Econometric Theory, **37**, 2021, 346–387.
doi:10.1017/S0266466620000171

ROBUSTIFIED EXPECTED MAXIMUM PRODUCTION FRONTIERS*

ABDELAATI DAOUIA AND JEAN-PIERRE FLORENS
University of Toulouse Capitole

LÉOPOLD SIMAR
Université Catholique de Louvain

The aim of this paper is to construct a robust nonparametric estimator for the production frontier. We study this problem under a regression model with one-sided errors, where the regression function defines the achievable maximum output, for a given level of inputs-usage, and the regression error defines the inefficiency term. The main tool is a concept of partial regression boundary defined as a special probability-weighted moment. This concept motivates a robustified unconditional alternative to the pioneering class of nonparametric conditional expected maximum production functions. We prove that both the resulting benchmark partial frontier and its estimator share the desirable monotonicity of the true full frontier. We derive the asymptotic properties of the partial and full frontier estimators, and unravel their behavior from a robustness theory point of view. We provide numerical illustrations and Monte Carlo evidence that the presented concept of unconditional expected maximum production functions is more efficient and reliable in filtering out noise than the original conditional version. The methodology is very easy and fast to implement. Its usefulness is discussed through two concrete datasets from the sector of Delivery Services, where outliers are likely to affect the traditional conditional approach.

1. INTRODUCTION

The conventional microeconomic theory of the firm is based on the assumption of optimizing behavior. It is assumed that producers optimize their production choices by avoiding wasting resources. Theoretically, producers shall operate somewhere on the upper boundary, rather than on the interior, of their production possibility set

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

*The authors would like to thank two referees and the Co-Editor for their valuable suggestions, which have significantly improved the paper. They acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010. This research was also supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project. Support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) is also gratefully acknowledged. Address correspondence to Léopold Simar, Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Belgium; e-mail: leopold.simar@uclouvain.be.

© Cambridge University Press 2020

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

The upper boundary of Ψ, referred to as production frontier or surface, represents the set of the most efficient firms. The economic performance of a firm is defined in terms of its ability to operate close to or on the production frontier. This efficient frontier is often described by the graph of the function $\varphi(x) = \sup\{y \,|\, (x,y) \in \Psi\}$, which gives the maximal level of output (e.g., a quantity of goods produced) attainable by a firm operating with a vector of inputs x (e.g., labor, energy, capital). The efficiency of a unit working at (x,y) may then be estimated via the distance between its production level y and the optimal level $\varphi(x)$. The standard Farrell–Debreu efficiency score is given by the ratio $y/\varphi(x)$, so that an efficiency equal to one corresponds to an output-efficient unit. More generally, the score $y/\varphi(x) \leqslant 1$ gives the increase of output that the firm should reach to be viewed as output-efficient.

The estimation of the frontier function φ from a random sample of production units $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ is thus of utmost importance in production econometrics. A large amount of literature is devoted to this problem, where two different approaches have been mainly developed: the deterministic frontier approach, which supposes that all the observations (X_i, Y_i) belong to Ψ with probability 1, and the stochastic frontier approach, where random noise allows some observations to be outside Ψ. The issue of stochastic frontier estimation goes back to the works of Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broek (1977). Typically, it is assumed that φ has a parametric structure like Cobb–Douglas or translog. The estimation techniques include modified least-squares and maximum likelihood methods, see, for instance, Greene (2008) for a survey. Some attempts have been proposed to relax the parametric restriction such as, for instance, Kumbhakar et al. (2007)) and Simar and Zelenyuk (2011), see also Kneip, Simar, and Van Keilegom (2015) and the references therein.

Our contribution in this paper is related to the context of inference for deterministic production frontiers, where it is assumed that φ is nondecreasing. A pioneering contribution in this area is due to Farrell (1957), who introduced Data Envelopment Analysis, based on either the conical hull or the convex hull of the data. This was further extended by Deprins, Simar, and Tulkens (1984) to the Free Disposal Hull (FDH) estimator, whose properties have been extensively discussed in the literature. See, for instance, Kneip, Simar, and Wilson (2008), Daouia, Florens, and Simar (2010), and Daouia, Girard, and Guillou (2014) for a recent survey of the available results. The most appealing characteristic of such frontier estimators is that they rely on very few assumptions, but they are by construction very sensitive to outliers. To remedy this vexing defect, robust extensions using a concept of partial production frontiers have been suggested. Instead of estimating the true full frontier φ itself, the idea is to first estimate a partial frontier of the production set Ψ and then shift the obtained estimator to the right place. Prominent among these developments are the concepts of conditional expected maximum production frontiers by Cazals, Florens, and Simar (2002) and conditional quantile-based frontiers by Aragon, Daouia, and Thomas-Agnan (2005) and Daouia and Simar (2007). Comparisons between the two concepts from

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

a robustness and an asymptotic point of view can be found in Daouia and Ruiz-Gazen (2006) and Daouia and Gijbels (2011). In particular, once the conditional quantile-based frontiers break down for large chosen tail probability levels, they become definitely less resistant to outliers than the conditional expected maximum output frontiers. Moreover, the latter class of partial production functions has the additional advantage to make more efficient use of the available data since it relies on the distance to observations, whereas quantiles only use the information on whether an observation is below or above the predictor.

Yet, the class of conditional expected maximum output frontiers is not without disadvantages. First, it is not constrained to inherit the requisite theoretical axiom of monotonicity of the true full production function $\varphi(x)$. Economic considerations lead actually to the general production axiom of free disposability of inputs and outputs, that is, if $(x,y) \in \Psi$, then $(x',y') \in \Psi$ for any $x' \geq x$ and $y' \leq y$. The monotonicity of $\varphi(x)$, referred to as nonnegative marginal productivity, is justified by the free disposability assumption and is a minimal requirement in production theory (see, e.g., Gijbels et al., 1999 and Park, Simar, and Weiner, 2000) . The conditional expected maximum production function enjoys the property of monotonicity if and only if the nonstandard conditional distribution function of Y given $X \leq x$ is nonincreasing in x (see Thm. A.3 in Cazals et al., 2002); this necessary and sufficient condition is referred to as tail monotonicity (see, e.g., Gijbels and Sznajder, 2013). Second and most importantly, even if the theoretical hypothesis of tail monotonicity is satisfied, the empirical estimators of the conditional expected maximum production function, needed to be used in practice, are not constrained to enjoy the property of monotonicity. Third, a desirable property of any benchmark partial frontier is to closely parallel the true production frontier, as argued by Wheelock and Wilson (2008) and Daouia, Simar, and Wilson (2017). However, by construction, both population and empirical conditional, expected maximum output frontiers diverge from the true full frontier as the input level increases (see, e.g., Daouia and Gijbels, 2011). In particular, similarly to the FDH boundary, the estimated partial frontiers tend to envelop production units with *"small"* inputs-usage including outliers, and are thus very nonrobust to such observations. However, in contrast to the FDH frontier, they may lie below some relatively inefficient production units with *"large"* inputs-usage. This opposite behavior for *"small"* and *"large"* inputs makes the selection in practice of an appropriate benchmark partial frontier a hard problem. Also, measuring the distance of production units relative to a conditional expected maximum production frontier may result in misleading efficiency scores accordingly. All of these limitations come from the reliance of expected maximum production functions on the conditioning by the event $\{X \leq x\}$, which involves a division by an estimate of $\mathbb{P}(X \leq x)$.

In this paper, we adopt a different strategy based on a *robustified unconditional* formulation of expected maximum production functions. This new formulation has an analogous interpretation to the original conditional concept and corrects all of its vexing defects. The proposed *unconditional* expected maximum output frontiers and their estimators share the desirable property of monotonicity without resorting

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

to the hypothesis of tail monotonicity or any other assumption. Another substantial advantage of these new partial production boundaries over the traditional *conditional* approach is that they do not suffer from border and divergence effects for small or large levels of inputs. Thanks to this benefit and because monotonicity eliminates sharp changes in the slope and curvature of the built unconditional partial frontiers, the selection problem of an appropriate benchmark frontier tends to be easier than *conditional unconstrained* partial boundaries. We derive the asymptotic distributional behavior of the resulting frontier estimators (both full and partial) by using simpler arguments relative to the standard *conditional* method. The superiority of our method is also established from a robustness theory point of view. To illustrate the discussed ideas, we use two concrete datasets from the sector of Delivery Services and a third dataset from the Ecuadorian manufacturing sector, where outliers are likely to affect the traditional conditional method. The first dataset involves 4,000 French post offices observed in 1994. It has been discussed in Cazals et al. (2002), Aragon et al. (2005), Daouia et al. (2010) and Daouia, Florens, and Simar (2012) among others. The second dataset comprises 2,326 European post offices observed in 2013. For each post office i, the input X_i represents the labor cost measured by the quantity of labor, and the output Y_i is the volume of delivered mail in number of objects. The third dataset from Daouia, Gardes, and Girard (2013) consists of 406 firms in the petroleum, chemical, and plastics industries in Ecuador in 2002. The scatterplots are given below in Figures 1, 2, and 7.

The paper is further organized as follows. In Section 2, we present a deeper discussion on the concept of expected maximum production functions. We provide the main results including robustness and asymptotic properties. In Section 3, we explore the estimation method through our motivating real data examples. Section 4 gives some numerical illustrations and Monte Carlo evidence. Section 5 concludes.

2. ROBUST BOUNDARY REGRESSION

2.1. Expected Maximum Production Frontiers

In the standard nonparametric frontier model, the data

$$Y_j = \varphi(X_j) - U_j, \quad j = 1, \ldots, n,$$

are observed, with $\varphi(\cdot)$ being the unknown nondecreasing production function and $U_j \geq 0$ being the inefficiency term such that the lower support boundary of the conditional distribution of U_j given X_j is zero for almost all values of X_j. The graph of φ is thus assumed to define the upper extremity of the joint support Ψ of (X, Y) (see, e.g., Gijbels et al., 1999). This means that the support Ψ, which defines the production possibility set, is of the form

$$\begin{aligned} \Psi = \{(x,y)|y \leq \varphi(x)\} &\supseteq \{(x,y)|f(x,y) > 0\}, \\ \{(x,y)|y > \varphi(x)\} &\subseteq \{(x,y)|f(x,y) = 0\}, \end{aligned}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

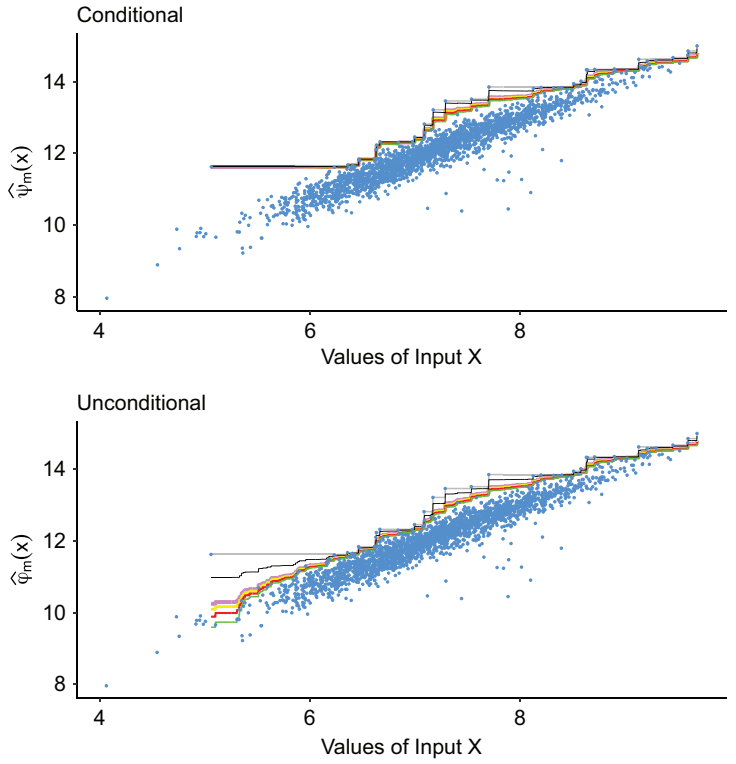


FIGURE 1. Scatterplot of the $n = 2,326$ delivery post offices (data in logarithms)—estimated expected maximum production frontiers $\widehat{\psi}_m$ (top) and $\widehat{\varphi}_m$ (bottom), with $m = 600, 700, 800, 900$, n and $m = \infty$ (FDH), respectively, in green, red, yellow, violet, black, and gray curves (see the online text for the coloured graphics).

where $f(\cdot,\cdot)$ stands for the joint density of (X, Y) (see, for instance, Daouia, Noh, and Park 2016). For a fixed level of inputs-usage $x \in \mathbb{R}_+^p$, a closed form expression of the frontier function $\varphi(x)$ has been suggested by Cazals et al. (2002) in terms of the nonstandard conditional distribution of Y given $X \le x$. If $F_{Y|X}(y|x) = \mathbb{P}(Y \le y \,|\, X \le x)$ denotes the distribution function of Y conditioned by $X \le x$, assuming $F_X(x) := \mathbb{P}(X \le x) > 0$, then $\varphi(x)$ can be characterized as the upper conditional endpoint

$$\varphi(x) = \sup\{y \ge 0 \,|\, F_{Y|X}(y|x) < 1\}. \tag{1}$$

This frontier function is isotonic nondecreasing in x. By substituting in (1), the empirical conditional distribution function $\widehat{F}_{Y|X}(y|x) = \sum_{i=1}^n \mathbb{1}(X_i \le x, Y_i \le y) / \sum_{i=1}^n \mathbb{1}(X_i \le x)$ in place of $F_{Y|X}(y|x)$, with $\mathbb{1}(\cdot)$ being the indicator function, Cazals et al. (2002) recover the usual FDH estimator

$$\widehat{\varphi}(x) = \sup\{y \ge 0 \,|\, \widehat{F}_{Y|X}(y|x) < 1\} = \max_{i:X_i \le x} Y_i.$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

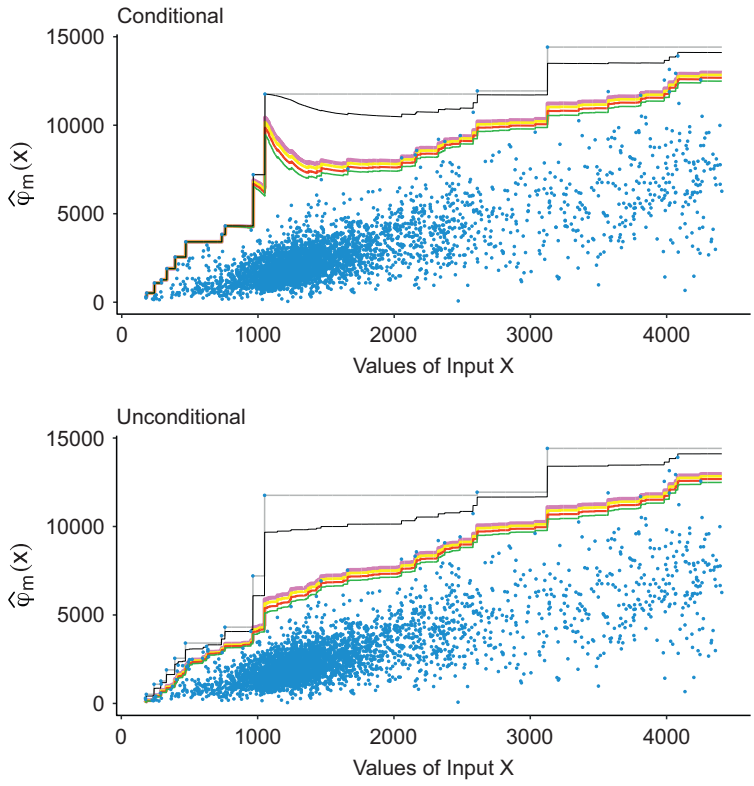


FIGURE 2. Scatterplot of the $n = 4,000$ delivery post offices—estimated expected maximum production frontiers $\widehat{\psi}_m$ (top) and $\widehat{\varphi}_m$ (bottom), with $m = 600, 700, 800, 900, n$ and $m = \infty$ (FDH), respectively, in green, red, yellow, violet, black, and gray curves (see the online text for the coloured graphics).

The graph of $\widehat{\varphi}$ being the lowest step and monotone surface which envelopes all the sample points (X_i, Y_i), it is very nonrobust to outliers. Instead, a practitioner can protect himself against outliers by estimating first an *anchor* partial frontier, well inside the cloud of data points, and then shifting the obtained estimate to the right place. For a given trimming number $m \in \{1, 2, \ldots\}$, Cazals et al. (2002) have suggested to use the concept of expected maximum output function of order m, defined as

$$\psi_m(x) = \mathbb{E}\big[\max(Y_x^1, \ldots, Y_x^m)\big] = \int_0^\infty \big(1 - [F_{Y|X}(y|x)]^m\big)\, dy,$$

where $(Y_x^1, \ldots, Y_x^m)$ are i.i.d. random variables generated by the conditional distribution of Y given $X \leq x$. The partial production function $\psi_m(x)$ converges to the true efficient frontier $\varphi(x)$ as $m \to \infty$. Likewise, for a fixed sample size n, the

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

empirical counterpart

$$\widehat{\psi}_m(x) = \int_0^\infty \left(1 - [\widehat{F}_{Y|X}(y|x)]^m\right) dy = \widehat{\varphi}(x) - \int_0^{\widehat{\varphi}(x)} [\widehat{F}_{Y|X}(y|x)]^m \, dy$$

achieves the envelopment FDH surface $\widehat{\varphi}(x)$ as $m \to \infty$. Top panels of Figures 1 and 2 display, respectively, the scatterplots of our motivating real datasets on the activity of $n = 2,326$ and $n = 4,000$ delivery post offices, along with the estimated expected maximum production frontiers of order $m = 600, 700, 800, 900, n$ and $m = \infty$ (FDH). We refer readers to the online text for the coloured graphics.

The strength of the partial frontier estimators $\widehat{\psi}_m(x)$ in terms of robustness has been established from a theoretical point of view by Daouia and Ruiz-Gazen (2006) and Daouia and Gijbels (2011). Yet, the conditioning by the event $\{X \leq x\}$ results in partial m-frontiers that can still be severely attracted by extreme and/or outlying observations with small X_i's, especially as the input level x decreases. This is visualized in the top panels of Figures 1 and 2, where the selected large m-frontiers $\widehat{\psi}_m(x)$ coincide with the nonrobust FDH estimator $\widehat{\varphi}(x)$ over an important range of values of x. Instead, we propose in the sequel to use a different formulation of expected maximum production functions without recourse to the conditioning by $X \leq x$.

2.2. Robustified Unconditional *m*-Frontiers

For a fixed level of inputs-usage $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we propose to first transform the $(p+1)$-dimensional random vector (X, Y) and the n-tuple $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ into the dimensionless random variables

$$Y^x = Y \mathbb{1}(X \leq x) \quad \text{and} \quad Y_i^x = Y_i \mathbb{1}(X_i \leq x), \quad i = 1, \ldots, n. \tag{2}$$

Their common distribution function $F_{Y^x}(\cdot)$ is closely related to the original conditional distribution function $F_{Y|X}(\cdot|x)$ since

$$F_{Y^x}(y) = \left\{1 - F_X(x)[1 - F_{Y|X}(y|x)]\right\} \mathbb{1}(y \geq 0).$$

A nice property of these transformed univariate random variables lies in the fact that

$$\begin{aligned} \varphi(x) &\equiv \sup\{y \geq 0 \,|\, F_{Y^x}(y) < 1\}, \\ \widehat{\varphi}(x) &\equiv \sup\{y \geq 0 \,|\, \widehat{F}_{Y^x}(y) < 1\} = \max(Y_1^x, \ldots, Y_n^x), \end{aligned}$$

where $\widehat{F}_{Y^x}(y) = (1/n) \sum_{i=1}^n \mathbb{1}(Y_i^x \leq y)$. We then introduce the alternative concept of expected maximum achievable level of production

$$\varphi_m(x) = \mathbb{E}\left[\max(Y_1^x, \ldots, Y_m^x)\right] = \int_0^\infty \left(1 - [F_{Y^x}(y)]^m\right) dy, \tag{3}$$

where $(Y_1^x, \ldots, Y_m^x)$ can be any m independent copies of Y^x such as, for instance, the Y_i^x's described in (2). Clearly, for any trimming number $m \geq 1$, the quantity $\varphi_m(x)$ is nothing but the expectation of the FDH estimator based on the m-tuple

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

$\{Y_i^x = Y_i \mathbb{1}(X_i \leq x)\}_{i=1,\ldots,m}$. Of particular interest is the limiting case where the partial frontier function

$$\varphi_m(x) = \int_0^{\varphi(x)} \left(1 - [F_{Y^x}(y)]^m\right) dy = \varphi(x) - \int_0^{\varphi(x)} [F_{Y^x}(y)]^m \, dy$$

converges monotonically to the true full frontier function $\varphi(x)$ as the trimming level $m \to \infty$.

Taking a closer look to $\varphi_m(x)$, we see that it can be defined equivalently as the following special probability-weighted moments.

PROPOSITION 1. *For all $m \geq 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we have*

$$\varphi_m(x) \equiv \mathbb{E}\left\{m \cdot [F_{Y^x}(Y^x)]^{m-1} \cdot Y^x\right\} \equiv \mathbb{E}\left\{J_m\left(F_{Y|X}(Y|x)\right) \cdot Y | X \leq x\right\},$$

where

$$\begin{aligned} J_m\left(F_{Y|X}(y|x)\right) &= mF_X(x)\left[1 - F_X(x)[1 - F_{Y|X}(y|x)]\right]^{m-1} \\ &= m\mathbb{P}(X \leq x)\left[1 - \mathbb{P}(X \leq x, Y > y)\right]^{m-1}. \end{aligned}$$

The probability weight $J_m\left(F_{Y|X}(y|x)\right)$ assigns bigger weights to relevant outputs. Like $\psi_m(x)$, $\varphi_m(x)$ achieves the optimal production frontier $\varphi(x)$ when the trimming number m tends to infinity. Likewise, its empirical version

$$\widehat{\varphi}_m(x) = \int_0^{\infty} \left(1 - [\widehat{F}_{Y^x}(y)]^m\right) dy = \widehat{\varphi}(x) - \int_0^{\widehat{\varphi}(x)} [\widehat{F}_{Y^x}(y)]^m \, dy \tag{4}$$

converges to the FDH frontier $\widehat{\varphi}(x)$ as $m \to \infty$. However, unlike $\widehat{\psi}_m(x)$, the weight-generating function defining $\widehat{\varphi}_m(x)$ is by construction appreciably less sensitive to border effects:

$$\widehat{\varphi}_m(x) = \sum_{i=1}^{n} Y_{(i)}^x \left\{\left(\frac{i}{n}\right)^m - \left(\frac{i-1}{n}\right)^m\right\}, \tag{5}$$

where $Y_{(i)}^x$ denotes the ith order statistic of the observations $Y_1^x, \ldots, Y_n^x$. This marks a substantial difference with $\widehat{\psi}_m(x)$ as can be visualized in the bottom panels of Figures 1 and 2 for both cases of postal services.

2.3. Monotonicity Requirement

From the point of view of the axiomatic foundation for production functions, nothing guarantees that the usual conditional expected maximum production function $\psi_m(x)$ and its estimator $\widehat{\psi}_m(x)$ satisfy the monotonicity requirement. By contrast, both our population and sample unconditional versions $\varphi_m(x)$ and $\widehat{\varphi}_m(x)$ enjoy the desirable axiom of monotonicity of the true efficient frontier $\varphi(x)$. Indeed, it is not hard to verify that

$$F_{Y^x}(y) \equiv \{1 - \mathbb{P}(X \leq x, Y > y)\}\, \mathbb{1}(y \geq 0).$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

Then, for all $y \geq 0$, the function $x \mapsto F_{Y^x}(y)$ is nonincreasing. Therefore, the unconditional partial frontier function $\varphi_m(x)$ defined in (3) is nondecreasing in x, for all $m \geq 1$. Likewise, it is easily seen that

$$\widehat{F}_{Y^x}(y) \equiv \left\{ 1 - \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i \leq x, Y_i > y) \right\} \mathbb{1}(y \geq 0)$$

is nonincreasing in x. Whence, the empirical estimator $\widehat{\varphi}_m(x)$ described in (4) is constrained to be nondecreasing in x, for all $m \geq 1$. This advantage of the new class of unconditional expected maxima $\{\widehat{\varphi}_m\}$ over the original concept of conditional versions $\{\widehat{\psi}_m\}$ is better illustrated by Figure 2 (top vs. bottom).

2.4. Asymptotic Properties

From the asymptotic point of view, we first establish the following representation.

PROPOSITION 2. *For all $m \geq 1$ and all $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we have*

$$\sqrt{n}\{\widehat{\varphi}_m(x) - \varphi_m(x)\} = \sqrt{n}\, \Phi_{m,n}(x) + o_p(1) \tag{6}$$

as $n \to \infty$, where $\Phi_{m,n}(x) = m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1} \left\{ F_{Y^x}(y) - \widehat{F}_{Y^x}(y) \right\} dy$.

An immediate consequence of this result is that $\sqrt{n}\left\{\widehat{\varphi}_m(x) - \varphi_m(x)\right\}$ is asymptotically normal with zero mean and variance

$$\begin{aligned} \sigma^2(x,m) &= \mathbb{E}\left\{ m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1} \left\{ \mathbb{1}(Y^x \leq y) - F_{Y^x}(y) \right\} dy \right\}^2 \\ &= m^2 \int_0^{\varphi(x)} \int_0^{\varphi(x)} \left[F_{Y^x}(y) F_{Y^x}(z) \right]^{m-1} \left\{ F_{Y^x}(y \wedge z) - F_{Y^x}(y) F_{Y^x}(z) \right\} dy dz. \end{aligned} \tag{7}$$

Even more strongly, we have the following functional central limit theorem.

THEOREM 1. *Suppose the support of Y is bounded. Then, for all $m \geq 1$ and any $\mathcal{X} \subset \mathbb{R}_+^p$ such that $\inf_{x \in \mathcal{X}} F_X(x) > 0$, (6) holds uniformly in $x \in \mathcal{X}$, that is,*

$$\{\sqrt{n}(\widehat{\varphi}_m(x) - \varphi_m(x));\ x \in \mathcal{X}\} = \{\sqrt{n}\Phi_{m,n}(x);\ x \in \mathcal{X}\} + o_p(1),$$

and $\sqrt{n}(\widehat{\varphi}_m - \varphi_m)$ converges in distribution in $L^\infty(\mathcal{X})$, as a process indexed by $x \in \mathcal{X}$, to the centered Gaussian process described in the proof.

Next, we show that $\sqrt{n}\left\{\widehat{\varphi}_m(x) - \varphi_m(x)\right\}$ also obeys a law of the iterated logarithm, which improves the order of convergence to $O(\sqrt{\log\log n})$ and even gives the proportionality constant.

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

THEOREM 2. *For all $m \geq 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we have almost surely, for either choice of sign,*

$$\limsup_{n\to\infty} \pm \frac{\sqrt{n}\{\widehat{\varphi}_m(x) - \varphi_m(x)\}}{(2\log\log n)^{1/2}} = \sigma(x,m).$$

2.5. Robustness Properties

From a robustness theory viewpoint, both the conditional expected maximum production function $\psi_m(x) \equiv T^{m,x}\big(F_{(X,Y)}\big)$ and its estimator $\widehat{\psi}_m(x) \equiv T^{m,x}\big(\widehat{F}_{(X,Y)}\big)$ are representable as a functional $T^{m,x}$ of the population and empirical distribution functions

$$F_{(X,Y)}(x,y) := \mathbb{P}(X \leq x, Y \leq y) \quad \text{and} \quad \widehat{F}_{(X,Y)}(x,y) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \leq x, Y_i \leq y),$$

respectively, where the statistical functional $T^{m,x}$ associates with a distribution function $F(\cdot,\cdot)$ on $\mathbb{R}_+^p \times \mathbb{R}_+$, such that $F(x,\infty) > 0$, the real value

$$T^{m,x}(F) = \int_0^\infty \left(1 - \left[\frac{F(x,y)}{F(x,\infty)}\right]^m\right) dy,$$

with the integrand being identically zero for $y \geq \inf\{y \in \mathbb{R}_+ | F(x,y)/F(x,\infty) = 1\}$. The richest quantitative robustness information is then provided by the influence function $(x_0,y_0) \mapsto \mathrm{IF}\big((x_0,y_0); T^{m,x}, F_{(X,Y)}\big)$ of $T^{m,x}$ at $F_{(X,Y)}$. It is defined as the first Gâteaux derivative of the functional $T^{m,x}$ at $F_{(X,Y)}$, where the point (x_0,y_0) plays the role of the coordinate in the infinite-dimensional space of probability distributions (see Hampel et al., 1986). The relevance of the influence function lies in its two main uses. First, it describes the effect of an infinitesimal contamination at the point (x_0,y_0) on the estimate, standardized by the mass of the contamination. Second, it allows one to assess the relative influence of individual observations $(x_0 = X_i, y_0 = Y_i)$ on the value of the estimate $\widehat{\psi}_m(x)$. An important robustness requirement is the B-robustness (Rousseeuw, 1981) which corresponds to a finite gross-error sensitivity. The maximum absolute value

$$\gamma^*\big(T^{m,x}, F_{(X,Y)}\big) = \sup_{(x_0,y_0)\in\mathbb{R}_+^{p+1}} \big|\mathrm{IF}\big((x_0,y_0); T^{m,x}, F_{(X,Y)}\big)\big|$$

defines the gross-error sensitivity of $T^{m,x}$ at $F_{(X,Y)}$. If this is unbounded, outliers can cause trouble. But according to Daouia and Ruiz-Gazen (2006), we have

$$\mathrm{IF}\big((x_0,y_0); T^{m,x}, F_{(X,Y)}\big) = \frac{m}{F_X(x)}\mathbb{1}(x_0 \leq x)\int_0^{\varphi(x)} F_{Y|X}^{m-1}(y|x)\big[F_{Y|X}(y|x) - \mathbb{1}(y_0 \leq y)\big]dy, \tag{8}$$

and hence $\gamma^*\big(T^{m,x}, F_{(X,Y)}\big) \leq \frac{m}{F_X(x)}\varphi(x)$. Even more precisely, we show here the following.

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

PROPOSITION 3. *For all $m \geq 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$,*

$$\begin{aligned}\gamma^*\left(T^{m,x}, F_{(X,Y)}\right) &= \frac{m}{F_X(x)} \max\left\{\int_0^{\varphi(x)} F_{Y|X}^m(y|x)dy, \int_0^{\varphi(x)} F_{Y|X}^{m-1}(y|x)\left[1 - F_{Y|X}(y|x)\right]dy\right\} \\ &\equiv \frac{m}{F_X(x)} \max\left\{\varphi(x) - \psi_m(x), \psi_m(x) - \psi_{m-1}(x)\right\}. \end{aligned} \tag{9}$$

The occurrence of the vexing border effect of the partial frontier estimators $\widehat{\psi}_m(x)$, due to the conditioning by the event $\{X \leq x\}$, is reflected by the presence of low values of $F_X(x)$ in the denominator of both (8) and (9).

Turning to the competing concept of unconditional expected maximum production functions, both $\varphi_m(x) \equiv T_m\left(F_{Y^x}\right)$ and $\widehat{\varphi}_m(x) \equiv T_m\left(\widehat{F}_{Y^x}\right)$ are representable as a functional T_m of the population and empirical transformed distribution functions F_{Y^x} and $\widehat{F}_{Y^x}$, respectively, where T_m associates with a univariate distribution function $F(\cdot)$ on $\mathbb{R}_+$ the real value

$$T_m(F) = \int_0^{\infty} \left(1 - [F(y)]^m\right)dy = \int_0^{F^{-1}(1)} \left(1 - [F(y)]^m\right)dy,$$

with the integrand being identically zero for $y \geq F^{-1}(1) := \inf\{y \in \mathbb{R} | F(y) = 1\}$. Following Hampel et al. (1986, Def. 1, p. 84) the corresponding influence function of T_m at F_{Y^x} is defined as the ordinary derivative

$$u \in \mathbb{R}_+ \mapsto \mathrm{IF}\left(u; T_m, F_{Y^x}\right) = \frac{d}{dt}_{|t=0} T_m\left((1-t)F_{Y^x} + t\delta_u\right).$$

In robust statistics, a small fraction of the observations would have a strong influence on the estimator if their values were equal to a u where the influence function is large.

PROPOSITION 4. *For all $m \geq 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we have*

$$\begin{aligned} IF\left(u; T_m, F_{Y^x}\right) &= -m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\left\{\delta_u(y) - F_{Y^x}(y)\right\}dy \\ &\equiv -m \int_0^{\varphi(x)} \left[1 - F_X(x) + F_{(X,Y)}(x,y)\right]^{m-1}\left\{\mathbb{1}(u \leq y) - 1 + F_X(x) - F_{(X,Y)}(x,y)\right\}dy. \end{aligned}$$

This closed form expression of the influence function indicates that the unconditional m-frontiers $\widehat{\varphi}_m(x) \equiv T_m\left(\widehat{F}_{Y^x}\right)$ do not suffer from the inherent border effects of the initial concept of conditional m-frontiers $\widehat{\psi}_m(x) \equiv T^{m,x}\left(\widehat{F}_{(X,Y)}\right)$. Moreover, by making use of the same technique of the proof of Proposition 3, it is easily seen that the gross-error sensitivity $\lambda^*\left(T_m, F_{Y^x}\right) := \sup_{u \geq 0} \left|\mathrm{IF}\left(u; T_m, F_{Y^x}\right)\right|$ satisfies

$$\begin{aligned}\lambda^*\left(T_m, F_{Y^x}\right) &= m \cdot \max\left\{\int_0^{\varphi(x)} F_{Y^x}^m(y)dy, \int_0^{\varphi(x)} F_{Y^x}^{m-1}(y)\left[1 - F_{Y^x}(y)\right]dy\right\} \\ &\equiv m \cdot \max\left\{\varphi(x) - \varphi_m(x), \varphi_m(x) - \varphi_{m-1}(x)\right\}\end{aligned}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

which, in contrast to $\gamma^*\big(T^{m,x}, F_{(X,Y)}\big)$, does not explode when x decreases. Also, as can be seen from (6) in Proposition 2, $\mathrm{IF}\big(Y_i^x; T_m, F_{Y^x}\big)$ represents the approximate contribution, or influence, of the observation (X_i, Y_i) toward the estimation error $\big\{\widehat{\varphi}_m(x) - \varphi_m(x)\big\}$, since

$$\sqrt{n}\{\widehat{\varphi}_m(x) - \varphi_m(x)\}$$
$$= \sqrt{n}\,\Phi_{m,n}(x) + o_p(1) \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathrm{IF}\big(Y_i^x; T_m, F_{Y^x}\big) + o_p(1), \quad n \to \infty.$$

Similarly, the influence function of the "conditional" partial frontier estimator, described in (8), measures the asymptotic bias caused by contamination in the observations (X_i, Y_i):

$$\sqrt{n}\big(\widehat{\psi}_m(x) - \psi_m(x)\big) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathrm{IF}\big((X_i, Y_i); T^{m,x}, F_{(X,Y)}\big) + o_p(1), \quad n \to \infty.$$

However, the consideration of the "unconditional" partial frontier estimator $\widehat{\varphi}_m(x)$, instead of the conditional frontier estimator $\widehat{\psi}_m(x)$, may result in a better asymptotic variance (7), especially when $\mathbb{P}(X \le x)$ is small.

2.6. Regularized Frontier Estimators

It should be clear that the estimation of a "partial" frontier φ_m, for a sufficiently large value of m, instead of the "full" frontier φ is mainly motivated by the search for a "robust" frontier estimator $\widehat{\varphi}_m$, which is well inside the cloud of data points $\{(X_i, Y_i), i = 1, \ldots, n\}$, but lies near the true upper support boundary. The robustness of $\widehat{\varphi}_m$ comes from its convergence monotonely from below to the smallest sample envelope (FDH) $\widehat{\varphi}$ as the trimming number m increases. When $m = m_n \to \infty$ at a fast rate as $n \to \infty$, the next theorem shows that the robust frontier $\widehat{\varphi}_{m_n}(x)$ estimates $\varphi(x)$ itself and converges to the same limit distribution as the FDH estimator $\widehat{\varphi}$ with the same scaling. Recall first that, following Daouia et al. (2010, Thm. 2.1), there exists $b_n > 0$ such that $b_n^{-1}(\widehat{\varphi}(x) - \varphi(x))$ converges to a non-degenerate distribution if and only if

$$F_X(x)[1 - F_{Y|X}(y|x)] = L_x\left(\{\varphi(x) - y\}^{-1}\right)(\varphi(x) - y)^{\rho_x} \quad \text{as} \quad y \uparrow \varphi(x),$$

for some constant $\rho_x > 0$, where $L_x(\cdot)$ is a slowly varying function, that is, $\lim_{t\uparrow\infty} L_x(tz)/L_x(t) = 1$ for all $z > 0$. The limit distribution function is identical to

$$F_{\rho_x}(y) = \exp\{-(-y)^{\rho_x}\} \quad \text{with support} \quad (-\infty, 0].$$

Under the sufficient condition that $L_x\left(\{\varphi(x) - y\}^{-1}\right) \sim \ell_x > 0$ as $y \uparrow \varphi(x)$, that is

Condition $\mathcal{C}(\boldsymbol{\rho_x}, \boldsymbol{\ell_x})$: For some constants $\rho_x > 0$ and $\ell_x > 0$,

$$F_X(x)[1 - F_{Y|X}(y|x)] = \ell_x\big(\varphi(x) - y\big)^{\rho_x} + o\big((\varphi(x) - y)^{\rho_x}\big) \quad \text{as} \quad y \uparrow \varphi(x),$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

it is shown in Daouia et al. (2010, Cor. 2.1), that $b_n \sim (n\ell_x)^{-1/\rho_x}$ and

$$(n\ell_x)^{1/\rho_x}\left\{\varphi(x)-\widehat{\varphi}(x)\right\} \xrightarrow{\mathcal{L}} \text{Weibull}(1,\rho_x) \quad \text{as} \quad n\to\infty,$$

where a random variable W is said to follow the distribution Weibull$(1,\rho_x)$ if W^{ρ_x} is Exponential with parameter 1. As described thoroughly in Rem. 2.3 of Daouia et al. (2010), the exponent ρ_x has the following intuitive meaning in terms of the density of (X,Y) and the dimension $(p+1)$: when $\rho_x > p+1$, the joint density decays to zero at a speed of power $\rho_x-(p+1)$ of the distance from the frontier point $\varphi(x)$. When $\rho_x = p+1$, the density has a sudden jump at the frontier. Finally, when $\rho_x < p+1$, the density rises up to infinity at a speed of power $\rho_x-(p+1)$ of the distance from the frontier.

THEOREM 3. *For $x\in\mathbb{R}_+^p$ such that $F_X(x)>0$, if $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$ holds and $m_n \geq \beta n\log n\{1+o(1)\}$ for some constant $\beta > \frac{1}{\rho_x}+1$, then*

$$(n\ell_x)^{1/\rho_x}\left\{\varphi(x)-\widehat{\varphi}_{m_n}(x)\right\} \xrightarrow{\mathcal{L}} \textit{Weibull}(1,\rho_x) \quad \textit{as} \quad n\to\infty.$$

By contrast, when $m=m_n\to\infty$ at a slow rate as $n\to\infty$, the robust frontier estimator $\widehat{\varphi}_{m_n}(x)$ becomes asymptotically Gaussian, as in the regular case of a fixed m.

THEOREM 4. *Let $x\in\mathbb{R}_+^p$ such that $F_X(x)>0$.*

(i) *If $m_n\to\infty$ and $\frac{m_n^2}{\sigma(x,m_n)} = O\left(\frac{\sqrt{n}}{\log\log n}\right)$ as $n\to\infty$, then*

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\left\{\widehat{\varphi}_{m_n}(x)-\varphi_{m_n}(x)\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad n\to\infty. \tag{10}$$

(ii) *Under the extreme-value condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$, we have*

$$c_x m^{1-2/\rho_x} \leq \sigma^2(x,m) \leq \tilde{c}_x m^{2-2/\rho_x} \quad \textit{as} \quad m\to\infty,$$

for some positive constants c_x and $\tilde{c}_x$.

(iii) *Also, under $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$, if $m_n\to\infty$ with $m_n = O\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}}$, then the asymptotic normality (10) is still valid.*

Note that the explicit condition $m_n\to\infty$ with $m_n = O\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}}$, in Theorem 4(iii), implies that $m_n/\sqrt{n}\to 0$ as $n\to\infty$. We would like also to comment on the speed of convergence $\sqrt{n}/\sigma(x,m_n)$, obtained in Theorem 4(i) and (iii), when the trimming level $m_n\to\infty$ at a slow rate so that $m_n = c\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}}$, for some constant $c>0$. By Theorem 4(ii), as $n\to\infty$, we get

$$k_1\sqrt{n}^{\frac{\frac{1}{2}+\frac{2}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}}(\log\log n)^{\frac{1-\frac{1}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}} \leq \frac{\sqrt{n}}{\sigma(x,m_n)} \leq k_2\sqrt{n}^{\frac{1+\frac{2}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}}(\log\log n)^{\frac{\frac{1}{2}-\frac{1}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

for some constants $k_1, k_2 > 0$. In the particular case $\rho_x = p+1$, often assumed in the literature of production econometrics, which corresponds to a joint density of (X, Y) having a jump at the frontier point $\varphi(x)$, we have

$$k_1 \left(n^{3/2} \log\log n\right)^{1/4} \leq \frac{\sqrt{n}}{\sigma(x, m_n)} \leq k_2 \sqrt{n} \qquad \text{as} \quad p = 1,$$

$$\text{and} \quad k_1 n^{1/6} (\log\log n)^{2/3} \leq \frac{\sqrt{n}}{\sigma(x, m_n)} \leq k_2 (n \log\log n)^{1/3} \qquad \text{as} \quad p \uparrow \infty.$$

Interestingly, even when the data dimension explodes, the speed of convergence does not deteriorate too much, thereby reducing the *curse of dimensionality* that is typical of many nonparametric frontier estimators such us, for instance, the FDH estimator.

2.7. Bias-Corrected Estimator of $\varphi(x)$

Under the extremal condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}}, \boldsymbol{\ell}_{\mathbf{x}})$, when $m_n \to \infty$ with $m_n = O\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}}$, Theorem 4(iii) actually indicates that $\widehat{\varphi}_{m_n}(x)$ estimates $\varphi(x)$ itself with the inherent bias $B_{m_n}(x) = \varphi(x) - \varphi_{m_n}(x)$ such that

$$\frac{\sqrt{n}}{\sigma(x, m_n)} \left\{\widehat{\varphi}_{m_n}(x) - \varphi(x) + B_{m_n}(x)\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad n \to \infty. \tag{11}$$

Recall that, in view of (3),

$$\varphi_{m_n}(x) = \mathbb{E}\left[\max(Y_1^x, \ldots, Y_{m_n}^x)\right]$$

is nothing but the expectation of the FDH estimator, $\max(Y_1^x, \ldots, Y_{m_n}^x) \equiv \max_{i:X_i \leq x} Y_i$, based on the m_n-tuple $\{Y_i^x = Y_i \mathbb{1}(X_i \leq x)\}_{i=1,\ldots,m_n}$. Under the sufficient condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}}, \boldsymbol{\ell}_{\mathbf{x}})$, the limit theorem of moments of the FDH estimator in Daouia et al. (2010, Thm. 2.1(iii)), shows that

$$\lim_{m_n \to \infty} \mathbb{E}\left[b_{m_n}^{-1} \left\{\varphi(x) - \max(Y_1^x, \ldots, Y_{m_n}^x)\right\}\right] = \Gamma(1 + 1/\rho_x),$$

where Γ is the gamma function, which entails that

$$\lim_{m_n \to \infty} b_{m_n}^{-1} \left\{\varphi(x) - \varphi_{m_n}(x)\right\} = \Gamma(1 + 1/\rho_x), \tag{12}$$

with $b_{m_n} \sim (m_n \ell_x)^{-1/\rho_x}$, or equivalently,

$$B_{m_n}(x) = \varphi(x) - \varphi_{m_n}(x) = (m_n \ell_x)^{-1/\rho_x} \Gamma(1 + 1/\rho_x) + o\left(m_n^{-1/\rho_x}\right), \quad n \to \infty. \tag{13}$$

Combining this with Theorem 4(ii), it follows that the introduced bias (normalized by the rate of convergence) is bounded below by

$$\frac{\sqrt{n}}{\sigma(x, m_n)} B_{m_n}(x) > \check{c}_x \left(\sqrt{n}^{\frac{1}{2}+\frac{1}{\rho_x}} \log\log n\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}},$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

for some constant $\check{c}_x > 0$. The normalized bias does not then vanish asymptotically, and hence one would use in practice the asymptotic approximation:

$$\varphi(x) - \widehat{\varphi}_{m_n}(x) \approx \mathcal{N}\left(B_{m_n}(x), \frac{\sigma^2(x, m_n)}{n}\right),$$

where $B_{m_n}(x)$ and $\sigma^2(x, m_n)$ have to be replaced by consistent estimators. The plugging version of $\sigma^2(x, m_n)$ in (7) provides a consistent estimator of this asymptotic variance. As for the bias term, a consistent estimator can be obtained through the leading part of (13) once ρ_x and ℓ_x are consistently estimated. One way of estimating these parameters is by adapting the ideas from Sect. 4.1 in Daouia et al. (2012). Given an integer $a \geq 2$, we have by (12) that

$$\lim_{n\to\infty} \frac{\varphi_{am_n}(x) - \varphi_{m_n}(x)}{\varphi_{a^2 m_n}(x) - \varphi_{am_n}(x)} = a^{1/\rho_x},$$

which motivates the estimator

$$\widehat{\rho}_x := \log(a)\left\{\log\left(\frac{\widehat{\varphi}_{am_n}(x) - \widehat{\varphi}_{m_n}(x)}{\widehat{\varphi}_{a^2 m_n}(x) - \widehat{\varphi}_{am_n}(x)}\right)\right\}^{-1}. \tag{14}$$

On the other hand, it follows from (13) that

$$\ell_x \sim \frac{1}{m_n}\left[\frac{(1 - a^{-1/\rho_x})\Gamma(1 + 1/\rho_x)}{\varphi_{am_n}(x) - \varphi_{m_n}(x)}\right]^{\rho_x}, \quad n \to \infty,$$

which suggests the estimator

$$\widehat{\ell}_x := \frac{1}{m_n}\left[\frac{(1 - a^{-1/\widehat{\rho}_x})\Gamma(1 + 1/\widehat{\rho}_x)}{\widehat{\varphi}_{am_n}(x) - \widehat{\varphi}_{m_n}(x)}\right]^{\widehat{\rho}_x}. \tag{15}$$

Both $\widehat{\rho}_x$ and $\widehat{\ell}_x$ are consistent estimators.

THEOREM 5. *Under the conditions of Theorem 4(iii),*

$$\widehat{\rho}_x \xrightarrow{p} \rho_x \text{ and } \widehat{\ell}_x \xrightarrow{p} \ell_x \text{ as } n \to \infty. \tag{16}$$

Let us now return to the starting point (4.3) to investigate the asymptotic normality of the bias-corrected estimator itself. This estimator is defined as

$$\widetilde{\varphi}_{m_n}(x) := \widehat{\varphi}_{m_n}(x) + \widehat{B}_{m_n}(x), \tag{17}$$

where, assuming for ease of presentation that ρ_x is given,

$$\widehat{B}_{m_n}(x) := (m_n \widehat{\ell}_x)^{-1/\rho_x}\Gamma(1 + 1/\rho_x)$$

is the plug-in version of the bias $B_{m_n}(x)$ obtained by replacing ℓ_x, in the leading part of (13), with its consistent estimate

$$\widehat{\ell}_x := \widehat{\ell}_x(\tilde{m}_n) = \frac{1}{\tilde{m}_n}\left[\frac{(1 - a^{-1/\rho_x})\Gamma(1 + 1/\rho_x)}{\widehat{\varphi}_{a\tilde{m}_n}(x) - \widehat{\varphi}_{\tilde{m}_n}(x)}\right]^{\rho_x}. \tag{18}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

Here, we shall distinguish between the trimming level m_n in the estimator $\widetilde{\varphi}_{m_n}(x)$ of the frontier function $\varphi(x)$ and the level $\tilde{m}_n$ used in the estimator $\widehat{\ell}_x$ of the parameter ℓ_x. Nothing guarantees that the two levels are necessarily the same. It should also be noted that, while the asymptotic normality of the *partial* frontier estimator $\widehat{\varphi}_{m_n}(x)$ in Theorem 4 hinges on the first-order representation (13), that is,

$$\varphi(x)-\varphi_{m_n}(x)=(m_n\ell_x)^{-1/\rho_x}\Gamma(1+1/\rho_x)+o\left(m_n^{-1/\rho_x}\right),\quad n\to\infty,$$

which is implied by the extremal condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$, the asymptotic normality of the *full* frontier estimator $\widetilde{\varphi}_{m_n}(x)$ requires the following second-order representation:

Condition $\mathcal{C}_2(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}},\boldsymbol{\alpha}_{\mathbf{x}})$: For some constants $\rho_x>0$, $\ell_x>0$, and $\alpha_x>0$,

$$\varphi(x)-\varphi_{m_n}(x)=(m_n\ell_x)^{-1/\rho_x}\Gamma(1+1/\rho_x)+o\left(m_n^{-(1+\alpha_x)/\rho_x}\right),\quad n\to\infty,$$

where the extra parameter α_x is needed to control the speed of convergence, in the first-order condition, of $(m_n\ell_x)^{1/\rho_x}\left\{\varphi(x)-\varphi_{m_n}(x)\right\}$ to $\Gamma(1+1/\rho_x)$.

THEOREM 6. *Let $x\in\mathbb{R}_+^p$ such that $F_X(x)>0$. Under $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$ and $\mathcal{C}_2(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}},\boldsymbol{\alpha}_{\mathbf{x}})$ with $\alpha_x>\rho_x+1$, if $m_n=c\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\varepsilon}$ and $\tilde{m}_n=\tilde{c}\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\tilde{\varepsilon}}$, for some constants $c,\tilde{c}>0$ and $0<\tilde{\varepsilon}<\left(\frac{3}{2}+\frac{1}{\rho_x}\right)^{-1}<2\tilde{\varepsilon}-\varepsilon$, such that $1-\frac{1}{2}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\varepsilon\right)-\frac{\alpha_x}{\rho_x}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\tilde{\varepsilon}\right)<0$, then*

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\left(\widetilde{\varphi}_{m_n}(x)-\varphi(x)\right)\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)\quad as\quad n\to\infty.$$

The condition $\alpha_x>\rho_x+1$ in Theorem 6 is needed to control the bias approximation error (driven by the last term in $\mathcal{C}_2(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}},\boldsymbol{\alpha}_{\mathbf{x}})$) so as to get $\frac{\sqrt{n}}{\sigma(x,m_n)}o\left(m_n^{-(1+\alpha_x)/\rho_x}\right)=o(1)$. The condition $\left(\frac{3}{2}+\frac{1}{\rho_x}\right)^{-1}<2\tilde{\varepsilon}-\varepsilon$ is required to select $\tilde{m}_n=o\left(m_n^{1/2}\right)$ in the estimator $\widehat{\ell}_x:=\widehat{\ell}_x(\tilde{m}_n)$ of ℓ_x. It is easily seen that the condition $\tilde{\varepsilon}<\left(\frac{3}{2}+\frac{1}{\rho_x}\right)^{-1}$ implies

$$1-\frac{1}{2}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\varepsilon\right)-\frac{\alpha_x}{\rho_x}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\tilde{\varepsilon}\right)<1-\frac{1}{2}(\tilde{\varepsilon}-\varepsilon).$$

Hence, the last condition of the theorem, that is $1-\frac{1}{2}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\varepsilon\right)-\frac{\alpha_x}{\rho_x}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}-\tilde{\varepsilon}\right)<0$ is satisfied if, for instance, $\tilde{\varepsilon}-\varepsilon>2$.

It should be noted that we restrict ourselves in Theorem 6 to the case where ρ_x is known. This corresponds, for instance, to the standard assumption in productivity and efficiency analysis that the joint density of data (X_i,Y_i) has jumps at the frontier, or equivalently $\rho_x=p+1$ (see the discussion above, Theorem 3). The

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

question of whether the asymptotic normality in Theorem 6 still holds when replacing ρ_x by its estimator $\widehat{\rho}_x$ is of interest. The complexity of using $\widehat{\rho}_x$ in place of ρ_x in the proof is that it adds two additional terms to the two terms I and II already in use in (A.8). Theoretical developments along these lines are left for future research.

3. TRIMMING SELECTION PROBLEM

We return here to our real data examples with a single input ($p = 1$) to explore in Section 3.1, the selection of the trimming level m_n in the partial frontier $\widehat{\varphi}_{m_n}$, before moving to the final bias-corrected frontier $\widetilde{\varphi}_{m_n}(x)$ in Section 3.2. We extend our discussion to multiple inputs ($p > 1$) in Section 3.3.

3.1. Selecting the Partial Frontier $\widehat{\varphi}_{m_n}$

In productivity and efficiency analysis where outliers are likely to affect traditional envelopment approaches, a common robust practice in operations research and applied work consists in using an empirical partial frontier as a benchmark to measure the efficiency of production units. Unfortunately, the chosen partial frontier is often based on an a priori selected order $m = m_n$ in the case of conditional expected maxima, or tail probability in the case of conditional quantiles. Here, we propose practical guidelines for a more justified selection from a robustness theory viewpoint.

As with any trimming techniques, the degree of truncation, here reflected through m selection, is a major issue in practice. But monotonicity itself is a rather powerful way of regularizing the estimated expected maximum production function. Because it eliminates sharp changes in the slope and curvature of the unconditional m-frontier function, the trimming selection problem tends to be easier than unconstrained conditional m-frontier estimation. Of course, if the model is known or believed to be nearly correct, then the use of the envelopment FDH estimator ($m = \infty$) is required. Otherwise, if the dataset contains suspicious isolated extreme observations, it is more prudent to seek for "robustification" via the choice of an adequate trimming level m. To verify the presence of such influential observations among the data (e.g., French and European postal datasets), a simple diagnostic tool is by using the gross-error sensitivity of the sequence $\{\widehat{\varphi}_m\}_m$, which corresponds to the maximum influence function. Figure 3 shows the sample gross-error sensitivity $x \mapsto \lambda^*\big(T_m, \widehat{F}_{Y^x}\big)$, for various values of $m = 100, 200, \ldots, 1{,}500$. For both postal services, the evolution of λ^* exhibits some slight and severe breakdowns at different values of x, especially in the case of French post offices (r.h.s.). This indicates the presence of isolated extreme and/or anomalous data. One way of choosing the trimming number m is then by looking to Figure 4 which indicates how the percentage of data points (X_i, Y_i) above the curve of $\widehat{\varphi}_m$ decreases with m. The basic idea is to choose values

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

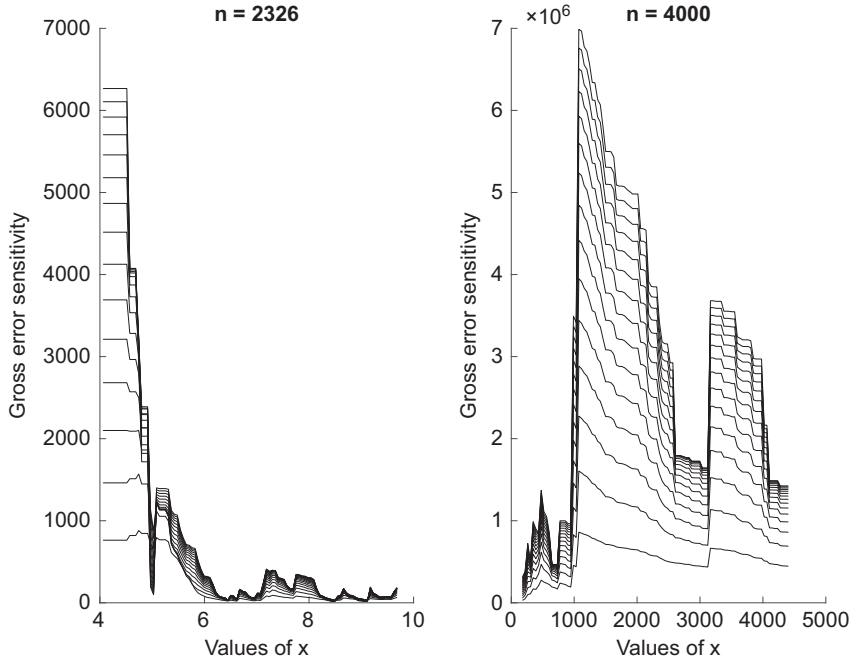


FIGURE 3. Plots of $x \mapsto \lambda^*\left(T_m, \widehat{F}_{Y^x}\right)$ for $m = 100, 200, \ldots, 1,500$. From left to right, the 2,326 and 4,000 post offices.

of m for which the frontier estimator $\widehat{\varphi}_m$ is sensitive to the magnitude of valuable extreme post offices while remaining resistant to isolated outliers.

The evolution of the percentage in both sectors of Delivery Services has clearly an "L" structure highlighted by a colour-scheme, ranging from dark red (high %) to dark violet (low %). We refer readers to the online text for the "colouring of the L evolution." Such an L deviation should appear for any other analyzed dataset since, by construction, the probability-weighted moments $\widehat{\varphi}_m$ steer an advantageous middle course between sensitivity and robustness to extreme values and/or outliers. In the case of 2,326 delivery post offices (top picture in Figure 4), the percentage first falls rapidly along the "red" part of the curve. This means that most of the observations lying above the corresponding m-frontiers are not extremes but interior points to the cloud of data points. Then the evolution of the percentage shows an "elbow effect" along the "orange" and "green" parts of the curve. This means that the observations outside the corresponding m-frontiers are no more inefficient, but still contain either relatively efficient post offices that are well inside the sample or top observations that are valuable post offices. In contrast, after the elbow effect, it may be seen that the percentage decreases very slowly along the "blue" part, say $850 \leq m \leq 1,250$, before becoming extremely stable along the "violet" part of the curve. This means that all observations left

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

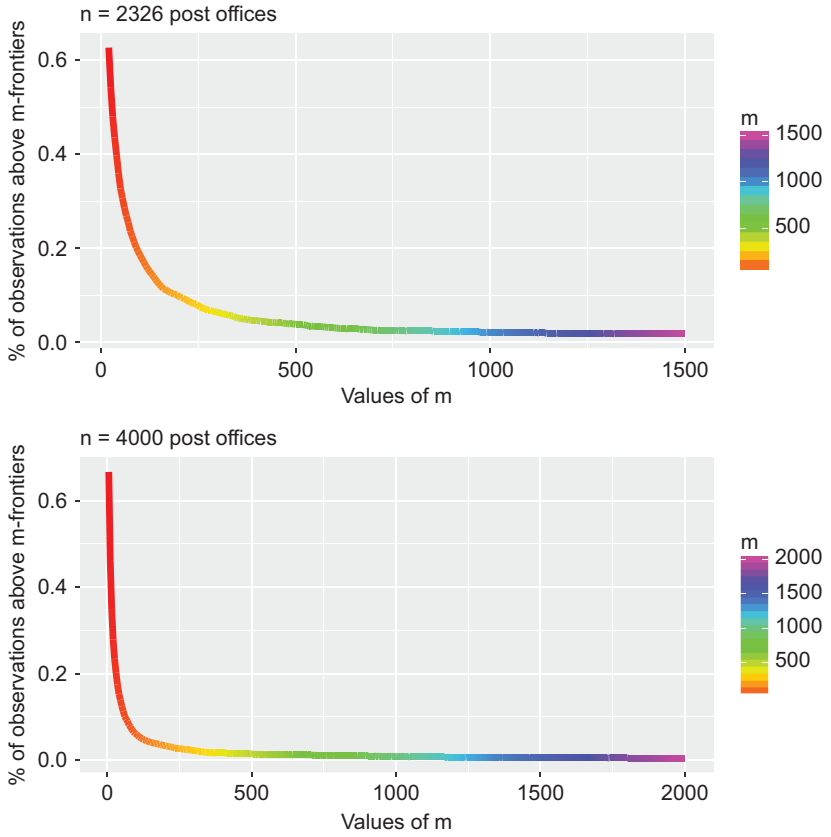


FIGURE 4. Evolution of the % of sample points outside the partial m-frontiers $\widehat{\varphi}_m$ (see the online text for a colour-scheme).

outside the partial frontier of order $m = 850$ are really very extreme in the Y-direction and could be outlying or perturbed by noise. This might suggest to select 850 as a potential lower value for m. On the other hand, the extreme stability of the percentage curve from $m = 1,250$ may indicate that the observations above the frontier $\widehat{\varphi}_{1,250}$ are really outlying or suspicious isolated extremes that deserve to be carefully examined. This might suggest to choose $1,250$ as a potential upper value for m. The two potential choices of the frontier estimator $\widehat{\varphi}_m$ are graphed in Figure 5 along with the FDH estimator.

As regards the $4,000$ delivery post offices (bottom picture in Figure 4), it may be seen that the "elbow effect" corresponds to the "orange" part of the percentage curve, and the desired range of values of m follows as the "green" part, say, $500 \leq m \leq 1,000$. The lower and upper selected prudential frontiers $\widehat{\varphi}_{500}$ and $\widehat{\varphi}_{1,000}$ are

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

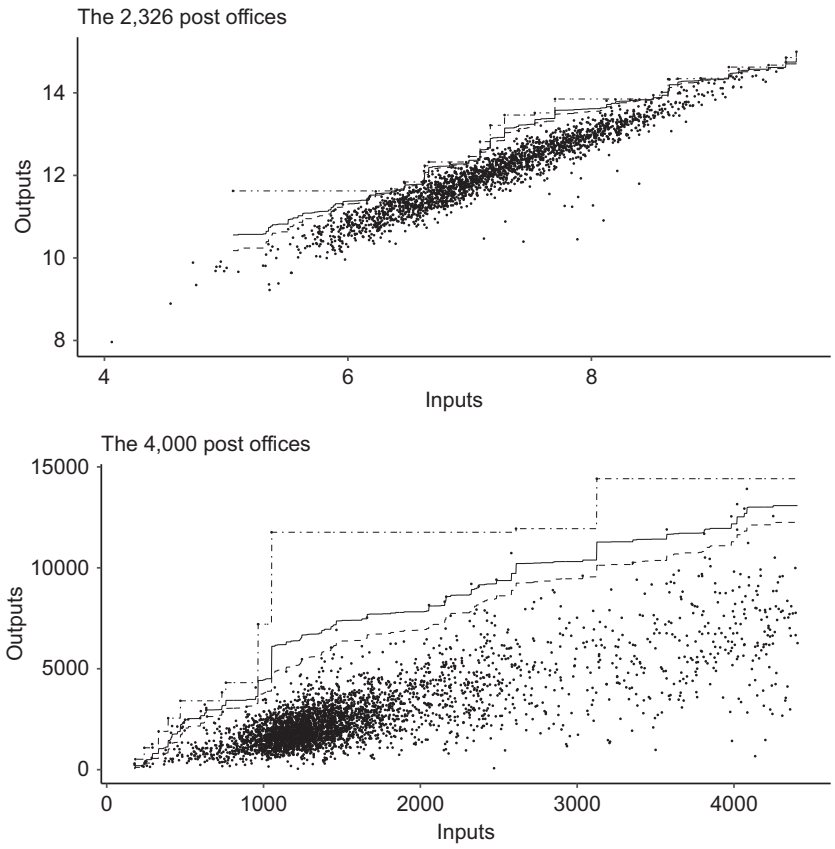


FIGURE 5. Selected (lower and upper) expected maximum production frontiers $\widehat{\varphi}_m$. Top—dataset of size 2,326 in logarithms, with $m = 1,250$ (upper) in solid line, $m = 850$ (lower) in dashed line, and $m = \infty$ (FDH) in dashdotted line. Bottom—dataset of size 4,000, with $m = 1,000$ (upper) in solid line, $m = 500$ (lower) in dashed line, and $m = \infty$ (FDH) in dashdotted line.

superimposed in Figure 5 along with the FDH estimator. Unsurprisingly, there are very few observations lying between the two partial frontiers.

3.2. The Final Bias-Corrected Frontier $\widetilde{\varphi}_{m_n}(x)$

Under the usual assumption in production econometrics that $\rho_x = p + 1 \equiv 2$, the final frontier estimator $\widetilde{\varphi}_{m_n}(x)$ has the closed form expression

$$\widetilde{\varphi}_{m_n}(x) = \widetilde{\varphi}_{m_n,\tilde{m}_n,a}(x) = \widehat{\varphi}_{m_n}(x) + (m_n\widehat{\ell}_x)^{-1/\rho_x}\Gamma(1+1/\rho_x)$$
$$\equiv \widehat{\varphi}_{m_n}(x) + \left(\frac{\tilde{m}_n}{m_n}\right)^{1/2} \frac{\widehat{\varphi}_{a\tilde{m}_n}(x) - \widehat{\varphi}_{\tilde{m}_n}(x)}{(1-a^{-1/2})},$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

where $m_n \in [850, 1,250]$ for the sample size 2,326 and $m_n \in [500, 1,000]$ for the sample size 4,000. For illustration purposes, we restrict to the upper selected prudential levels $m_n = 1,250$ for $n = 2,326$ and $m_n = 1,000$ for $n = 4,000$. Our experience with these data indicates that $\widetilde{\varphi}_{m_n}(x)$ is not sensitive to the choice of the tuning parameter $a \geq 2$. For example, the frontier estimates obtained for all values of a in $[2, 10]$ appear to be very similar. However, the estimates seem to be more sensitive to the choice of $\tilde{m}_n$. This is illustrated in Figure 6 for both datasets, where the final bias-corrected frontiers $x \mapsto \widetilde{\varphi}_{m_n}(x)$ are plotted for $a = 2$ and two different values of $\tilde{m}_n = m_n^{0.005}$ (dashed) and $\tilde{m}_n = m_n^{0.2}$ (solid), along with the nonrobust FDH frontier (dashdotted). Although the resulting (blue and red) frontiers for both values of $\tilde{m}_n$ are very close for the largest dataset of size $n = 4,000$ (bottom panel), it may be seen that they are quite different in the case $n = 2,326$ (top panel). We do not enter here into the question of optimal selection of $\tilde{m}_n$, but it is clearly of genuine interest and is still open for future research.

3.3. Extension to Multiple Inputs

It should be clear that, thanks to the dimensionless transformation adopted in (2), the practical guidelines described above evidently apply to higher dimension $p > 1$. For our illustration purposes, we consider here a real data example in the case $p = 2$, where the dataset consists of $n = 406$ firms in the petroleum, chemical and plastics industries in Ecuador in 2002. For each firm, we have information on the capital K in thousands of USD, the average number of employees L and the value-added real output Y in thousands of USD. The scatterplot of the 406 observations (in logarithm scale) is displayed in Figure 7. In this particular example, the efficient FDH surface is determined by only 12.56% of the firms, and some of these extremal FDH firms are outlying as can be seen from Daouia et al. (2013). The latter authors used the "conditional" partial $m-$frontiers $\{\widehat{\psi}_m\}_m$, rather than the unreliable FDH frontier, as a robust benchmark for the assessment of the production performance of firms. The objective here is to compare their method with our alternative proposal of "unconditional" partial m-frontiers $\{\widehat{\varphi}_m\}_m$, for a suitable choice of the trimming levels m.

Figure 8 shows the evolution of the percentage of sample points left outside both partial m-frontiers $\widehat{\psi}_m$ (pink curve) and $\widehat{\varphi}_m$ (rainbow curve); we refer readers to the online text for the coloured graphics. The decrease of the percentage corresponding to the "unconditional" partial m-frontiers $\widehat{\varphi}_m$ is clearly slower than the one corresponding to the "conditional" versions $\widehat{\psi}_m$. This reflects the resistance of the "unconditional" partial m-frontiers to the magnitude of extremes and/or outliers. It may also be seen that the decrease of the percentage becomes very slow from $m = 183$ for the pink curve (indicated by the vertical pink line) and from $m = 336$ for the rainbow curve (indicated by the vertical blue line). Figure 9 (top panel) shows the resulting values $\widehat{\psi}_{183}(x_i)$ for 20 randomly chosen grid inputs $x_i = (K_i, L_i)$. As is to be expected in the case of conditional expected maxima, there are many violations of monotonicity by the multiargument function $\widehat{\psi}_{183}(x_i)$ (with

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

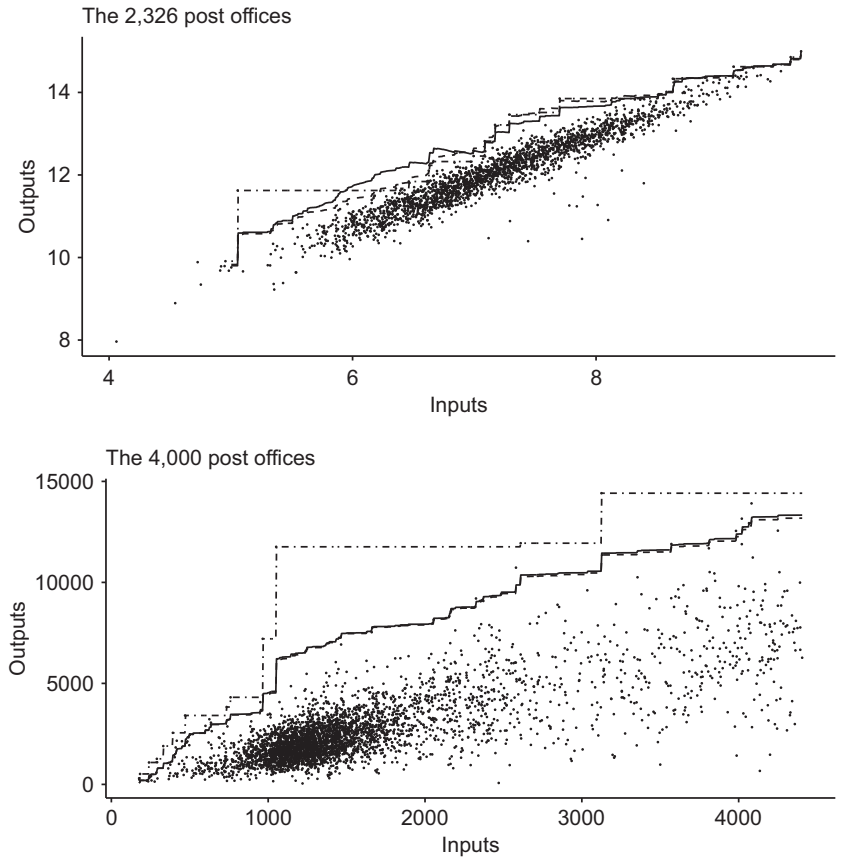


FIGURE 6. Final bias-corrected frontiers $x \mapsto \widetilde{\varphi}_{m_n,\tilde{m}_n,a}(x)$ for $a = 2$ and two different values of $\tilde{m}_n = m_n^{0.005}$ (dashed) and $\tilde{m}_n = m_n^{0.2}$ (solid), along with the FDH frontier (dashdotted). Top—dataset of size 2,326 in logarithms, with $m_n = 1250$. Bottom—dataset of size 4,000, with $m_n = 1,000$.

respect to the partial order induced by "$\leq$"). Figure 9 (bottom panel) displays the values of $\widehat{\varphi}_{336}(x_i)$ for the same selected 20 points, showing that the unconditional expected maximum production function is well isotonic nondecreasing. When taking larger trimming levels m (in the stable regions starting from the vertical dashed lines), the lessons were the same in terms of robustness and monotonicity.

4. NUMERICAL ILLUSTRATIONS

In this section, we illustrate our procedure through two standard examples with simulated data. We consider the same data generating processes traditionally used in the literature of nonparametric frontier estimation such as, for instance, Gijbels et al. (1999), Cazals et al. (2002), Simar (2003), Florens and Simar (2005), Daouia

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

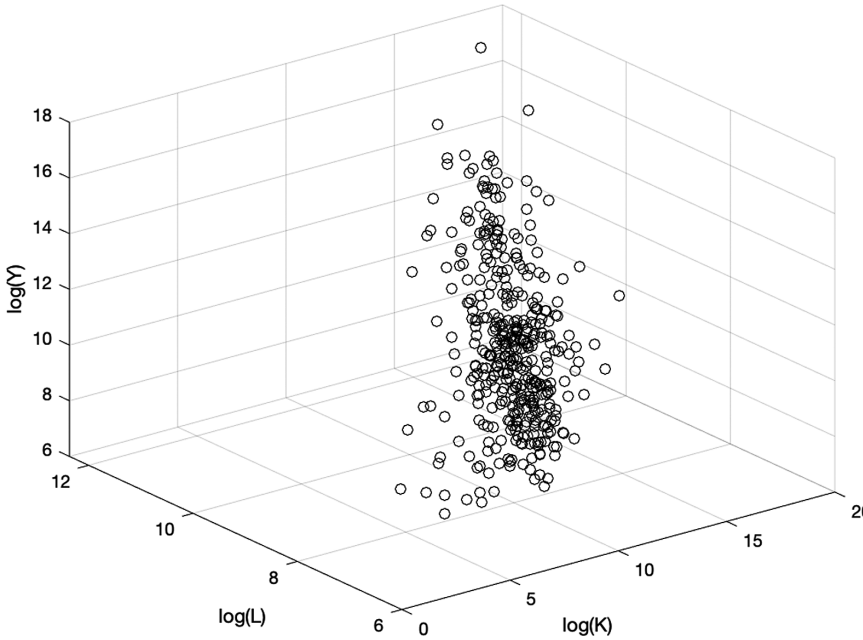


FIGURE 7. The capital K (in log), the average number of employees L (in log) and the value-added real output Y (in log).

and Simar (2005), Daouia and Ruiz-Gazen (2006), Daouia and Gijbels (2011), and Noh (2014).

Example 1. *We first consider a situation where the upper extremity of the joint support of* (X,Y) *is linear. We choose* (X,Y) *uniformly distributed over the triangle* $\{(x,y)\in[0,1]^2 : y\le x\}$. *Here, the true full frontier function is* $\varphi(x)=x$, *and the conditional distribution function is* $F_{Y|X}(y|x)=2x^{-1}y-x^{-2}y^2$, *for* $0<x\le 1$ *and* $0\le y\le\varphi(x)$. *The partial conditional order-m frontier function is*

$$\psi_m(x)=\varphi(x)-\sum_{k=0}^{m}\binom{m}{k}2^{m-k}(-1)^k x/(m+k+1).$$

Its unconditional analogue for the same order m is given by

$$\varphi_m(x)=\varphi(x)-\sum_{k=0}^{m}\binom{m}{k}(-1)^k x^{2k+1}/(2k+1).$$

Example 2. *We now consider a more realistic example from the point of view of production econometrics. We choose a nonlinear production frontier given by the Cobb–Douglas model* $Y=X^{1/2}\exp(-U)$, *where X is uniform on* $[0,1]$ *and U, independent of X, is exponential with mean* $1/3$. *Here, the full production function*

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

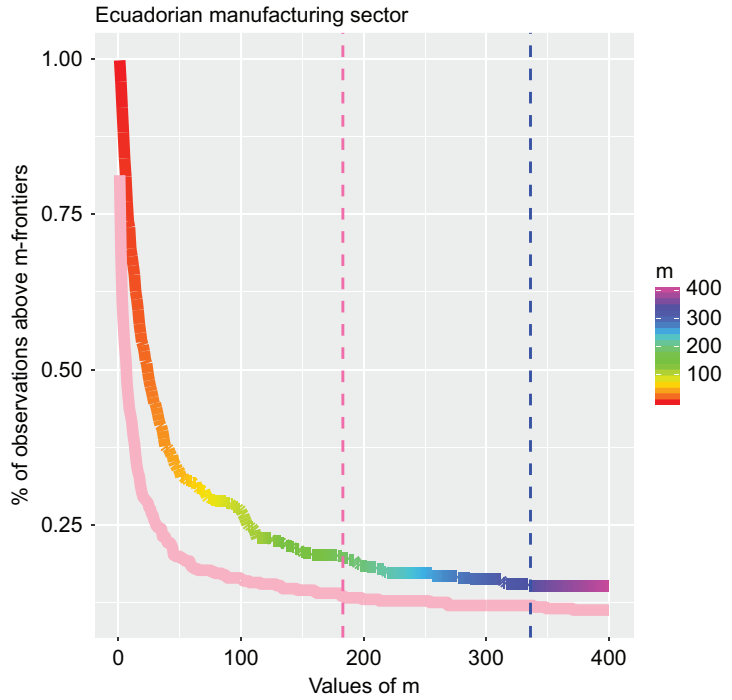


FIGURE 8. Evolution of the % of sample points outside the partial m-frontiers $\widehat{\psi}_m$ (pink) and $\widehat{\varphi}_m$ (rainbow), see the online text for the coloured graphics. The vertical dashed lines correspond to $m = 183$ (pink) and $m = 336$ (blue).

is $\varphi(x) = x^{1/2}$*, and the conditional distribution function is* $F_{Y|X}(y|x) = 3x^{-1}y^2 - 2x^{-3/2}y^3$*, for* $0 < x \leq 1$ *and* $0 \leq y \leq \varphi(x)$*. The partial order-m frontier functions have the following closed form expressions:*

$$\psi_m(x) = \varphi(x) - \sum_{k=0}^{m} \binom{m}{k} 3^{m-k}(-2)^k \sqrt{x}/(2m+k+1),$$

$$\varphi_m(x) = \varphi(x) - \sum_{k=0}^{m} \binom{m}{k} x^{k+1/2}(-1)^k \sum_{j=0}^{k} \binom{k}{j} \sum_{i=0}^{j} \binom{j}{i} (-3)^{j-i} 2^i/(2j+i+1).$$

4.1. Comparison of Population m−Frontiers

For both examples, the graphs of ψ_m and φ_m are superimposed in Figures 10 and 11, for three values of $m = 1, 10, 25$, along with the true support boundary φ. First, it may be seen from the plots that the conditional m-frontiers $\psi_m(x)$ (dotted curves) diverge from the true frontier $\varphi(x)$ (solid curve) as x increases. Whereas the new

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

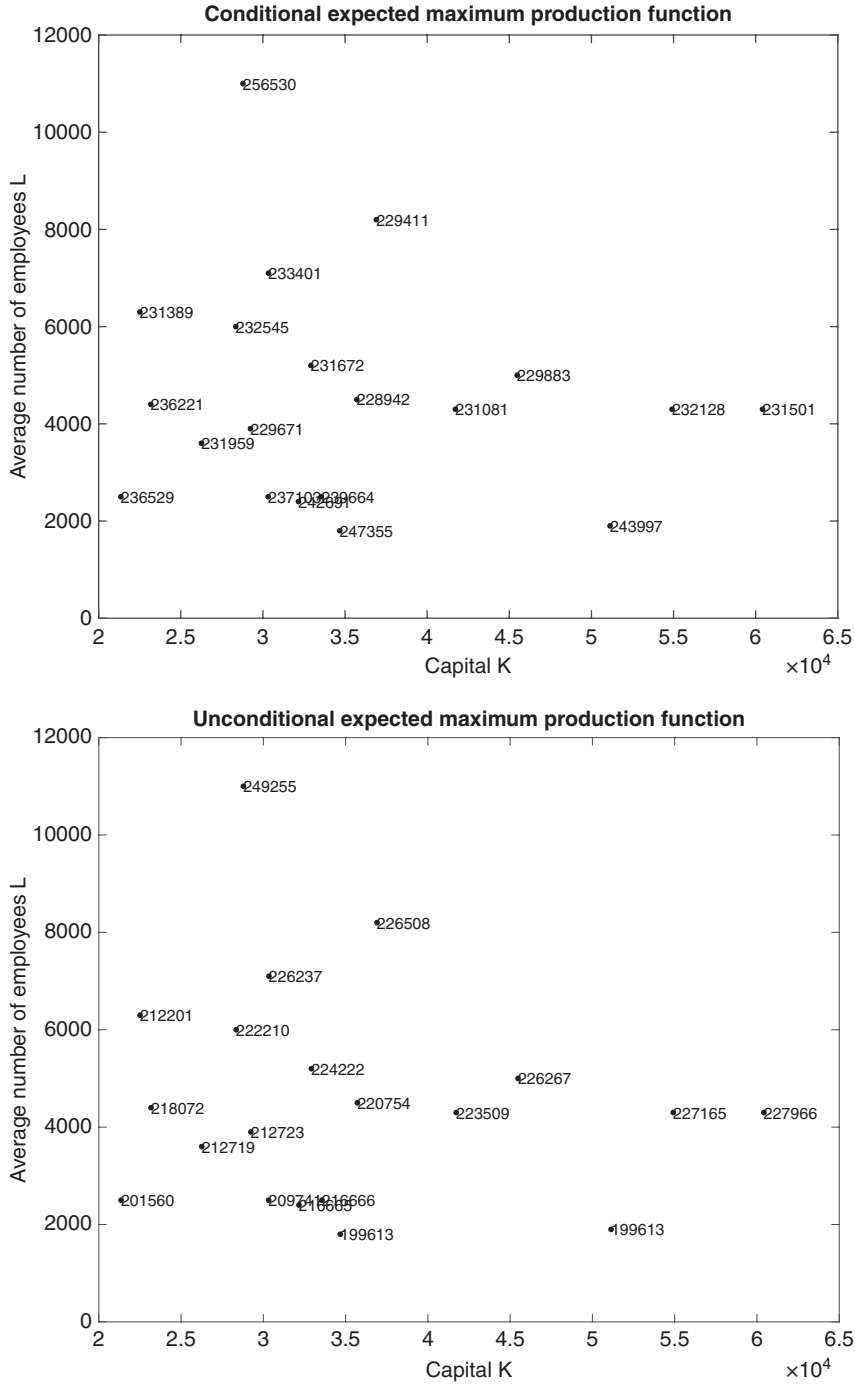


FIGURE 9. In both pictures, the points indicate 20 randomly chosen grid inputs $x_i = (K_i, L_i)$. The number associated to each point x_i indicates the conditional expected maximum $\widehat{\psi}_{183}(x_i)$ (top) and the unconditional expected maximum $\widehat{\varphi}_{336}(x_i)$ (bottom).

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

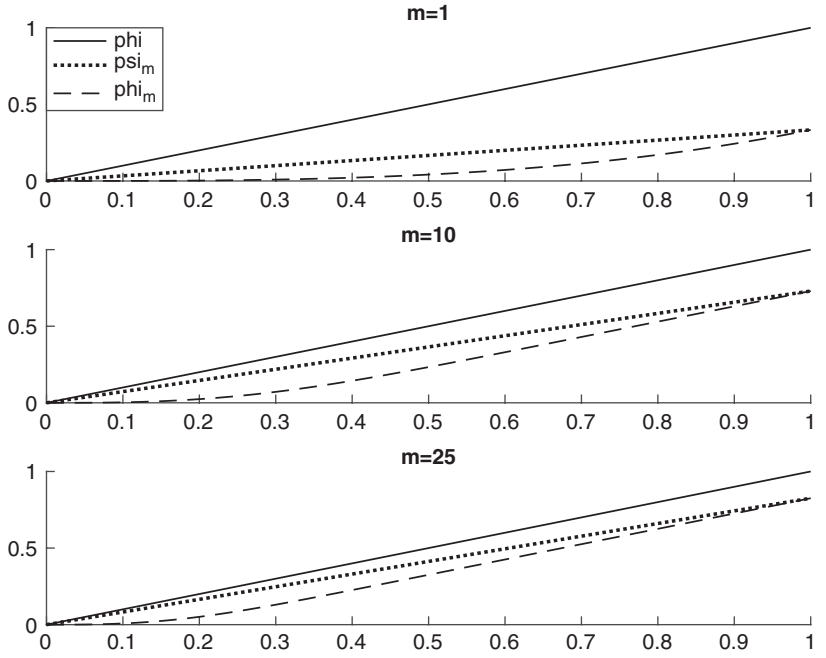


FIGURE 10. Uniform triangle example—graphs of φ in solid line, ψ_m in dotted line, and φ_m in dashed line.

unconditional m-frontiers $\varphi_m(x)$ (dashed curves) tend to be more parallel to the full frontier $\varphi(x)$. Second, the partial conditional m-frontiers approach rapidly the full frontier as m increases, while the convergence of the unconditional m-frontiers seems to be slower. Already these substantial differences indicate the usefulness of the new concept of unconditional expected maximum production m-frontiers.

Moreover, the new unconditional m-frontier φ_m can be viewed as a "robustified" alternative to the original conditional m-frontier ψ_m, for each trimming level m. This is visualised in Figures 12 and 13, where the gross-error sensitivities $\gamma^*\left(T^{m,x}, F_{(X,Y)}\right)$ of $\psi_m(x)$ and $\lambda^*\left(T_m, F_{Y^x}\right)$ of $\varphi_m(x)$ are plotted against m, for various values of $x \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. According to Hampel et al. (1986, p. 43), the most important quantitative robustness requirement is a low gross-error sensitivity. From this basis, it is clear that the new class of unconditional m-frontiers affords more reliability since the corresponding gross-error sensitivity λ^* (dashed line) is overall smaller than γ^* (solid line). Of interest is the limit case $m \nearrow \infty$, where γ^* explodes especially for low inputs-usage x, whereas λ^* remains appreciably small and stable whatever the value of x. This indicates that the sequence of empirical unconditional m-frontiers $\{\widehat{\varphi_m}(x)\}_n$ is more resistant to extreme values and/or outliers than its conditional analogue $\{\widehat{\psi_m}(x)\}_n$ for estimating the true full frontier $\varphi(x) = \lim_{m\to\infty} \varphi_m(x) = \lim_{m\to\infty} \psi_m(x)$. The lack of robustness of $\{\widehat{\psi_m}(x)\}_n$, for small values of x, is due to its construction via the conditioning by $X \leq x$.

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

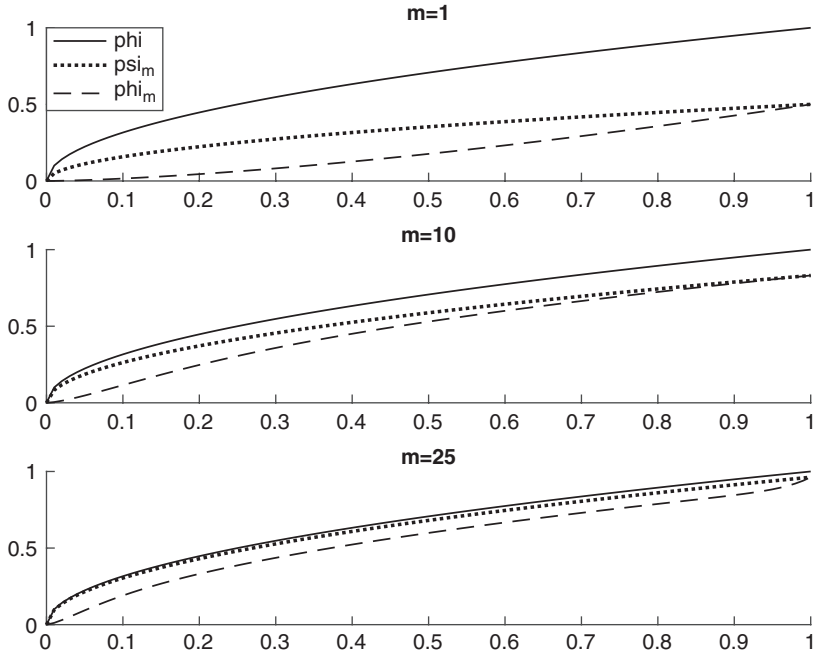


FIGURE 11. Cobb–Douglas example—same graphs as before.

4.2. Biased Frontier Estimators

To evaluate finite-sample performance of $\widehat{\psi}_m(\cdot)$ and $\widehat{\varphi}_m(\cdot)$, as robust estimators of $\varphi(\cdot)$, we have undertaken some simulation experiments. All the experiments were performed over 1,000 simulations for the sample sizes $n = 100, 500, 1,000$. Three outliers were added in each simulated dataset: $\{(0.1, 0.6), (0.35, 0.8), (0.6, 1)\}$ for both uniform-triangle and Cobb–Douglas examples. The measures of efficiency for each simulation used were the mean squared error and the bias

$$\mathrm{MSE}\{\widehat{\psi}_m\} = \frac{1}{L}\sum_{\ell=1}^{L}\left\{\widehat{\psi}_m(x_\ell) - \varphi(x_\ell)\right\}^2, \qquad \mathrm{Bias}\{\widehat{\psi}_m\} = \frac{1}{L}\sum_{\ell=1}^{L}\left\{\widehat{\psi}_m(x_\ell) - \varphi(x_\ell)\right\}$$

$$\mathrm{MSE}\{\widehat{\varphi}_m\} = \frac{1}{L}\sum_{\ell=1}^{L}\{\widehat{\varphi}_m(x_\ell) - \varphi(x_\ell)\}^2, \qquad \mathrm{Bias}\{\widehat{\varphi}_m\} = \frac{1}{L}\sum_{\ell=1}^{L}\{\widehat{\varphi}_m(x_\ell) - \varphi(x_\ell)\}$$

with the x_ℓ's being $L = 100$ points regularly distributed in $[\wedge X_i, \vee X_i]$. To guarantee a fair comparison among the two rival estimation methods, we used for each estimator the optimal parameter m minimizing its MSE over the wide range $\{1, \ldots, n\}$. The resulting values of MSE and bias are averaged on the 1,000 Monte Carlo replications and reported in Tables 1 and 2, along with the average $\overline{m}$ of the optimal 1,000 trimming levels m. The obtained estimates provide Monte Carlo

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

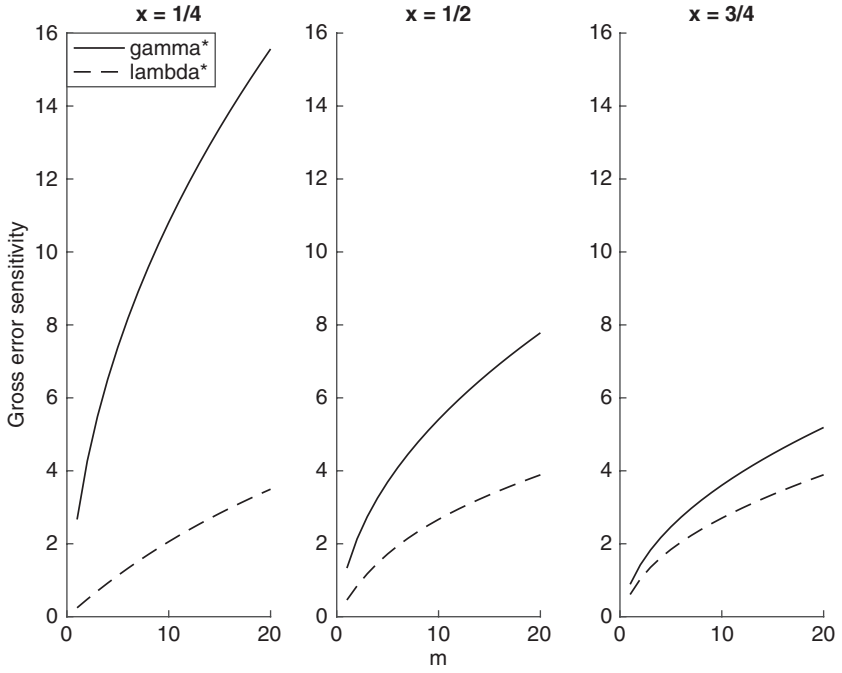


FIGURE 12. Uniform triangle example—gross-error sensitivities $m \mapsto \gamma^*\left(T^{m,x}, F_{(X,Y)}\right)$ in solid line and $m \mapsto \lambda^*\left(T_m, F_{Y^x}\right)$ in dashed line.

evidence that the new class of partial m-frontiers $\{\widehat{\varphi}_m\}_m$ is more efficient and robust relative to $\{\widehat{\psi}_m\}_m$ for estimating φ. A typical realization of the experiment in each simulated scenario with $n = 100$ is shown in Figure 14, where the optimal parameter m of each frontier estimator was chosen in such a way to minimize its MSE.

4.3. Bias-Corrected Frontier Estimators

This section provides Monte Carlo evidence on the usefulness of the proposed "unconditional" expected maximum output frontiers relative to their "conditional" competitors in terms of average lengths and achieved coverages of the corresponding asymptotic confidence intervals. More specifically, Theorem 4 indicates that $\widehat{\varphi}_m(x)$ estimates $\varphi(x)$ itself with the inherent bias $B_m(x) = \varphi(x) - \varphi_m(x)$ such that

$$\frac{\sqrt{n}}{\sigma(x,m)}\{\widehat{\varphi}_m(x) - \varphi(x) + B_m(x)\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad n \to \infty,$$

for a suitable choice of $m = m_n \to \infty$ as $n \to \infty$. In our experiments, we used the true value of the bias $B_m(x)$ and the empirical counterpart $\hat{\sigma}^2(x,m)$ of $\sigma^2(x,m)$. As

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

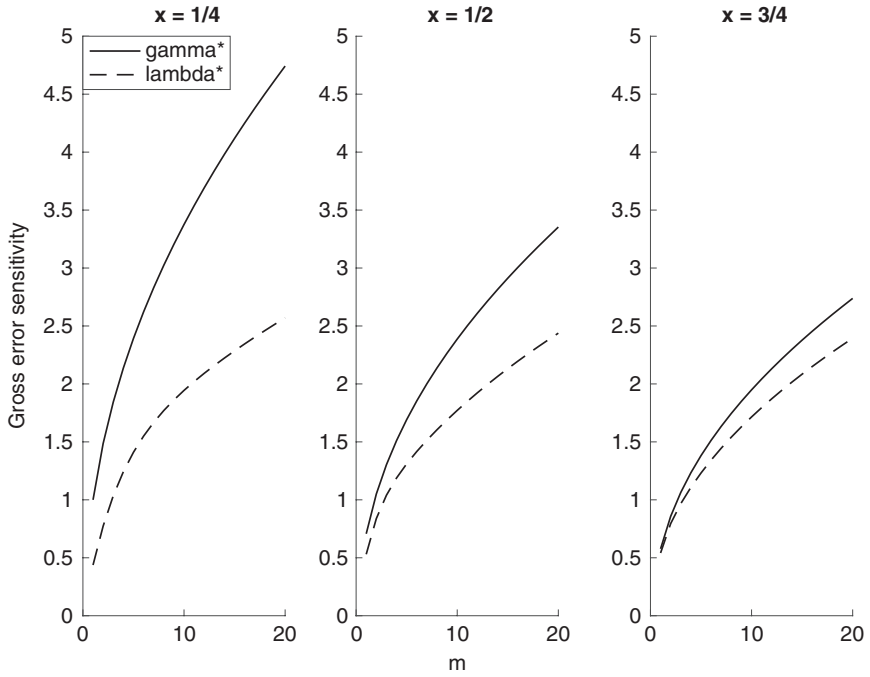


FIGURE 13. Cobb–Douglas example—gross-error sensitivities plots as before.

TABLE 1. *Uniform triangle example—results averaged on 1,000 simulations.*

	MSE		Bias		$\overline{m}$	
n	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$
100	0.0414	0.0031	0.0169	−0.0103	7.90	31.76
500	0.0240	0.0014	−0.0219	−0.0104	15.71	100.61
1,000	0.0175	0.0010	−0.0312	−0.0095	21.02	163.09

TABLE 2. *Cobb–Douglas example—results averaged on 1,000 simulations.*

	MSE		Bias		$\overline{m}$	
n	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$	$\{\widehat{\psi}_m\}$	$\{\widehat{\varphi}_m\}$
100	0.0050	0.0019	−0.0104	−0.0101	21.19	51.24
500	0.0023	0.0006	−0.0147	−0.0074	51.42	150.73
1,000	0.0016	0.0004	−0.0139	−0.0062	76.65	239.33

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

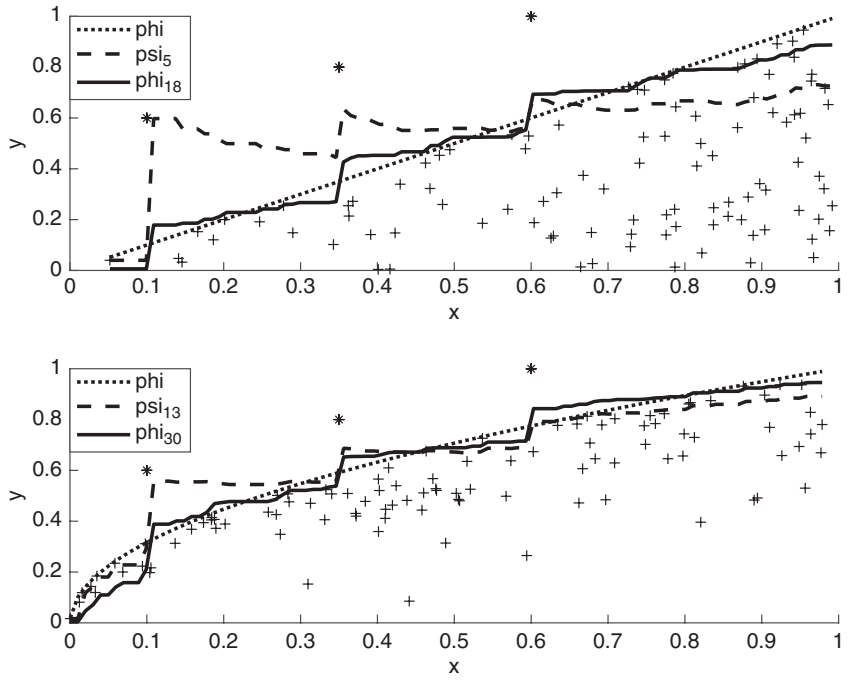


FIGURE 14. Typical realizations for simulated samples of size $n = 100$. Top—uniform triangle example. Bottom—Cobb–Douglas example. True frontier φ in dotted line with its optimal m-frontier estimators $\widehat{\psi}_m$ in dashed line and $\widehat{\varphi}_m$ in solid line.

for the conditional competitor $\widehat{\psi}_m$, we have by Thm. 3.1 in Daouia et al. (2012) that

$$\frac{\sqrt{n}}{s(x,m)}\left\{\widehat{\psi}_m(x)-\varphi(x)+b_m(x)\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad n\to\infty,$$

where $b_m(x)=\varphi(x)-\psi_m(x)$ and

$$s^2(x,m)=\frac{2m^2}{F_X(x)}\int_0^{\varphi(x)}\int_0^{\varphi(x)} F^m(y|x)F^{m-1}(u|x)[1-F(u|x)]\mathbb{1}(y\le u)dydu.$$

To guarantee a fair comparison with $\widehat{\varphi}_m(x)$, we used the true value of the bias $b_m(x)$ and the empirical counterpart $\hat{s}^2(x,m)$ of $s^2(x,m)$. For each *pseudo*-bias-corrected estimator

$$\widetilde{\varphi}_m(x):=\widehat{\varphi}_m(x)+B_m(x), \quad \widetilde{\psi}_m(x):=\widehat{\psi}_m(x)+b_m(x),$$

and each simulated sample, we used the optimal parameter m which minimizes the corresponding MSE over the range $\{1,\dots,\lfloor\sqrt{n}\rfloor\}$. The asymptotic confidence intervals with confidence level $100\alpha\%$ have the form

$$\mathrm{CI}_{\widetilde{\varphi}_m(x)}:=\left[\widetilde{\varphi}_m(x)\pm z_{(1+\alpha)/2}\frac{\hat{\sigma}(x,m)}{\sqrt{n}}\right], \quad \mathrm{CI}_{\widetilde{\psi}_m(x)}:=\left[\widetilde{\psi}_m(x)\pm z_{(1+\alpha)/2}\frac{\hat{s}(x,m)}{\sqrt{n}}\right],$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

TABLE 3. *Uniform triangle example—average lengths and coverages.*

$x = 0.25$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	0.2981	0.0248	**0.9930**	0.9940
500	0.0753	0.0064	0.9820	0.9730
1,000	0.0434	0.0039	**0.9590**	0.9630
$x = 0.50$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	0.1404	0.0508	0.8820	0.9460
500	0.0471	0.0178	0.9380	0.9550
1,000	0.0314	0.0122	0.9340	0.9510
$x = 0.75$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	0.1099	0.0828	0.8800	0.9010
500	0.0432	0.0330	0.9380	0.9450
1,000	0.0299	0.0230	0.9390	0.9400

TABLE 4. *Cobb–Douglas example—average lengths and coverages.*

$x = 0.25$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	0.1029	0.0539	0.9630	0.9630
500	0.0419	0.0245	0.9610	0.9550
1,000	0.0293	0.0176	0.9710	0.9510
$x = 0.50$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	0.0950	0.0870	0.9410	0.9570
500	0.0406	0.0382	0.9650	0.9400
1,000	0.0287	0.0269	0.9680	0.9670
$x = 0.75$				
n	$avl_{\widetilde{\psi}_m(x)}$	$avl_{\widetilde{\varphi}_m(x)}$	$cov_{\widetilde{\psi}_m(x)}$	$cov_{\widetilde{\varphi}_m(x)}$
100	**0.0935**	0.1031	0.9180	0.9430
500	**0.0404**	0.0449	**0.9590**	0.9600
1,000	**0.0285**	0.0316	**0.9750**	0.9780

with $z_{(1+\alpha)/2}$ being the $(1+\alpha)/2-$quantile of the standard Gaussian distribution. The average lengths and the achieved coverages of the 95% asymptotic confidence intervals $\mathrm{CI}_{\widetilde{\varphi}_m(x)}$ and $\mathrm{CI}_{\widetilde{\psi}_m(x)}$ are displayed in Tables 3 and 4, for $x \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. It may

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

be seen that the "unconditional" *pseudo*-bias-corrected estimator $\widetilde{\varphi}_m(x)$ globally performs better than the "conditional" variant $\widetilde{\psi}_m(x)$ in terms of both average lengths and achieved coverages. The few cases where $\widetilde{\psi}_m(x)$ is the winner are indicated in bold.

5. CONCLUSION

In this paper, we suggest a new approach to estimate nonparametrically and in a robust way the upper extremity of the joint support of a random vector $(X, Y) \in \mathbb{R}_+^p \times \mathbb{R}_+$. For a prespecified level of inputs-usage x interior to the marginal support of X, the basic idea is to first transform the $(p+1)$-dimensional vector (X, Y) into a dimensionless random variable $Y^x = Y\mathbb{1}(X \leq x)$, and then to define a concept of partial m-frontier $\varphi_m(x) = \mathbb{E}\big[\max(Y_1^x, \ldots, Y_m^x)\big]$ as the expected maximum of m independent copies of Y^x. In other words, we characterize $\varphi_m(x)$ as the expectation of the popular envelopment FDH estimator of the true full frontier $\varphi(x)$ based on the m-tuple of observations $Y_i^x = Y_i\mathbb{1}(X_i \leq x)$, $i = 1, \ldots, m$. We get robust estimators of the partial m-frontier functions φ_m as well as the full production function φ (corresponding to the limiting case $m \to \infty$). We derive their asymptotic distributions and robustness properties, and show their superiority over the pioneering class of *conditional* expected maximum production frontiers initiated by Cazals et al. (2002) and popularized by Daouia and Simar (2005), Florens and Simar (2005), Daouia and Ruiz-Gazen (2006), Daouia and Gijbels (2011), Daouia et al. (2012), to name a few. The merits and usefulness of our new class of *unconditional* expected maximum output frontiers are explored through two concrete datasets on delivery offices in the sector of postal services. The question of estimating both φ_m and φ in a stochastic frontier model, where the regression errors are assumed to be composite, is a topic of interest for future research.

REFERENCES

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

Aragon, Y., A. Daouia & C. Thomas-Agnan (2005) Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory* 21, 358–389.

Cazals, C., J.P. Florens & L. Simar (2002) Nonparametric frontier estimation: A robust approach. *Journal of Econometrics* 106, 1–25.

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

Daouia, A., J.P. Florens & L. Simar (2012) Regularization of non-parametric frontier estimators. *Journal of Econometrics* 168, 285–299.

Daouia, A., L. Gardes & S. Girard (2013) On kernel smoothing for extremal quantile regression. *Bernoulli* 19, 2557–2589.

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

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

Daouia, A., S. Girard & A. Guillou (2014) A Γ moment approach to monotonic boundary estimation. *Journal of Econometrics* 78, 727–740.

Daouia, A., H. Noh & B.U. Park (2016) Data envelope fitting with constrained polynomial splines. *Journal of the Royal Statistical Society: Series B* 78(Part 1), 3–30.

Daouia, A. & A. Ruiz-Gazen (2006) Robust nonparametric frontier estimators: Qualitative robustness and influence function. *Statistica Sinica* 16, 1233–1253.

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

Daouia, A. & L. Simar (2007) Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics* 140, 375–400.

Daouia, A., L. Simar & P. Wilson (2017) Measuring firm performance by using nonparametric quantile-type distances. *Econometric Reviews* 36, 156–181.

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

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

Florens, J.P. & L. Simar (2005) Parametric approximations of nonparametric frontier. *Journal of Econometrics* 124(1), 91–116.

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

Gijbels, I. & D. Sznajder (2013) Testing tail monotonicity by constrained copula estimation. *Insurance: Mathematics and Economics* 52(2), 338–351.

Greene, W.H. (2008) The econometric approach to efficiency analysis. In Harold Fried, C.A. Knox Lovell & Shelton Schmidt (eds.), *The Measurement of Productive Efficiency*, 2nd edn. Oxford University Press.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw & W.A. Stahel (1986) *Robust Statistics: The Approach Based on Influence Functions*. Wiley.

Kneip, A., L. Simar & I. Van Keilegom (2015) Frontier estimation in the presence of measurement error with unknown variance. *Journal of Econometrics* 184, 379–393.

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

Kumbhakar, S.C. , B.U. Park, L. Simar & E.G. Tsionas (2007) Nonparametric stochastic frontiers: A local likelihood approach. *Journal of Econometrics* 137(1), 1–27.

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

Noh, H. (2014) Frontier estimation using kernel smoothing estimators with data transformation. *Journal of the Korean Statistical Society* 43, 503–512.

Park, B. L. Simar & C. Weiner (2000) The FDH estimator for productivity efficiency scores: Asymptotic properties. *Econometric Theory* 16, 855–877.

Rousseeuw, P.J. (1981) A new infinitesimal approach to robust estimation. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 56, 127–132.

Serfling, R.T. (1980) *Approximation Theorems of Mathematical Statistics*. Wiley.

Simar, L. (2003) Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis* 20, 391–424.

Simar, L. & V. Zelenyuk (2011) Stochastic FDH/DEA estimators for frontier analysis. *Journal of Productivity Analysis* 36, 1–20.

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

van der Vaart, A.W. & J.A. Wellner (1996) *Weak Convergence and Empirical Processes. With Applications to Statistics. Springer Series in Statistics.* Springer-Verlag.

Wheelock, D.C. & P.W. Wilson (2008) Non-parametric, unconditional quantile estimation for efficiency analysis with an application to federal reserve check processing operations. *Journal of Econometrics* 145(1–2), 209–225.

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

Appendix: Proofs

Proof of Proposition 1. By definition (3), we have $\varphi_m(x) = \mathbb{E}(W_m)$, where $W_m = \max(Y_1^x, \ldots, Y_m^x)$. Hence $\varphi_m(x) = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}\{(W_m - \theta)^2\}$. On the other hand, it is easily seen that

$$\mathbb{E}\{(W_m - \theta)^2\} = \mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1} \cdot (Y^x - \theta)^2\}.$$

Therefore, $\varphi_m(x) = \arg\min_{\theta \in \mathbb{R}} \mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1} \cdot (Y^x - \theta)^2\}$. The first-order necessary condition for the optimality leads to the solution

$$\varphi_m(x) = \mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1} \cdot Y^x\} / \mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1}\} = \mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1} \cdot Y^x\}.$$

The last equality follows from the fact that $\mathbb{E}\{m[F_{Y^x}(Y^x)]^{m-1}\} = 1$. ■

To prove Proposition 2 and Theorem 1, the basic arguments go as those of the proof of proposition 4.1 in Daouia and Gijbels (2011). Fix $m \geq 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$. Define the domain $\mathbb{D}_x$ to be the set of joint distribution functions $G(\cdot, \cdot)$ on $\mathbb{R}_+^p \times \mathbb{R}_+$ such that

$$G(x, \infty) > 0 \quad \text{and} \quad G_x^{-1}(1) \leq \varphi(x), \tag{A.1}$$

where $G_x^{-1}(1) := \inf\{y \geq 0 |\ G_x(y) = 1\}$ stands for the right-endpoint of the support of the transformed distribution function

$$G_x(y) = \left\{1 - G(x, \infty)\left[1 - \frac{G(x, y)}{G(x, \infty)}\right]\right\} \mathbb{1}(y \geq 0). \tag{A.2}$$

For any $G \in \mathbb{D}_x$, define

$$\overset{m,x.}{\phi}(G) = \int_0^\infty \left(1 - [G_x(y)]^m\right) dy \equiv \int_0^{G_x^{-1}(1)} \left(1 - [G_x(y)]^m\right) dy.$$

It follows from (A.1) that $\overset{m,x.}{\phi}(G) = \int_0^{\varphi(x)} (1 - [G_x(y)]^m) dy$, for all $G \in \mathbb{D}_x$. Note also that $F_{(X,Y)} \in \mathbb{D}_x$, $\overset{m,x.}{\phi}(F_{(X,Y)}) = \int_0^{\varphi(x)} (1 - [F_{Y^x}(y)]^m) dy \equiv \varphi_m(x)$ and $\overset{m,x.}{\phi}(\widehat{F}_{(X,Y)}) = \int_0^{\widehat{\varphi}(x)} (1 - [\widehat{F}_{Y^x}(y)]^m) dy = \widehat{\varphi}_m(x) \overset{a.s.}{=} \int_0^{\varphi(x)} (1 - [\widehat{F}_{Y^x}(y)]^m) dy$, since $\widehat{\varphi}(x) \leq \varphi(x)$ with probability 1.

The following lemma will be useful for the proof of Proposition 2.

LEMMA A.1. *The map* $\overset{m,x.}{\phi} : \mathbb{D}_x \subset L^\infty(\overline{\mathbb{R}}^{p+1}) \longrightarrow [0, \varphi(x)]$ *is Hadamard-differentiable at* $F_{(X,Y)}$ *with derivative* $(\overset{m,x.}{\phi})'_{F_{(X,Y)}} : h \in L^\infty(\overline{\mathbb{R}}^{p+1}) \longmapsto (\overset{m,x.}{\phi})'_{F_{(X,Y)}}(h)$, *where*

$$(\overset{m,x.}{\phi})'_{F_{(X,Y)}}(h) = m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1} (h(x, \infty) - h(x, y)) dy.$$

Proof. Let $F := F_{(X,Y)}$, $h \in L^\infty(\overline{\mathbb{R}}^{p+1})$ and $h_t \to h$ uniformly in $L^\infty(\overline{\mathbb{R}}^{p+1})$, where $F + th_t \in \mathbb{D}_x$ for all small $t > 0$. Write $\varphi_{mt}(x) := \overset{m,x.}{\phi}(F + th_t)$. Following the definition of the

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

Hadamard differentiability (see van der Vaart, 1998, p. 296), we shall show that $(\varphi_{mt}(x) - \varphi_m(x))/t$ converges to $(\overset{m,x.}{\phi})'_F(h)$ as $t \downarrow 0$. We have

$$\varphi_{mt}(x) - \varphi_m(x) = \int_0^{\varphi(x)} \left([F_{Y^x}(y)]^m - [(F + th_t)_x(y)]^m\right) dy,$$

where $(F + th_t)_x(y)$ is described in (A.2) for $G = F + th_t$. By Taylor's formula, for any $y \in [0, \varphi(x)]$, there exists a point $\zeta_{t,x}(y)$ interior to the interval joining $F_{Y^x}(y)$ and $(F + th_t)_x(y)$ such that

$$[F_{Y^x}(y)]^m - [(F + th_t)_x(y)]^m = mt \left[\zeta_{t,x}(y)\right]^{m-1} (h_t(x, \infty) - h_t(x, y)),$$

whence

$$\frac{\varphi_{mt}(x) - \varphi_m(x)}{t} = m \int_0^{\varphi(x)} \left[\zeta_{t,x}(y)\right]^{m-1} (h_t(x, \infty) - h_t(x, y))\, dy. \tag{A.3}$$

It follows from the definition of $\zeta_{t,x}(y)$ and the uniform convergence $h_t \to h$ in $L^\infty(\overline{\mathbb{R}}^{p+1})$ that $\left[\zeta_{t,x}(y)\right]^{m-1} (h_t(x, \infty) - h_t(x, y))$ converges to $[F_{Y^x}(y)]^{m-1} (h(x, \infty) - h(x, y))$ uniformly in y as $t \downarrow 0$. Therefore, we obtain that $(\varphi_{mt}(x) - \varphi_m(x))/t \to (\overset{m,x.}{\phi})'_F(h)$ as $t \downarrow 0$. ■

Proof of Proposition 2. It is well known that the empirical process $\sqrt{n}(\widehat{F}_{(X,Y)} - F_{(X,Y)})$ converges in distribution in $L^\infty(\overline{\mathbb{R}}^{p+1})$ to $\mathbb{F}$, a $p+1$ dimensional $F_{(X,Y)}$-Brownian bridge (see van der Vaart and Wellner, 1996, p. 82). $\mathbb{F}$ is a Gaussian process with zero mean and covariance function $\mathbb{E}[\mathbb{F}(t_1)\mathbb{F}(t_2)] = F_{(X,Y)}(t_1 \wedge t_2) - F_{(X,Y)}(t_1)F_{(X,Y)}(t_2)$, for all $t_1, t_2 \in \overline{\mathbb{R}}^{p+1}$. Then, by applying the functional delta method (see Thm. 20.8 in van der Vaart, 1998, p. 297] in conjunction with Lemma A.1, we obtain that

$$\begin{aligned}\sqrt{n}(\widehat{\varphi}_m(x) - \varphi_m(x)) &\equiv \sqrt{n}\left(\overset{m,x.}{\phi}\left(\widehat{F}_{(X,Y)}\right) - \overset{m,x.}{\phi}\left(F_{(X,Y)}\right)\right) \\ &= (\overset{m,x.}{\phi})'_{F_{(X,Y)}}\left(\sqrt{n}\left(\widehat{F}_{(X,Y)} - F_{(X,Y)}\right)\right) + o_p(1),\end{aligned}$$

where

$$\begin{aligned}&(\overset{m,x.}{\phi})'_{F_{(X,Y)}}\left(\sqrt{n}\left(\widehat{F}_{(X,Y)} - F_{(X,Y)}\right)\right) \\ &= m\sqrt{n}\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\left\{\left(\widehat{F}_{(X,Y)}(x, \infty) - F_{(X,Y)}(x, \infty)\right) - \left(\widehat{F}_{(X,Y)}(x, y) - F_{(X,Y)}(x, y)\right)\right\} dy \\ &= m\sqrt{n}\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\left\{F_{Y^x}(y) - \widehat{F}_{Y^x}(y)\right\} dy.\end{aligned}$$

This ends the proof. ■

Let us now consider $\sqrt{n}(\widehat{\varphi}_m(x) - \varphi_m(x))$ as a process indexed by x in an arbitrarily fixed set $\mathcal{X}$ such that $\inf_{x \in \mathcal{X}} F_X(x) > 0$. As before, $m \geq 1$ is fixed. Define the domain $\mathbb{D}_{\mathcal{X}}$ to be the set of joint distribution functions G on $\mathbb{R}_+^{p+1}$ such that $G \in \mathbb{D}_x$ for all $x \in \mathcal{X}$. Let $\tau_Y < \infty$ be the right-endpoint of the support of Y and define, for any $G \in \mathbb{D}_{\mathcal{X}}$, the map

$$\overset{m}{\phi}(G) : x \mapsto \overset{m,x.}{\phi}(G) \quad \text{as a map} \quad \mathcal{X} \longrightarrow [0, \tau_Y].$$

Here, the functional $\overset{m}{\phi}: G \mapsto \overset{m}{\phi}(G)$ is defined as a map $\mathbb{D}_{\mathcal{X}} \subset L^\infty(\overline{\mathbb{R}}^{p+1}) \to L^\infty(\mathcal{X})$. Note that $\overset{m}{\phi}(F_{(X,Y)}) := \{\overset{m,x.}{\phi}(F_{(X,Y)}); x \in \mathcal{X}\} = \{\varphi_m(x); x \in \mathcal{X}\} = \{\int_0^{\varphi(x)} (1 - [F_{Y^x}(y)]^m) dy;$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

$x \in \mathcal{X}\}$ and $\overset{m}{\phi}(\widehat{F}_{(X,Y)}) := \{\overset{m,x.}{\phi}(\widehat{F}_{(X,Y)}); x \in \mathcal{X}\} = \{\widehat{\varphi}_m(x); x \in \mathcal{X}\} \overset{a.s.}{=} \{\int_0^{\varphi(x)} (1-[\widehat{F}_{Y^x}(y)]^m)dy; x \in \mathcal{X}\}$ since $\mathbb{P}[\widehat{\varphi}(x) \le \varphi(x), \forall x \in \mathcal{X}] = 1$. The following lemma will be useful for the proof of Theorem 1.

LEMMA A.2. *$\overset{m}{\phi}$ is Hadamard-differentiable at $F_{(X,Y)} \in \mathbb{D}_{\mathcal{X}}$ with derivative*

$$(\overset{m}{\phi})'_{F_{(X,Y)}}(h) : x \in \mathcal{X} \mapsto (\overset{m,x.}{\phi})'_{F_{(X,Y)}}(h),$$

for any $h \in L^\infty(\overline{\mathbb{R}}^{p+1})$.

Proof. The basic idea is to make the proof of Lemma A.1 uniform in $x \in \mathcal{X}$ by using the same notation. Here, $h_t \to h$ in $L^\infty(\overline{\mathbb{R}}^{p+1})$ as $t \downarrow 0$, with $F + th_t \in D_{\mathcal{X}}$ for all small t. To establish the Hadamard differentiability (see van der Vaart, 1998, p. 296), we have to show that $\sup_{x\in\mathcal{X}} |(\varphi_{mt}(x) - \varphi_m(x))/t - (\overset{m,x.}{\phi})'_F(h)| \to 0$ as $t \downarrow 0$.

By the uniform convergence of h_t to h and the definition of $\zeta_{t,x}(y)$ described in Lemma A.1, we have $\sup_{x\in\mathcal{X}, y\in\overline{\mathbb{R}}} |\zeta_{t,x}^{m-1}(y) - F_{Y^x}^{m-1}(y)| \to 0$ as $t \downarrow 0$. By using this and applying again the uniform convergence of h_t, it is easily seen that $\left[\zeta_{t,x}(y)\right]^{m-1}(h_t(x,\infty) - h_t(x,y))$ converges to $[F_{Y^x}(y)]^{m-1}(h(x,\infty) - h(x,y))$ uniformly in (x,y) as $t \downarrow 0$. Finally, since $\sup_{x\in\mathcal{X}} \varphi(x) \le \tau_Y$, we get $\sup_{x\in\mathcal{X}} |(\varphi_{mt}(x) - \varphi_m(x))/t - (\overset{m,x.}{\phi})'_F(h)| \to 0$ as $t \downarrow 0$. ■

Proof of Theorem 1. By applying the functional delta method in conjunction with Lemma A.2, we get the convergence in distribution of $\sqrt{n}\{\overset{m}{\phi}(\widehat{F}_{(X,Y)}) - \overset{m}{\phi}(F_{(X,Y)})\}$ in $L^\infty(\mathcal{X})$ to the linear transformation $(\overset{m}{\phi})'_{F_{(X,Y)}}(\mathbb{F})$ of the Gaussian process $\mathbb{F}$ described in the proof of Proposition 2. Note that the linear operator $(\overset{m}{\phi})'_{F_{(X,Y)}}(\cdot)$ is defined and continuous on the space $L^\infty(\overline{\mathbb{R}}^{p+1})$ since

$$||(\overset{m}{\phi})'_{F_{(X,Y)}}(h)||_{L^\infty(\mathcal{X})} = \sup_{x\in\mathcal{X}} |(\overset{m,x.}{\phi})'_{F_{(X,Y)}}(h)| \le 2m\tau_Y ||h||_{L^\infty(\overline{\mathbb{R}}^{p+1})}$$

for any $h \in L^\infty(\overline{\mathbb{R}}^{p+1})$. Therefore, according to Thm. 20.8 in van der Vaart (1998, p. 297), we have that $\sqrt{n}\{\overset{m}{\phi}(\widehat{F}_{(X,Y)}) - \overset{m}{\phi}(F_{(X,Y)})\} = (\overset{m}{\phi})'_{F_{(X,Y)}}(\sqrt{n}(\widehat{F}_{(X,Y)} - F_{(X,Y)})) + o_p(1)$. ■

Proof of Theorem 2. By Taylor's formula, for any $y \in [0, \varphi(x)]$, there exists a point $\eta_{x,n}(y)$ interior to the interval joining $F_{Y^x}(y)$ and $\widehat{F}_{Y^x}(y)$ such that $[\widehat{F}_{Y^x}(y)]^m - [F_{Y^x}(y)]^m = m[F_{Y^x}(y)]^{m-1}\{\widehat{F}_{Y^x}(y) - F_{Y^x}(y)\} + (m/2)(m-1)[\eta_{x,n}(y)]^{m-2}\{\widehat{F}_{Y^x}(y) - F_{Y^x}(y)\}^2$. By using the fact that $\{\widehat{\varphi}_m(x) - \varphi_m(x)\} \overset{a.s.}{=} \int_0^{\varphi(x)} \left([F_{Y^x}(y)]^m - [\widehat{F}_{Y^x}(y)]^m\right)dy$, we get

$$\begin{aligned}\{\widehat{\varphi}_m(x) - \varphi_m(x)\} - m\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\{F_{Y^x}(y) - \widehat{F}_{Y^x}(y)\}dy \qquad \textbf{(A.4)}\\ \overset{a.s.}{=} -(m/2)(m-1)\int_0^{\varphi(x)} [\eta_{x,n}(y)]^{m-2}\{\widehat{F}_{Y^x}(y) - F_{Y^x}(y)\}^2 dy.\end{aligned}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

On the other hand, we have by the law of the iterated logarithm (LIL) for empirical processes

$$\sup_y \left|\widehat{F}_{Y^x}(y) - F_{Y^x}(y)\right| = O\left((\log\log n/n)^{1/2}\right), \tag{A.5}$$

with probability 1. It follows that $\sup_y\{\sqrt{n}[\widehat{F}_{Y^x}(y) - F_{Y^x}(y)]^2\} \xrightarrow{a.s.} 0$ as $n \to \infty$. Finally, since $0 \le \eta_{x,n}(y) \le 1$ for all y, we arrive at

$$R_{m,n}(x) := \sqrt{n}\left(\left\{\widehat{\varphi}_m(x) - \varphi_m(x)\right\} - m\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\left\{F_{Y^x}(y) - \widehat{F}_{Y^x}(y)\right\}dy\right) \xrightarrow{a.s.} 0.$$

By applying again the classical LIL (see, e.g., Serfling, 1980, Thm. A, p. 35), we obtain

$$\limsup_{n\to\infty} \pm \frac{\sqrt{n}m}{(2\log\log n)^{1/2}} \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\left\{F_{Y^x}(y) - \widehat{F}_{Y^x}(y)\right\}dy = \sigma(x,m)$$

for either choice of sign, with probability 1. By combining this result with the fact that $R_{m,n}(x)/(2\log\log n)^{1/2} \xrightarrow{a.s.} 0$ as $n \to \infty$, we get the desired LIL. ■

Proof of Proposition 3. We have

$$\begin{aligned}\gamma^*\left(T^{m,x}, F_{(X,Y)}\right) &= \frac{m}{F_X(x)} \sup_{y_0 \ge 0} \left| \int_0^{\varphi(x)} F_{Y|X}^{m-1}(y|x)\left[\mathbb{1}(y_0 \le y) - F_{Y|X}(y|x)\right]dy \right| \\ &= \frac{m}{F_X(x)} \max\left\{ \int_0^{\varphi(x)} F_{Y|X}^{m}(y|x)dy, \sup_{0\le y_0 \le \varphi(x)} H(y_0) \right\},\end{aligned}$$

where $H(y_0) := \int_0^{y_0} F_{Y|X}^m(y|x)dy + \int_{y_0}^{\varphi(x)} F_{Y|X}^{m-1}(y|x)\left[1 - F_{Y|X}(y|x)\right]dy$. The function $H(\cdot)$ being convex and continuous on $[0, \varphi(x)]$, it achieves its supremum at $y_0 = 0$ or $y_0 = \varphi(x)$. The conclusion is then immediate. ■

Proof of Proposition 4. Putting $F_t = (1-t)F_{Y^x} + t\delta_u$ and $F_t^{-1}(1) = \inf\{y | F_t(y) = 1\}$, we have

$$\mathrm{IF}\left(u; T_m, F_{Y^x}\right) = \frac{d}{dt}_{|t=0} T_m(F_t) = \frac{d}{dt}_{|t=0} \int_0^{F_t^{-1}(1)} [1 - F_t^m(y)]dy.$$

Since $F_t(y) \to F_{Y^x}(y)$ as $t \to 0$ for every $y \in \mathbb{R}$, we obtain the weak convergence of the distribution functions $F_t \rightsquigarrow F_{Y^x}$, which in turn implies the weak convergence of the underlying quantile functions as $t \to 0$ in view of a van der Vaart's lemma (1998, Lem. 21.2, p. 305). In particular, $F_t^{-1}(1) \to F_{Y^x}^{-1}(1) \equiv \varphi(x)$ as $t \to 0$. Then for any $\nu > \varphi(x)$, we have $F_t^{-1}(1) < \nu$ as $t \to 0$. Therefore,

$$\mathrm{IF}\left(u; T_m, F_{Y^x}\right) = \frac{d}{dt}_{|t=0} \int_0^{\nu} [1 - F_t^m(y)]dy = -m\int_0^{\nu} [F_{Y^x}(y)]^{m-1}\left\{\delta_u(y) - F_{Y^x}(y)\right\}dy,$$

for any $\nu > \varphi(x)$. Taking the limit as $\nu \to \varphi(x)$ ends the proof. ■

Proof of Theorem 3. According to Daouia et al. (2010, Cor. 2.1), we have under $\boldsymbol{\mathcal{C}}(\boldsymbol{\rho_x}, \boldsymbol{\ell_x})$ that

$$(n\ell_x)^{1/\rho_x}\left\{\varphi(x) - \widehat{\varphi}(x)\right\} \xrightarrow{\mathcal{L}} \mathrm{Weibull}(1, \rho_x) \quad \text{as} \quad n \to \infty.$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

The basic idea of proof is then to consider the following decomposition

$$(n\ell_x)^{1/\rho_x}\{\varphi(x)-\widehat{\varphi}_m(x)\}=(n\ell_x)^{1/\rho_x}\{\varphi(x)-\widehat{\varphi}(x)\}+(n\ell_x)^{1/\rho_x}\{\widehat{\varphi}(x)-\widehat{\varphi}_m(x)\}$$

and show that the second term on the right-hand side $(n\ell_x)^{1/\rho_x}\{\widehat{\varphi}(x)-\widehat{\varphi}_m(x)\}=o_p(1)$ as $n\to\infty$. It follows from (5) that

$$\widehat{\varphi}(x)-\widehat{\varphi}_m(x)=Y^x_{(n)}-\widehat{\varphi}_m(x)=\sum_{i=1}^{n-1}(i/n)^m\{Y^x_{(i+1)}-Y^x_{(i)}\}.$$

The support of Y^x being bounded (included in $[0,\varphi(x)]$), we have with probability 1 that

$$\widehat{\varphi}(x)-\widehat{\varphi}_m(x)=O\left(\sum_{i=1}^{n-1}(i/n)^m\right)=O\left(n\left[1-\frac{1}{n}\right]^m\right).$$

Hence, for the term $(n\ell_x)^{1/\rho_x}\{\widehat{\varphi}(x)-\widehat{\varphi}_m(x)\}$ to be $o_p(1)$, it is sufficient to choose $m=m_n$ such that $n^{\frac{1}{\rho_x}+1}\left[1-\frac{1}{n}\right]^{m_n}\to 0$ as $n\to\infty$. To achieve this, it suffices to have $\left[1-\frac{1}{n}\right]^{m_n}=O(n^{-\beta})$, or equivalently, $\left[1-\frac{1}{n}\right]^{m_n}\le(Cn)^{-\beta}$ for some constants $\beta>\frac{1}{\rho_x}+1$ and $C>0$. This condition reduces to $m_n\ge\beta n\log n\{1+o(1)\}$ by using the fact that $\log(1-1/n)\sim -1/n$ as $n\to\infty$. ■

Proof of Theorem 4. Here we employ similar arguments of proof as in Thm. 3.1 and Lem. 3.1 of Daouia et al. (2012). We know by (A.4) that

$$\frac{\sqrt{n}}{\sigma(x,m)}\{\widehat{\varphi}_m(x)-\varphi_m(x)\}\overset{a.s.}{=}\frac{\sqrt{n}}{\sigma(x,m)}m\int_0^{\varphi(x)}[F_{Y^x}(y)]^{m-1}\{F_{Y^x}(y)-\widehat{F}_{Y^x}(y)\}dy$$
$$-\frac{\sqrt{n}}{\sigma(x,m)}(m/2)(m-1)\int_0^{\varphi(x)}[\eta_{x,n}(y)]^{m-2}\{\widehat{F}_{Y^x}(y)-F_{Y^x}(y)\}^2dy,$$

and that $\sup_y\left|\widehat{F}_{Y^x}(y)-F_{Y^x}(y)\right|\overset{a.s.}{=}O\left((\log\log n/n)^{1/2}\right)$ in view of (A.5). For any $y\in]0,\varphi(x)[$ we have $0<\eta_{x,n}(y)<1$ and so $[\eta_{x,n}(y)]^{m-2}\overset{a.s.}{\to}0$ when $n\to\infty$. Hence, using the dominated convergence theorem, we have $\int_0^{\varphi(x)}[\eta_{x,n}(y)]^{m-2}dy\overset{a.s.}{\to}0$. Since $\sqrt{n}m(m-1)/\sigma(x,m)=O(n/\log\log n)$, we get

$$\frac{\sqrt{n}}{\sigma(x,m)}\frac{m}{2}(m-1)\int_0^{\varphi(x)}[\eta_{x,n}(y)]^{m-2}\{\widehat{F}_{Y^x}(y)$$
$$-F_{Y^x}(y)\}^2dy\overset{a.s.}{=}O(1)\int_0^{\varphi(x)}[\eta_{x,n}(y)]^{m-2}dy\xrightarrow{a.s.}0.$$

On the other hand,

$$\frac{\sqrt{n}}{\sigma(x,m)}m\int_0^{\varphi(x)}[F_{Y^x}(y)]^{m-1}\{F_{Y^x}(y)-\widehat{F}_{Y^x}(y)\}dy=\sum_{i=1}^n\frac{Z_{n,i}}{\sqrt{n}\sigma(Z_{n,i})},\tag{A.6}$$

where $Z_{n,i}=m\int_0^{\varphi(x)}[F_{Y^x}(y)]^{m-1}\{F_{Y^x}(y)-\mathbb{1}(Y^x_i\le y)\}dy$ and its variance is $\sigma^2(Z_{n,i})=\sigma^2(x,m)$. We have $n\mathbb{E}[|Z_{n,1}|^3]/\{n\sigma^2(Z_{n,1})\}^{3/2}\le m\varphi(x)/\sqrt{n}\sigma(Z_{n,1})\to 0$ since $m/\sqrt{n}\sigma$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

$(x,m) \to 0$. Hence the leading term (A.6) converges in distribution to $\mathcal{N}(0,1)$ by Lyapounov's Theorem. Therefore, $\sqrt{n}\sigma^{-1}(x,m)\{\widehat{\varphi}_m(x) - \varphi_m(x)\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

In what concerns the second assertion, it is easily seen that

$$\begin{aligned}\sigma^2(x,m) &= 2m^2 \int_0^{\varphi(x)} \int_0^{\varphi(x)} F_{Y^x}^m(y) F_{Y^x}^{m-1}(z)[1 - F_{Y^x}(z)] \mathbb{1}(y \le z) dy dz \\ &= 2m^2 \int_0^{\varphi(x)} F_{Y^x}^{m-1}(z) S_{Y^x}(z) \left(\int_0^z F_{Y^x}^m(y) dy \right) dz. \end{aligned} \tag{A.7}$$

Then, for all $\delta > 0$ sufficiently small, we have

$$\begin{aligned}\sigma^2(x,m) &\ge 2m^2 \int_{\varphi(x)-\delta}^{\varphi(x)} F_{Y^x}^{m-1}(z) S_{Y^x}(z) \left(\int_{z-\delta}^z F_{Y^x}^m(y) dy \right) dz \\ &\ge 2m^2 \delta \int_{\varphi(x)-\delta}^{\varphi(x)} F_{Y^x}^{m-1}(z) S_{Y^x}(z) F_{Y^x}^m(z-\delta) dz \\ &\ge 2m^2 \delta \int_{\varphi(x)-\delta}^{\varphi(x)} F_{Y^x}^{2m}(z-\delta) S_{Y^x}(z) dz \\ &\ge 2m^2 \delta F_{Y^x}^{2m}(\varphi(x) - 2\delta) \int_{\varphi(x)-\delta}^{\varphi(x)} S_{Y^x}(z) dz.\end{aligned}$$

It follows from the regularity condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}}, \boldsymbol{\ell}_{\mathbf{x}})$ that

$$\sigma^2(x,m) \ge m^2 \delta F_{Y^x}^{2m}(\varphi(x) - 2\delta) \ell_x \delta^{\rho_x + 1} / (\rho_x + 1), \quad \delta \to 0.$$

We also have by $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}}, \boldsymbol{\ell}_{\mathbf{x}})$ that

$$\begin{aligned}F_{Y^x}^{2m}(\varphi(x) - 2\delta) &\ge \left\{1 - 2\ell_x (2\delta)^{\rho_x}\right\}^{2m} \\ &= \exp\left[2m \log\left\{1 - 2\ell_x (2\delta)^{\rho_x}\right\}\right] \ge e^{-8m\ell_x (2\delta)^{\rho_x}}, \quad \delta \to 0.\end{aligned}$$

Thus, for $\delta = (1/m)^{1/\rho_x}$, we get

$$\sigma^2(x,m) \ge m^2 \delta^{\rho_x + 2} e^{-8m\ell_x (2\delta)^{\rho_x}} \ell_x / (\rho_x + 1) \ge c_x m^{1 - 2/\rho_x}, \quad m \to \infty,$$

for some constant $c_x > 0$. Whence

$$m(m-1)/\sigma(x,m) \le c_x^{-1/2} m(m-1) m^{-\frac{1}{2} + \frac{1}{\rho_x}}, \quad m \to \infty.$$

Hence, if $m = O\left(\sqrt{n}/\log\log n\right)^{\frac{1}{\frac{3}{2} + \frac{1}{\rho_x}}}$, it is immediate that $\frac{m(m-1)}{\sigma(x,m)} = O\left(\sqrt{n}/\log\log n\right)$ as $n \to \infty$, and so the asymptotic normality holds.

It remains to show that $\sigma^2(x,m) \le \tilde{c}_x m^{2-2/\rho_x}$ as $m \to \infty$, for some positive constant $\tilde{c}_x$. It follows from (A.7), in conjunction with the identity $\varphi_m(x) = \int_0^\infty \left(1 - [F_{Y^x}(y)]^m\right) dy = \varphi(x) - \int_0^{\varphi(x)} [F_{Y^x}(y)]^m dy$, that

$$\begin{aligned}\sigma^2(x,m) &\le 2m^2 \int_0^{\varphi(x)} F_{Y^x}^{m-1}(z) S_{Y^x}(z) dz \left(\int_0^{\varphi(x)} F_{Y^x}^m(y) dy \right) \\ &= 2m^2 \left[(\varphi(x) - \varphi_{m-1}(x)) - (\varphi(x) - \varphi_m(x))\right] (\varphi(x) - \varphi_m(x))\end{aligned}$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

$$= 2m^2\,(\varphi(x)-\varphi_m(x))^2\left[\frac{\varphi(x)-\varphi_{m-1}(x)}{\varphi(x)-\varphi_m(x)}-1\right].$$

By the very definition (3), $\varphi_m(x)\equiv\mathbb{E}\big[\max(Y_1^x,\ldots,Y_m^x)\big]$ is nothing but the expectation of the FDH estimator based on the m-tuple $\{Y_i^x, i=1,\ldots,m\}$. Hence, the limit theorem of moments of the FDH estimator, established in Daouia et al. (2010, Thm. 2.1(iii)), yields $\lim_{m\to\infty} b_m^{-1}\,(\varphi(x)-\varphi_m(x))=\Gamma(1+1/\rho_x)$, where $b_m\sim(m\ell_x)^{-1/\rho_x}$ under the sufficient condition $\mathcal{C}(\boldsymbol{\rho}_{\mathbf{x}},\boldsymbol{\ell}_{\mathbf{x}})$, and Γ is the gamma function. It follows that $\left[\frac{\varphi(x)-\varphi_{m-1}(x)}{\varphi(x)-\varphi_m(x)}-1\right]\to 0$ and $(\varphi(x)-\varphi_m(x))\le 2(m\ell_x)^{-1/\rho_x}\Gamma(1+1/\rho_x)$, as $m\to\infty$. Therefore,

$$\sigma^2(x,m)\le m^2\,(\varphi(x)-\varphi_m(x))^2\le 4m^2(m\ell_x)^{-2/\rho_x}\Gamma^2(1+1/\rho_x):=\tilde{c}_x m^{2-2/\rho_x}$$

as $m\to\infty$, which ends the proof. ■

Proof of Theorem 5. By Theorem 4(ii) and (iii), we have

$$\widehat{\varphi}_{m_n}(x)-\varphi_{m_n}(x)=O_p\big(\sigma(x,m_n)/\sqrt{n}\big)=O_p\Big(m^{1-1/\rho_x}/\sqrt{n}\Big).$$

It follows from (13) that

$$\begin{aligned}\varphi(x)-\widehat{\varphi}_{m_n}(x)&=(\varphi(x)-\varphi_{m_n}(x))+(\varphi_{m_n}(x)-\widehat{\varphi}_{m_n}(x))\\&=\left(\frac{1}{m_n\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)+o\left(m_n^{-1/\rho_x}\right)+O_p\Big(\frac{m^{1-1/\rho_x}}{\sqrt{n}}\Big).\end{aligned}$$

Similarly, we have for all $a\ge 2$,

$$\varphi(x)-\widehat{\varphi}_{am_n}(x)=\left(\frac{1}{am_n\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)+o\left(m_n^{-1/\rho_x}\right)+O_p\Big(\frac{m^{1-1/\rho_x}}{\sqrt{n}}\Big),$$

$$\varphi(x)-\widehat{\varphi}_{a^2m_n}(x)=\left(\frac{1}{a^2m_n\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)+o\left(m_n^{-1/\rho_x}\right)+O_p\Big(\frac{m^{1-1/\rho_x}}{\sqrt{n}}\Big).$$

The differences lead to

$$m_n^{1/\rho_x}\big(\widehat{\varphi}_{am_n}(x)-\widehat{\varphi}_{m_n}(x)\big)=\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\left[1-1/a^{1/\rho_x}\right]+o(1)+O_p\Big(\frac{m_n}{\sqrt{n}}\Big),$$

$$(am_n)^{1/\rho_x}\big(\widehat{\varphi}_{a^2m_n}(x)-\widehat{\varphi}_{am_n}(x)\big)=\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\left[1-1/a^{1/\rho_x}\right]+o(1)+O_p\Big(\frac{m_n}{\sqrt{n}}\Big),$$

which gives

$$\frac{\widehat{\varphi}_{am_n}(x)-\widehat{\varphi}_{m_n}(x)}{\widehat{\varphi}_{a^2m_n}(x)-\widehat{\varphi}_{am_n}(x)}=a^{1/\rho_x}\frac{\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\left[1-1/a^{1/\rho_x}\right]+o(1)+O_p\left(\frac{m_n}{\sqrt{n}}\right)}{\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\left[1-1/a^{1/\rho_x}\right]+o(1)+O_p\left(\frac{m_n}{\sqrt{n}}\right)}.$$

Since $m_n/\sqrt{n}\to 0$ as $n\to\infty$, we get

$$\frac{\widehat{\varphi}_{am_n}(x)-\widehat{\varphi}_{m_n}(x)}{\widehat{\varphi}_{a^2m_n}(x)-\widehat{\varphi}_{am_n}(x)}\xrightarrow{p}a^{1/\rho_x},$$

whence $\widehat{\rho}_x\xrightarrow{p}\rho_x$. On the other hand, by applying again

$$m_n^{1/\rho_x}\big(\widehat{\varphi}_{am_n}(x)-\widehat{\varphi}_{m_n}(x)\big)=\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\left[1-1/a^{1/\rho_x}\right]+o(1)+O_p\Big(\frac{m_n}{\sqrt{n}}\Big),$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

in conjunction with $m_n/\sqrt{n} \to 0$ and $\widehat{\rho}_x \xrightarrow{p} \rho_x$ as $n \to \infty$, we get

$$\widehat{\ell}_x \equiv \frac{1}{m_n}\left[\frac{(1-a^{-1/\widehat{\rho}_x})\Gamma(1+1/\widehat{\rho}_x)}{\widehat{\varphi}_{am_n}(x)-\widehat{\varphi}_{m_n}(x)}\right]^{\widehat{\rho}_x} \xrightarrow{p} \ell_x,$$

which ends the proof. ■

Proof of Theorem 6. By Theorem 4(iii), we have

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\left\{\widehat{\varphi}_{m_n}(x)-\varphi(x)+B_{m_n}(x)\right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \quad n\to\infty.$$

It remains to show that

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\left\{\widehat{B}_{m_n}(x)-B_{m_n}(x)\right\} \xrightarrow{p} 0, \quad n\to\infty,$$

or equivalently, as $n \to \infty$,

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\left\{\left(\frac{1}{m_n\widehat{\ell}_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)-\left(\frac{1}{m_n\ell_x}\right)^{1/\rho_x}\Gamma\left(1+\frac{1}{\rho_x}\right)\right.$$
$$\left.+o\left(m_n^{-(1+\alpha_x)/\rho_x}\right)\right\} \xrightarrow{p} 0.$$

By Theorem 4(ii) and the condition $m_n = O\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}}}$, we have

$$\frac{\sqrt{n}}{\sigma(x,m_n)}o\left(m_n^{-(1+\alpha_x)/\rho_x}\right) = \sqrt{n}m_n^{-\frac{1}{2}-\frac{\alpha_x}{\rho_x}}o(1) = \sqrt{n}m_n^{-\frac{1}{2}-\frac{\alpha_x}{\rho_x}}o(1)$$
$$= \sqrt{n}\left(\frac{\sqrt{n}}{\log\log n}\right)^{\frac{-\frac{1}{2}-\frac{\alpha_x}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}}o(1) = \sqrt{n}^{\frac{\rho_x+1-\alpha_x}{\frac{3}{2}\rho_x+1}}(\log\log n)^{\frac{\frac{1}{2}+\frac{\alpha_x}{\rho_x}}{\frac{3}{2}+\frac{1}{\rho_x}}}o(1),$$

where the last term tends to zero since $\alpha_x > \rho_x + 1$. Hence, it suffices to show that

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\Gamma\left(1+\frac{1}{\rho_x}\right)\frac{1}{m_n^{1/\rho_x}}\left\{\left(\frac{1}{\widehat{\ell}_x}\right)^{1/\rho_x}-\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\right\} \xrightarrow{p} 0,$$

or equivalently,

$$\frac{\sqrt{n}}{\sigma(x,m_n)}\Gamma\left(1+\frac{1}{\rho_x}\right)\frac{1}{m_n^{1/\rho_x}}\left[\left(\frac{1}{\widehat{\ell}_x}\right)^{1/\rho_x}-\left(\frac{1}{\widetilde{\ell}_x}\right)^{1/\rho_x}\right]$$
$$+\frac{\sqrt{n}}{\sigma(x,m_n)}\Gamma\left(1+\frac{1}{\rho_x}\right)\frac{1}{m_n^{1/\rho_x}}\left[\left(\frac{1}{\widetilde{\ell}_x}\right)^{1/\rho_x}-\left(\frac{1}{\ell_x}\right)^{1/\rho_x}\right]$$
$$:= \mathrm{I}+\mathrm{II} \xrightarrow{p} 0, \tag{A.8}$$

where, similarly to $\widehat{\ell}_x = \frac{1}{\tilde{m}_n}\left[\frac{(1-a^{-1/\rho_x})\Gamma(1+1/\rho_x)}{\widehat{\varphi}_{a\tilde{m}_n}(x)-\widehat{\varphi}_{\tilde{m}_n}(x)}\right]^{\rho_x}$, we set $\widetilde{\ell}_x := \frac{1}{\tilde{m}_n}\left[\frac{(1-a^{-1/\rho_x})\Gamma(1+1/\rho_x)}{\varphi_{a\tilde{m}_n}(x)-\varphi_{\tilde{m}_n}(x)}\right]^{\rho_x}$. The first term I can be expressed as

$$\mathrm{I} = \frac{\sqrt{n}}{\sigma(x,m_n)}\frac{1}{m_n^{1/\rho_x}}\frac{\tilde{m}_n^{1/\rho_x}}{(1-a^{-1/\rho_x})}\left[\left(\widehat{\varphi}_{a\tilde{m}}(x)-\varphi_{a\tilde{m}}(x)\right)-\left(\widehat{\varphi}_{\tilde{m}}(x)-\varphi_{\tilde{m}}(x)\right)\right]$$

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press

$$= \frac{\sqrt{n}}{\sigma(x,m_n)} \frac{\widetilde{m}_n^{1/\rho_x}}{m_n^{1/\rho_x}} \frac{\sigma(x,\tilde{m}_n)}{\sqrt{n}} O_p(1) = O_p\left(\frac{\sigma(x,\tilde{m}_n)}{\sigma(x,m_n)} \frac{\widetilde{m}_n^{1/\rho_x}}{m_n^{1/\rho_x}} \right),$$

where the term $O_p(1)$ follows from Theorem 4(iii). Now, by Theorem 4(ii), we have

$$\frac{\sigma(x,\tilde{m}_n)}{\sigma(x,m_n)} \frac{\widetilde{m}_n^{1/\rho_x}}{m_n^{1/\rho_x}} \leq \left(\frac{\tilde{c}_x}{c_x} \right)^{1/2} \frac{\widetilde{m}_n}{m_n^{1/2}}, \quad n \to \infty.$$

The ratio $\widetilde{m}_n / m_n^{1/2}$ tends to zero as soon as $\left(\frac{3}{2} + \frac{1}{\rho_x} \right)^{-1} < 2\tilde{\varepsilon} - \varepsilon$. Whence $\mathrm{I} = o_p(1)$. The second term II in (A.8) can be written explicitly as

$$\begin{aligned}
\mathrm{II} &= \frac{\sqrt{n}}{\sigma(x,m_n)} \Gamma\Big(1 + \frac{1}{\rho_x}\Big) \frac{\widetilde{m}_n^{1/\rho_x}}{m_n^{1/\rho_x}} \left[\frac{\varphi_{a\widetilde{m}}(x) - \varphi_{\widetilde{m}}(x)}{\Gamma\big(1+1/\rho_x\big)\big(1-a^{-1/\rho_x}\big)} - \Big(\frac{1}{\widetilde{m}_n \ell_x} \Big)^{1/\rho_x} \right] \\
&= \frac{\sqrt{n}}{\sigma(x,m_n)} \Gamma\Big(1 + \frac{1}{\rho_x}\Big) \frac{\widetilde{m}_n^{1/\rho_x}}{m_n^{1/\rho_x}} o\left(\widetilde{m}_n^{-(1+\alpha_x)/\rho_x} \right) \\
&= \frac{\sqrt{n}}{\sigma(x,m_n)} o\Big(\frac{\widetilde{m}_n^{-\alpha_x/\rho_x}}{m_n^{1/\rho_x}} \Big),
\end{aligned}$$

where the second equality follows from $\mathcal{C}_2(\boldsymbol{\rho}_\mathbf{x}, \boldsymbol{\ell}_\mathbf{x}, \boldsymbol{\alpha}_\mathbf{x})$. Applying again Theorem 4(ii), we get

$$\begin{aligned}
|\mathrm{II}| &\leq \frac{\sqrt{n}}{m_n^{1/2-1/\rho_x}} \frac{\widetilde{m}_n^{-\alpha_x/\rho_x}}{m_n^{1/\rho_x}} o(1) \\
&\leq \sqrt{n} \left(\frac{\sqrt{n}}{\log\log n} \right)^{-\frac{1}{2}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}} - \varepsilon \right) - \frac{\alpha_x}{\rho_x}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}} - \tilde{\varepsilon} \right)} o(1).
\end{aligned}$$

The condition $1 - \frac{1}{2}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}} - \varepsilon \right) - \frac{\alpha_x}{\rho_x}\left(\frac{1}{\frac{3}{2}+\frac{1}{\rho_x}} - \tilde{\varepsilon} \right) < 0$ ensures that the term on the right-hand side is $o(1)$. This completes the proof of the theorem. ■

https://doi.org/10.1017/S0266466620000171 Published online by Cambridge University Press