Advances in Statistical Measurement in Efficiency Analysis:

From nonparametric to parametric models

UPEG (EERC- KIEV), July 5, 2005

Léopold Simar
Institut de Statistique
Université Catholique de Louvain, Belgium

Contents -1-

- Parametric models allows to investigate production processes
  - understand the structure of the technology (returns to scale, elasticities, elasticities of substitution, ...)
  - estimate and investigate efficiency scores
- Standard parametric approches (MOLS, COLS, MLE) have serious drawbacks:
  - strong restrictive assumptions on the stochastic part of the models
  - if deterministic, sensitive to extreme/outliers
  - if stochastic models, identification of noise
  - most are regression based models and capture the shape of the cloud of points near its center (not at the efficient boundary)
  - multivariate versions (on distance functions) have the same drawback.

Contents -2-

- Recent developments (Florens-Simar, 2005), Daouia, Florens and Simar (2005) offer an alternative (two stage semiparametric approach)
  - First estimate the efficient frontier using nonparametric and robust nonparametric methods
  - them fit, by standard OLS, the approriate parametric model on the obtained nonparametric frontier
- This new approach provide much more sensible estimator of the parametric frontier model and allows for some noise by tunning the robustness parameter.
- The paper:
  - robust nonparametric methods (order-m and order-α quantile frontiers)
  - parametric adjustement
  - multivariate extensions
  - some examples

Basics: output orientation $y \in \mathbb{R}$

- The production set is the set:

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

- The production process is defined by the joint cdf of (X,Y) on $\mathbb{R}_+^p \times \mathbb{R}_+$:

$$\begin{aligned} F(x,y) = \text{Prob}(X \le x, Y \le y) &= \text{Prob}(Y \le y | X \le x)\text{Prob}(X \le x) \\ &= F_{Y|X}(y|x)F_X(x), \end{aligned}$$

  where $F_{Y|X}(y|x)$ is a nonstandard conditional cdf (conditionned on $X \le x$). The support of $F(x,y)$ is $\mathbf{\Psi}$.

- If $\mathbf{\Psi}$ is free disposal, the Farrell-Debreu efficient frontier function is for all x, s.t. $F_X(x) > 0$:

$$\varphi(x) = \{y | (x, \lambda y) \notin \mathbf{\Psi}, \forall \lambda > 1\} \equiv \inf\{y | F_{Y|X}(y|x) = 1\}.$$

- $\mathbf{\Psi}$ is unknown: has to be estimated from $\mathcal{X}_n = \{(x_i, y_i) | i = 1, \ldots, n\}$.

Nonparametric estimators -1-

- The most flexible nonparametric estimator, Deprins, Simar and Tulkens (1984), is the Free Disposal Hull (FDH) estimator:

$$\widehat{\boldsymbol{\Psi}}_{\text{FDH}} = \left\{(x, y) \in \mathbb{R}_+^{p+q} \,|\, y \leq y_i,\ x \geq x_i, \quad i = 1, \ldots, n\right\}.$$

- Frontier estimator:

$$\hat{\varphi}_{FDH,n}(x) = \max_{i|x_i \leq x}(y_i) \equiv \inf\{y|\widehat{F}_{n,Y|X}(y|x) = 1\}$$

where

$$\widehat{F}_{n,Y|X}(y|x) = \frac{\sum_{i=1}^n \mathbb{I}(x_i \leq x, y_i \leq y)}{\sum_{i=1}^n \mathbb{I}(x_i \leq x)}.$$

- Sampling properties (Park, Simar and Weiner, 2000):

$$n^{1/(p+1)}(\hat{\varphi}_{FDH,n}(x) - \varphi(x)) \xrightarrow{\mathcal{L}} Weibull(\cdot)$$

Nonparametric estimators -2-

- advantage: flexibility
  - no parametric assumption for the frontier
  - no parametric assumption for the stochastics of the DGP
- drawback:
  - curse of dimensionality
  - no interpretation of the frontier (elasticities,. . .)
  - sensitivity to extreme or outliers.

Parametric estimators -1-

- Parametric model for the frontier

$$y = \varphi(x;\theta) - u,$$

  where $u \geq 0$ and the frontier function can be written with specific analytical function of x depending on a finite set of parameters $\theta \in \mathbb{R}^k$.

- Regression type estimators (Greene, 1980)
  - If u is independent of x: $E(y|x) = \varphi(x;\theta) - E(u)$, where $E(u) = \mu$ is a constant.
  - At a shift μ, $\varphi(x;\theta)$ is the regression of y on x.

- Advantages:
  - Estimation of θ is easy (MOLS, COLS, MLE)
  - $\sqrt{n}$-consistency (even for frontier level for COLS and MLE)
  - θ has economic interpretation (return to scale, elasticities,. . .)

Parametric estimators -2-

- Drawbacks:
  - whatever the method used, $\varphi(x;\hat{\theta})$ will capture the shape of the "middle" of the cloud of points, whereas the frontier (and its characteristics) are properties of the boundary of the observed cloud of points.
  - requires strong stochastic assumptions:
    - independence between u and x
    - distributional parametric assumption for COLS and MLE.
  - sensitivity to extreme or outliers.
- Stochastic parametric frontier models:
  - Main advantage: allows noise (so less sensitive to outliers or extreme)
  - Same drawbacks, but in addition:
    - more parametric assumptions
    - identification and computational problems (estimation of θ and of individual efficiencies) (see Simar and Wilson, 2005)

Partial order frontiers -1-

- Economic interpretation: a new benchmark frontier less extreme than the "full frontier".
  - Order-m
    - a unit (x, y) is benchmarked against the average maximal output reached by m peers randomly drawn from the population of units using less input than x.
    - As $m \to \infty$, order-m frontier converges to the full-frontier.
  - Order-α
    - a unit (x, y) is benchmarked against the output level not exceeded by $100(1 - \alpha)\%$ of firms in the population of units using less input than x.
    - As $\alpha \to 1$, order-α frontier converges to the full-frontier.

Partial order frontiers -2-

- In place of looking for "full frontier"

$$\varphi(x) = \inf\{y | F_{Y|X}(y|x) = 1\}.$$

define a less extreme concept:

– Order-m frontier (Cazals, Florens, Simar, 2002)

$$\begin{aligned}\varphi_m(x) &= E\left[\max(Y^1, \ldots, Y^m) | X \leq x\right] \\ &= \int_0^{\infty} \left(1 - [F_{Y|X}(y|x)]^m\right) dy\end{aligned}$$

– Order-α quantile frontier (Aragon, Daouia, Thomas-Aignan, 2002)

$$\begin{aligned}\varphi_\alpha(x) &= F^{-1}_{Y|X}(\alpha|x) \\ &= \inf\{y \in \mathbb{R}_+ | F_{Y|X}(y|x) \geq \alpha\}\end{aligned}$$

- Property (details in CFS and ADT)

$$\text{as } m \to \infty,\ \varphi_m(x) \to \varphi(x) \quad \text{and as} \quad \alpha \to 1,\ \varphi_\alpha(x) \to \varphi(x)$$

Partial order frontiers -3-

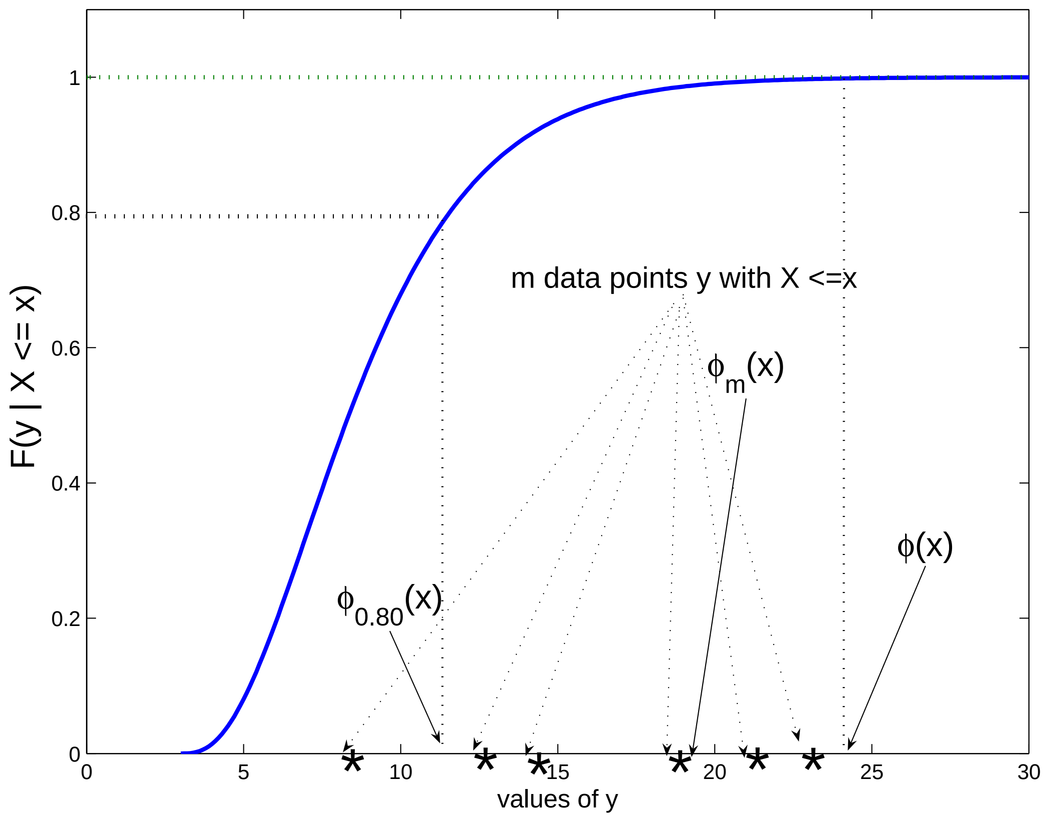


Figure 1: *Illustration of full and partial frontiers:* $m = 6$ *and* $\alpha = 0.80$

Nonparametric estimators of partial order frontier

- Plug-in principle:

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

$$\hat{\varphi}_{\alpha,n}(x) = \inf\{y \in \mathbb{R}_+ | \widehat{F}_{n,Y|X}(y|x) \geq \alpha\}$$

- Properties
  - $\sqrt{n}$-consistency and asymptotic normality when estimating partial frontier

$$\sqrt{n}(\hat{\varphi}_{m,n}(x) - \varphi_m(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_m^2(x))$$

$$\sqrt{n}(\hat{\varphi}_{\alpha,n}(x) - \varphi_\alpha(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\alpha^2(x))$$

  - Convergence to FDH estimator:

as $m \to \infty,\ \hat{\varphi}_{m,n}(x) \to \hat{\varphi}_{FDH,n}(x)$ and as $\alpha \to 1,\ \hat{\varphi}_{\alpha,n}(x) \to \hat{\varphi}_{FDH,n}(x)$

Robust estimator of "full frontier" $\varphi(x)$

- When $m \to \infty$ or $\alpha \to 1$, the partial frontiers and their nonparametric estimator converge to full frontier and to the FDH frontier respectively.

  Theorem 1. *If $m = m(n)$ is such that $m(n) = O(n \log(n))$ when $n \to \infty$ and if $\alpha = \alpha(n)$ is such that $n^{(p+2)/(p+1)}(1 - \alpha(n)) \to 0$ as $n \to \infty$, then*

$$n^{1/(p+1)}(\hat{\varphi}_{m(n),n}(x) - \varphi(x)) \xrightarrow{\mathcal{L}} Weibull(\cdot)$$
$$n^{1/(p+1)}(\hat{\varphi}_{\alpha(n),n}(x) - \varphi(x)) \xrightarrow{\mathcal{L}} Weibull(\cdot)$$

  (see CFS and ADT for details)

- Same asymptotic properties that the FDH frontier, but, for finite n, $\hat{\varphi}_{m(n),n}(x)$ and $\hat{\varphi}_{\alpha(n),n}(x)$ provide estimators of $\varphi(x)$ that will not envelop all the data points and so, are more robust to extreme or outliers.

- In practice, m and α are chosen as tuning parameters that tune the percentage of points left out the obtained partial frontier estimate.

Parametric approximations of nonparametric frontier

- The idea (Simar, 1992 and Florens and Simar, 2005):
  - Consider a parametric family of functions defined on $\mathbb{R}^p$, $\{\varphi(\cdot;\theta) | \theta \in \Theta \subset \mathbb{R}^k\}$. We want to estimate θ providing the best parametric approximation of the frontier function $\varphi(\cdot)$.
  - Pseudo-true value of θ:
$$\theta_0 = \arg\min_\theta \left[\int (\varphi(x) - \varphi(x;\theta))^2 w(x) dx \right].$$
  - Often $w(x) = f_n(x)$ (mass $1/n$ at each observed x_i):
$$\theta_0 = \arg\min_\theta \left[\sum_{i=1}^n (\varphi(x_i) - \varphi(x_i;\theta))^2 \right].$$
- Same approach for partial frontiers $\varphi_m(\cdot)$ or $\varphi_\alpha(\cdot)$. Example:

$$\theta_0^m = \arg\min_\theta \left[\sum_{i=1}^n (\varphi_m(x_i) - \varphi(x_i;\theta))^2 \right] \text{ and } \theta_0^\alpha = \arg\min_\theta \left[\sum_{i=1}^n (\varphi_\alpha(x_i) - \varphi(x_i;\theta))^2 \right]$$

Estimators of $\theta_0, \theta_0^m, \theta_0^\alpha$

- Idea (Simar, 1992): plug-in nonparametric estimators in place of unknown $\varphi, \varphi_m, \varphi_\alpha$.

$$\hat{\theta}_n = \arg\min_\theta \left[\sum_{i=1}^n (\widehat{\varphi}_{FDH,n}(x_i) - \varphi(x_i;\theta))^2 \right]$$

$$\hat{\theta}_n^m = \arg\min_\theta \left[\sum_{i=1}^n (\widehat{\varphi}_{m,n}(x_i) - \varphi(x_i;\theta))^2 \right]$$

$$\hat{\theta}_n^\alpha = \arg\min_\theta \left[\sum_{i=1}^n (\widehat{\varphi}_{\alpha,n}(x_i) - \varphi(x_i;\theta))^2 \right]$$

- Properties (Florens and Simar, 2005 and Daouia, Florens and Simar, 2005)

$$\hat{\theta}_n \xrightarrow{p} \theta_0$$

$$\sqrt{n}(\hat{\theta}_n^m - \theta_0^m) \xrightarrow{\mathcal{L}} \mathcal{N}_k(0, V_m)$$

$$\sqrt{n}(\hat{\theta}_n^\alpha - \theta_0^\alpha) \xrightarrow{\mathcal{L}} \mathcal{N}_k(0, V_\alpha)$$

Robust Estimators of θ_0 for full-frontier

- We have seen that $\widehat{\varphi}_{m(n),n}$ and $\widehat{\varphi}_{\alpha(n),n}$ are robust estimator of the full frontier φ, if $m(n) \to \infty$ and $\alpha(n) \to 1$ when $n \to \infty$.
  - Daouia, Florens, Simar (2005): if $m(n)$ and $\alpha(n)$ are such that

$$\lim_{n\to\infty} m(n) = \infty, \quad \lim_{n\to\infty} m(n)(\log n/n)^{1/2} = 0, \quad \lim_{n\to\infty} n(1-\alpha(n)) = 0$$

  - Then, as $n \to \infty$:

$$\hat{\theta}_n \xrightarrow{co.} \theta_0$$
$$\hat{\theta}_n^{m(n)} \xrightarrow{co.} \theta_0$$
$$\hat{\theta}_n^{\alpha(n)} \xrightarrow{co.} \theta_0$$

- For finite n, partial frontier estimators provide more robust estimators of θ_0.

Simulated examples -1a- (from DFS, 2005)

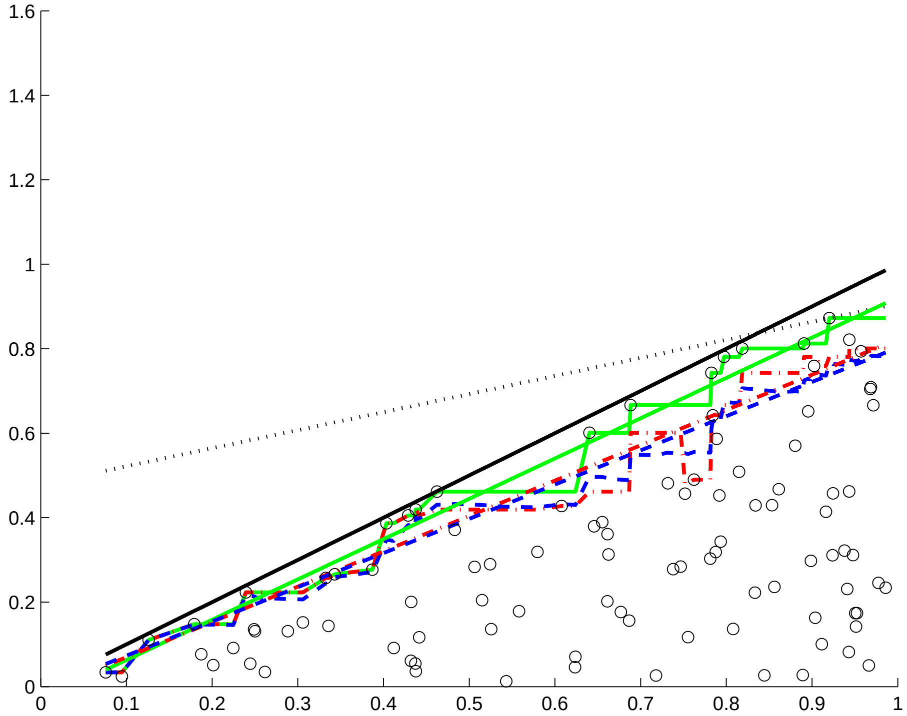


Figure 2: *In black line, the true frontier $y = x$. In cyan, the FDH frontier estimate, in blue dashed the order-m frontier and in dash-dot red the order-α frontier $\widehat{q}_{\alpha,n}$. Here, $m = 20$ and $\alpha = .9622$. In black dotted, the shifted OLS estimate.*

Simulated examples -1b-

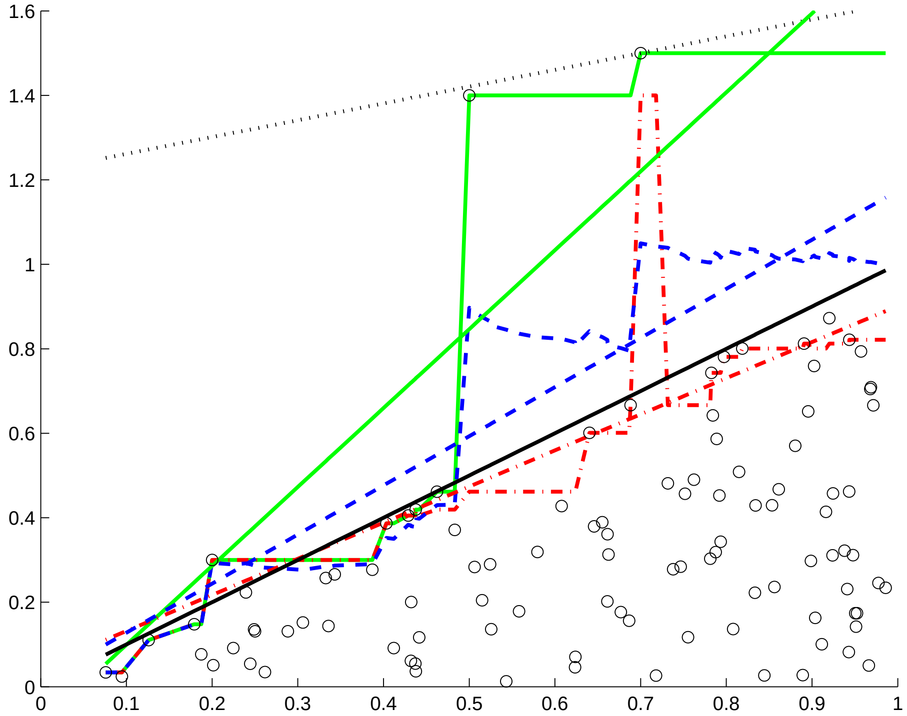


Figure 3: *Same with 3 outliers included.*

Simulated examples -2a-

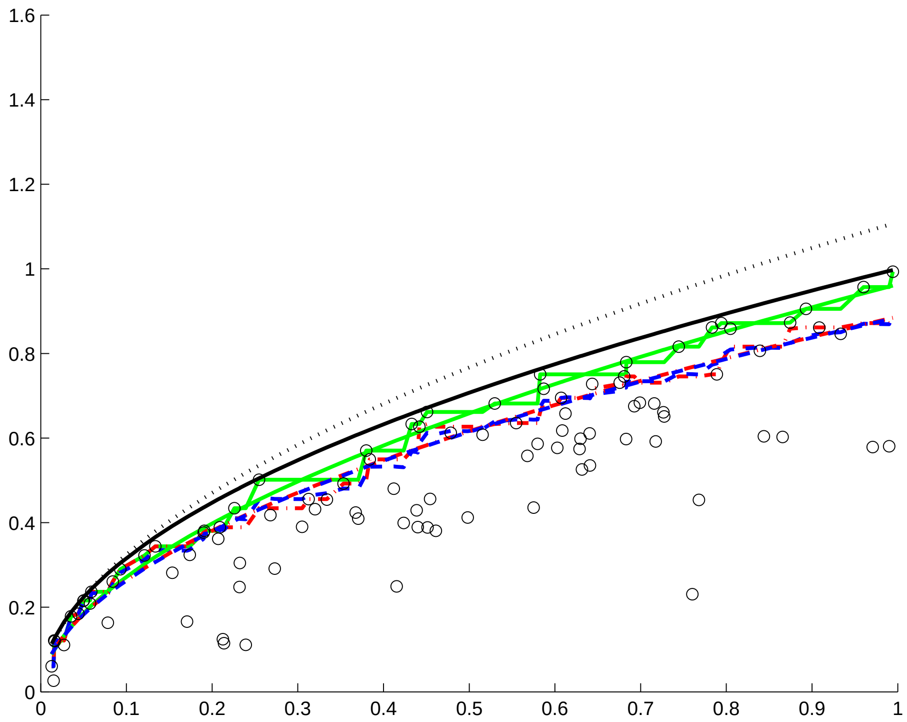


Figure 4: *In solid black line, the true frontier* $y = x^{0.5}$. *In cyan solid, the FDH frontier, in blue dashed the order-m frontier and in dash-dot red the order-*α *frontier. Here,* $m = 20$ *and* $\alpha = .9622$. *In black dotted, the shifted OLS estimate.*

Simulated examples -2b-

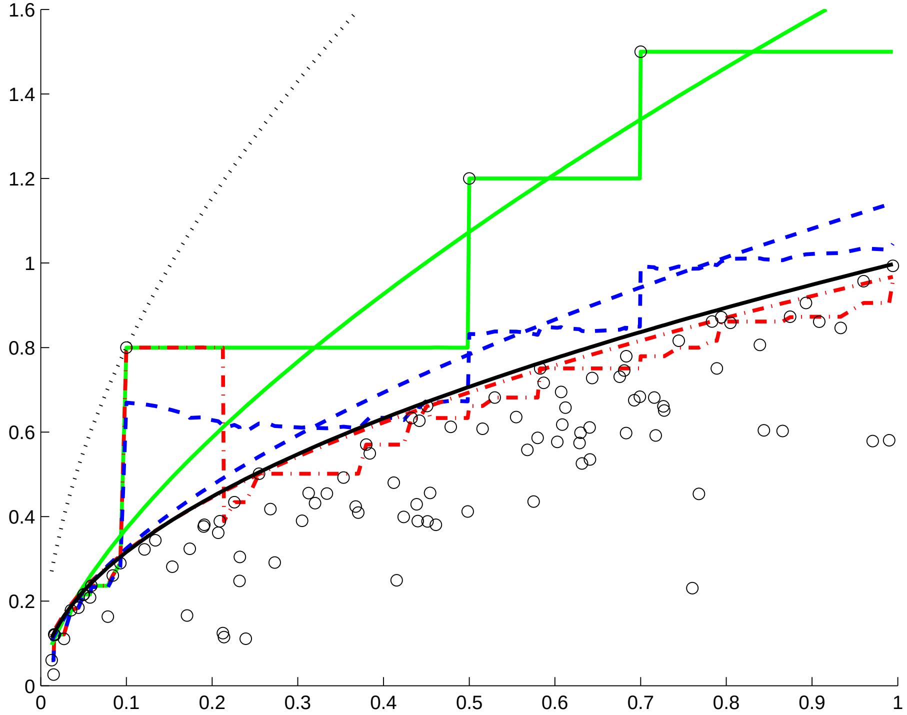


Figure 5: *Same with 3 outliers included.*

Simulated examples -3a-

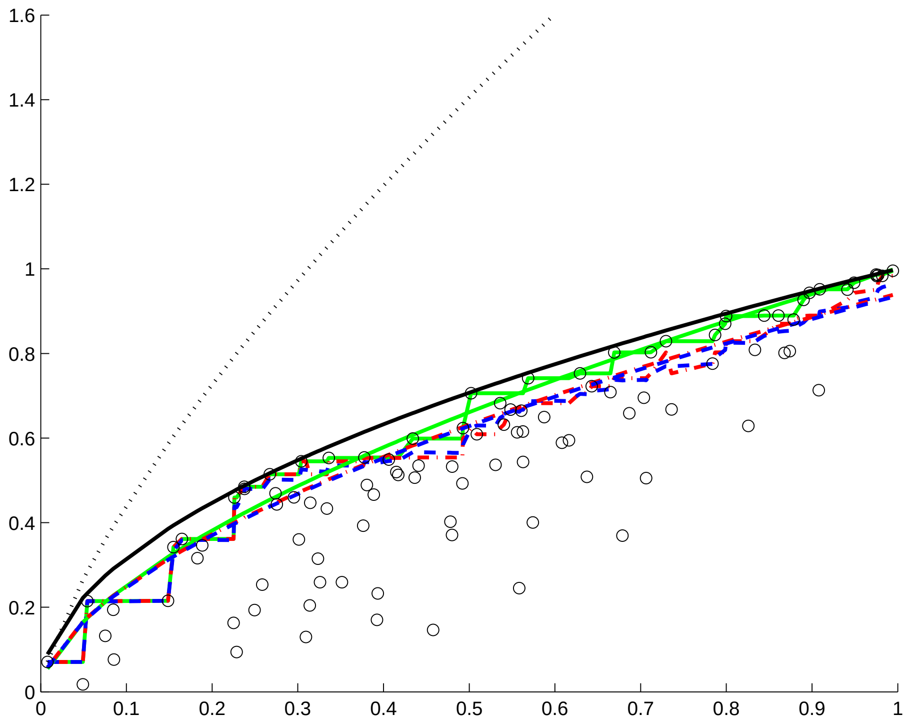


Figure 6: *Same with heteroscedasticity. In cyan solid, the FDH frontier estimate, in blue dashed the order-m frontier and in dash-dot red the order-α frontier. Here, $m = 20$ and $\alpha = .9622$. In black dotted, the shifted OLS estimate.*

Simulated examples -3b-

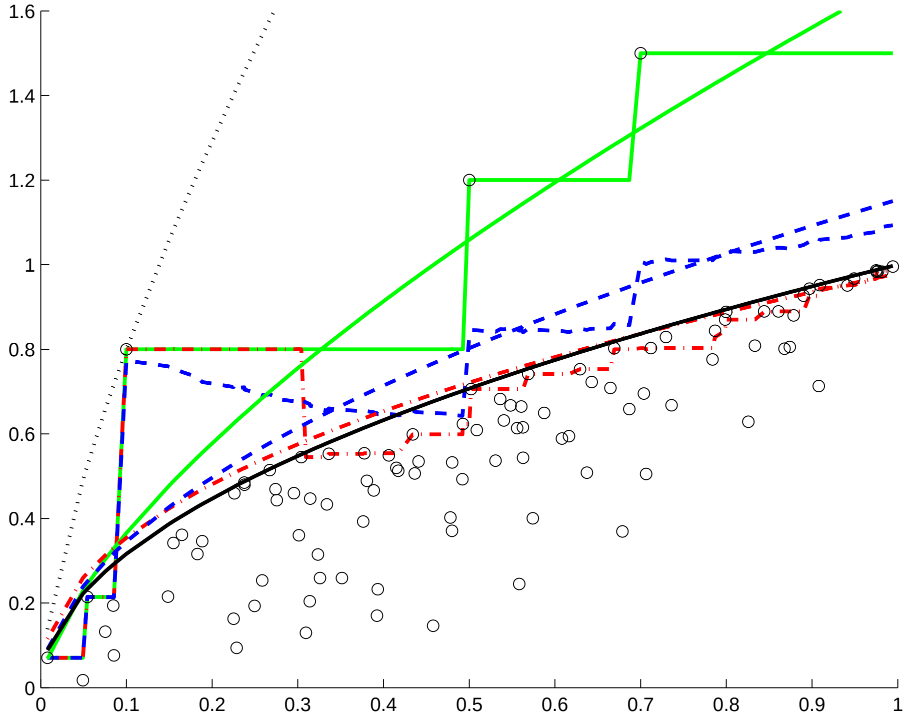
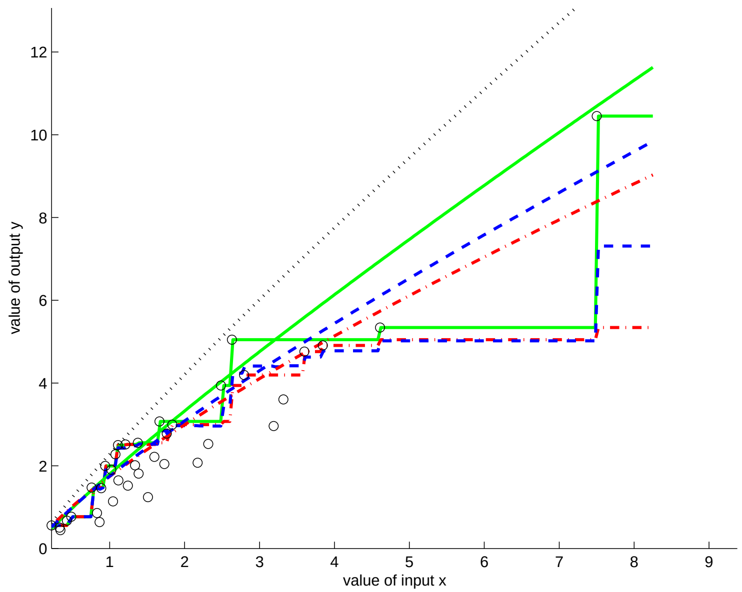


Figure 7: *Same with 3 outliers included.*

Air controlers example (DFS, 2005)



method	Intercept	Elasticity
Shifted-OLS (black-dotted)	0.8238	0.8833
FDH (cyan)	0.5886	0.8838
α-quantile (red)	0.5554	0.7798
order-m (blue)	0.5583	0.8191

Multivariate extension $(x, y) \in \mathbb{R}^p_+ \times \mathbb{R}^q_+$ -1-

- General idea: fit a parametric model to (output) distance function: Shephard (Grosskopf, Hayes, Taylor and Weber, 1997 and Coelli, 2000).
  - Same drawback as above (restricted stochastics assumptions,...)
  - Plus additional consitency issues (Coelli, 2000)
- Our approach:
  - Let $\delta(x, y)$ be the output distance function (inverse of Farrell efficiency score):

$$\delta(x, y) = \inf\{\delta | (x, y/\delta) \in \mathbf{\Psi}\}$$

  - Properties:
    * for all $(x, y) \in \mathbf{\Psi}$, $\delta(x, y) \leq 1$;
    * $\delta(x, y) = 1$, iff (x, y) is on the effcient boundary of $\mathbf{\Psi}$;
    * $\delta(x, y)$ is homogeneous of degree one in y: $\delta(x, \eta y) = \eta\delta(x, y)$.
- Also order-m ($\delta_m(x, y)$) and order-α ($\delta_\alpha(x, y)$) concepts (CFS, 2002 and Daouia, Simar, 2004)

Multivariate extension -2-

- Consider a parametric family of functions defined on $\mathbb{R}_+^p \times \mathbb{R}_+^q$, $\{\varphi(\cdot,\cdot;\theta)|\theta \in \Theta \subset \mathbb{R}^k\}$, such that

$$\forall (x,y) \text{ and } \forall \eta > 0, \quad \varphi(x, \eta y; \theta) = \eta \varphi(x,y;\theta)$$

  We want to estimate θ providing the best parametric approximation of the distance function $\delta(\cdot,\cdot)$. (same idea as in Simar and Wilson, 2003)

- Pseudo-true value of θ:

$$\theta_0 = \arg\min_\theta \left[\int (\delta(x,y) - \varphi(x,y;\theta))^2 w(x,y) dx dy \right].$$

- Often $w(x,y) = f_n(x,y)$ (mass $1/n$ at each observed (x_i, y_i)):

$$\theta_0 = \arg\min_\theta \left[\sum_{i=1}^n (\delta(x_i,y_i) - \varphi(x_i,y_i;\theta))^2 \right].$$

Multivariate extension -3-

- Estimation:

$$\hat{\theta}_n = \arg\min_{\theta} \left[\sum_{i=1}^{n} (\hat{\delta}_{FDH}(x_i, y_i) - \varphi(x_i, y_i; \theta))^2 \right].$$

- Same with partial frontiers and robust estimators of θ_0, by using $\hat{\delta}_m$ or $\hat{\delta}_\alpha$ in place of $\hat{\delta}_{FDH}$.
- By Florens-Simar (2005) and Daouia-Florens-Simar (2005): same statistical properties
  - Conistency for full frontier
  - $\sqrt{n}$ and asymptotic normality for partial frontiers parameters.

Multivariate extension: a simple example -1a-

- Generalized Cobb-Douglas parametric model for $\ln \delta(x, y)$:
  - The model:

$$\ln \delta(x, y) \approx \ln \varphi(x, y; \theta) = \alpha_0 + \alpha' \ln x + \beta' \ln y$$

  - Homogeneity of order one in y: for all $\eta > 0$,

$$\ln \varphi(x, \eta y; \theta) = \ln \eta + \ln \varphi(x, y; \theta) \quad \text{so that} \quad \beta' i_q = 1$$

- Pseudo-true value $\theta_0 = (\alpha_0, \alpha, \beta)$:

$$\theta_0 = \arg \min_{\alpha_0, \alpha, \beta_2} \left[\sum_{i=1}^{n} \left(\ln \delta(x_i, y_i) - [\alpha_0 + \alpha' \ln x_i + \beta_2' \ln y_i^2 + (1 - \beta_2' i_{q-1}) \ln y_i^1] \right)^2 \right].$$

where y^1 is the first component of y and y^2 and β_2 are the$((q-1) \times 1)$ vectors of the last $q - 1$ components of y and β. Then $\beta_1 = 1 - \beta_2' i_{q-1}$.

Multivariate extension: a simple example -1b-

- Equivalently:

$$\theta_0 = \arg \min_{\alpha_0,\alpha,\beta_2} \left[\sum_{i=1}^{n} \left(\ln(1/y_i^1)\delta(x_i, y_i) - [\alpha_0 + \alpha' \ln x_i + \beta_2' \ln \tilde{y}_i^2] \right)^2 \right].$$

where $\tilde{y}_i^2 = y_i^2 / y_i^1$

- Finally:

$$\theta_0 = \arg \min_{\alpha_0,\alpha,\beta_2} \left[\sum_{i=1}^{n} \left(- \ln y_i^{*,1} - [\alpha_0 + \alpha' \ln x_i + \beta_2' \ln \tilde{y}_i^2] \right)^2 \right].$$

where y^* is the value of y projected on the output efficient frontier: $y^* = y/\delta(x, y)$. Note that $\tilde{y}^2 \equiv \tilde{y}^{*,2} = y^{*,2}/y^{*,12}$.

- Estimation: replace the unknown $y_i^{*,1}$ by their nonparametric estimates:

$$\hat{y}^* = y/\hat{\delta}_{FDH,n}(x, y) \text{ or } \hat{y}^* = y/\hat{\delta}_{m,n}(x, y) \text{ or } \hat{y}^* = y/\hat{\delta}_{\alpha,n}(x, y)$$

Multivariate extension: a simple example -1c-

- For example, for the full-frontier approach, we compute the estimators as follows:

$$\hat{\theta}_n = \arg\min_{\alpha_0,\alpha,\beta_2} \left[\sum_{i=1}^{n} \left(-\ln \hat{y}_i^{*,1} - [\alpha_0 + \alpha' \ln x_i + \beta_2' \ln \tilde{y}_i^2] \right)^2 \right],$$

and then $\hat{\beta}_{1,n} = 1 - \hat{\beta}_{2,n}' i_{q-1}$ (the same could be done for $\hat{\theta}_n^m$ and $\hat{\theta}_n^\alpha$).

- Classical approach estimate the model (by COLS, MOLS or MLE)

$$-\ln \hat{y}_i^1 = \alpha_0 + \alpha' \ln x_i + \beta_2' \ln \tilde{y}_i^2 - u_i,$$

where $u_i > 0$ is considered as the inefficiency term. Drawbacks:

  - restrictive and non realistic assumptions on u (homoscedasticity,...)
  - problems of consistency (u in reality is not independent of $\tilde{y}^2$), see Coelli (2000)

Multivariate extension: another example -2a-

- The same can also be done for Translog model for $\ln \delta(x, y)$

$$\ln \delta(x, y) \approx \varphi(x, y; \theta) = \alpha_0 + \alpha' \ln x + \beta' \ln y + \frac{1}{2}[\ln x' \ln y'] \, \Gamma \, \begin{bmatrix} \ln x \\ \ln y \end{bmatrix}.$$

where $\Gamma = \Gamma'$ is symmetric.

- Homogeneity of degree one in y impose $p + q + 1$ constraints:

$$\begin{aligned} \beta' \, i_q &= 1, & &\text{1 constraint} \\ \Gamma_{12} \, i_q &= 0, & &p \text{ constraints} \\ \Gamma_{22} \, i_q &= 0, & &q \text{ constraints} \end{aligned}$$

$$\text{where } \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}.$$

Multivariate extension: another example -2b-

- Same as the Cobb-Douglas case, pseudo true value of θ:

$$\theta_0 = \arg \min_{\alpha_0,\alpha,\beta_2,\Gamma_{11},\widetilde{\Gamma}_{12},\widetilde{\Gamma}_{22}} \left[\sum_{i=1}^{n} \Big(-\ln y_i^{*,1} - [\alpha_0 + \alpha' \ln x_i + \tilde{\beta}_2' \ln \tilde{y}_i^2 + \frac{1}{2} \ln x_i' \, \Gamma_{11} \, \ln x_i + \ln x_i' \, \widetilde{\Gamma}_{12} \, \ln \tilde{y}_i^2 + \frac{1}{2} \ln \tilde{y}_i^{',2} \, \widetilde{\Gamma}_{22} \, \ln \tilde{y}_i^{',2}] \Big)^2 \right].$$

where: $\beta = (\beta_1 \quad \tilde{\beta}_2')'$, $\Gamma_{12} = \begin{pmatrix} a & \widetilde{\Gamma}_{12} \end{pmatrix}$ and $\Gamma_{22} = \begin{pmatrix} c_{11} & c_2' \\ c_2 & \widetilde{\Gamma}_{22} \end{pmatrix}$

with $a \in \mathbb{R}^p$, $\widetilde{\Gamma}_{12}$ is $(p \times (q-1))$, $c = (c_{11} \quad c_2')' \in \mathbb{R}^q$ and $\widetilde{\Gamma}_{22}$ is $((q-1) \times (q-1))$.

Multivariate extension: another example -2c-

- Estimation of $\alpha_0, \alpha, \beta_2, \Gamma_{11}, \widetilde{\Gamma}_{12}, \widetilde{\Gamma}_{22}$: replace the unknown $y_i^{*,1} = y_i^1/\delta(x_i, y_i)$ by the nonparametric estimates

$$\hat{y}_i^{*,1} = y_i^1/\hat{\delta}_{FDH,n}(x_i, y_i)$$

or by $y_i^1/\hat{\delta}_{m,n}(x_i, y_i)$ or by $y_i^1/\hat{\delta}_{\alpha,n}(x_i, y_i))$

- The estimator for (β_1, a, c) will be recovered from the homogeneity constraints:

$$\beta' i_q = 1 \Longrightarrow \hat{\beta}_1 = 1 - \hat{\beta}_2' i_{q-1}$$

$$\Gamma_{12}\, i_q = 0 \Longrightarrow \hat{a} = -\widehat{\widetilde{\Gamma}}_{12}\, i_{q-1}$$

$$\Gamma_{22}\, i_q = 0 \Longrightarrow \begin{cases} \hat{c}_2 = -\widehat{\widetilde{\Gamma}}_{22}\, i_{q-1} \\ \hat{c}_{11} = -\hat{c}_2\, i_{q-1} \end{cases}$$

- So we have now the full parameter estimates: $\left(\hat{\alpha}_0, \hat{\alpha}, \hat{\beta}, \widehat{\Gamma}\right)$.

Conclusions

- Parametric frontier models are useful
- Classical Estimation Methods have serious drawbacks
- Use rather the two-stage semi-parametric approach:
  - more robust to stochastic assumptions of the DGP
  - more robust to outliers (if partial frontiers are used)
  - all the asymptotic is established
- Multivariate extensions are available.

Main References

- Aigner, D.J. and S.F. Chu, 1968, On estimating the industry production function, American Economic Review, 58, 826–839.
- Aragon, Y. and Daouia, A. and Thomas-Agnan, C. (2002), Nonparametric Frontier Estimation: A Conditional Quantile-based Approach, Discussion paper, GREMAQ et LSP, Université de Toulouse. To appear in *Econometric Theory*.
- Cazals, C. Florens, J.P. and L. Simar (2002), Nonparametric Frontier Estimation: a Robust Approach , in *Journal of Econometrics*, 106, 1–25.
- Coelli, T. (2000), On the econometric estimation of the distance function representation of a production technology, CORE discussion paper 2000/42, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- Daouia, A. and L. Simar (2004), Nonparametric efficiency analysis: a multivariate conditional quantile approach, Discussion paper 0419, Institut de Statistique, UCL.

- Daouia, A., J.P. Florens and L. Simar (2005), Semi-parametric estimation of the upper boundary of a multivariate support, Discussion paper to appear, Institut de Statistique, UCL.
- 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.
- Farrell, M.J. (1957), The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A*, 120, 253–281.
- Florens, J.P. and L. Simar, (2005), Parametric Approximations of Nonparametric Frontier, *Journal of Econometrics*, vol 124, 1, 91–116
- Greene, W.H., 1980, Maximum likelihood estimation of econometric frontier, Journal of Econometrics, 13, 27–56.
- Grosskopf, S., Hayes, K., Taylor, L. and W. Weber (1997), Budget constrained frontier measures of fiscal equality and efficiency in schooling, *Review of Economics and Statistics*, 79, 116–124.

- Shephard, R.W. (1970). *Theory of Cost and Production Function.* Princeton University Press, Princeton, New-Jersey.
- Simar, L. (1992), Estimating efficiencies from frontier models with panel data: a comparison of parametric, non-parametric and semi-parametric methods with bootstrapping, *Journal of Productivity Analysis*, 3, 167-203.
- Simar, L and P. Wilson (2003), Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes, Discussion paper 0307, Institut de Statistique, Université Catholique de Louvain, Belgium, to appear in *Journal of Econometrics.*
- Simar, L. and P.W. Wilson (2005), Statistical Inference in Nonparametric Frontier Models: recent Developments and Perspectives, forthcoming in *The Measurement of Productive Efficiency*, 2nd Edition, Harold Fried, C.A.Knox Lovell and Shelton Schmidt, editors, Oxford University Press, 2005.