WORKING PAPERS

N° 18-890

Februray 2018

"Robustified expected maximum production frontiers"

Abdelaati Daouia, Jean-Pierre Florens and Léopold Simar

Toulouse School of Economics

Robustified expected maximum production frontiers

Abdelaati Daouia*
abdelaati.daouia@tse-fr.eu

Jean-Pierre Florens*
jean-pierre.florens@tse-fr.eu

Léopold Simar*,§
leopold.simar@uclouvain.be

January 31, 2018

Abstract

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.

Key words : Boundary regression, Expected maximum, Nonparametric estimation, Production function, Robustness.

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

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

*Toulouse School of Economics, France. This research was supported by the Toulouse School of Economics Individual Research Fund (IRF/Daouia-20125), and by the French "Agence Nationale pour la Recherche" under grant ANR-08-BLAN-0106-01/EPI project.

§Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Belgium. Support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) is gratefully acknowledged.

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 *et al.* (1977) and Meeusen and van den Broeck (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 *et al.* (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 monotone nondecreasing. A pioneering contribution in this area is due to Farrell (1957), who introduced Data Envelopment Analysis (DEA), based on either the conical hull or the convex hull of the data. This was further extended by Deprins *et al.* (1984) to the Free Disposal Hull (FDH) estimator, whose properties have been extensively discussed in the literature. See for instance Kneip *et al.* (2008) and Daouia *et al.* (2010, 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 expected maximum production frontiers by Cazals *et al.* (2002) and quantile-based frontiers by Aragon *et al.* (2005) and Daouia and Simar

(2007). Comparisons between the two concepts from a robustness and an asymptotic points of view can be found in Daouia and Ruiz-Gazen (2006) and Daouia and Gijbels (2011). In particular, once the 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 its 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 $x' \geqslant x$ and $y' \leqslant y$. The monotonicity of $\varphi(x)$, referred to as non-negative 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 *et al.* (2000)]. The partial expected maximum production functions enjoy the property of monotonicity if and only if the hypothesis of tail monotonicity holds [see Theorem A.3 in Cazals *et al.* (2002)]. Second and most importantly, even if this theoretical hypothesis is satisfied, the empirical estimators of expected maximum production functions, 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 *et al.* (2017). However, by construction, both population and empirical 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 non-robust 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 of an appropriate benchmark partial frontier in practice 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.

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 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 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, 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), and Daouia *et al.* (2010, 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 scatterplots are given below in Figures 1 and 2.

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 as well as our motivating real data examples. Section 3 gives some numerical illustrations and Monte Carlo evidence. Section 4 concludes.

2 Robust boundary regression

Let us revisit the popular free disposal hull (FDH) frontier estimator in Section 2.1 and the concept of expected maximum production frontiers in Section 2.2, before moving to the main conceptual results in Section 2.3. Practical guidelines to effect the necessary parameter selection are described in Section 2.4.

2.1 Setting and objective

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 production function and $U_j \geqslant 0$ being the inefficiency term. 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 non-standard conditional distribution of Y given $X \leqslant x$. If $(\Omega, \mathcal{A}, \mathbb{P})$ denotes the probability space on which the random vector $(X, Y) \in \mathbb{R}_+^p \times \mathbb{R}_+$ is defined and

$$F_{Y|X}(y|x) = \mathbb{P}(Y \leqslant y \,|\, X \leqslant x)$$

is the distribution function of Y conditioned by $X \leqslant x$, assuming $F_X(x) := \mathbb{P}(X \leqslant x) > 0$, then $\varphi(x)$ can be characterized as the upper conditional endpoint

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

Generally speaking, $\varphi(x)$ is not the upper extremity of the support of (X, Y) at $X = x$, say $\varphi^u(x)$, but equals $\sup_{x' \leqslant x} \varphi^u(x')$. Therefore, it is isotonic nondecreasing in x and envelops the true upper support boundary. In the case where the frontier function φ^u is nondecreasing, which is a minimal requirement in production econometrics, φ coincides with φ^u. Then, consideration of the constrained envelop $\varphi(x)$ is advantageous as it affords estimation at a faster rate than $\varphi^u(x)$, see Daouia and Park (2013). Because of the local nature of $\varphi^u(x)$, one can use only the data points in a local strip around x to estimate it. In contrast, by substituting the empirical conditional distribution function

$$\widehat{F}_{Y|X}(y|x) = \sum_{i=1}^{n} 1\!\!1(X_i \leqslant x, Y_i \leqslant y) / \sum_{i=1}^{n} 1\!\!1(X_i \leqslant x)$$

in place of $F_{Y|X}(y|x)$ in (1), with $1\!\!1(\cdot)$ being the indicator function, Cazals *et al.* (2002) recover the usual FDH estimator

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

which defines the smallest step and monotone surface lying above the sample points (X_i, Y_i). Park *et al.* (2000) have determined its limit distribution under the condition that the density of data is strictly positive at the boundary. More recently, Daouia *et al.* (2010, 2014) have elucidated its full asymptotic theory in a general setting from the perspective of extreme value theory. In particular, 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, *i.e.*, $\lim_{t\uparrow\infty} L_x(tz)/L_x(t) = 1$ 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

$$F_X(x)[1 - F_{Y|X}(y|x)] = \ell_x\big(\varphi(x) - y\big)^{\rho_x} + o\big((\varphi(x) - y)^{\rho_x}\big) \quad \text{as} \quad y \uparrow \varphi(x), \tag{2}$$

it is shown in Daouia *et al.* (2010, Corollary 2.1) that

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

where a random variable W is said to follow the distribution Weibull$(1, \rho_x)$ if W^{ρ_x} is Exponential with parameter 1. 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.

In absence of information on whether the available data are measured accurately, it would look awkward for practitioners to assume that only the FDH frontier points $(X_i, Y_i \equiv \widehat{\varphi}(X_i))$ contain valuable information about the true efficient support extremity. In many empirical applications, some FDH observations may appear so isolated that they hardly seem related to the sample. They may be outliers resulting from data corruption due to by reporting, transcription, or other errors. Other top observations, well inside the sample, could help the practitioners to achieve their objective of 'robustification'. The underlying idea is to estimate an *anchor* partial frontier well inside the production set but near the true full frontier, and then to shift it to the right place. As suggested by Cazals *et al.* (2002) and Daouia *et al.* (2012), a practitioner can protect himself by specifying a trimming number $m \in \{1, 2, \ldots\}$ and considering a notion of expected maximum achievable level of output among m firms drawn in the population of firms using less than a given level of inputs. Next, we introduce formally this robust concept of expected maximum output frontier and propose a new and more valuable variant based on an unconditional dimensionless characterization of the production process.

2.2 Expected maximum production frontiers

For a given level of inputs-usage x such that $F_X(x) > 0$, the expected maximum output function of order m is defined as

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

where $(Y^1, \ldots, Y^m)$ are i.i.d. random variables generated by the conditional distribution of Y given $X \leqslant x$. The quantity $\psi_m(x)$ gives the expected maximum achievable production among a fixed number of m firms drawn from the population of production units using less than x as inputs. For a particular firm operating at (x, y), comparing its output y with the benchmark value $\psi_m(x)$ gives a clear indication of how well this firm is performing compared with m production units using less inputs than x.

It is easily seen that $\psi_m(x) = \varphi(x) - \int_0^{\varphi(x)}[F_{Y|X}(y|x)]^m\, dy$, and hence the partial production frontier $\psi_m(x)$ converges to the true efficient frontier $\varphi(x)$ itself as $m \to \infty$. Likewise, for a fixed sample size n, the 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 $m \to \infty$. Putting $N_x = \sum_{i=1}^n \mathbb{1}(X_i \leqslant x)$ and taking $\mathcal{Y}_1^x, \ldots, \mathcal{Y}_{N_x}^x$ to be the Y_i's such that $X_i \leqslant x$, the exact value of $\widehat{\psi}_m(x)$ can be computed as

$$\widehat{\psi}_m(x) = \sum_{i=1}^{N_x} \mathcal{Y}_{(i)}^x \left\{ \left(\frac{i}{N_x}\right)^m - \left(\frac{i-1}{N_x}\right)^m \right\}, \tag{3}$$

where $\mathcal{Y}_{(i)}^x$ denotes the ith order statistic of the points $\mathcal{Y}_1^x, \ldots, \mathcal{Y}_{N_x}^x$ [see Equation (2.4) in Daouia and Gijbels (2011)]. Figure 1 (top panel) and Figure 2 (top panel) 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 and $m = \infty$ (FDH).

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). In particular, both $\widehat{\psi}_m(x) \equiv T^{m,x}\big(\widehat{F}_{(X,Y)}\big)$ and $\psi_m(x) \equiv T^{m,x}\big(F_{(X,Y)}\big)$ are representable as a functional $T^{m,x}$ of the empirical and population distribution functions

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

respectively, where the statistical functional $T^{m,x}$ associates to 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 $y \geqslant \inf\{y \in \mathbb{R}_+ | F(x,y)/F(x,\infty) = 1\}$. 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

$$\mathrm{IF}\big((x_0,y_0); T^{m,x}, F_{(X,Y)}\big) = \frac{m}{F_X(x)} 1\!\!I(x_0 \leqslant x) \int_0^{\varphi(x)} F^{m-1}_{Y|X}(y|x)\left[F_{Y|X}(y|x) - 1\!\!I(y_0 \leqslant y)\right] dy, \tag{4}$$

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

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

$$\begin{aligned}
\gamma^*\big(T^{m,x}, F_{(X,Y)}\big) &= \frac{m}{F_X(x)} \max\left\{\int_0^{\varphi(x)} F^m_{Y|X}(y|x)dy, \int_0^{\varphi(x)} F^{m-1}_{Y|X}(y|x)\left[1 - F_{Y|X}(y|x)\right] dy\right\} \\
&\equiv \frac{m}{F_X(x)} \max\left\{\varphi(x) - \psi_m(x), \psi_m(x) - \psi_{m-1}(x)\right\}. \qquad (5)
\end{aligned}$$

Also, it follows from the functional convergence theorem of Cazals *et al.* (2002) that

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

Thus the influence function also measures the asymptotic bias caused by contamination in the observations (X_i, Y_i), $i = 1, \ldots, n$. More recently, under regularity conditions on $F_{Y|X}(\cdot|x)$, Daouia *et al.* (2012) have used $\widehat{\psi}_m(x)$ to estimate the full frontier $\varphi(x)$ itself, with $m = m_n \to \infty$ at a slow rate as $n \to \infty$. The estimator $\widehat{\psi}_{m_n}(x)$ is then corrected for its inherent bias to obtain a final regularized frontier more robust than the traditional FDH curve to extreme values and outliers.

Yet, the conditioning by the event $\{X \leqslant x\}$ 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. The occurence of this vexing border effect is reflected by the presence of low values of $F_X(x)$ in the denominator of (4) and (5). This is visualized in Figure 1 (top panel) and Figure 2 (top panel), where the selected large m-frontiers $\widehat{\psi}_m(x)$ coincide with the non-robust 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 $X \leqslant x$.

2.3 Robustified unconditional m-frontiers

For a fixed level of inputs-usage $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, we propose in this paper to 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

$$Y^x = Y\,\mathbb{1}(X \leqslant x) \quad \text{and} \quad Y_i^x = Y_i\,\mathbb{1}(X_i \leqslant x)\,, \quad i = 1, \ldots, n. \tag{6}$$

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

$$F_{Y^x}(y) = \big\{1 - F_X(x)[1 - F_{Y|X}(y|x)]\big\}\,\mathbb{1}(y \geqslant 0).$$

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

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

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

$$\varphi_m(x) = \mathbb{E}\big[\max(Y_1^x, \ldots, Y_m^x)\big] = \int_0^\infty \big(1 - [F_{Y^x}(y)]^m\big)\,dy, \tag{7}$$

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 (6). Clearly, for any trimming number $m \geqslant 1$, the quantity $\varphi_m(x)$ is nothing but the expectation of the FDH estimator based on the m-tuple $\{Y_i^x = Y_i\,\mathbb{1}(X_i \leqslant x)\,, i = 1, \ldots, m\}$.

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

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

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

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

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

$$\widehat{\varphi}_m(x) = \int_0^\infty \big(1 - [\widehat{F}_{Y^x}(y)]^m\big)\,dy = \widehat{\varphi}(x) - \int_0^{\widehat{\varphi}(x)} [\widehat{F}_{Y^x}(y)]^m\,dy \tag{8}$$

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:

$$\begin{aligned} \widehat{\varphi}_m(x) &= \sum_{i=1}^{n} Y^x_{(i)} \left\{ \left(\frac{i}{n}\right)^m - \left(\frac{i-1}{n}\right)^m \right\} \\ &= \sum_{i=1}^{N_x} \mathcal{Y}^x_{(i)} \left\{ \left(\frac{n-N_x+i}{n}\right)^m - \left(\frac{n-N_x+i-1}{n}\right)^m \right\}, \end{aligned} \tag{9}$$

where $Y^x_{(i)}$ denotes the ith order statistic of the observations $Y^x_1, \ldots, Y^x_n$. This marks a substantial difference with $\widehat{\psi}_m(x)$ as can be seen from (3) and visualized in Figure 1 (bottom panel) and Figure 2 (bottom panel) in both cases of postal services.

From a robustness theory viewpoint, both $\varphi_m(x) \equiv T_m\big(F_{Y^x}\big)$ and $\widehat{\varphi}_m(x) \equiv T_m\big(\widehat{F}_{Y^x}\big)$ 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 to a univariate distribution function $F(\cdot)$ on $\mathbb{R}_+$ the real value

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

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

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

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 3. *For all $m \geqslant 1$ and $x \in \mathbb{R}^p_+$ such that $F_X(x) > 0$, we have*

$$IF\big(u; T_m, F_{Y^x}\big) = -m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1} \big\{\delta_u(y) - F_{Y^x}(y)\big\} dy$$

$$\equiv -m \int_0^{\varphi(x)} \big[1 - F_X(x) + F_{(X,Y)}(x,y)\big]^{m-1} \big\{ 1\!\!1(u \leqslant y) - 1 + F_X(x) - F_{(X,Y)}(x,y)\big\}\, dy.$$

This closed form expression of the influence function indicates that the unconditional m-frontiers $\widehat{\varphi}_m(x) \equiv T_m\big(\widehat{F}_{Y^x}\big)$ do not suffer from the inherent border effects of the initial concept of conditional m-frontiers $\widehat{\psi}_m(x) \equiv T^{m,x}\big(\widehat{F}_{(X,Y)}\big)$. Moreover, by making use of the same technique of proof of Proposition 1, it is easily seen that the gross-error sensitivity $\lambda^*\big(T_m, F_{Y^x}\big) := \sup_{u \geqslant 0} \big|\mathrm{IF}\big(u; T_m, F_{Y^x}\big)\big|$ satisfies

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

Also, as can be seen from the next proposition, $\mathrm{IF}\big(Y_i^x; T_m, F_{Y^x}\big)$ represents the approximate contribution, or influence, of the observation (X_i, Y_i) toward the estimation error $\big\{\widehat{\varphi}_m(x) - \varphi_m(x)\big\}$.

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

$$\begin{aligned}\sqrt{n}\big\{\widehat{\varphi}_m(x) - \varphi_m(x)\big\} &= -m\sqrt{n}\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\big\{\widehat{F}_{Y^x}(y) - F_{Y^x}(y)\big\}dy + o_p(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} IF\big(Y_i^x; T_m, F_{Y^x}\big) + o_p(1) \quad as \quad n \to \infty.\end{aligned}$$

Another immediate consequence of this proposition is that $\sqrt{n}\big\{\widehat{\varphi}_m(x)-\varphi_m(x)\big\}$ is asymptotically normal with zero mean and variance

$$\begin{aligned}\sigma^2(x,m) &= \mathbb{E}\left\{m\int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1}\big\{1\!\!I(Y^x \leqslant y) - F_{Y^x}(y)\big\}dy\right\}^2 \\ &= m^2\int_0^{\varphi(x)}\int_0^{\varphi(x)} \big[F_{Y^x}(y)F_{Y^x}(z)\big]^{m-1}\big\{F_{Y^x}(y \wedge z) - F_{Y^x}(y)F_{Y^x}(z)\big\}dydz.\end{aligned}$$

Next, we show that $\sqrt{n}\big\{\widehat{\varphi}_m(x) - \varphi_m(x)\big\}$ 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.

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

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

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 construction of 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. It is then natural to verify whether the asymptotic normality of the *anchor* production frontier $\widehat{\varphi}_m(x)$ is still valid when it is shifted to the right place for $m = m_n \to \infty$ at a slow rate as $n \to \infty$.

Theorem 2. *For $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$, if $m_n \to \infty$ and $\frac{m_n(m_n-1)}{\sigma(x,m_n)} = O\left(\frac{\sqrt{n}}{\log\log n}\right)$ as $n \to \infty$, then*

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

Also, under the extreme-value condition (2), *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 this asymptotic normality is still valid.*

When the trimming level $m = m_n \to \infty$ fast enough as $n \to \infty$, the frontier estimator $\widehat{\varphi}_{m_n}(x)$ estimates $\varphi(x)$ itself and converges to the same limit distribution as the FDH estimator with the same scaling.

Theorem 3. *For* $x \in \mathbb{R}_+^p$ *such that* $F_X(x) > 0$, *if* (2) *holds and* $m_n \geqslant \beta n \log(Cn)\{1 + o(1)\}$ *for some constants* $\beta > \frac{1}{\rho_x} + 1$ *and* $C > 0$, *then*

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

It should also be clear that, from the point of view of the axiomatic foundation for production functions, nothing guarantees that the conditional expected maximum production function $\psi_m(x)$ and its estimator $\widehat{\psi}_m(x)$ satisfy the monotonicity requirement. In contrast, both 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

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

Then, for all $y \geqslant 0$, the function $x \mapsto F_{Y^x}(y)$ is monotone nonincreasing. Therefore, the unconditional partial frontier function $\varphi_m(x)$ defined in (7) is monotone nondecreasing in x, for all $m \geqslant 1$. Likewise, it is easily seen that

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

is monotone nonincreasing in x. Whence, the empirical estimator $\widehat{\varphi}_m(x)$ described in (8) is constrained to be monotone nondecreasing in x, for all $m \geqslant 1$.

2.4 Trimming selection problem

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

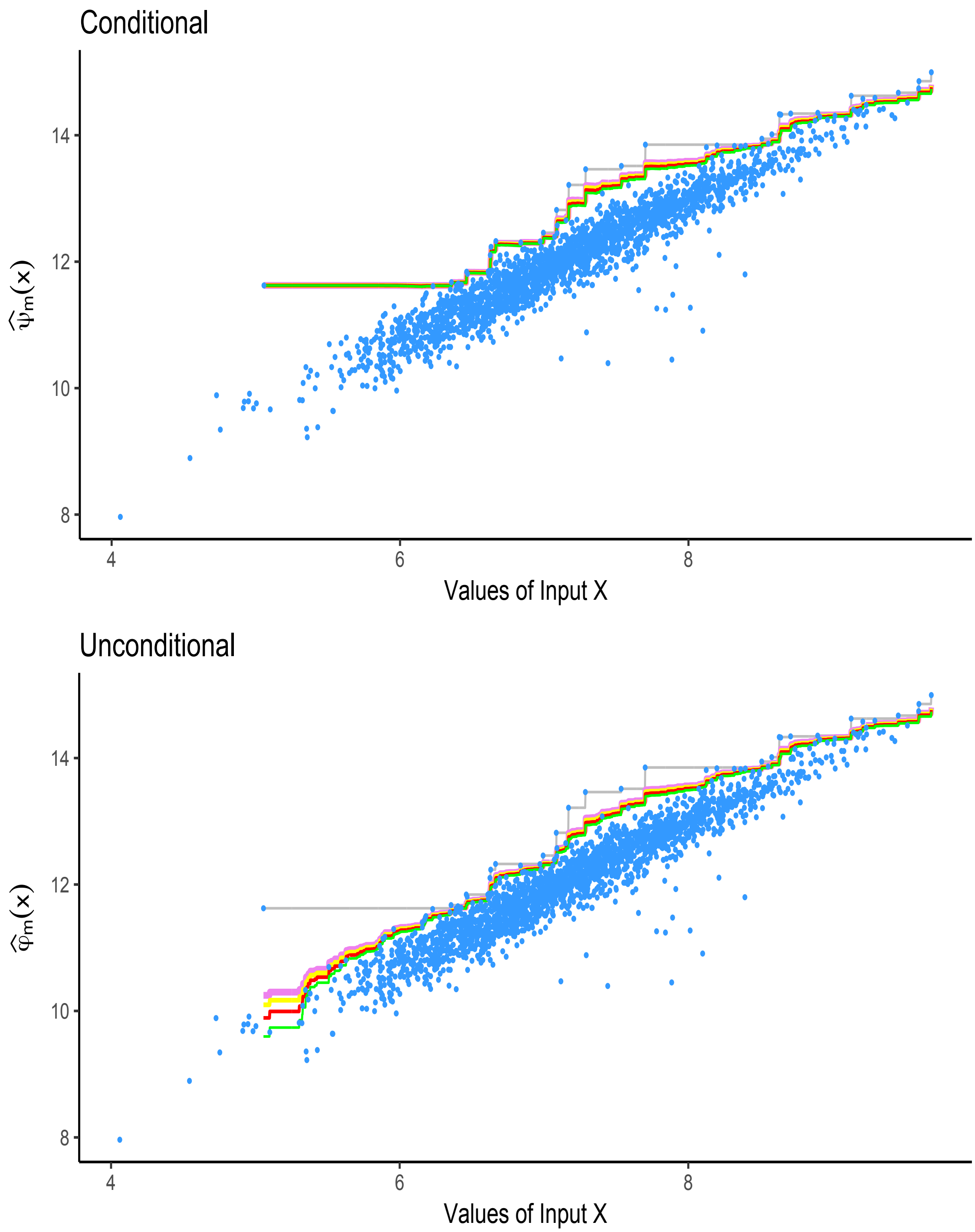

Figure 1: *Scatterplot of the* $n = 2,326$ *delivery post offices (data in logarithms)— Estimated expected maximum production frontiers* $\widehat{\psi}_m$ *(top) and* $\widehat{\varphi}_m$ *(bottom), with* $m = 600, 700, 800, 900$ *and* $m = \infty$ *(FDH), respectively, in green, red, yellow, violet and gray curves.*

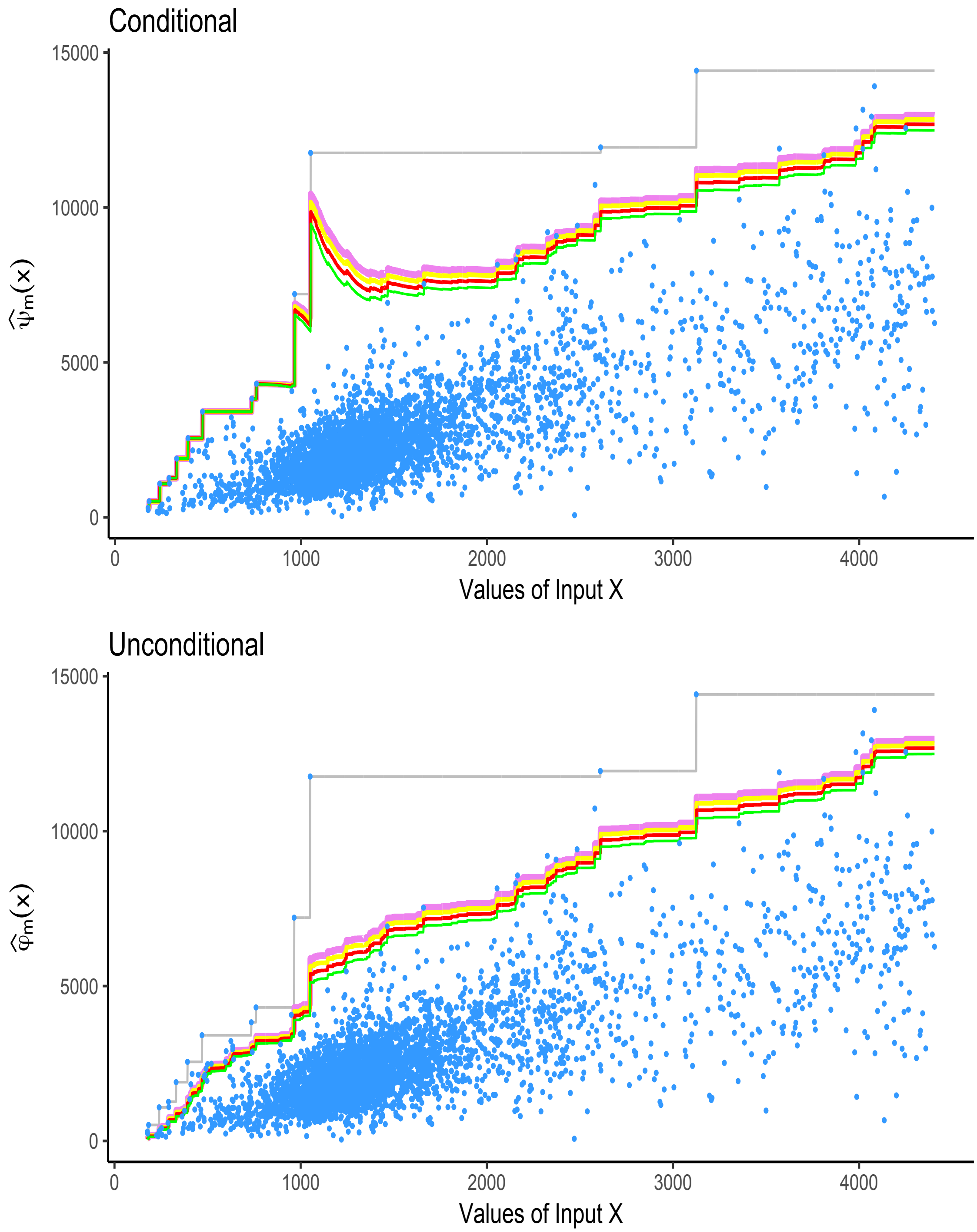

Figure 2: *Scatterplot of the* $n = 4,000$ *delivery post offices—Estimated expected maximum production frontiers* $\widehat{\psi}_m$ *(top) and* $\widehat{\varphi}_m$ *(bottom), with* $m = 600$, 700, 800, 900 *and* $m = \infty$ *(FDH), respectively, in green, red, yellow, violet and gray curves.*

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 $m = 100, 200, \ldots, 1500$. 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

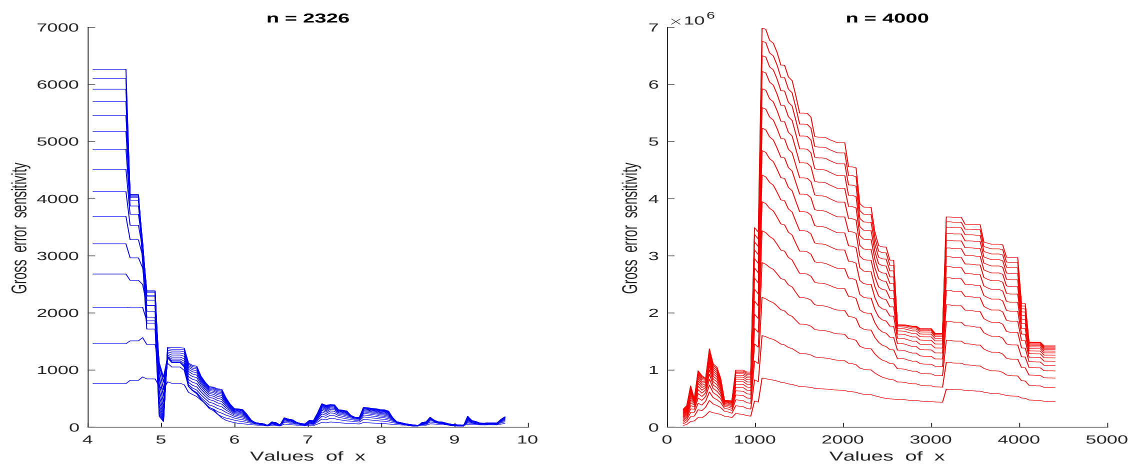

Figure 3: *Plots of* $x \mapsto \lambda^*\big(T_m, \widehat{F}_{Y^x}\big)$ *for* $m = 100, 200, \ldots, 1500$. *From left to right, the 2,326 and 4,000 post offices.*

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 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. Such an L deviation should appear for any other analyzed data set 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

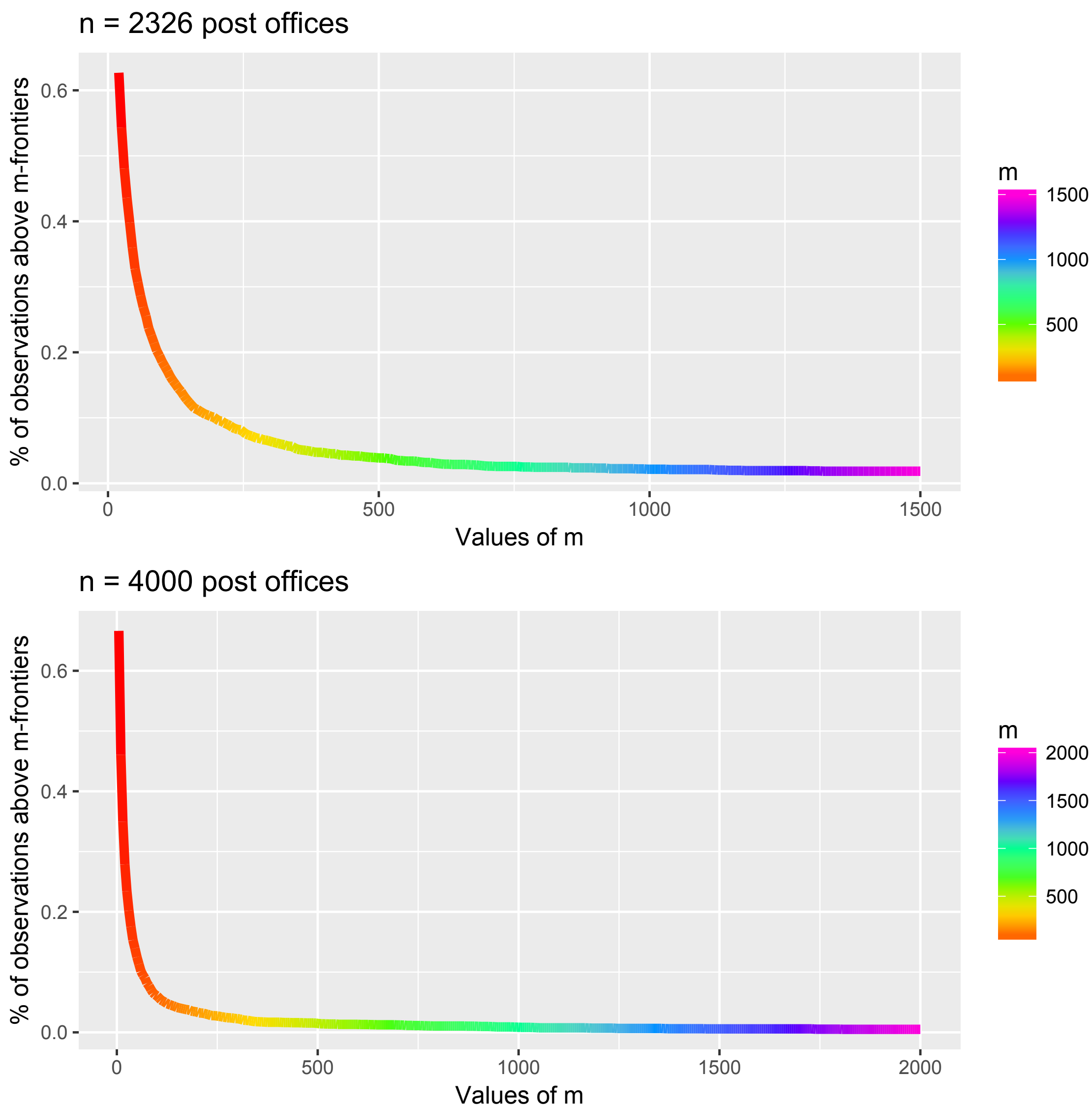

Figure 4: *Evolution of the % of sample points outside the partial m-frontiers* $\widehat{\varphi}_m$.

effect, it may be seen that the percentage decreases very slowly along the 'blue' part, say $850 \leqslant m \leqslant 1250$, before to become extremely stable along the 'violet' part of the curve. This means that all observations left 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 $m = 1250$ may indicate that the observations above the frontier $\widehat{\varphi}_{1250}$ are really outlying or suspicious isolated extremes that deserve to be carefully examined. This might suggest to choose 1250 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, $500 \leqslant m \leqslant 1000$. The lower and upper selected prudential frontiers $\widehat{\varphi}_{500}$ and $\widehat{\varphi}_{1000}$ are superimposed in Figure 5 along with the FDH estimator. Unsurprisingly, there are very few observations lying between the two partial frontiers.

3 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 *et al.* (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 $\{(x, y) \in [0, 1]^2 : y \leqslant x\}$. 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 $0 < x \leqslant 1$ and $0 \leqslant y \leqslant \varphi(x)$. 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 non-linear 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 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 $0 < x \leqslant 1$ and $0 \leqslant y \leqslant \varphi(x)$. The partial order-$m$ frontier functions have the following closed form expressions:

$$\begin{aligned}
\psi_m(x) &= \varphi(x) - \sum_{k=0}^{m} \binom{m}{k} 3^{m-k}(-2)^k \sqrt{x}/(2m + k + 1), \\
\varphi_m(x) &= \varphi(x) - \sum_{k=0}^{m} \binom{m}{k} x^{k+1/2}(-1)^k \sum_{j=0}^{k} \binom{k}{j} \sum_{i=0}^{j} \binom{j}{i} (-3)^{j-i} 2^i/(2j + i + 1).
\end{aligned}$$

For both examples, the graphs of ψ_m and φ_m are superimposed in Figures 6 and 7, for three values of $m = 1, 10, 25$, along with the true support boundary φ. First, it may be

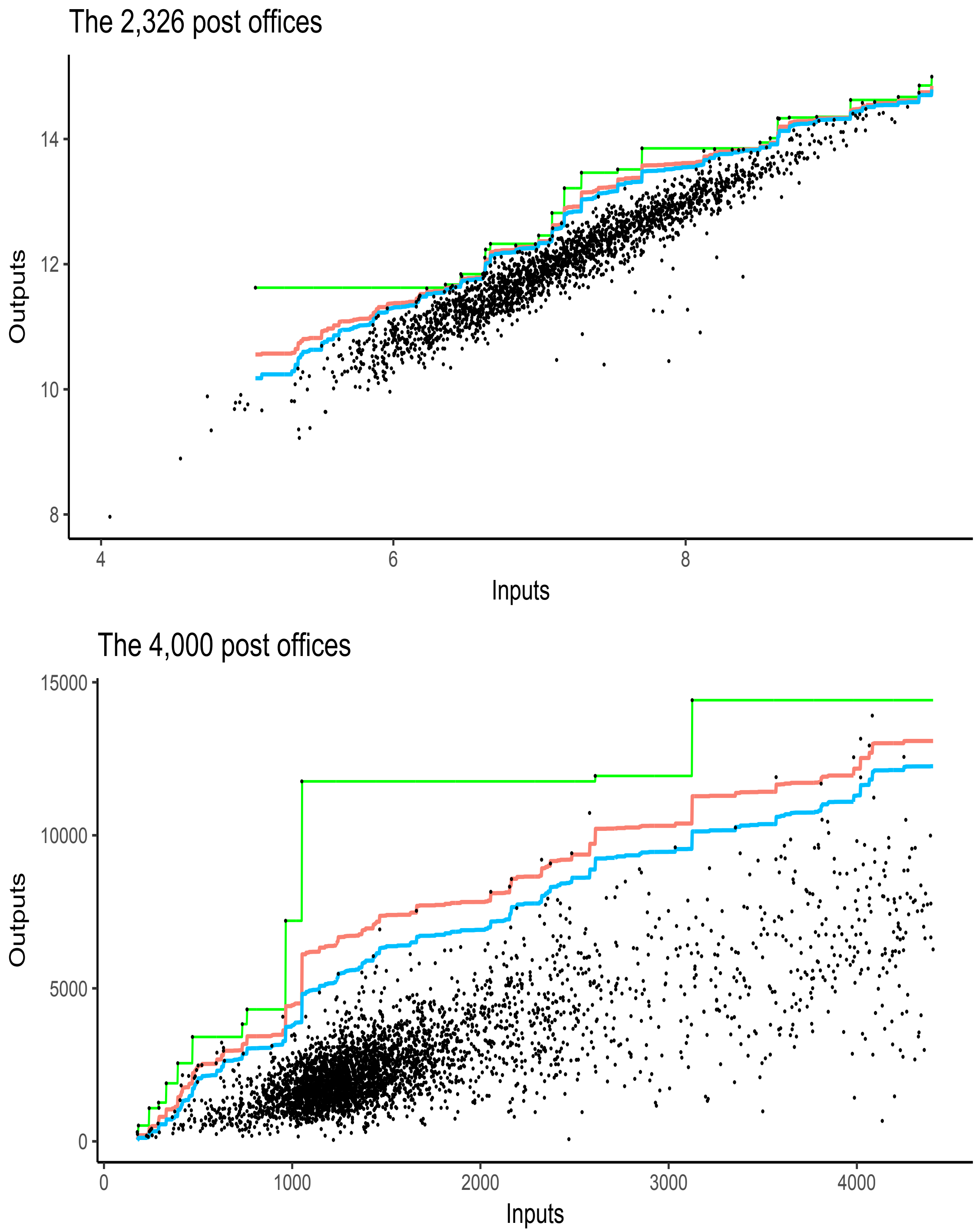

Figure 5: *Selected (lower and upper) expected maximum production frontiers* $\widehat{\varphi}_m$. *Top—dataset of size* $2,326$ *in logarithms, with* $m = 1250$ *(upper) in red line,* $m = 850$ *(lower) in blue line, and* $m = \infty$ *(FDH) in green line. Bottom—dataset of size* $4,000$*, with* $m = 1000$ *(upper) in red line,* $m = 500$ *(lower) in blue line, and* $m = \infty$ *(FDH) in green line.*

seen from the plots that the conditional $m-$frontiers $\psi_m(x)$ [dotted red curves] diverge from the true frontier $\varphi(x)$ [solid green curve] as x increases. Whereas the new unconditional $m-$frontiers $\varphi_m(x)$ [dashed blue 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 8 and 9, where the gross-error sensitivities $\gamma^*\big(T^{m,x}, F_{(X,Y)}\big)$ of $\psi_m(x)$ and $\lambda^*\big(T_m, F_{Y^x}\big)$ 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, Ronchetti, Rousseeuw and Stahel (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 magenta] is overall smaller than γ^* [solid cyan]. 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 $X \leqslant x$.

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 $n = 100, 500, 1000$. Three outliers were added in each simulated data set: $\{(0.1, 0.6), (0.35, 0.8), (0.6, 1)\}$ for both uniform-triangle and Cobb-Douglas examples. The measures of efficiency for each simulation used were the mean squared error and the bias

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

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

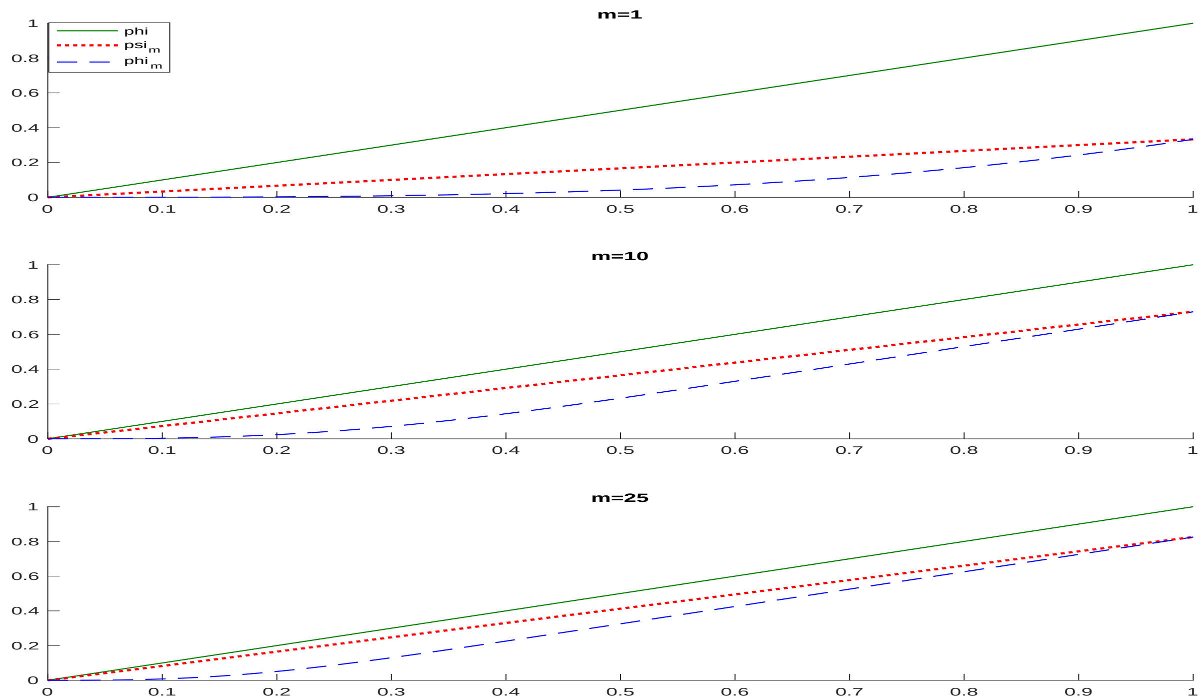

Figure 6: *Uniform triangle example—Graphs of φ in solid line, ψ_m in dotted line, and φ_m in dashed line.*

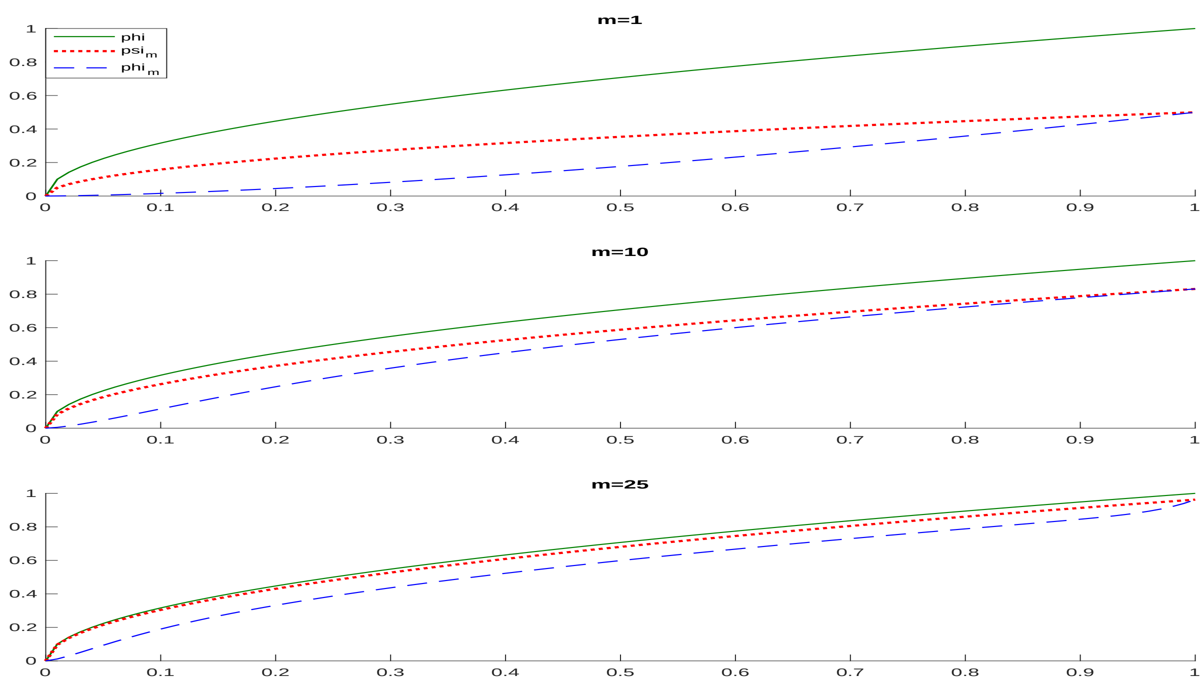

Figure 7: *Cobb-Douglas example—Same graphs as before.*

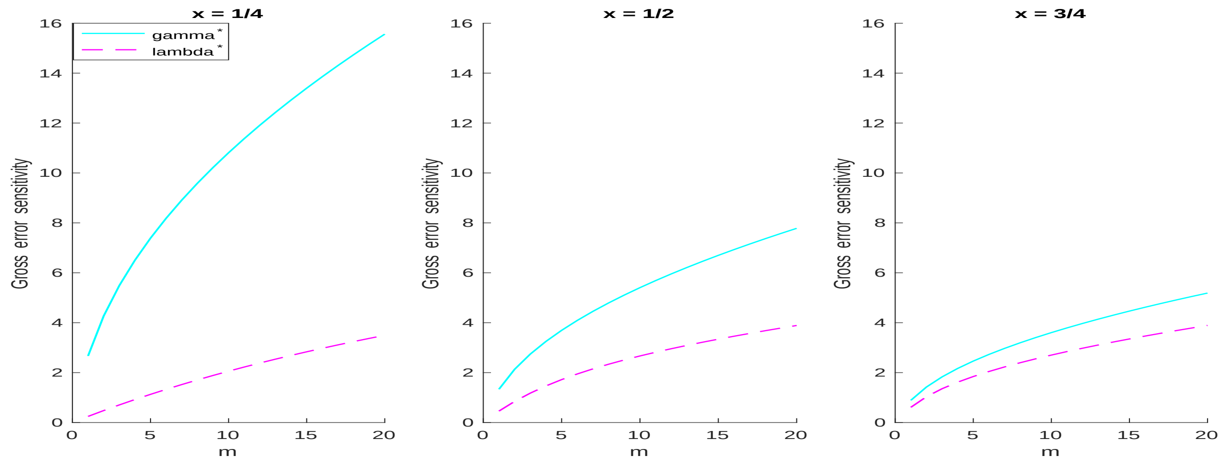

Figure 8: *Uniform triangle example—Gross-error sensitivities* $m \mapsto \gamma^*\big(T^{m,x}, F_{(X,Y)}\big)$ *in solid line and* $m \mapsto \lambda^*\big(T_m, F_{Y^x}\big)$ *in dashed line.*

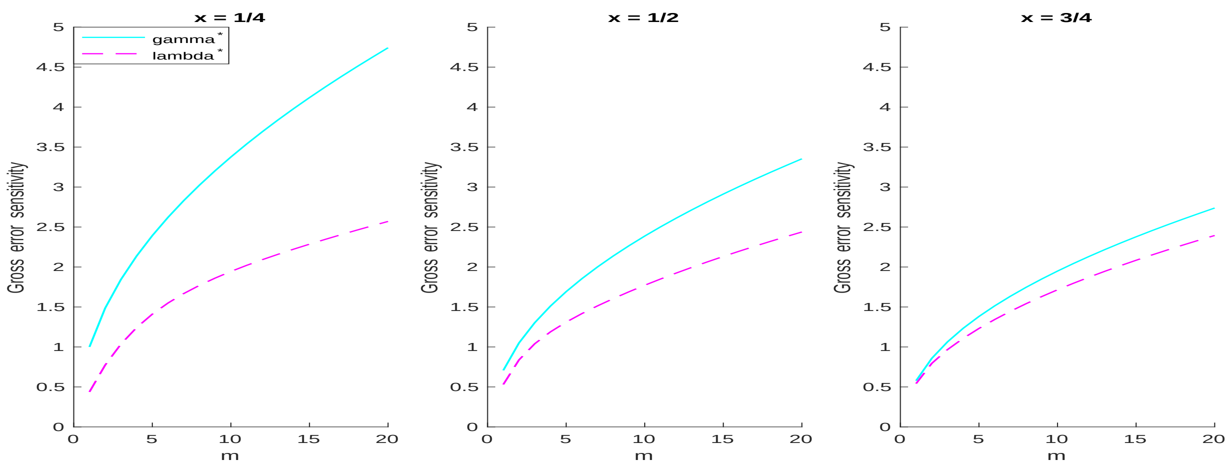

Figure 9: *Cobb-Douglas example—Gross-error sensitivities plots as before.*

and 2, along with the average $\overline{m}$ of the optimal 1,000 trimming levels m. The obtained estimates provide Monte Carlo 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 10, where the optimal parameter m of each frontier estimator was chosen in such a way to minimize its MSE.

4 Conclusions

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

	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
1000	0.0175	0.0010		-0.0312	-0.0095	21.02	163.09

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	0.0050	0.0019		-0.0104	-0.0101	21.19	51.24
500	0.0023	0.0006		-0.0147	-0.0074	51.42	150.73
1000	0.0016	0.0004		-0.0139	-0.0062	76.65	239.33

Table 2: *Cobb-Douglas example—Results averaged on 1,000 simulations.*

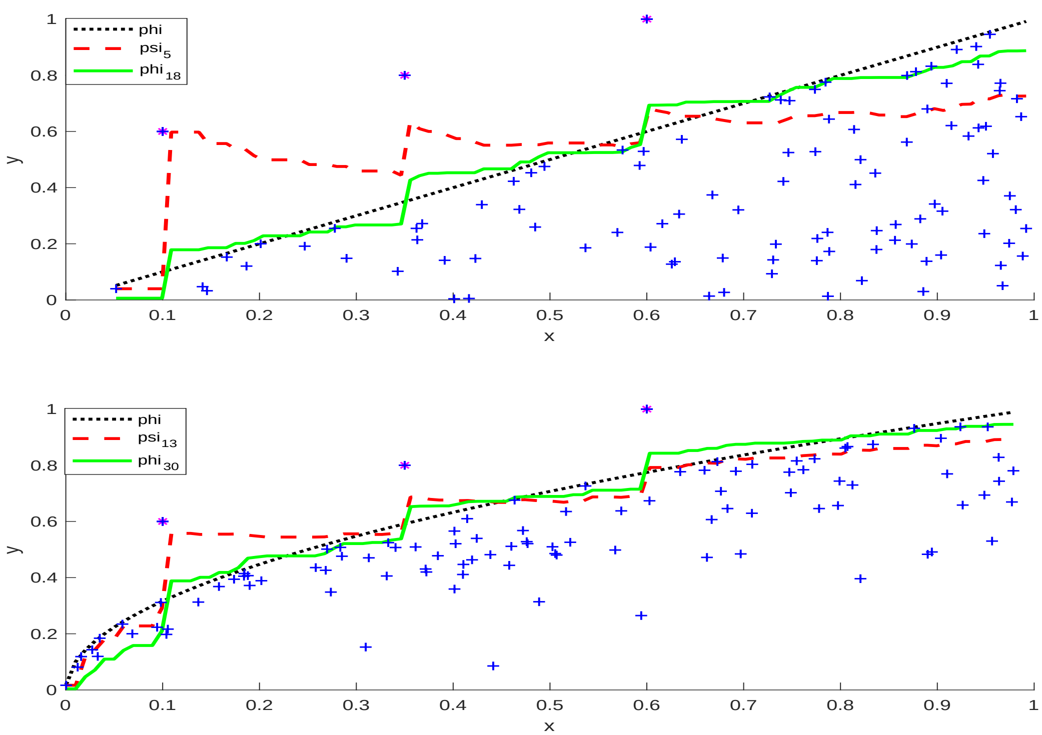

Figure 10: *Typical realizations for simulated samples of size* $n = 100$. *Top—Uniform triangle example. Bottom—Cobb-Douglas example. True frontier* φ *in dotted line with its optimal* m*−frontier estimators* $\widehat{\psi}_m$ *in dashed line and* $\widehat{\varphi}_m$ *in solid line.*

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 $Y^x = Y 1\!\!1(X \leqslant x)$, and then to define a concept of partial m-frontier $\varphi_m(x) =$

$\mathbb{E}\big[\max(Y_1^x,\dots,Y_m^x)\big]$ 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 $Y_i^x = Y_i \mathbb{1}(X_i \leqslant x)$, $i = 1,\dots,m$. 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.

Appendix: Proofs

Proof of Proposition 1. We have

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

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

Proof of Proposition 2. By definition (7) we have $\varphi_m(x) = \mathbb{E}(W_m)$, where $W_m = \max(Y_1^x,\dots,Y_m^x)$. Hence $\varphi_m(x) = \arg\min_{\theta\in\mathbb{R}} \mathbb{E}\big\{\,(W_m - \theta)^2\,\big\}$. On the other hand, it is easily seen that

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

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

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

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

Proof of Proposition 3. Putting $F_t = (1-t)F_{Y^x} + t\delta_u$ and $F_t^{-1}(1) = \inf\{y|F_t(y) = 1\}$, we have

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

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, Lemma 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

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

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

The influence function in Proposition 3 coincides with the first Gâteaux derivative of the functional T_m at F_{Y^x}. To prove Proposition 4 we need the stronger concept of Hadamard-differentiability.

Fix $m \geqslant 1$ and $x \in \mathbb{R}_+^p$ such that $F_X(x) > 0$. Define the domain $\mathbb{D}_x$ to be the set of distribution functions $F(\cdot)$ on $\mathbb{R}_+$ whose right endpoint $F^{-1}(1) := \inf\{y \geqslant 0|F(y) = 1\}$ satisfies $F^{-1}(1) \leqslant \varphi(x)$. Then, for any $F \in \mathbb{D}_x$, we have

$$T_m(F) = \int_0^{\infty} [1 - F^m(y)]dy = \int_0^{F^{-1}(1)} [1 - F^m(y)]dy = \int_0^{\varphi(x)} [1 - F^m(y)]dy.$$

In particular, we have $T_m\big(F_{Y^x}\big) \equiv \varphi_m(x)$ and $T_m\big(\widehat{F}_{Y^x}\big) \equiv \widehat{\varphi}_m(x) = \int_0^{\hat{\varphi}(x)}(1 - [\widehat{F}_{Y^x}(y)]^m)dy \overset{a.s.}{=} \int_0^{\varphi(x)}(1 - [\widehat{F}_{Y^x}(y)]^m)dy$ since $\hat{\varphi}(x) \leqslant \varphi(x)$ with probability 1.

Lemma 1. *The map $T_m : \mathbb{D}_x \subset L^\infty(\bar{\mathbb{R}}) \longrightarrow [0, \varphi(x)]$ is Hadamard-differentiable at F_{Y^x} with derivative*

$$(T_m)'_{F_{Y^x}} : h \in L^\infty(\bar{\mathbb{R}}) \longmapsto (T_m)'_{F_{Y^x}}(h) = -m \int_0^{\varphi(x)} [F_{Y^x}(y)]^{m-1} h(y)dy.$$

Proof. Let $h \in L^\infty(\bar{\mathbb{R}})$ and $h_t \to h$ uniformly in $L^\infty(\bar{\mathbb{R}})$, where $F_{Y^x} + th_t \in \mathbb{D}_x$ for all small $t > 0$. Write $\varphi_{mt}(x) := T_m(F_{Y^x} + th_t)$. Following the definition of the Hadamard differentiability [see van der Vaart (1998), p.296], we shall show that $(\varphi_{mt}(x) - \varphi_m(x))/t$ converges to $(T_m)'_{F_{Y^x}}(h)$ as $t \downarrow 0$. We have

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

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 $[F_{Y^x}(y) + th_t(y)]$ such that

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

Hence

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

It follows from the definition of $\zeta_{t,x}(y)$ and the uniform convergence $h_t \to h$ in $L^\infty(\bar{\mathbb{R}})$ that $[\zeta_{t,x}(y)]^{m-1}h_t(y)$ converges to $[F_{Y^x}(y)]^{m-1}h(y)$ uniformly in y as $t \downarrow 0$. Therefore, we obtain $\lim_{t\downarrow 0}(\varphi_{mt}(x) - \varphi_m(x))/t = (T_m)'_{F_{Y^x}}(h)$. $\square$

Proof of Proposition 4. By the Donsker Theorem, the empirical process $\sqrt{n}(\widehat{F}_{Y^x} - F_{Y^x})$ converges in distribution in $L^\infty(\overline{\mathbb{R}})$ to an F_{Y^x}-Brownian bridge $\mathbb{F}_{Y^x}$, a Gaussian process with zero mean and covariance function $E(\mathbb{F}_{Y^x}(t_1)\mathbb{F}_{Y^x}(t_2)) = F_{Y^x}(t_1 \wedge t_2) - F_{Y^x}(t_1)F_{Y^x}(t_2)$, for all $t_1, t_2 \in \overline{\mathbb{R}}$. Then, by applying the functional delta method [see van der Vaart (1998), Theorem 20.8, p.297] in conjunction with Lemma 1, we get $\sqrt{n}\{T_m(\widehat{F}_{Y^x}) - T_m(F_{Y^x}\} = (T_m)'_{F_{Y^x}}(\sqrt{n}(\widehat{F}_{Y^x} - F_{Y^x})) + o_p(1)$. $\blacksquare$

Proof of Theorem 1. 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 $\left\{\widehat{\varphi}_m(x) - \varphi_m(x)\right\} \overset{a.s.}{=} \int_0^{\varphi(x)} \left([F_{Y^x}(y)]^m - [\widehat{F}_{Y^x}(y)]^m\right) dy$, we get

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

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

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

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 $0 \leqslant \eta_{x,n}(y) \leqslant 1$ 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), Theorem A, p.35], we obtain

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

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 Theorem 2. Here we employ similar arguments of proof as in Theorem 3.1 and Lemma 3.1 of Daouia *et al.* (2012). We know by (A.1) that

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

and that $\sup_y\big|\widehat{F}_{Y^x}(y)-F_{Y^x}(y)\big| \overset{a.s.}{=} O\left((\log\log n/n)^{1/2}\right)$ in view of (A.2). 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,

$$\frac{\sqrt{n}}{\sigma(x,m)}m\int_0^{\varphi(x)}[F_{Y^x}(y)]^{m-1}\big\{F_{Y^x}(y)-\widehat{F}_{Y^x}(y)\big\}dy = \sum_{i=1}^n\frac{Z_{n,i}}{\sqrt{n}\sigma(Z_{n,i})} \tag{A.3}$$

where $Z_{n,i} = m\int_0^{\varphi(x)}[F_{Y^x}(y)]^{m-1}\big\{F_{Y^x}(y)-1\!\!I(Y_i^x\leqslant y)\big\}dy$ and its variance is $\sigma^2(Z_{n,i}) = \sigma^2(x,m)$. We have $n\mathbb{E}[|Z_{n,1}|^3]/\{n\sigma^2(Z_{n,1})\}^{3/2}\leqslant m\varphi(x)/\sqrt{n}\sigma(Z_{n,1})\to 0$ since $m/\sqrt{n}\sigma(x,m)\to 0$. Hence the leading term (A.3) converges in distribution to $\mathcal{N}(0,1)$ by Lyapounov's Theorem. Therefore $\sqrt{n}\sigma^{-1}(x,m)\big\{\widehat{\varphi}_m(x)-\varphi_m(x)\big\}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$.

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

$$\begin{aligned}\sigma^2(x,m) &= 2m^2\int_0^{\varphi(x)}\int_0^{\varphi(x)}F^m_{Y^x}(y)F^{m-1}_{Y^x}(z)[1-F_{Y^x}(z)]1\!\!I(y\leqslant z)dydz\\ &= 2m^2\int_0^{\varphi(x)}F^{m-1}_{Y^x}(z)S_{Y^x}(z)\left(\int_0^z F^m_{Y^x}(y)dy\right)dz.\end{aligned}$$

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

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

It follows from the regularity condition (2) that

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

We also have by (2) that

$$F^{2m}_{Y^x}(\varphi(x)-2\delta) \geqslant \{1-2\ell_x(2\delta)^{\rho_x}\}^{2m} = \exp\left[2m\log\{1-2\ell_x(2\delta)^{\rho_x}\}\right] \geqslant e^{-8m\ell_x(2\delta)^{\rho_x}}, \quad \delta \to 0.$$

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

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

for some constant $c_x > 0$. Whence

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

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. ■

Proof of Theorem 3. According to Daouia *et al.* (2010, Corollary 2.1), we have under (2) that

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

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

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

and show that the second term on the right-hand side $(n\ell_x)^{1/\rho_x}\left\{\widehat{\varphi}(x)-\widehat{\varphi}_m(x)\right\} = o_p(1)$ as $n \to \infty$. It follows from (9) 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 $(n\ell_x)^{1/\rho_x}\left\{\widehat{\varphi}(x)-\widehat{\varphi}_m(x)\right\}$ 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, $\left[1-\frac{1}{n}\right]^{m_n} \leqslant (Cn)^{-\beta}$ for some constants $\beta > \frac{1}{\rho_x}+1$ and $C > 0$. This condition reduces to $m_n \geqslant \beta n \log(Cn)\{1+o(1)\}$ by using the fact that $\log(1-1/n) \sim -1/n$ as $n \to \infty$.

■

References

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

[2] Aragon, Y., A. Daouia and C. Thomas-Agnan (2005), Nonparametric Frontier Estimation: A Conditional Quantile-based Approach, *Econometric Theory*, 21, 358–389.

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

[4] Daouia, A., J.P. Florens and L. Simar (2010), Frontier estimation and Extreme value theory, *Bernoulli*, Vol. 16, No. 4, 1039–1063.

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

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

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

[8] Daouia, A. and B.U. Park (2013), On Projection-Type Estimators of Multivariate Isotonic Functions, *Scandinavian Journal of Statistics*, 40, 363–386.

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

[10] Daouia, A. and L. Simar (2005), Robust Nonparametric Estimators of Monotone Boundaries, *Journal of Multivariate Analysis*, 96, 311–331.

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

[12] Daouia, A., Simar, L. and Wilson, P. (2017) Measuring Firm Performance by using Nonparametric Quantile-type Distances. *Econometric Reviews*, 36, 156–181.

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

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

[15] Florens, J.P. and L. Simar, (2005), Parametric Approximations of Nonparametric Frontier, *Journal of Econometrics*, vol 124, 1, 91–116

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

[17] Greene, W.H. (2008), The Econometric Approach to Efficiency Analysis, in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A.Knox Lovell and Shelton Schmidt, editors, Oxford University Press.

[18] Hampel, F.R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986), *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New-York.

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

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

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

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

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

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

[25] Rousseeuw, P.J. (1981), A new infinitesimal approach to robust estimation, *Z. Wahrsch. verw. Geb.* **56**, 127–132.

[26] Serfling, R.T. (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley, New-York.

[27] Simar, L. (2003), Detecting Outliers in Frontier Models: A Simple Approach. *Journal of Productivity Analysis*, 20, 391–424.

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

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

[30] Wheelock, D.C., and Wilson, P.W. (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.