X_i : X_i \le x\})) = F_X(x) > 0$ and by Lemma 1 $E(\hat{F}(x,q_\alpha(x))) \to F(x,q_\alpha(x))$. Hence, $\left(\frac{s_n(x)}{\hat{F}_X(x)}\right)^{-1}\sqrt{n}C_n \xrightarrow{d} N(0,1)$ since $\hat{F}(x) \xrightarrow{p} F_X(x)$.

Theorem 3 . *Proof* From Theorem 2, we have that for given x^l and $\alpha \in (0,1)$,

$$\sqrt{n}(\hat{q}_{\alpha,n}(x^l) - q_\alpha(x^l) - \frac{A_n(x^l)}{f(q_\alpha(x^l)/x^l)}) = \frac{\sqrt{n}}{f(q_\alpha(x^l)/x^l)}C_n(x^l) + o_p(1)$$

where $A_n(x^l) = -\frac{1}{F_X(x^l)}\frac{h_n^2}{2}\int_{g^{-1}([q_\alpha(x^l),g(x^l)])} f^{(1)}(X, q_\alpha(x^l))dX\sigma_K^2 + o(h_n^2)$ and

$$\begin{aligned} C_n(x^l) &= -\left(\frac{1}{\hat{F}(x^l)} - \frac{1}{F_X(x^l)} + \frac{1}{F_X(x^l)}\right)\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) \right.\\ &- \left. I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right). \end{aligned}$$

Since, $\frac{1}{\hat{F}(x^l)} - \frac{1}{F_X(x^l)} = o_p(1)$ and

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) - I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right) = O_p(1),$$

we have

$$\begin{aligned} C_n(x^l) &= -\frac{1}{F_X(x^l)}\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) - I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right)\\ &+ o_p(n^{-\frac{1}{2}}). \end{aligned}$$

For $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_r\}'$ let $W_n = \Lambda'\left[\sqrt{n}\frac{C_n(x^1)}{f(q_\alpha(x^1)/x^1)}, \sqrt{n}\frac{C_n(x^2)}{f(q_\alpha(x^2)/x^2)}, \cdots, \sqrt{n}\frac{C_n(x^r)}{f(q_\alpha(x^r)/x^r)}\right]'$. By the Cramer-Wold theorem, if $W_n \xrightarrow{d} N(0, \Lambda' Q\Lambda)$, then $\sqrt{n}\left(\frac{C_n(x^1)}{f(q_\alpha(x^1)/x^1)}, \frac{C_n(x^2)}{f(q_\alpha(x^2)/x^2)}, \cdots, \frac{C_n(x^r)}{f(q_\alpha(x^r)/x^r)}\right)' \xrightarrow{d} N(0, Q)$, which proves the theorem given the asymptotic representation of $A_n(x^l)$. Now,

$$\begin{aligned} W_n &= \sqrt{n}\sum_{l=1}^r \frac{\lambda_l}{f(q_\alpha(x^l)/x^l)}C_n(x^l)\\ &= \sum_{i=1}^n -\frac{1}{\sqrt{n}}\sum_{l=1}^r\frac{\lambda_l}{f(q_\alpha(x^l)/x^l)F_X(x^l)}\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) \right.\\ &- \left. I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right) + o_p(1) = \sum_{i=1}^n E(Z_{in}) + o_p(1) \end{aligned}$$

Note that $E(Z_{in}) = 0$ and let $S_n^2 = \sum_{i=1}^n E(Z_{in}^2) = S_{1n} + S_{2n}$. From theorem 2,

$$\begin{aligned} S_{1n} &= \sum_{l=1}^r \frac{\lambda_l^2}{f^2(q_\alpha(x^l)/x^l)F_X^2(x^l)}E\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) \right.\\ &- \left. I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right)^2\\ &= \sum_{l=1}^r \frac{\lambda_l^2}{f^2(q_\alpha(x^l)/x^l)F_X^2(x^l)}\left[\alpha(1-\alpha)F_X(x^l) - 2h_n\sigma_\kappa\int_{g^{-1}([q_\alpha(x^l),g(x^l)])} f(X, q_\alpha(x^l)) + o(h_n)\right) \end{aligned}$$

and $S_{2n} = \sum_{l=1}^r\sum_{m=1,l\neq m}^r \frac{\lambda_l\lambda_m}{f(q_\alpha(x^l)/x^l)F_X(x^l)f(q_\alpha(x^m)/x^m)F_X(x^m)}S(l,m)$ where

$$S(l,m) = E\left(\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right)dzI(\{X_i : X_i \le x^l\}) - I(\{X_i : X_i \le x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right)\right.$$

$$
\begin{aligned}
&\times \left(\frac{1}{n}\int_0^{q_\alpha(x^m)} K\left(\frac{Y_i - z}{h_n}\right) dz I(\{X_i : X_i \le x^m\}) - I(\{X_i : X_i \le x^m\})\frac{E(\hat{F}(x^m, q_\alpha(x^m)))}{F_X(x^m)}\right)\Bigg) \\
&= E\left(\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right) dz I(\{X_i : X_i \le x^l\})\right)\right. \\
&\times \left.\left(\frac{1}{n}\int_0^{q_\alpha(x^m)} K\left(\frac{Y_i - z}{h_n}\right) dz I(\{X_i : X_i \le x^m\})\right)\right) \\
&- \frac{1}{F_X(x^m)} E(\hat{F}(x^m, q_\alpha(x^m))) E\left(\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right) dz I(\{X_i : X_i \le x^l, X_i \le x_m\})\right) \\
&- \frac{1}{F_X(x^l)} E(\hat{F}(x^l, q_\alpha(x^l))) E\left(\frac{1}{n}\int_0^{q_\alpha(x^m)} K\left(\frac{Y_i - z}{h_n}\right) dz I(\{X_i : X_i \le x^l, X_i \le x_m\})\right) \\
&+ \frac{1}{F_X(x^l)F_X(x^m)} E(\hat{F}(x^l, q_\alpha(x^l))) E(\hat{F}(x^m, q_\alpha(x^m))) E(I(\{X_i : X_i \le x^l, X_i \le x_m\})) \\
&= S_{21} + S_{22} + S_{23} + S_{24}
\end{aligned}
$$

We analyze each term separately. First, consider

$$
\begin{aligned}
S_{21} &= E\left(\kappa\left(\frac{q_\alpha(x^l) - Y_i}{h}\right)\kappa\left(\frac{q_\alpha(x^m) - Y_i}{h}\right) I(\{X_i : X_i \le x^{lm}\})\right) \\
&= \int_{g^{-1}([0, g(x^{lm})])} \int_0^{g(X)} \kappa\left(\frac{q_\alpha(x^l) - Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) f(X, Y) dY dX
\end{aligned}
$$

By definition of x^{lm}, for $\alpha \in (0,1)$, $q_\alpha(x^l)$ or $q_\alpha(x^m)$ could be smaller than $g(x^{lm})$, but it cannot be that $q_\alpha(x^l) \ge g(x^{lm})$ and $q_\alpha(x^m) \ge g(x^{lm})$ simultaneously. Hence, we consider the following four possible situations: (1)$q_\alpha(x^l) \le q_\alpha(x^m) \le g(x^{lm})$; (2) $q_\alpha(x^l) \le g(x^{lm}) \le q_\alpha(x^m)$; (3) $q_\alpha(x^m) \le q_\alpha(x^l) \le g(x^{lm})$; (4) $q_\alpha(x^m) \le g(x^{lm}) \le q_\alpha(x^l)$. Under (1), we have

$$
\begin{aligned}
S_{21} &= \int_{g^{-1}([0, q_\alpha(x^l)] \cup [q_\alpha(x^l), q_\alpha(x^m)] \cup [q_\alpha(x^m), g(x^{lm})])} \int_0^{g(X)} \kappa\left(\frac{q_\alpha(x^l) - Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) f(X, Y) dY dX \\
&= S_{211} + S_{212} + S_{213}
\end{aligned}
$$

where $S_{211} = \int_{g^{-1}([0, q_\alpha(x^l)])} \int_0^{g(X)} \kappa\left(\frac{q_\alpha(x^l) - Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) f(X, Y) dY dX$, as $X \in g^{-1}([0, q_\alpha(x^l)])$, where $0 \le Y_i \le g(X_i) \le q_\alpha(x^l) \le q_\alpha(x^m)$, and $\kappa\left(\frac{q_\alpha(x^l) - Y}{h_n}\right) = 1$ and $\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) = 1$, and therefore $S_{211} = \int_{g^{-1}([0, q_\alpha(x^l)])} \int_0^{g(X)} f(X, Y) dY dX$. Now,

$$
\begin{aligned}
S_{212} &= \int_{g^{-1}([q_\alpha(x^l), q_\alpha(x^m)])} \int_0^{q_\alpha(x^l)} \kappa\left(\frac{q_\alpha(x^l) - Y_i}{h}\right)\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) f(X, Y) dY dX \\
&+ \int_{g^{-1}([q_\alpha(x^l), q_\alpha(x^m)])} \int_{q_\alpha(x^l)}^{g(X)} \kappa\left(\frac{q_\alpha(x^l) - Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m) - Y}{h_n}\right) f(X, Y) dY dX = S_{212a} + S_{212b}
\end{aligned}
$$

For S_{212a}, $Y \leq q_\alpha(x^l) \leq q_\alpha(x^m)$, hence there exists $N \in \Re_+$ such that for all $n > N$ $\kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right) = 1$ and $\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) = 1$ and therefore $S_{212a} \to \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^m)])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX$. For S_{212b}, $q_\alpha(x^l) \leq Y \leq g(X) \leq q_\alpha(x^m)$, hence there exists $N \in \Re_+$ such that for $n > N$ we have $\kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right) = 0$ and $\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) = 1$, and therefore $S_{212b} \to 0$ and $S_{212} \to \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^m)])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX$. Now,

$$
\begin{aligned}
S_{213} &= \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} \kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) f(X,Y)dYdX \\
&+ \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_{q_\alpha(x^l)}^{q_\alpha(x^m)} \kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) f(X,Y)dYdX \\
&+ \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_{q_\alpha(x^m)}^{g(X)} \kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) f(X,Y)dYdX \\
&= S_{213a} + S_{213b} + S_{213c}
\end{aligned}
$$

Similar arguments show that $S_{213a} \to \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX$, $S_{213b} \to 0$ and $S_{213c} \to 0$, hence $S_{213} \to \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX$. In all, $S_{21} \to \int_{g^{-1}([0,q_\alpha(x^l)])} \int_0^{g(X)} f(X,Y)dYdX + \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX$. Under (2),

$$
\begin{aligned}
S_{21} &= \int_{g^{-1}([0,q_\alpha(x^l)])} \int_0^{g(X)} \kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) f(X,Y)dYdX \\
&+ \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^{lm})])} \int_0^{g(X)} \kappa\left(\frac{q_\alpha(x^l)-Y}{h_n}\right)\kappa\left(\frac{q_\alpha(x^m)-Y}{h_n}\right) f(X,Y)dYdX \\
&= S_{211} + S_{212}
\end{aligned}
$$

Applying arguments such as those used in (1), we obtain $S_{211} \to \int_{g^{-1}([0,q_\alpha(x^l)])} \int_0^{g(X)} f(X,Y)dYdX$ and $S_{212} \to \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^{lm})])} \int_0^{g(x^l)} f(X,Y)dYdX$. Hence, under (1) and (2) we have

$$
S_{21} \to \int_{g^{-1}([0,q_\alpha(x^l)])} \int_0^{g(X)} f(X,Y)dYdX + \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX.
$$

Similarly, under (3) and (4) we have

$$
S_{21} \to \int_{g^{-1}([0,q_\alpha(x^m)])} \int_0^{g(X)} f(X,Y)dYdX + \int_{g^{-1}([q_\alpha(x^m),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^m)} f(X,Y)dYdX.
$$

From Lemma 1,

$$
\begin{aligned}
S_{22} &= -\frac{1}{F_X(x^m)} E(\hat{F}(x^m, q_\alpha(x^m))) E(\hat{F}(x^{lm}, q_\alpha(x^l))) \\
&= -\frac{1}{F_X(x^m)} F(x^m, q_\alpha(x^m)) F(x^{lm}, q_\alpha(x^l)) + o(h_n) = -\alpha F(x^{lm}, q_\alpha(x^l)) + o(h_n)
\end{aligned}
$$

and $S_{23} = -\frac{1}{F_X(x^l)} E(\hat{F}(x^l, q_\alpha(x^l)))E(\hat{F}(x^{lm}, q_\alpha(x^m))) = -\alpha F(x^{lm}, q_\alpha(x^m)) + o(h_n)$, and since $E(I(\{X_i : X_i \leq x^{lm}\})) = F_X(x^{lm})$, we have $S_{24} = \alpha^2 F_X(x^{lm}) + O(h_n^2)$. Now, for $q_\alpha(x^l) \leq q_\alpha(x^m)$,

$$\int_{g^{-1}([0,q_\alpha(x^l)])} \int_0^{g(X)} f(X,Y)dYdX + \int_{g^{-1}([q_\alpha(x^l),q_\alpha(x^{lm})])} \int_0^{q_\alpha(x^l)} f(X,Y)dYdX = F(x^{lm}, q_\alpha(x^l)),$$

and collecting all items analyzed above, we obtain $S_n^2 \to \Lambda' Q \Lambda$.

By Liapounov's CLT, $\frac{W_n}{S_n} = \sum_{i=1}^n \frac{Z_{in}}{S_n} \xrightarrow{d} N(0,1)$ or $W_n \xrightarrow{d} N(0, \Lambda' Q\Lambda)$, provided that

$$\lim_{n\to\infty} \sum_{i=1}^n E\left|\frac{Z_{in}}{S_n}\right|^{2+\delta} = 0$$

for some $\delta > 0$. By the c_r inequality,

$$\begin{aligned}\sum_{i=1}^n E\left|\frac{Z_{in}}{S_n}\right|^{2+\delta} \quad &\leq \quad (S_n^2)^{-1-\frac{\delta}{2}} n^{-\frac{\delta}{2}} r^{1+\delta} \sum_{l=1}^r \frac{|\lambda_l|^{2+\delta}}{|f(q_\alpha(x^l)/x^l)F_X(x^l)|^{2+\delta}} \\ &\times \quad E\left|\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y-z}{h_n}\right) dzI(\{X_i : X_i \leq x^l\}) - I(\{X_i : X_i \leq x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right|^{2+\delta}\end{aligned}$$

Since r is finite, $S_n^2 \to \Lambda' Q\Lambda$, and from Theorem 2

$$n^{-\frac{\delta}{2}} E\left|\frac{1}{n}\int_0^{q_\alpha(x^l)} K\left(\frac{Y_i - z}{h_n}\right) dzI(\{X_i : X_i \leq x^l\}) - I(\{X_i : X_i \leq x^l\})\frac{E(\hat{F}(x^l, q_\alpha(x^l)))}{F_X(x^l)}\right|^{2+\delta} \to 0,$$

which completes the proof.

Theorem 4 . *Proof* a) Recall that by definition $q_{1,n}(x) = inf\{y \in \Re_+ : \hat{F}(y/x) = 1\}$, i.e., $q_{1,n}(x)$ is the greatest lower bound for the set under the constraint that $(nh_n)^{-1} \sum_{i=1}^n \int_0^y K\left(\frac{Y_i - \gamma}{h_n}\right) d\gamma I(\{X_i : X_i \leq x\}) = n^{-1}\sum_{i=1}^n I(\{X_i : X_i \leq x\})$. Under the assumption that $min_{\{i:X_i\leq x\}} Y_i \geq h_n B_K$, there exists $N(x) \in R_+$ such that for all $n > N(x)$, we have that the equality holds for all $y \geq \max_{\{i:X_i\leq x\}} Y_i + h_n B_K$, and it is false for all $y < \max_{\{i:X_i \leq x\}} Y_i + h_n B_K$. Hence, $q_{1,n}(x) = max_{\{i:X_i\leq x\}} Y_i + h_n B_K$ for all $n > N(x)$. b) Now, from Park et al. (2000) we have that the FDH estimator is defined as $\theta_{FDH}(x) = max_{\{i:X_i \leq x\}} Y_i$ and under the assumptions in the theorem they show that $n^{1/(d+1)}(q_1(x) - \theta_{FDH}(x)) \xrightarrow{d} Weibull(\mu_x^{d+1}, d+1)$. Consequently, provided that $nh_n^{d+1} = O(1)$ we have that $n^{1/(d+1)}(q_1(x) - q_{1,n}(x) + h_n B_K \theta_{FDH}(x)) \xrightarrow{d} Weibull(\mu_x^{d+1}, d+1)$.

Appendix 2 - Tables and Figures

Table 1 bias($\times 10^{-1}$)(B) and RMSE(R) for α-frontier estimators with $g(x) = \sqrt{x}$

n	$x_1 = 9$				$x_2 = 16$				$x_3 = 25$			
= 100	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	-.004	.126	-.103	.135	-.007	.130	-.093	.138	-.013	.134	-.197	.142
0.97	-.168	.116	-.070	.122	-.135	.126	-.071	.138	-.089	.120	-.627	.147
0.98	-.269	.121	-.489	.131	-.167	.124	-.642	.150	-.066	.117	-.706	.149
0.99	-.427	.129	-1.044	.160	-.176	.124	-.414	.135	-.031	.117	-.931	.161
1	-.838	.156	-2.322	.262	-.350	.140	-2.398	.272	.371	.160	-2.327	.264

n	$x_1 = 9$				$x_2 = 16$				$x_3 = 25$			
= 200	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	.020	.093	-.042	.098	-.025	.092	-.079	.098	-.021	.100	-.121	.104
0.97	-.139	.089	-.265	.098	-.066	.087	-.171	.097	-.001	.089	-.301	.102
0.98	-.143	.086	-.137	.097	-.098	.087	-.332	.101	.018	.087	-.326	.103
0.99	-.207	.090	-.399	.099	-.114	.087	-.540	.110	.042	.084	-.443	.107
1	-.511	.107	-1.677	.191	-.028	.129	-1.682	.192	.562	.143	-1.594	.181

n	$x_1 = 9$				$x_2 = 16$				$x_3 = 25$			
= 400	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	.016	.064	-.004	.067	.010	.067	-.010	.070	.015	.069	-.047	.071
0.97	-.024	.062	-.043	.069	-.016	.061	-.048	.067	.027	.063	-.119	.069
0.98	-.042	.061	-.034	.068	-.014	.058	-.047	.064	.028	.061	-.173	.070
0.99	-.057	.057	-.042	.066	-.023	.056	-.207	.066	.021	.060	-.248	.071
1	-.228	.066	-1.151	.130	.111	.065	-1.154	.130	.506	.090	-1.144	.130

Table 2 bias($\times 10^{-1}$)(B) and RMSE(R) for α-frontier estimators with $g(x) = x^3$

n	$x_1 = 1.33$				$x_2 = 1.66$				$x_3 = 2$			
= 100	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	-.075	.123	-.157	.129	-.088	.218	-.233	.226	-.168	.334	-.695	.352
0.97	-.319	.133	-.348	.148	-.312	.225	-.439	.254	-.242	.369	-1.697	.426
0.98	-.374	.137	-.300	.140	-.455	.223	-.931	.251	-.301	.369	-2.071	.445
0.99	-.638	.152	-.985	.168	-.447	.213	-.425	.220	-.329	.367	-2.784	.480
1	-1.400	.203	-2.604	.294	-1.594	.332	-4.413	.492	.638	1.069	-6.974	.786

n	$x_1 = 1.33$				$x_2 = 1.66$				$x_3 = 2$			
= 200	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	-.010	.088	-.039	.092	.005	.148	-.059	.153	-.051	.225	-.305	.234
0.97	-.098	.093	-.117	.100	-.092	.162	-.122	.173	-.236	.261	-1.004	.294
0.98	-.150	.093	-.315	.103	-.126	.161	-.095	.170	-.264	.260	-1.189	.297
0.99	-.157	.094	-.116	.098	-.196	.161	-.561	.182	-.311	.264	-1.633	.324
1	-.786	.128	-1.737	.199	-.823	.207	-2.987	.343	.843	.725	-5.020	.573

n	$x_1 = 1.33$				$x_2 = 1.66$				$x_3 = 2$			
= 400	Smooth		Empirical		Smooth		Empirical		Smooth		Empirical	
α	B	R	B	R	B	R	B	R	B	R	B	R
0.75	-.016	.063	-.037	.065	-.035	.106	-.066	.110	.079	.163	-.050	.167
0.97	-.094	.069	-.106	.073	-.015	.116	-.042	.123	-.106	.187	-.507	.202
0.98	-.117	.068	-.101	.072	-.011	.118	-.141	.126	-.073	.184	-.584	.203
0.99	-.142	.068	-.286	.076	-.059	.118	-.058	.127	-.084	.186	-.759	.215
1	-.467	.083	-1.216	.138	-.374	.139	-2.091	.238	.736	.373	-3.429	.390

Table 3 Empirical Coverage Probability for α-frontier estimators
Smooth (S) and Empirical (E)

n	$g(x) = \sqrt{x}$						$g(x) = x^3$					
$= 100$	$x_1 = 9$		$x_2 = 16$		$x_3 = 25$		$x_1 = 1.33$		$x_2 = 1.66$		$x_3 = 2$	
α	S	E	S	E	S	E	S	E	S	E	S	E
0.75	.919	.887	.952	.938	.953	.933	.897	.877	.918	.907	.929	.896
0.97	.895	.895	.940	.915	.957	.865	.867	.811	.906	.870	.917	.837
0.98	.871	.815	.935	.804	.954	.841	.835	.828	.905	.835	.928	.816
0.99	.820	.631	.897	.849	.951	.751	.762	.647	.893	.867	.928	.732
n	$g(x) = \sqrt{x}$						$g(x) = x^3$					
$= 200$	$x_1 = 9$		$x_2 = 16$		$x_3 = 25$		$x_1 = 1.33$		$x_2 = 1.66$		$x_3 = 2$	
α	S	E	S	E	S	E	S	E	S	E	S	E
0.75	.925	.911	.951	.941	.950	.937	.922	.894	.937	.915	.952	.939
0.97	.927	.865	.951	.905	.958	.915	.915	.881	.926	.903	.934	.886
0.98	.917	.871	.943	.880	.961	.895	.900	.846	.924	.908	.935	.872
0.99	.875	.804	.925	.799	.958	.854	.863	.841	.910	.846	.922	.794
n	$g(x) = \sqrt{x}$						$g(x) = x^3$					
$= 400$	$x_1 = 9$		$x_2 = 16$		$x_3 = 25$		$x_1 = 1.33$		$x_2 = 1.66$		$x_3 = 2$	
α	S	E	S	E	S	E	S	E	S	E	S	E
0.75	.942	.930	.950	.945	.953	.949	.925	.914	.942	.931	.954	.947
0.97	.943	.920	.959	.939	.966	.931	.917	.892	.939	.924	.948	.918
0.98	.942	.918	.962	.948	.970	.914	.907	.893	.934	.906	.949	.908
0.99	.937	.892	.967	.901	.967	.898	.912	.852	.930	.906	.945	.871

Figure 1: A typical simulated dataset of sample size 100 with $g(x) = \sqrt{x}$. Thin solid lines are true $\alpha-$frontiers, thick solid lines are smooth α-frontiers and dashed lines the empirical α-frontiers. Upper lines are for $\alpha = 1$ and lower ones for $\alpha = 0.75$.

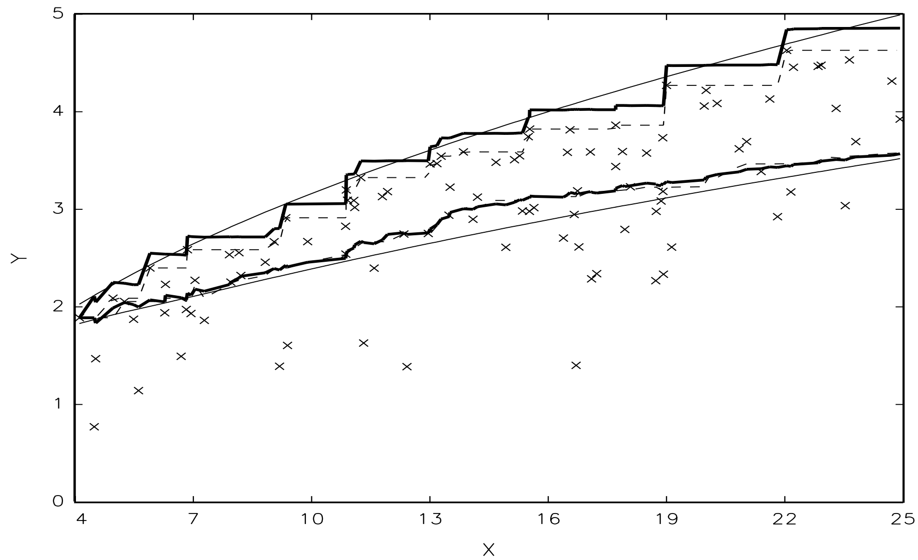


Figure 2: A typical simulated dataset of sample size 100 with $g(x) = x^3$. Thin solid lines are true $\alpha-$frontiers, thick solid lines are smooth α-frontiers and dashed lines the empirical α-frontiers. Upper lines are for $\alpha = 1$ and lower ones for $\alpha = 0.75$.

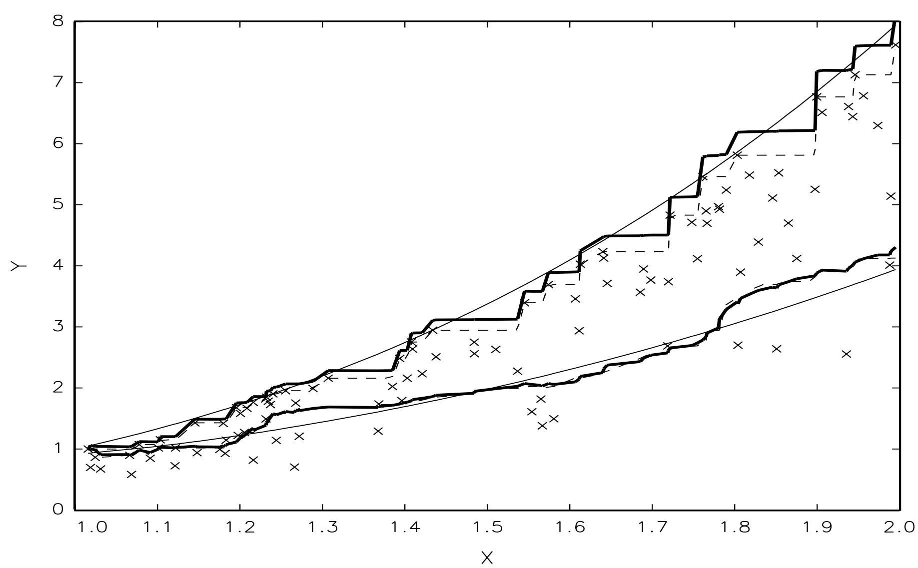


Figure 3: Plot of true $\alpha-$frontiers with estimated smooth and empirical α-frontiers, for $n = 50$, $g_1(x) = g_1(25)$ and α ranging over $0.02, 0.04, \cdots, 1$

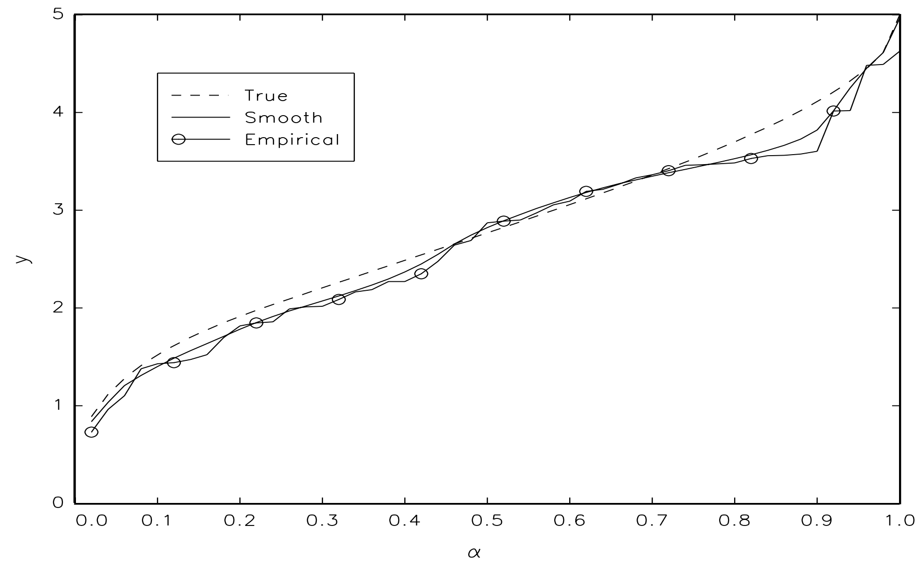


Figure 4: Kernel density estimates for the smooth(S) and empirical(E) α-frontier estimators evaluated at $x_0 = 25$ centered around the true value $q_{0.99}(25)$, the $\alpha = 0.99$ frontier function. The kernel density estimates are based on 1000 simulations from $g_1(x)$ of sample sizes $n = 100$ and 400

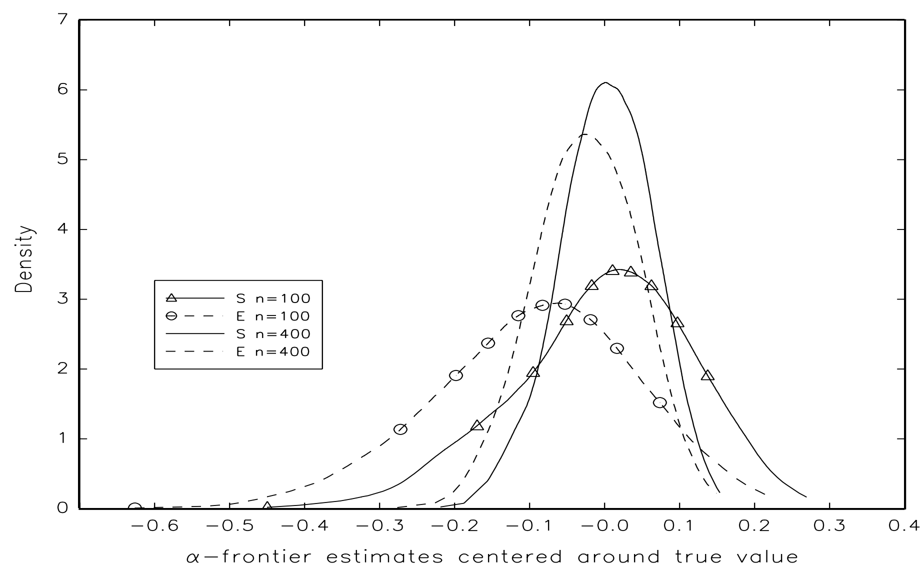


Figure 5: Plot of the American Electric Utility data together with estimated smooth (S) and empirical (E) frontiers for $\alpha = 0.99$ and 0.90

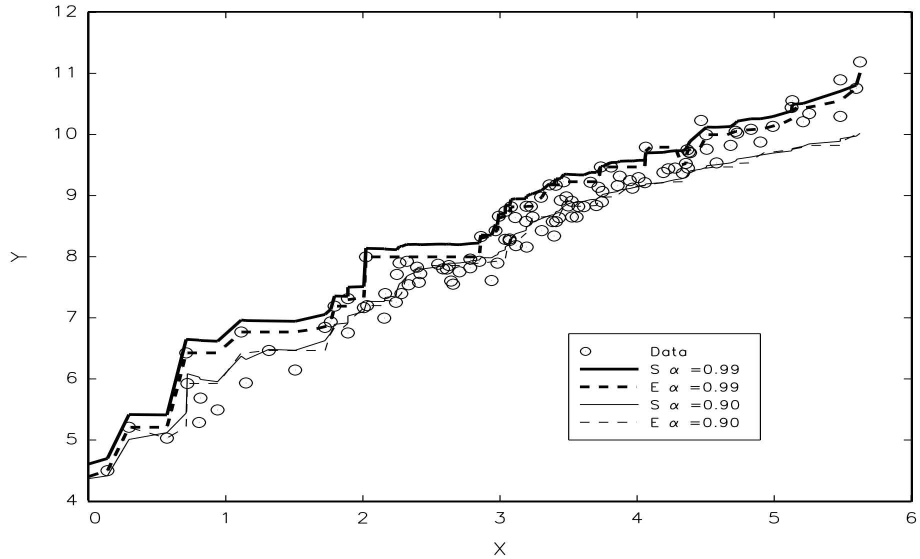
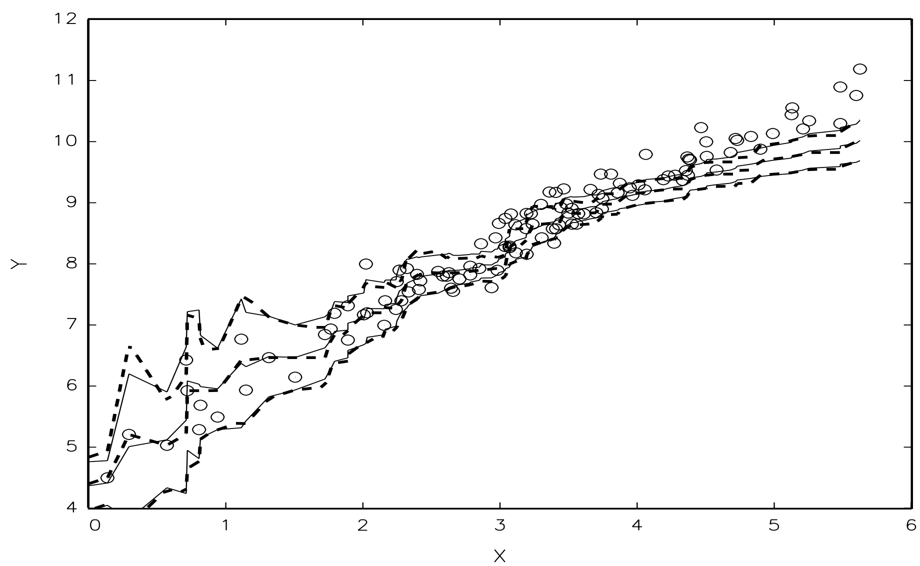


Figure 6: 95% confidence intervals for $\alpha = 0.90$ frontiers with smooth (solid) and empirical (dashed) estimates using American Electric Utility data



References

[1] Aragon, Y., A. Daouia, C. Thomas-Agnan, 2005, Nonparametric frontier estimation: a conditional quantile-based approach. Econometric Theory 21, 358-389.

[2] Azzalini, A., 1981, A note on the estimation of a distribution function and quantiles by a kernel method. Biometrika 68, 326-328.

[3] Banker, R., 1993, Maximum likelihood, consistency and data development analysis: a statistical foundation. Management Science 39, 1265-1273.

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

[5] Charnes, A., W. Cooper and E. Rhodes, 1978, Measuring the efficiency of decision making units. European Journal of Operational Research 2, 429-444.

[6] Christensen, L. R. and W.H. Greene, 1976, Economies of scale in U.S. electric power generation. Journal of Political Economy 84, 655-676.

[7] Davidson, J., 1994, Stochastic Limit Theory. Oxford University Press, New York.

[8] Deprins, D., L. Simar and H. Tulkens, 1984, Measuring labor inefficiency in post offices, in: M. Marchand, P. Pestiau and H. Tulkens, (Eds.), The performance of public enterprises: concepts and measurements. North Holland, Amsterdam.

[9] Falk, M., 1985, Asymptotic normality of the kernel quantile estimator. The Annals of Statistics 13, 428-433.

[10] Farrell, M., 1957, The measurement of productive efficiency. Journal of the Royal Statistical Society A 120, 253-290.

[11] Geffroy, J., 1964, Sur un problème d'estimation géométrique. Publications de l'Institut de Statistique de l'Université de Paris XIII, 191-200.

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

[13] Girard, S. and P. Jacob, 2004, Extreme values and kernel estimates of point processes boundaries. ESAIM: Probability and Statistics 8, 150-168.

[14] Greene, W.H., 1990, A gamma-distributed stochastic frontier model. Journal of Econometrics 46, 141-163.

[15] Hall, P., B. Park and S. Stern, 1998, On polynomial estimators of frontiers and boundaries. Journal of Multivariate Analysis 66, 71-98.

[16] Knight, K., 2001, Limiting distributions of linear programming estimators. Extremes, 4, 87-103.

[17] Korostelev, A. P., L. Simar and A. B. Tsybakov, 1995, Efficient estimation of monotone boundaries. Annals of Statistics 23, 476-489.

[18] Martins-Filho, C. and F. Yao, 2006, Nonparametric frontier estimation via local linear regression, working paper, Department of Economics, Oregon State University.

[19] Nadaraya, E., 1964, Some new estimate for distribution functions. Theory of Probability and Applications 15, 497-500.

[20] Park, B.U., L. Simar and C. Weiner, 2000, The FDH estimator for for productivity efficiency scores. Econometric Theory 16, 855-877.

[21] Rudin, W., 1976, Principles of Mathematical Analysis, McGraw Hill, New York.

[22] Simar, L. and P. Wilson, 2006, Statistical inference in nonparametric frontier models: recent developments and perspectives, in: H. Fried, C.A.K. Lovell, and S.S. Schmidt, (Eds.), The Measurement of Productive Efficiency, 2nd edition. Oxford University Press, Oxford.

[23] Silverman, B.W., 1986, Density estimation for statistics and data analysis. Chapman and Hall, London.

[24] Yang, S., 1985, A Smooth nonparametric estimator of a quantile function. Journal of the American Statistical Association 80, 1004-1011.