European Journal of Operational Research 223 (2012) 818–833

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

How to measure the impact of environmental factors in a nonparametric production model

Luiza Bădin [a,b], Cinzia Daraio [c], Léopold Simar [d,*]

[a] Department of Applied Mathematics, Bucharest Academy of Economic Studies, Bucharest, Romania
[b] Gh. Mihoc – C. Iacob Institute of Mathematical Statistics and Applied Mathematics, Bucharest, Romania
[c] Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), Sapienza University of Rome, Italy
[d] Institut de Statistique, Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Voie du Roman Pays 20, B 1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 14 July 2011
Accepted 18 June 2012
Available online 26 June 2012

Keywords:
Data Envelopment Analysis (DEA)
Free Disposal Hull (FDH)
Conditional efficiency measures
Nonparametric frontiers
Managerial efficiency
Two-stage efficiency analysis

ABSTRACT

The measurement of technical efficiency allows managers and policy makers to enhance existing differentials and potential improvements across a sample of analyzed units. The next step involves relating the obtained efficiency estimates to some external or environmental factors which may influence the production process, affect the performances and explain the efficiency differentials. Recently introduced conditional efficiency measures (Daraio and Simar, 2005, 2007a,b), including conditional FDH, conditional DEA, conditional order-m and conditional order-α, have rapidly developed into a useful tool to explore the impact of exogenous factors on the performance of Decision Making Units in a nonparametric framework. This paper contributes in a twofold fashion. It first extends previous studies by showing that a careful analysis of both full and partial conditional measures allows the disentangling of the impact of environmental factors on the production process in its two components: impact on the attainable set and/or impact on the distribution of the efficiency scores. The authors investigate these interrelationships, both from an individual and a global perspective. Second, this paper examines the impact of environmental factors on the production process in a new two-stage type approach but using conditional measures to avoid the flaws of the traditional two-stage analysis. This novel approach also provides a measure of inefficiency whitened from the main effect of the environmental factors allowing a ranking of units according to their managerial efficiency, even when facing heterogeneous environmental conditions. The paper includes an illustration on simulated samples and a real data set from the banking industry.

1. Introduction and basic notations

Productivity analysis is concerned with the evaluation of the performances of firms to identify inefficient units wherein improvements could help increase profitability or reduce costs. Most of the efficiency analysis literature has focused on the estimation of the production frontier that provides the benchmark against which the economic producers are evaluated. Nevertheless, a very important component that concerns recent studies is the explanation of efficiency differentials by including in the analysis exogenous variables or environmental factors that cannot be controlled by the producer but may influence the production process. From a managerial point of view, it is important to identify the particularities of the production process or the economic conditions that might be responsible for inefficiency as well as to detect and analyze possible influential factors that can determine changes in productivity patterns. The choice of the environmental variables has to be done on a case-by-case basis by taking into account the economic field of application.

We will first introduce the notations and the basic assumptions on the Data Generating Process (DGP) characterizing the production process in the presence of environmental factors. Let $X \in \mathbb{R}^p_+$ denote the vector of inputs and let $Y \in \mathbb{R}^q_+$ denote the vector of outputs. We consider a vector of environmental factors $Z \in \mathcal{Z} \subset \mathbb{R}^r$ that may influence the process and the productivity patterns. Firms transform quantities of inputs into outputs, but the environmental variables may affect this process. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be the probability space on which the random variables are defined; we denote by $\mathcal{P}$ the support of the joint distribution of (X,Y,Z); and we denote a particular DGP by $P \in \mathbb{P}$.

The impact and influence of Z on the production process may be multiple and can be quite different from one application to another. The effect of Z on the production may either affect the range of achievable values for the couples (X,Y), including the shape of the boundaries of the attainable set, or it may only affect the distribution of the inefficiencies inside a set with boundaries not depending on Z (only the probability of being more or less far from the efficient frontier may depend on Z), or it can affect both. Finally,

* Corresponding author.
E-mail addresses: luiza.badin@csie.ase.ro (L. Bădin), daraio@dis.uniroma1.it (C. Daraio), leopold.simar@uclouvain.be (L. Simar).

0377-2217/$ - see front matter
http://dx.doi.org/10.1016/j.ejor.2012.06.028

the environmental factors Z may also be completely independent of (X, Y).

Daraio and Simar (2005, 2007a,b), extending previous work of Cazals et al. (2002), provide a quite general and unrestricted framework to investigate the joint behavior of (X,Y,Z) from a productivity point of view. They consider a probability model that generates the variables (X,Y,Z) where the conditional distribution of (X,Y) given a particular value of Z will be of particular interest. This conditional process can be described by

$$H(x,y|z) = \mathrm{Prob}(X \leqslant x, Y \geqslant y|Z = z), \tag{1.1}$$

or any equivalent variation of it (the joint conditional density function or the joint conditional cumulative distribution function, etc.). The function $H(x,y|z)$ is simply the probability for a unit operating at level (x,y) to be dominated by firms facing the same environmental conditions z. Given that $Z = z$, the range of possible combinations of inputs $\times$ outputs, Ψ^z, is the support of $H(x,y|z)$:

$$\Psi^z = \{(x,y)|\ Z = z,\ x \text{ can produce } y\}. \tag{1.2}$$

If $H(x,y)$ denotes the unconditional probability of being dominated, we have

$$H(x,y) = \int_{\mathcal{Z}} H(x,y|z)\, f_Z(z)\, dz, \tag{1.3}$$

having support Ψ, the marginal (unconditional) attainable set defined as

$$\Psi = \{(x,y)|\ x \text{ can produce} y\} = \bigcup_{z\in\mathcal{Z}} \Psi^z. \tag{1.4}$$

Remember that the joint support of the variables (X,Y,Z) is denoted by $\mathcal{P}$. It is clear that, by construction, for all $z \in \mathcal{Z},\ \Psi^z \subseteq \Psi$.

A large part of the literature on this topic has focused on so-called two-stage analysis, where typically, some first stage estimates of the efficiency of the firms are regressed in a second stage on these additional factors to investigate their effect on efficiency. [1] Simar and Wilson (2007) clarified that these two-stage approaches are restricted to models where these factors do not influence the shape of the production set. This is the *separability* condition which states that the support of (X,Y) is not dependent of Z, equivalently

$$\text{separability condition}: \quad \Psi^z = \Psi, \quad \text{for all } z \in \mathcal{Z}. \tag{1.5}$$

In this latter case, the support of (X,Y,Z) can be written as $\mathcal{P} = \Psi \times \mathcal{Z}$, where $\times$ represents the cartesian product. As clearly illustrated by Figs. 1 and 2 in Simar and Wilson (2011b), it is important to understand the implications of condition (1.5). If the condition is verified, the only potential remaining impact of the environmental factors on the production process may be on the distribution of the efficiencies. This justifies the use of two-stage approaches as illustrated in Simar and Wilson (2007). If the condition (1.5) is not verified, the measure of the distance of a unit (x,y) to the boundary of Ψ, even if it can be well defined and estimated (details on this follow), has little economic interest, because it ignores the heterogeneity introduced by Z on the attainable set of values for (X,Y). Whether or not Ψ^z is independent of z is an empirical issue, and Daraio et al. (2010) provide a statistical procedure to test this hypothesis. The test is a global test of separability since it tests the null hypothesis $\Psi^z = \Psi, \quad \forall z \in \mathcal{Z}$ against its complement: $\exists z \in \mathcal{Z}$ such that $\Psi^z \neq \Psi$.

[1] Simar and Wilson (2007) cite more than 50 articles that use this approach. As pointed out by Simar and Wilson (2011b), a Google Scholar search returned about 1590 articles after a search on "efficiency", "two-stage", and "dea" for the period 2007–2010. A large number of these papers incorrectly use either ordinary least squares (OLS) or tobit regression in the second stage and rely on conventional methods for inference.

Banker and Natarajan (2008) suggest another model where a two-stage approach is valid, but the model heavily depends on quite restrictive and unrealistic assumptions on the production process, as described and commented in detail in Simar and Wilson (2011b). If the two-stage approach is validated by the appropriate test (Daraio et al., 2010), one can indeed estimate in the first stage the efficiency scores of the units relative to the boundary of the unconditional attainable set in the inputs $\times$ outputs space and then regress, in a second stage, the obtained efficiencies on the environmental factors. We know that even if an appropriate model is used (e.g. Logit, Truncated Normal, nonparametric truncated regression), the inference on the impact of Z on the efficiency measures has to be carefully conducted, using adapted bootstrap techniques (i.e. traditional inference is wrong: see Simar and Wilson, 2007, 2011b for details).

Another line of research, that we could call *one-stage* approaches, has been proposed in the literature (see e.g. Banker and Morey, 1986a,b for categorical external factors; Färe et al., 1989; or Färe et al., 1994, p. 223–6). Here the factors Z are considered as free disposal inputs and/or outputs which contribute to defining the attainable set $\Psi \subset \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}^r$, but which are not active in the optimization process defining the efficiency scores. In this case, and for giving just an example in the input-oriented case, the efficiency score of a unit facing external conditions z, would be defined as:

$$\theta(x,y,z) = \inf\{\theta|(\theta x,y,z) \in \Psi\}. \tag{1.6}$$

The estimator of Ψ would be obtained as usual, by adding the variables Z in defining the Free Disposal Hull (FDH) and/or the Data Envelopment Analysis (DEA) enveloping set, with a variable Z being considered as an input if it is favorable to efficiency and as an output if it is detrimental to efficiency. The drawback of this approach is twofold: first we have to know a priori Z's role on the production process, and second we assume the free disposability (and eventually convexity, if DEA is used) of the corresponding attainable extended set Ψ. Finally, this approach only admits variable Z having a monotone effect on the production process (no U-shape, or inversed-U shaped impacts are allowed; see Daraio and Simar, 2005 for further discussion).

As described, e.g. in Daraio and Simar (2007a), the two measures $H(x,y|z)$ and $H(x,y)$ allow us to define conditional and marginal efficiency scores that can be estimated by nonparametric methods. The comparison of the conditional and marginal efficiency scores can be used to investigate the impact of Z on the production process. One of the objectives of this paper is to clarify what can be learned from the analysis of these conditional efficiency scores. We will also focus on the particular role of efficiency scores relative to partial order frontiers (order-m frontiers from Cazals et al., 2002 and order-α quantile type frontiers from Daouia and Simar, 2007), that not only provide robust versions of the efficient frontier, but also allow us to investigate different aspects of the role of Z on the production process: impact on the attainable set and/or impact on the distribution of efficiencies.

In this paper, we also propose a regression-type procedure allowing us to make inference on the impact of Z on the conditional efficiency scores. Confidence intervals for the local impact of Z will be obtained by adapting subsampling ideas from Simar and Wilson (2011a). The latter analysis can be seen as a two-stage method, as described earlier, but with the great difference that here, the object regressed on Z (the conditional efficiency) is economically meaningful. The unexplained part of the conditional efficiencies can then be interpreted as a measure of managerial efficiency allowing us to rank the performance of firms facing different environmental conditions.

Summing up, in this paper, we develop a nonparametric production model where the role of these environmental factors is explicitly introduced in a non restrictive way, through conditional efficiency measures. The paper contributes in a twofold fashion. It first extends previous research on conditional measures by showing that a careful analysis of both full and partial conditional measures allows us to disentangle the impact of environmental factors on the production process in its two components: impact on the attainable set in the input $\times$ output space, and/or impact on the distribution of the efficiency level. To the best of our knowledge this was never developed before. Inference tools are then proposed by adapting appropriate bootstrap algorithms. Second, for the first time, the impact of environmental factors on the production process is also examined by using a novel *two-stage* type approach on conditional measures of efficiency to avoid the limitations of the traditional two-stage analysis. Our approach also provides a measure of inefficiency whitened from the main effect of the environmental factors. This allows us to rank the firms according to their managerial efficiency, even when facing heterogeneous environmental conditions.

The paper is organized as follows. Section 2 reviews the basic definition of marginal and conditional efficiency scores, with respect to full frontier and also to more robust partial frontiers. Then Section 3 explains how we can disentangle the impact of external factors on the production process (i.e. impact on the support of the production set and impact on the distribution of the inefficiency scores) by the analysis and comparison of conditional and unconditional efficiencies. In addition, we propose a flexible model to try to whiten the conditional efficiencies from the effect of Z, in order to derive a measure of managerial efficiency. Section 4 provides the various nonparametric estimates of the quantities of interest and offers useful guidelines to conduct inference, by using the bootstrap. We illustrate the procedure with a simulated data set, and we apply the approach to a real data set from the banking sector in Section 5.2. Section 6 summarizes the main findings and concludes the paper.

2. Marginal and conditional efficiency measures

2.1. Farrell efficiency scores

The literature on efficiency analysis proposes several ways for measuring the distance of a firm operating at the level (x_0,y_0) to the efficient boundary of the attainable set. In the lines of the pioneering work of Debreu (1951), Farrell (1957) and Shephard (1970), radial distances became very popular in the efficiency literature. They can be input or output oriented (maximal radial contraction of the inputs or maximal radial expansion of the outputs to reach the efficient boundary). Färe et al. (1985) introduced hyperbolic radial distances that avoid some of the ambiguity in choosing output or input orientation. In this case, input and output levels are adjusted simultaneously. We can define these radial measures as follows:

$$\theta(x_0, y_0) = \inf\{\theta > 0 | (\theta x_0, y_0) \in \Psi\} \tag{2.1}$$

$$\lambda(x_0, y_0) = \sup\{\lambda > 0 | (x_0, \lambda y_0) \in \Psi\} \tag{2.2}$$

$$\gamma(x_0, y_0) = \sup\{\gamma > 0 | (\gamma^{-1} x_0, \gamma y_0) \in \Psi\}. \tag{2.3}$$

In this section, we limit the technical presentation to the output orientation, but it is easy to adapt the formulae to the input-oriented and hyperbolic cases. From Cazals et al. (2002) and Daraio and Simar (2005), we know that under the assumption of free disposability of the inputs and of the outputs, these measures can be characterized by some appropriate probability function determined by $H(x,y)$. We have, for the marginal Farrell output measure of efficiency,

$$\lambda(x_0, y_0) = \sup\{\lambda > 0 | S_{Y|X}(\lambda y_0 | X \leqslant x_0) > 0\}, \tag{2.4}$$

where $S_{Y|X}(y_0|X \leqslant x_0) = \text{Prob}(Y \geqslant y_0 | X \leqslant x_0) = \frac{H(x_0,y_0)}{H(x_0,0)}$ is the (nonstandard) conditional survival function of Y, nonstandard because the condition is $X \leqslant x_0$ and not $X = x_0$.

If the firm is facing environmental factors $Z = z_0$, then Daraio and Simar (2005) define the conditional Farrell output measure of efficiency as

$$\begin{aligned} \lambda(x_0, y_0 | z_0) &= \sup\{\lambda > 0 | (x_0, \lambda y_0) \in \Psi^{z_0}\} && (2.5)\\ &= \sup\{\lambda > 0 | S_{Y|X,Z}(\lambda y_0 | X \leqslant x_0, Z = z_0) > 0\}, && (2.6) \end{aligned}$$

where

$S_{Y|X,Z}(y_0|X \leqslant x_0, Z = z_0) = \text{Prob}(Y \geqslant y_0 | X \leqslant x_0, Z = z_0) = \frac{H(x_0,y_0|z_0)}{H(x_0,0|z_0)}$ is the conditional survival function of Y, here we condition on $X \leqslant x_0$ and $Z = z_0$. Since for all $z_0 \in \mathcal{Z}$, $\Psi^{z_0} \subseteq \Psi$, we have for all $(x_0, y_0, z_0) \in \mathcal{P}$ the relations $1 \leqslant \lambda(x_0,y_0|z_0) \leqslant \lambda(x_0,y_0)$.

Daraio et al. (2010) use these two measures to conduct a global test of separability. In their approach, using unconditional and conditional efficiency measures, they propose to estimate (by using FDH or DEA techniques) a mean integrated square difference between the boundaries of $\mathcal{P}$ and $\Psi \times \mathcal{Z}$. This provides a test statistic with a sampling distribution approximated by the bootstrap.

2.2. Partial order frontiers

The literature has proposed partial frontiers and the resulting partial efficiency scores to provide robust measures of efficiencies, robust to extreme data points or outliers (a survey and a detailed analysis of these approaches can be found in Daraio and Simar, 2007a). In our setup here, this remains true when we will use partial frontiers of extreme orders, as explained later. However, when using partial frontiers of lower order, we will see that we obtain useful complementary information on the impact of Z on the distribution of the inefficiencies inside the attainable set. To save space, we limit the presentation to the output-oriented case and to the order-α quantile frontiers. The extension to other orientations (input and hyperbolic) is immediate. The case of the partial order-m frontier is described in Cazals et al. (2002) and in Daraio and Simar (2005); see also Daraio and Simar (2007a) for a general presentation and applications to real data.

2.2.1. Order-α quantile frontiers

Daouia and Simar (2007) define for any $\alpha \in (0,1]$ the order-α output efficiency score as

$$\lambda_\alpha(x_0, y_0) = \sup\{\lambda > 0 | S_{Y|X}(\lambda y_0 | X \leqslant x_0) > 1 - \alpha\}. \tag{2.7}$$

We see that if $\alpha \to 1$, $\lambda_\alpha(x_0,y_0) \to \lambda(x_0,y_0)$. If $\lambda_\alpha(x_0,y_0) = 1$, the point (x_0,y_0) belongs to the order-α quantile frontier, meaning that only $(1-\alpha) \times 100\%$ of the firms using fewer resources than x_0, dominate the unit (x_0,y_0). A value $\lambda_\alpha(x_0,y_0) < 1$ indicates a firm producing more than the level determined by the order-α frontier at x_0.

By conditioning on $Z = z_0$, Daouia and Simar (2007) similarly define the conditional order-α output efficiency score of (x_0,y_0) as

$$\lambda_\alpha(x_0, y_0 | z_0) = \sup\{\lambda > 0 | S_{Y|X,Z}(\lambda y_0 | X \leqslant x_0, Z = z_0) > 1 - \alpha\}. \tag{2.8}$$

Again, if $\alpha \to 1$, $\lambda_\alpha(x_0,y_0|z_0) \to \lambda(x_0,y_0|z_0)$.

3. What do we learn by the analysis of conditional and unconditional efficiency scores?

3.1. Individual analysis

The individual efficiency scores $\lambda(x,y)$ and $\lambda(x,y|z)$ have their usual interpretation: they measure the radial feasible proportionate increase of output a unit operating at the level (x,y) should perform to reach the efficient boundary of Ψ and Ψ^z, respectively. In case the environmental factor Z has an effect on this boundary, the first measure $\lambda(x,y)$ suffers from a lack of economic sounding, because, facing the external conditions z, this firm may not be able to reach the frontier of Ψ, that may be quite different from the one of Ψ^z. So, the conditional measure is more appropriate to evaluate the effort a firm must exert to be considered efficient. Note, however, that ranking firms according to these conditional measures can always be done, but as far as managerial efficiency is concerned, this ranking is meaningless because firms face different operating conditions, and, some external conditions may be easier (or harder) to handle than others to reach the frontier. We will see later how to derive a measure of managerial efficiency allowing us to rank the units even when they face different environmental conditions.

The analysis of the individual ratios may also be of interest: they allow us to measure, for a unit (x,y), the local effect of Z on the reachable frontier, independently of the inherent inefficiency of the unit (x,y). Indeed, $R_O(x,y|z) = \lambda(x,y|z)/\lambda(x,y) \leqslant 1$ is the ratio of the radial distances of (x,y) to the two frontiers. The inherent level of inefficiency of the unit (x,y) has been cleaned off, in the following sense:

$$R_O(x,y|z) = \frac{\lambda(x,y|z)}{\lambda(x,y)} = \frac{\|y\|\lambda(x,y|z)}{\|y\|\lambda(x,y)} = \frac{\|y_x^{\partial,z}\|}{\|y_x^{\partial}\|} \tag{3.9}$$

where $\|y\|$ is the modulus (Euclidean norm) of y and y_x^{∂} and $y_x^{\partial,z}$ are the projections of (x,y) on the efficient frontiers (unconditional and conditional, respectively), along the ray y and orthogonally to x. Clearly $\|y_x^{\partial,z}\|$ and $\|y_x^{\partial}\|$ are both independent of the inherent inefficiency of the unit (x,y). So, the ratio measures the shift of the frontier in the output direction, due to the particular value of z, along the ray y and for an input level x, whatever being the modulus of y.

This becomes even easier to see if we consider the particular case of univariate y. To be specific, in this case, the efficient boundaries can be described by maximal production functions:

$$\varphi(x) = \sup\{y|S_{Y|X}(y|X \leqslant x) > 0\} \tag{3.10}$$

$$\varphi(x|z) = \sup\{y|S_{Y|X,Z}(y|X \leqslant x, Z = z) > 0\}. \tag{3.11}$$

Here we have $R_O(x,y|z) = \varphi(x|z)/\varphi(x) \leqslant 1$, and we note that $\varphi(x) = \sup_z \varphi(x|z)$. We observe that the ratio is indeed independent of the level of output y. So, to summarize, these ratios allow us to investigate the local effect of Z of the attainable frontier itself, for a given x and a given output mix. Using efficiency scores is particularly useful when y is multidimensional.

The same can be said for the input orientation, where $R_I(x,y|z) = \theta(x,y|z)/\theta(x,y) \geqslant 1$. In the particular case where x is univariate, the efficient boundaries can be described by the minimal input functions:

$$\phi(y) = \inf\{x \in \mathbb{R}_+|F_{X|Y}(x|Y \geqslant y) > 0\}, \tag{3.12}$$

$$\phi(y|z) = \inf\{x \in \mathbb{R}_+|F_{X|Y,Z}(x|Y \geqslant y, Z = z) > 0\}, \tag{3.13}$$

where the notation introduced here is unambiguous. In this case, the ratio can be written as $R_I(x,y|z) = \phi(y|z)/\phi(y) \geqslant 1$, with the relation $\phi(y) = \inf_z \phi(y|z)$. The same analysis as the one described earlier, can be done, mutatis mutandis, for the input orientation. The top panel of Fig. 9, reported in Appendix A, illustrates possible behaviors of these minimal input functions, conditional and unconditional, in the simple case of $p = q = r = 1$.

3.2. Global analysis

The first important global analysis required in this setup is the one provided by Daraio et al. (2010), where conditional and unconditional efficiency scores are used to build some test statistics to test the separability condition (1.5). This statistic is built by measuring, in some way, the difference between the two efficient boundaries. The bootstrap is then used to find critical values. To save space, we refer the reader to Daraio et al. (2010) for the details.

Besides a global test of separability, the comparison of the individual ratios of conditional to unconditional scores as a function of Z may also be useful. However, this comparison could be misleading when wrongly conducted. In this section, we clarify exactly what can be done and how to interpret the resulting pictures, extending the previous methodologies suggested in Daraio and Simar (2005, 2007a) to more general setups.

Indeed, Daraio and Simar described the usefulness of analyzing the ratios considered as a function of z. This allows us to capture the marginal effect of Z on the frontier shifts, but this effect may change according to the level of the inputs, when frontier output ratios are considered or according to the level of the outputs, when frontier input ratios are analyzed.

This situation is explained and illustrated in detail in Appendix A, for a simple univariate scenario in the input-oriented framework. To summarize the Appendix, the interpretation of the ratios as a function of z exclusively can always be done to explore the marginal effect of z on the frontier shifts, but the picture might be rather difficult to interpret when some dependence exists between Z and both the efficient input levels and the outputs Y. Therefore, in the absence of any information, it is better to first analyze the behavior of the ratios $R_I(x,y|z)$ as a function of z, for fixed levels of the outputs y (multivariate analysis), or as illustrated in the applications below, as a joint function of both y and z. Of course, if Y is independent of Z, or in a less restrictive way, under the assumption that the shape of the boundaries of $\mathcal{P}$ in the sections $Y = y$ (in the (X,Z) space) would not change with the level y (which formally defines what we call *partial separability*), the analysis of the ratios $R_I(x,y|z)$ as a function of z is largely simplified.

For the output orientation, and for the same reasons, the analysis of the ratios $R_O(x,y|z)$ as a function of z should first be conducted for fixed levels of the inputs X. Here, for given values of the inputs x, an increasing shape for $R_O(x,y|z)$ as a function of z, would correspond to a favorable effect of Z (higher values of Z allow us to reach higher outputs, Z is acting as a freely available input) and the opposite for a decreasing shape (Z is acting as an undesirable output). Here again, under the additional assumption of partial separability, i.e. the shape of the boundaries of $\mathcal{P}$ in the sections $X = x$ (in the (Y,Z) space) would not change with the level x, the ratios $R_O(x,y|z)$ would have the same shape for all values of x, and so the analysis of the effect of Z on the efficient frontier, as a function of z only, would be simplified.

3.3. Full frontier or partial frontier?

Partial frontiers are very popular nowadays because they produce robust estimators of the efficient frontiers and of the efficiency scores, sharing nice statistical properties. Here, robustness refers to outliers or extreme data points, and we know that sometimes, outliers can mask the effect of Z on the production process (see Daraio and Simar, 2007a, Section 5.4.1 for details). In our setup here, we clarify what the partial scores and their corresponding ratios, e.g. $R_{O,\alpha}(x,y|z) = \lambda_\alpha(x,y|z)/\lambda_\alpha(x,y)$ can add to the analysis of the effect of Z on the production process. Here we will focus the presentation on the output orientation.

First, as mentioned earlier, it is important to remember that the ratios $R_O(x,y|z)$, when defined relative to the full frontiers, only bring information on potential differences between the boundaries of Ψ and Ψ^z. They are not sensitive to changes in the distribution of inefficiencies. We saw earlier that the measure $R_O(x,y|z) \leqslant 1$ for a fixed point (x,y) only depends on the relative position of the boundaries of Ψ and Ψ^z (in the radial direction given by y). It is no more the case for partial frontiers: the values of $\lambda_\alpha(x,y)$ and $\lambda_\alpha(x,y|z)$ do not depend only on the boundary, they also depend on the effect of Z on the distribution of the output Y inside Ψ^z, conditionally to $X \leqslant x$. It is easy to see that the ratios $R_{O,\alpha}(x,y|z) = \lambda_\alpha(x,y|z)/\lambda_\alpha(x,y)$ could be either $\leqslant 1$ or $\geqslant 1$, depending on the actual effect of Z on the distribution of Y given $X \leqslant x$ (when conditioning on $Z = z$). We illustrate these facts in Appendix B, in the simple case of a univariate output y and a univariate z.

So, to summarize the analysis of Appendix B, we see that if α is near 1, the partial measures provide the same information as the full measure but they use more robust frontiers. Using small values of α could be misleading, without having a clear picture of the separability condition. Under the latter, the analysis with small values of α (e.g. $\alpha = 0.5$: median frontier) provides complementary information on the effect of Z on the distribution of the inefficiencies (its median value when $\alpha = 0.5$). The same analysis could be done, mutatis mutandis, for the input-orientation and for partial order-m frontiers.

3.4. Second-stage regression and managerial efficiency

The idea of regressing efficiency scores on the environmental variables to estimate the average effect of Z on the efficiency is quite old. However, as pointed out in Section 1 earlier, and in details in Simar and Wilson (2007, 2011b), this second-stage regression is meaningless, or at minimum difficult to interpret, when using the unconditional efficiency scores $\lambda(x,y)$. Indeed, if the separability assumption (1.5) is not verified, the unconditional efficiencies are relative to the boundary of Ψ defined by (1.4), which has no economic meaning for firms facing different environmental conditions, i.e. facing different attainable sets Ψ^z.[2] Note that comparing two firms facing different values of z by comparing their value $\lambda(x,y|Z = z)$ is unfair.

It is therefore much more meaningful to analyze the average behavior of $\lambda(x,y|z)$ as a function of z, to capture the main effect of Z on these conditional measures.[3] It is clear that here too, the conditional measures $\lambda(x,y|Z = z)$ may vary with both x and z. However, here we want to capture the marginal effect of Z on the efficiency scores, so it is legitimate to analyze the regression $\mathbb{E}(\lambda(X,Y|Z)|Z=z)$ as a function of z. We suggest using a flexible regression model defining $\mu(z) = \mathbb{E}(\lambda(X,Y|Z)|Z=z)$, and $\sigma^2(z) = \mathbb{V}(\lambda(X,Y|Z)|Z=z)$ thus, we may write

$$\lambda(X,Y|Z=z) = \mu(z) + \sigma(z)\varepsilon, \tag{3.14}$$

where $\mathbb{E}(\varepsilon|Z=z) = 0$ and $\mathbb{V}(\varepsilon|Z=z) = 1$. In the next section we will briefly address the problem of estimating the functions $\mu(z)$ and $\sigma(z)$ in a nonparametric way, with some guidelines for a bootstrap procedure for getting confidence interval for $\mu(z)$, at any given value z. Whereas $\mu(z)$ measures the average effect of z on the efficiency, $\sigma(z)$ provides additional information on the dispersion of the efficiency distribution as a function of z.

Another important result of this approach is the analysis of the residuals. For a particular given unit (x,y,z), we can define the error term

$$\varepsilon = \frac{\lambda(x,y|z) - \mu(z)}{\sigma(z)}. \tag{3.15}$$

This can be viewed as the unexplained part of the conditional efficiency score. If Z and ε do not show a strong correlation in (3.14), this quantity can be interpreted as a pure efficiency measure of the unit (x,y).[4] If Z and ε show some correlation, still the quantity defined in (3.15) can be used as a proxy for measuring the managerial efficiency, since it is the remaining part of the conditional efficiency after removing the location and scale effect due to Z. We call it managerial because it depends only upon the managers' ability and not upon the environmental factors. What we have done is a kind of whitening of the conditional efficiency scores from the effects due to the environmental conditions Z. We can use these quantities, which are standardized (mean zero and variance one), to compare the firms between them: a large value of ε indicates a unit which has poor performance, even after eliminating the main effects of the environmental factors. A small (negative) value, on the contrary, indicates very good managerial performance of the firm (x,y,z). It allows us to rank the firms facing different environmental conditions, because the main effects of these factors have been eliminated. Extreme (unexpected) values of ε would also warn for potential outliers.

This analysis could also be performed using partial efficiency scores, like $\lambda_\alpha(X,Y|Z = z)$. When α is near 1, it would provide a robust version of this analysis. We will also later review that the quality of the estimation is better when using partial efficiency measures (better rates of convergence).

4. Nonparametric estimator

4.1. Efficiency estimators

Nonparametric estimators of the conditional and unconditional efficiency scores are very easy to obtain. We summarize the notations and properties here as to what is needed for the remainder paper (details can be found in Daraio and Simar, 2007a, or Simar and Wilson, 2008). We will denote $\mathcal{S}_n = \{(X_i,Y_i,Z_i) \mid i = 1,\ldots,n\}$ the sample of n iid observations on (X,Y,Z) generated in $\mathcal{P}$ according to the DGP $P \in \mathbb{P}$. If we plug nonparametric estimators of $S_{Y|X}$ and $S_{Y|X,Z}$ into all the formulae mentioned earlier, we obtain very natural nonparametric estimators of the efficiencies. For the $S_{Y|X}$ we can use the empirical probabilities

$$\widehat{S}_{Y|X}(y_0|X \leqslant x_0) = \frac{1/n\sum_{i=1}^n \mathbb{I}(X_i \leqslant x_0, Y_i \geqslant y_0)}{1/n\sum_{i=1}^n \mathbb{I}(X_i \leqslant x_0)}, \tag{4.1}$$

where $\mathbb{I}(\cdot)$ is the indicator function. This provides the popular FDH estimator of $\lambda(x_0,y_0)$

$$\hat{\lambda}(x_0,y_0) = \max_{\{i|X_i \leqslant x_0\}} \left\{ \min_{j=1,\ldots,q} \frac{Y_i^j}{y_0^j} \right\}, \tag{4.2}$$

whose statistical properties are well known (see e.g. Simar and Wilson, 2008). To summarize, under mild regularity conditions:

$$n^{1/(p+q)}(\lambda(x_0,y_0) - \hat{\lambda}(x_0,y_0)) \xrightarrow{\mathcal{L}} \text{Weibull}(\mu_0^{p+q}, p+q), \tag{4.3}$$

where μ_0 is a constant depending on the DGP $P \in \mathbb{P}$ that is described in Park et al. (2000). For the conditional (conditional to $Z = z_0$) some smoothing techniques are required. We have the estimator

$$\widehat{S}_{Y|X,Z}(y_0|X \leqslant x_0, Z = z_0) = \frac{1/n\sum_{i=1}^n \mathbb{I}(X_i \leqslant x_0, Y_i \geqslant y_0)K((z_0 - Z_i)/b)}{1/n\sum_{i=1}^n \mathbb{I}(X_i \leqslant x_0)K((z_0 - Z_i)/b)}, \tag{4.4}$$

where for simplicity, we wrote the expression for a univariate Z. Here $K(\cdot)$ is a kernel with compact support and $b > 0$ is the bandwidth. For the general multivariate case, see Daraio and Simar (2007a). In the general multivariate setup, an optimal bandwidth selection procedure has been suggested in Bădin et al. (2010), based on a least-squares cross validation technique. This leads to the conditional efficiency estimator

[2] This is a relevant empirical issue due to the great number of papers that have appeared in recent years. See e.g. Kao and Hwang (2008); Chen et al. (2009a,b); and Zha and Liang (2010), just to cite a few of the most recent.

[3] As a matter of fact, since $\lambda(X,Y|Z = z)$ is a ratio of two output levels, and since an additive model will be suggested in (3.14), it might be more appropriate to perform this second-stage regression on the log $\lambda(X,Y|Z = z)$. This is an empirical issue, and we come back to this in the empirical illustrations.

[4] Our approach may be seen as a recent advanced and robust interpretation of Leibenstein's (1966, 1979) X-inefficiency theory which has among its proximate causes those related to the performance of management (see also Leibenstein and Maital, 1992).

$$\hat{\lambda}(x_0, y_0|z_0) = \max_{\{i|X_i \leqslant x_0, \|Z_i - z_0\| \leqslant b\}} \left\{ \min_{j=1,\dots,q} \frac{Y_i^j}{y_0^j} \right\}. \tag{4.5}$$

So, it appears that the estimation of the conditional efficiency score is a kind of restricted FDH program (restricted to data points having $\|Z_i - z_0\| \leqslant b$). The statistical properties of the estimators of the conditional measures have been determined in Jeong et al., 2010. To summarize and roughly speaking, these estimators maintain similar properties as the FDH estimator but with an effective sample size depending on the bandwidth: n is replaced by nb^r, where r is the dimension of Z. In practice, since the optimal bandwidth has a size $n^{-1/(r+4)}$ (see Bădin et al., 2010, for details), this gives a rate of convergence for the conditional measures estimators of $n^{4/((r+4)(p+q))}$ in place of the better rate $n^{1/(p+q)}$ achieved by the FDH estimators. It is important to report these rates to ensure consistent subsequent bootstrap algorithms.

The nonparametric partial frontier efficiency estimates are obtained in a similar way, by plugging the estimators $\widehat{S}_{Y|X}$ and $\widehat{S}_{Y|X,Z}$ in the expressions defining the partial efficiency measures: algorithms have been proposed in Cazals et al. (2002), Daraio and Simar (2005, 2007a) for the order-m case and in Daouia and Simar (2007) and Daraio and Simar (2007a) for the order-α quantile case. Their statistical properties have been also established. Under mild regularity conditions, we have for instance

$$\sqrt{n}\left(\lambda_\alpha(x_0, y_0) - \hat{\lambda}_\alpha(x_0, y_0)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(\alpha, x_0)), \tag{4.6}$$

where an expression for $\sigma^2(\alpha,x_0)$ is given in Daraio and Simar (2006). A similar result holds for the order-m case (see Cazals et al., 2002).

For the estimators of the conditional partial measures, we have similar results where the rate of convergence $\sqrt{n}$ deteriorates to $\sqrt{nb^r} = n^{2/(r+4)}$ when the optimal bandwidth of Bădin et al. (2010) described here is used.

4.1.1. Robust estimators of the full frontier

As explained earlier, the partial frontiers have their particular usefulness providing less extreme surfaces to benchmark individual units and allowing us to investigate the impact of Z on the distribution of the efficiencies. In particular, for $m = 1$, the order-m frontier is not looking to an optimal behavior but rather to an average behavior of firms (the same is true for the order-α frontier with $\alpha = 0.50$).

But as discussed and illustrated in Daraio and Simar (2007a), it may happen that outliers or extreme data points can hide the real effect of environmental factors. So, in this case, it is particularly useful to build robust estimators of the full frontier. This can be achieved by using partial order frontier with extreme orders.

Indeed, if we let $\alpha = \alpha(n) \to 1$ (or $m = m(n) \to \infty$) when $n \to \infty$ fast enough (see Cazals et al., 2002 and Daraio and Simar (2006), for details), the respective partial frontier estimators will converge to the full frontier sharing the same properties as the FDH estimator (with the same limiting Weibull distribution). But for finite n (as we use in practice), $\alpha(n)$ will be less than 1 (and $m(n)$ will be less than infinity) and so the corresponding estimate of the full frontier will not envelop all the data points being more robust and resistant to outliers and extreme values than the standard envelopment estimators like FDH or DEA.

Simar (2003) and Daraio and Simar (2007a) have suggested some data-driven techniques to select reasonable values of α and m by analyzing the proportion of data points remaining outside the corresponding partial frontiers over a grid of values of the orders. This allows us to detect potential outliers. Daouia and Gijbels (2011a) provide a theoretical background for the comparison of both partial frontiers in terms of their robustness properties; see also Daouia and Gijbels (2011b), where a theoretical rule is given to select the appropriate order of the partial frontiers for obtaining robust estimators of the full frontier in the presence of outliers.

4.1.2. Estimation of the ratios $R_O(x,y|z)$ and $R_{O,\alpha}(x,y|z)$

Consistent estimators of the ratios are directly obtained by plugging the nonparametric estimators derived earlier into the corresponding formulae. So we have

$$\widehat{R}_O(x, y|z) = \frac{\widehat{\lambda}(x, y|z)}{\widehat{\lambda}(x, y)}. \tag{4.7}$$

For any given point (x,y,z), it is easy to prove that $\widehat{R}_O(x,y|z)$ is a consistent estimator of $R_O(x,y|z)$, sharing the worst rate of convergence of its components, i.e. the numerator. The limiting distribution of the error is rather complicated, but it can be shown (see Daraio et al., 2010, for details) that

$$n^{\kappa}(\widehat{R}_O(x, y|z) - R_O(x, y|z)) \xrightarrow{\mathcal{L}} Q_P^z(\cdot), \tag{4.8}$$

where the rate of convergence was given, $\kappa = 4/((r + 4)(p + q))$ and where Q_P^z is a nondegenerate distribution (i.e. it is not a Dirac distribution with mass 1 at one single value) depending on the current value of z and on characteristics of the DGP P.

The partial measures will benefit from the better rate of convergence of the individual efficiency estimators. We have

$$\widehat{R}_{O,\alpha}(x, y|z) = \frac{\widehat{\lambda}_\alpha(x, y|z)}{\widehat{\lambda}_\alpha(x, y)} \tag{4.9}$$

and the asymptotic distribution of the error of estimation follows

$$n^{\gamma}\left(\widehat{R}_{O,\alpha}(x, y|z) - R_{O,\alpha}(x, y|z)\right) \xrightarrow{\mathcal{L}} Q_P^{z,\alpha}(\cdot), \tag{4.10}$$

where here the rate is $\gamma = 2/(r + 4)$ and where $Q_P^{z,\alpha}$ is another nondegenerate limiting distribution.

Note that the unit (x,y,z) of interest can be any point in $\mathcal{P}$, even if in practice we will be interested to estimate these ratios at the observed data points (X_i,Y_i,Z_i).

Since the limiting distributions are unknown, the bootstrap is the only available route to draw inference on these individual ratios. Here we can directly apply the subsampling procedure described in Simar and Wilson (2011a) to derive confidence intervals for $R_O(x,y|z)$ (and for the partial correspondents). To save space we refer simply to the algorithms described in Simar and Wilson's paper (Section 4.2), the adaptation being straightforward. Just to avoid misunderstandings, we sketch the subsampling algorithm:

[1] First, we compute from the sample $\mathcal{S}_n = \{(X_i, Y_i, Z_i)| i = 1, \dots, n\}$ the efficiency score $\hat{\lambda}(x, y)$ and its conditional version $\hat{\lambda}(x, y|z)$. By doing so, we compute $b_{n,z}$ the optimal bandwidth for the conditional survival function at z (we do this by using the Bădin et al. (2010) approach). We compute the ratio $\widehat{R}_O(x,y|z)$.

[2] For a given value of $m < n$, we repeat the next steps [2.1]–[2.2] L times, for $\ell = 1,\dots,L$, where L is large enough (say, $L = 2000$).

  [2.1] Draw a random sample $\mathcal{S}^*_{m,\ell} = \{(X_i^{*,\ell}, Y_i^{*,\ell}, Z_i^{*,\ell})| i = 1, \dots, m\}$ without replacement from $\mathcal{S}_n$.

  [2.2] Compute the ratio $\widehat{R}_O^{*,\ell}(x, y|z)$ by the same techniques as in [1]. Note that here we have to rescale the corresponding bandwidth at the appropriate size. So we will use the bandwidths $b_{m,z} = (n/m)^{1/(r+4)} b_{n,z}$ for computing the conditional scores with the bootstrap sample $\mathcal{S}^*_{m,\ell}$.[5]

[5] As pointed out in Jeong and Simar (2006), if the point of interest (x,y,z) is rather extreme with respect to the cloud of data $\mathcal{S}_n$, some of the FDH estimators (conditional and unconditional) may be undefined when computed relative to a bootstrap sample $\mathcal{S}^*_m$. This is a small sample issue, and this event should disappear asymptotically. In this case, as Jeong and Simar recommend, we define the estimators as being equal to 1. This does not alter the asymptotic consistency of the bootstrap.

[3] From the collection of L values $\widehat{R}_O^{*,\ell}(x,y|z)$ with $\widehat{R}_O(x,y|z)$, build the confidence interval obtained for this particular value of m.
[4] Select the appropriate value of the subsample size m, which will correspond to a value where the results show low volatility with respect to m (see Simar and Wilson, 2011a).

4.2. Second-stage regression of the conditional efficiency scores

To save space we present only the full-frontier case, where we want to estimate $\mu(z)$ and $\sigma(z)$ in model (3.14) by using basic tools from the nonparametric econometrics literature. Our presentation is for continuous Z.[6]

Several flexible nonparametric estimators could be provided when working with the full-frontier efficiency scores. We know indeed that, by definition, $\lambda(X,Y|Z=z) \geqslant 1$ with probability one. So, we could for instance estimate in a first step $\mu(z)$, by local constant methods (Pagan and Ullah, 1999) or local exponential smoothing (see Ziegelmann, 2002) on the values $\lambda(X,Y|Z=z) - 1$. The estimation of the variance function $\sigma^2(z)$ is rather standard (see Fan and Gijbels, 1996; Fan and Yao, 1998; Pagan and Ullah, 1999; and the references therein) and is obtained by regressing the squares of the residuals obtained from the first step, on Z. Here again, local constant or local exponential methods can be used. Once $\hat{\mu}(z)$ and $\hat{\sigma}^2(z)$ are obtained, we can compute the residuals by applying (3.15).[7]

Whatever the selected nonparametric estimators, they share typically similar asymptotic properties with the same rate of convergence. As pointed out in Simar and Wilson (2007), the main statistical issue in this second-stage regression comes from the fact that we have neither observations of $\lambda(X_i,Y_i|Z=z)$ nor observations $\lambda(X_i,Y_i|Z_i)$ because the lambda's are unknown. What we have is the set of the n estimators $\hat{\lambda}(X_i,Y_i|Z_i)$, obtained from the sample $\mathcal{S}_n$. So we have a sample of n pairs $\left(Z_i, \widehat{\lambda}(X_i,Y_i|Z_i)\right)$, $i = 1,\ldots,n$ from which we will estimate $\mu(z)$ and $\sigma(z)$, by one of the methods described earlier. For the regression, the resulting estimator will be written $\hat{\mu}_n(z)$. All these techniques involve smoothing (localizing) techniques requiring the selection of bandwidths h_z, for the Z variables. Bandwidths h_z with appropriate size (i.e. $h_z = c\, n^{-1/(r+4)}$) can be obtained by least-squares cross validation criterion (see e.g. Li and Racine, 2007 for details).

If the true but unavailable independent $\lambda(X_i,Y_i|Z_i)$ would be used as dependent variable in the regression, standard tools would provide an estimator $\tilde{\mu}_n(z)$ with standard properties, i.e., as $n \to \infty$ and $h_z \to 0$ with $nh_z^r \to \infty$

$$\sqrt{nh_z^r}\left(\tilde{\mu}_n(z) - \mu(z) - h_z^2 B^z\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, V^z), \tag{4.11}$$

where the bias B^z and the variance V^z are bounded constants depending on the model and on the chosen estimator. Balancing between bias and variance, the optimal bandwidth should indeed be of the order $h_z = cn^{-1/(r+4)}$, providing the asymptotic result

$$n^{2/(r+4)}(\tilde{\mu}_n(z) - \mu(z)) \xrightarrow{\mathcal{L}} \mathcal{N}(B^z, V^z). \tag{4.12}$$

When replacing $\lambda(X_i,Y_i|Z_i)$ by $\hat{\lambda}(X_i,Y_i|Z_i)$, which are estimators with n^{κ} rate of convergence, we obtain the estimator $\hat{\mu}_n(z)$. By using the same arguments as in Kneip et al. (2012), it can be shown that it is a consistent estimator of $\mu(z)$ but at the rate of convergence $n^{4/((r+4)(p+q))}$, which becomes lower as soon as $p + q > 2$. Asymptotic behavior of the error $(\hat{\mu}_n(z) - \mu(z))$ should also be available along the recent theoretical results developed in Kneip et al. (2012), allowing us to develop bootstrap algorithms and derive confidence intervals for $\mu(z)$, but this is rather technical and beyond the scope of this paper.

5. Numerical illustrations

5.1. Simulated examples

To illustrate how the procedure can work in practice, we first introduce some simulated examples, because there we know what we expect to find. We will use, as simulated scenario, an example inspired from Simar and Wilson (2011b) where we see clearly the two different ways an environmental factor can influence the production process. We analyze the three following different DGPs:

$$Y = g(X)e^{-U} \tag{5.1}$$

$$Y^* = g(X)e^{-U|Z-2|} \tag{5.2}$$

$$Y^{**} = g(X)(1 + |Z-2|/2)^{1/2}\ e^{-U}, \tag{5.3}$$

where $g(X) = [1 - (X-1)^2]^{1/2}$ with $X \sim U(0,1)$ and $Z \sim U(0,4)$. Finally $U \geqslant 0$ with $U \sim \mathcal{N}^+(0,\sigma_U^2)$, and we choose for the illustration $\sigma_U^2 = 0.05$.

In DGP1 (5.1), Z has no effect on the production process (Z is independent of (X,Y)). In DGP2 (5.2), we have the separability condition $\Psi^z \equiv \Psi$, $\forall z$ but Z influences the distribution of the inefficiencies (higher probability of being inefficient when $|Z-2|$ increases). In DGP3 (5.3), the effect of Z is only on the boundary of the attainable (X,Y), violating the separability condition, the shift (increasing the level of the attainable frontier) is multiplicative and more important when $|Z-2|$ increases.

5.1.1. Analysis of the ratios $\widehat{R}_O(x,y|z)$

We first investigate how the ratios $\widehat{R}_O(x,y|z)$ can inform us of the potential shifts of the frontier due to the environmental factor Z. We look at $\widehat{R}_O(x,y|z)$ as a function of Z and of X. Fig. 1 displays a summary of the results for the case $n = 200$ (the case $n = 100$ gave similar plots with more sampling noise). Here, DGP1 corresponds to the top panels, then DGP2 is in the middle and DGP3 in the bottom panels. Since it is not easy to visualize three-dimensional pictures without rotating the pictures (as most up-to-date software allows on screen), we display in the right panels the two marginal views from the X-side (marginal effect of X) and from the Z-side (marginal effect of Z). The results are as we expected from comments in Section 3.2, taking into account the fact that we use estimates with low rates of convergence ($n^{4/((r+4)(p+q))} = n^{2/5}$) in place of true values. For the three DGPs, the clouds of points are flat from the perspective of X, because in the three cases, the effect of Z on the efficient frontier, if any, is independent of X. For the first two DGPs the separability condition is verified, the cloud is really flat with respect to the two dimensions. For the DGP3, the U-shape with respect to Z that appears in both figures is exactly what we expected from (5.3). To conclude, the pictures of these ratios as functions of x and z are clearly informative. In our examples here, a marginal analysis of the effect of Z on the shifts of the efficient frontiers would also provide meaningful interpretations.

The analysis of the partial ratios, $\widehat{R}_{O,\alpha}(x,y|z)$, with α not far from 1, would provide the same pictures (we do not have outliers here). However, it is interesting to look at these ratios for smaller values of α to detect potential effects of Z on the distribution of the inefficiencies. Fig. 2 does this, for the median value ($\alpha = 0.5$). Here again the results confirm mostly what we expected (again, spurious unexpected behaviors could come from the fact that we use

[6] For more details on how to handle discrete variables Z in a similar framework, see Bădin, 2011 and the references therein.

[7] As mentioned earlier, in some applications, the additive model (3.14) could be more appropriate for the log $\lambda(X,Y|Z=z)$. In a particular application, it is important to use some standard checks to see if the hypothesis of the independence between ε and Z, is more reasonable or not in the log scale (see the empirical illustrations that follow in this paper).

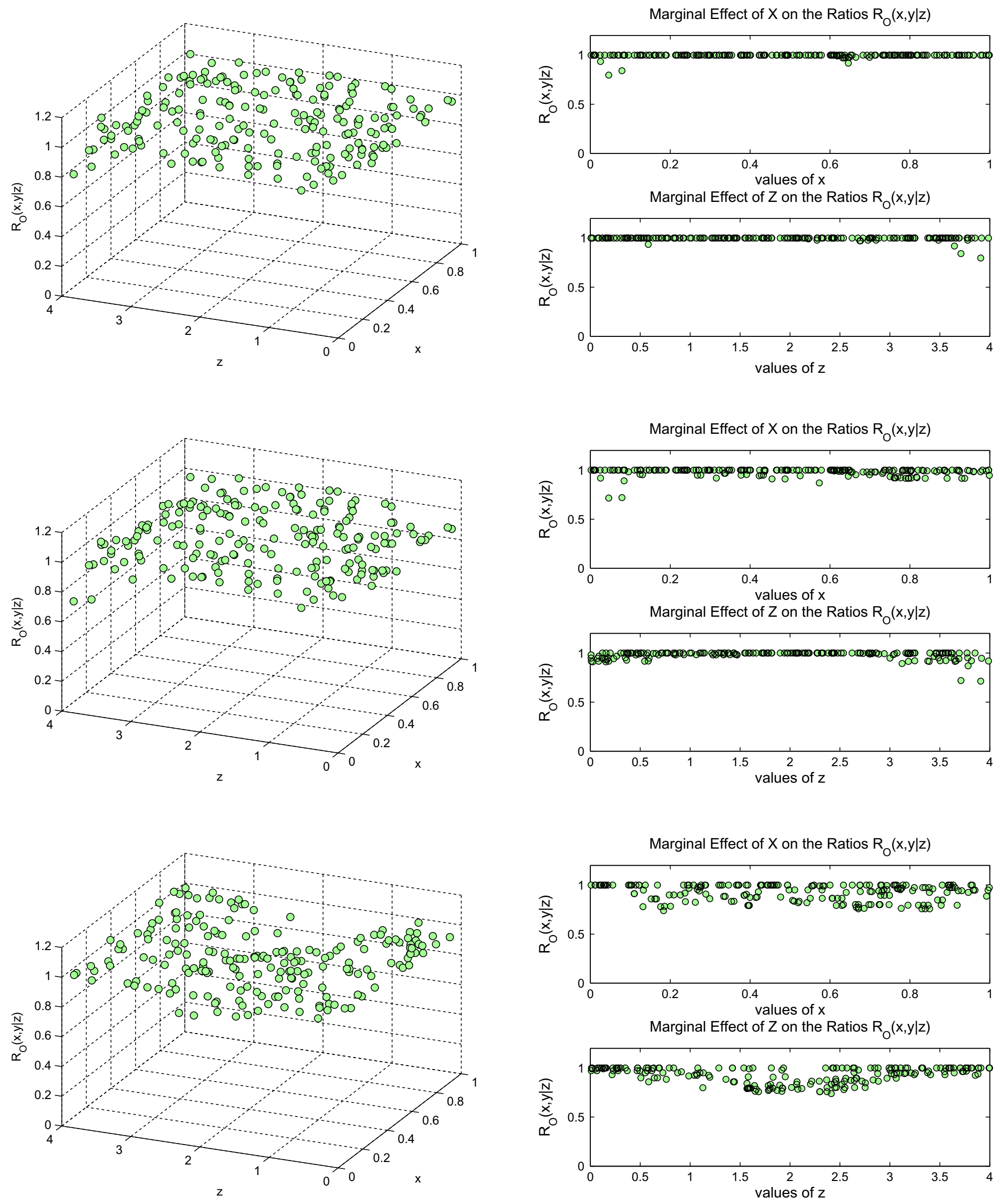

Fig. 1. Effect of X and Z on the ratios $\widehat{R}_O(x, y|z)$. From top to bottom: DGP1, DGP2 and DGP3. Here $n = 200$ and the circles are the estimated ratios.

estimates with low rates of convergence). For DGP1, top panel, the cloud of points is flat: we see no effect of Z on the efficiencies. In the case of DGP2, where we have the separability, we see some curvature in the z direction and flat behavior in the x direction; indeed, the marginal and conditional frontiers have the same support but the distribution of the inefficiencies is changing with the value of z. Near the center ($z = 2$), the sampling variations of $\widehat{R}_{O,\alpha}(x, y|z)$ are near 1, and for larger values of $|z - 2|$ we have more values smaller than 1. For DGP3, we reproduce for the median the U-shaped effect of the shift of the frontier we have observed by looking to the bottom panel of Fig. 1.

5.1.2. Second-stage regression

Fig. 3 summarises the results of the analysis of the second stage regression of $\log \hat{\lambda}(x, y|z)$ on z. The analysis done with $\hat{\lambda}(x, y|z)$ gave very similar results. For each DGP, we show on the left the results of the nonparametric regression for $\mu(z)$ and $\sigma(z)$, in the middle the histogram of the resulting residuals that can be interpreted as managerial efficiency and on the right, the clouds of n points $(Z_i, \hat{\varepsilon}_i)$ to check if any pattern is still apparent after whitening the effect of Z on the conditional efficiencies.

Again, the results are as we expected. We do not see any effect of z for DGP1, for DGP2, we observe a visible U-shaped effect for both $\mu(z)$ and $\sigma(z)$, confirming that the distribution of the inefficiencies varies with z. For DGP3, a small spurious effect (due to sampling uncertainties) shows, but it still maintains, roughly a stable $\mu(z)$ and a constant $\sigma(z)$, as it should (the distribution of the inefficiencies does not depend on z). Note that the histograms of the managerial efficiencies $\hat{\varepsilon}_i$, recover in the three cases the shape of the half-normal distribution that has been simulated for U. The correlations between U_i and $\hat{\varepsilon}_i$ are quite high in each case: 0.94, 0.87, 0.84 for DGP1, DGP2 and DGP3, respectively. So, it seems legitimate and meaningful to use these residuals to rank the firms according their managerial inefficiencies. The scatter plots between $\hat{\varepsilon}_i$ and Z_i do not show a particular structure; the correlation between the two variables are indeed very low: -0.05, 0.02, 0.01 for DGP1, DGP2 and DGP3, respectively.

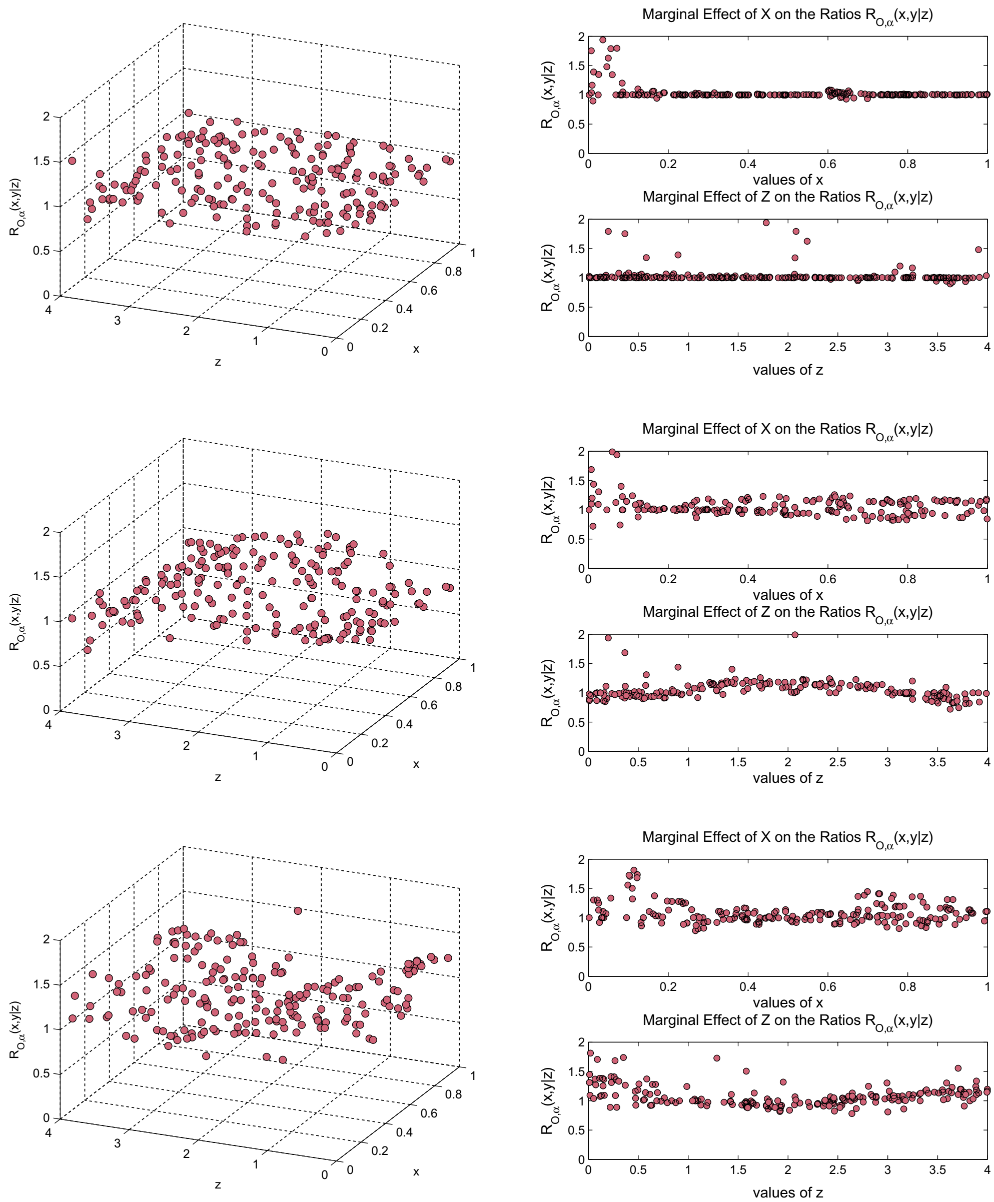

Fig. 2. Effect of X and Z on the ratios $\widehat{R}_{O,\alpha}(x,y|z)$, with $\alpha = 0.5$. From top to bottom: DGP1, DGP2 and DGP3. Here $n = 200$ and the circles are the estimated ratios.

5.2. Efficiency in the banking sector

Simar and Wilson (2007) include an empirical example based on Aly et al. (1990) using data on 6.955 US commercial banks observed at the end of the fourth quarter, 2002.[8] They run a truncated regression on the input-oriented DEA estimates of efficiency in a second stage (as suggested in Aly et al. (1990)). Daraio et al. (2010) used the same data set to test the separability condition which was rejected at any reasonable level, indicating that any two-stage procedure is meaningless for this dataset. This was a global test; we will proceed here to a local analysis and try to detect the size and direction of the detected effect.

The original data set contains three inputs (purchased funds, core deposits and labor) and four outputs (consumer loans, business loans, real estate loans, and securities held) for banks. Aly et al. (1990) considered two continuous environmental factors, the size of the banks Z_1, and a measure of the diversity of the services proposed by the banks Z_2 (see Aly et al. (1990), for details) and one binary variable indicating if the banks belong to a Metropolitan Statistical Area (MSA). We use, as in Simar and Wilson (2007), a measure of the size of the banks by the log of the total assets, rather than the total deposits as in Aly et al. (1990). For simplifying the presentation, we will illustrate our procedure with a subsample of 322 banks (also used in Simar and Wilson, 2007).

Some prior exploratory data analysis indicates that the three inputs are highly correlated and the same is true for the four outputs. So, due to the dimensionality of the problem (three inputs, four outputs, and three environmental factors) with the limited sample used here (322 units), we first reduce the dimension in the input $\times$ output space by using the methodology suggested in Daraio and Simar (2007a).

Since the radial measures are scale invariant, we divide each input and output by their mean (to be unit free) and replace the three scaled inputs by their best (non-centered) linear combination (we use here a kind of non-centered PCA, as explained in details in Daraio and Simar, 2007a), and we check that we did not loose much

[8] We would like to thank Paul W. Wilson who provided us this data set.

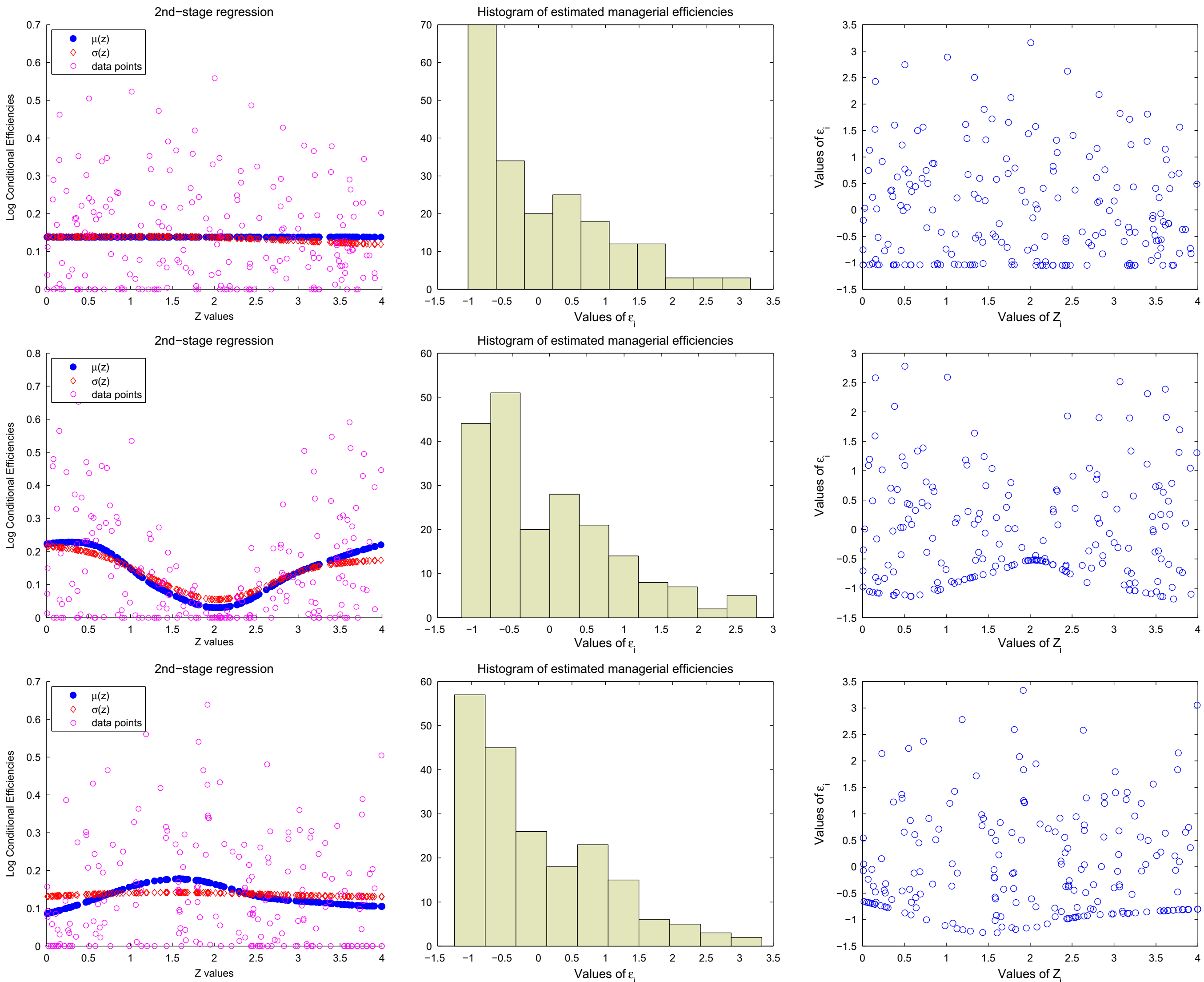

Fig. 3. Results of second-stage regression. From left to right: location and scale estimates of log $\widehat{\lambda}(x,y|z)$ as a function of z, histograms of managerial efficiencies, scatter plot of $\hat{\varepsilon}_i$ against Z_i. From top to bottom: DGP1, DGP2 and DGP3. Here $n = 200$ and the circles are the estimated ratios.

information by doing so, and that the resulting univariate input factor is highly correlated with the three original inputs. We follow the same procedure with the four outputs. The results are

$$IF = 0.5707X_1 + 0.5731X_2 + 0.5881X_3,$$
$$OF = 0.4851Y_1 + 0.4875Y_2 + 0.5095Y_3 + 0.5172Y_4,$$

indicating that both the input and the output factor represent a kind of average of the scaled inputs and outputs, respectively (the weights are equal). We obtain the following correlations $\hat{\rho}_{IF,X_j} = (0.972,\ 0.971,\ 0.996)$ for $j = 1, 2, 3$ and IF explains 96% of total inertia of the original data (X_1,X_2,X_3). We obtain similar results when reducing the dimension in the output space: $\hat{\rho}_{OF,Y_j} = (0.924, 0.938, 0.975, 0.990)$ for $j = 1,\dots,4$, and OF explains 92% of total inertia of the original data $(Y_1,\dots,Y_4)$. Hence, we can conclude that we do not loose much information by this dimension reduction, and the factors IF and OF are good representatives of the input and output activities of the banks.

Remember that with the full data set and with all the original variables, Daraio et al. (2010) rejected the null hypothesis of global separability. We will here illustrate in our simplified version of the examples what we can learn by the methodology we proposed in this paper.

We first investigate the effect of Z_1 = SIZE on the production process. It is clear that in this example Z_1 is highly correlated with Y (the linear correlation is 0.57, but the Spearman rank correlation is 0.97). Fig. 4 shows the ratios as function of Y and Z_1. Here it is the input orientation: $\widehat{R}_I(X_i, Y_i|Z_i)$ and for the partial frontiers $\widehat{R}_{I,\alpha}(X_i, Y_i|Z_i)$ with $\alpha = 0.95$, to see if some extreme data points could hide some effect and with $\alpha = 0.5$ to investigate the effect on the middle of the distribution of the inefficiencies. Without being able to rotate the three-dimensional figures on the left panels, we have an idea of what happens complementing the left picture with the two marginal views. It is not clear from the full-frontier ratios if Y has some effect on the frontier levels, but looking to the picture for the extreme quantile 0.95, it is more clear. For Z_1 it is also clear for the three pictures that Z_1 has a negative (unfavorable) effect on the frontier levels. When $\alpha = 0.5$, the effect is also visible, confirming the effect of the shift of the frontier. This short descriptive analysis also confirms that the separability condition for Z_1 seems unrealistic.

Fig. 5 shows the results of the analysis for the full-frontier conditional efficiencies as a function of Z_1 (the analysis was done

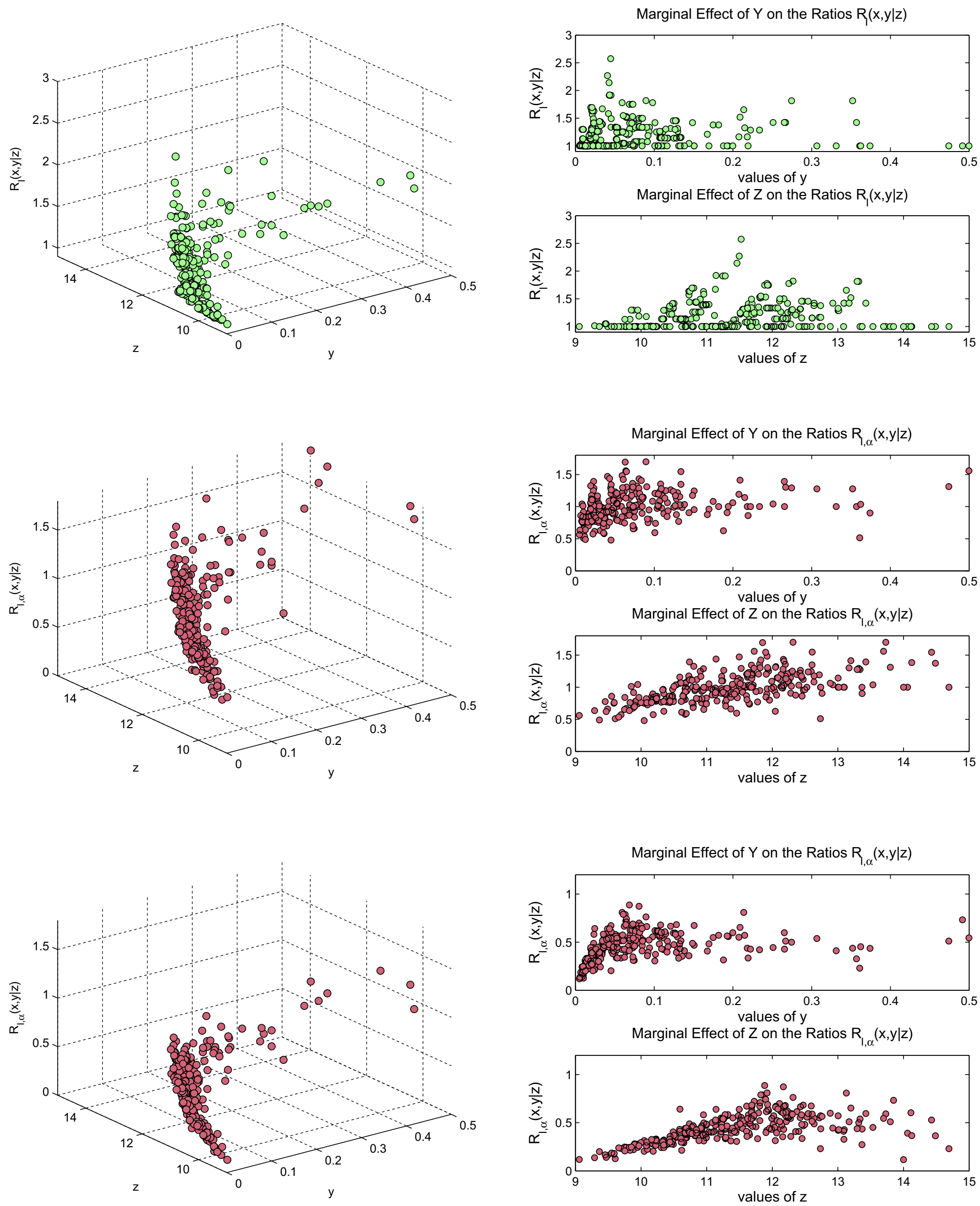

Fig. 4. Effect of Y and Z_1 = SIZE on the ratios $\widehat{R}_I(X_i, Y_i|Z_i)$ (top panel) and $\widehat{R}_{I,\alpha}(X_i, Y_i|Z_i)$ (middle panel $\alpha = 0.95$ and bottom panel $\alpha = 0.5$).

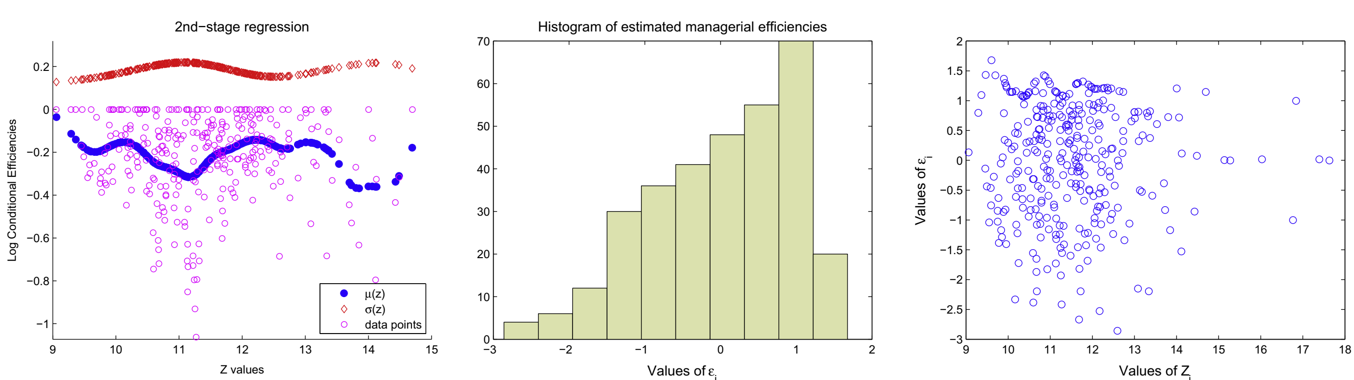

Fig. 5. Effect of Z_1 = SIZE on conditional efficiencies $\log \widehat{\lambda}(x, y|z)$, histograms of managerial efficiencies, scatter plot of $\hat{\varepsilon}_i$ against Z_1.

on the *logs*, but the picture in original units is very similar). The regression line $\mu(z)$ has a global shape not far from the horizontal line, indicating that the effect of Z_1 on the distribution of the efficiencies is rather low. This is confirmed by the rather constant

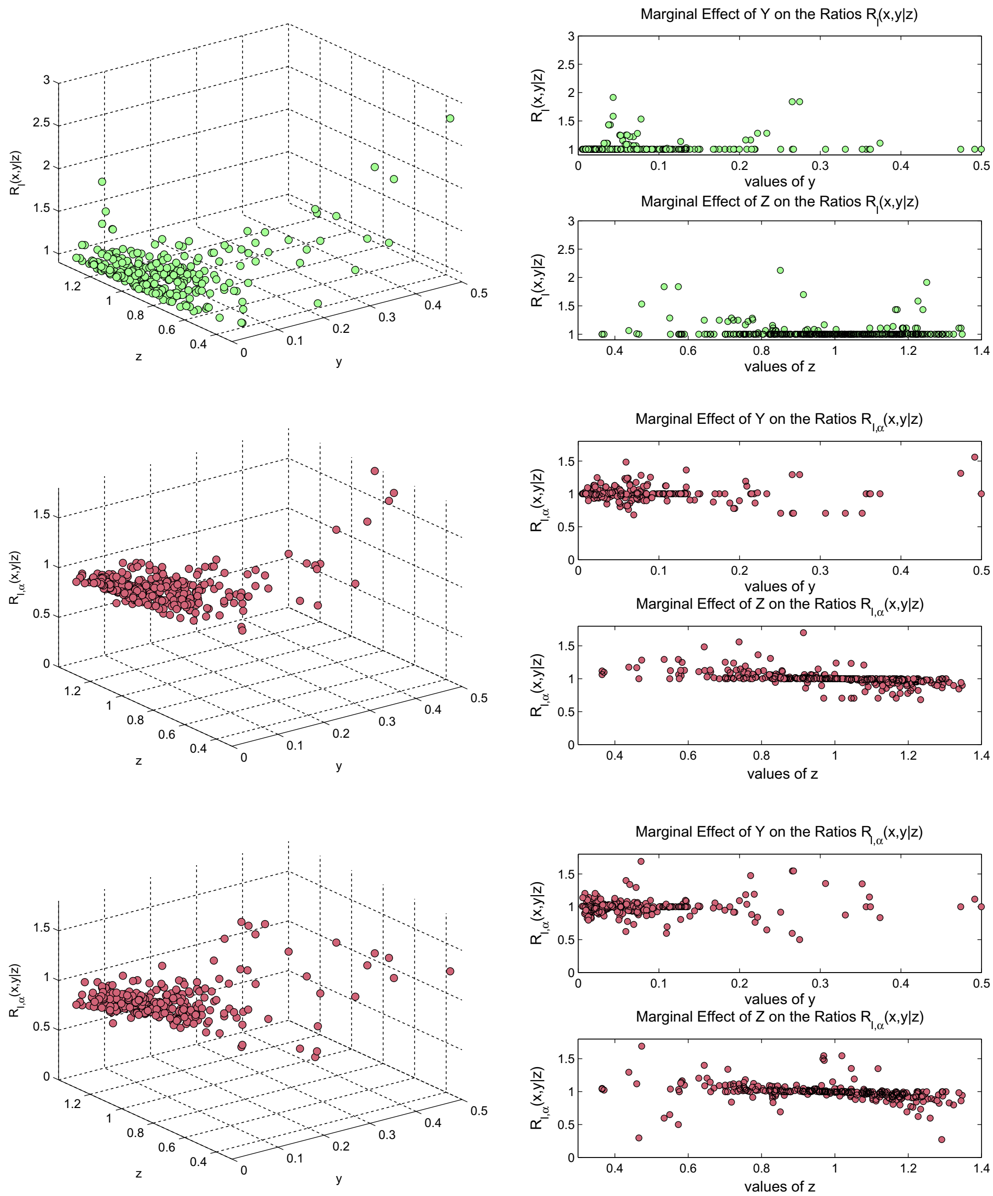

Fig. 6. Effect of Y and Z_2 = DIVERSE on the ratios $\widehat{R}_l(X_i, Y_i|Z_i)$ (top panel) and $\widehat{R}_{l,\alpha}(X_i, Y_i|Z_i)$ (middle panel $\alpha = 0.95$ and bottom panel $\alpha = 0.5$).

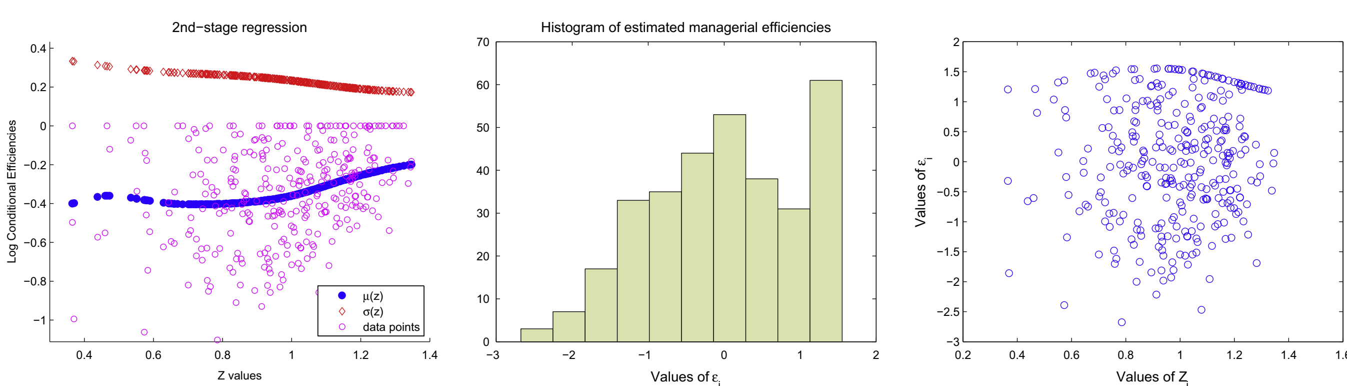

Fig. 7. Effect of Z_2 = DIVERSE on conditional efficiencies $\log \widehat{\lambda}(x, y|z)$, histograms of managerial efficiencies, scatter plot of $\hat{\varepsilon}_i$ against Z_2.

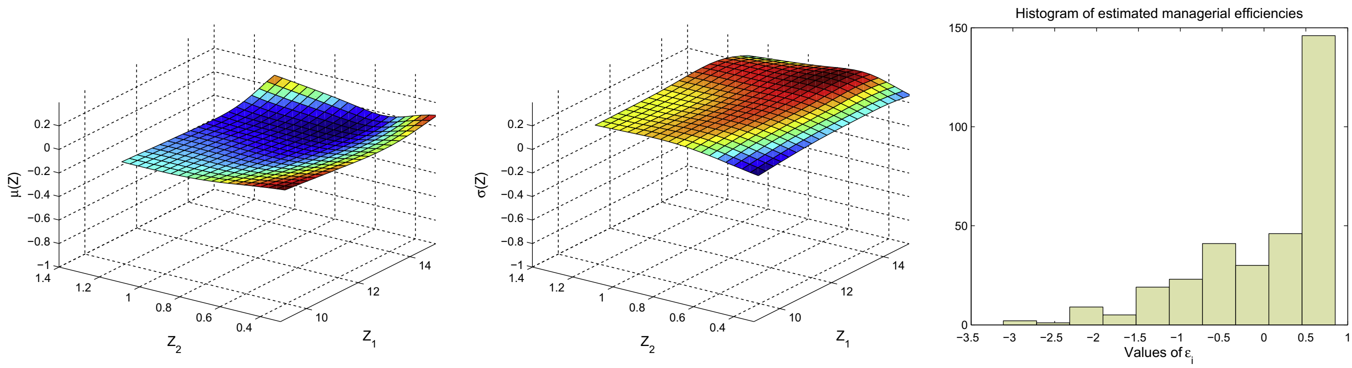

Fig. 8. Effect of $Z = (Z_1,Z_2)$ on conditional efficiencies $\log\hat{\theta}(x,y|z)$ (left panel $\mu(z)$, middle panel $\sigma(z)$) and the histogram of estimated managerial efficiencies.

shape of $\sigma(z)$. The distribution of managerial efficiencies has a reasonable shape (typically not far from a half-normal). The effect of Z_1 on the conditional efficiency scores has been nicely whitened: the Pearson linear correlation between Z_i and $\hat{\varepsilon}_i$ is -0.009 and the Spearman rank correlation is -0.008. So the ranking of the banks according to $\hat{\varepsilon}_i$ is cleaned from the effect of the size variable Z_1. Note that the resulting ranking is different from the ranking obtained by the marginal FDH scores $\hat{\lambda}(X_i, Y_i)$, but since the effect of Z_1 is not a big effect, there remains some correlation (the correlation between the two rankings is 0.64).

We did the same univariate exercise to investigate the marginal effect of the variable Z_2 (DIVERSE: a measure of the diversity of the products of the Banks). Figs. 6 and 7 display the results. We summarize very shortly the conclusions and let the reader complete the analysis. The effect of Z_2 seems quite small (see top panel of Fig. 6). Z_2 is not responsible for the rejection of the separability condition. However, we can see a small effect on the partial ratios $\widehat{R}_{I,\alpha}(X_i, Y_i|Z_i)$, indicating a small effect on the distribution of the efficiencies, but a favorable one: banks having more diversity seems to have a distribution of their efficiency slightly more concentrated near the efficient frontier. This is confirmed by looking at Fig. 7, where $\mu(z)$ is slightly increasing, combined with a slightly decreasing $\sigma(z)$. The effect of Z_2 on the conditional efficiencies has been well removed, and the correlation between $\hat{\varepsilon}_i$ and Z_i is 0.06 (the right panel of Fig. 7 reveals no clear remaining pattern).

The bivariate analysis would consist of using the location-scale regression model with $Z = (Z_1,Z_2)$. Pictures to see the joint effect of Z on the frontier levels for fixed levels of the outputs are difficult to display (four dimensions), but we know that the assumption of separability was rejected in Daraio et al. (2010). The analysis of the conditional efficiency scores as a function of Z is similar to what we did earlier for one dimension. Fig. 8 displays the results for the surfaces $\mu(z)$ and $\sigma(z)$. To summarize, we confirm typically the two marginal analysis we performed, and we do not see any interaction between Z_1 and Z_2 in the effect on the conditional efficiencies. The resulting managerial estimates have correlation -0.0461 and -0.0339 with Z_1 and Z_2, respectively, so most of the effects of Z have been removed by our location-scale model. The right panel of Fig. 8 shows the resulting histogram of these residuals. The shape looks very similar to those obtained by the marginal analysis completed earlier, but we have as expected more mass near the efficient frontier, because we explain the conditional efficiencies by two environmental factors here.

In the case of the discrete Z_3 = MSA, the analysis could be performed separately on the two groups, or Z_3 could be introduced in these models by using appropriate kernels (see, for details). The procedure would go along the same lines as the one illustrated above. Of course, increasing the number of variables will give estimators with less precision and the descriptive tools presented earlier are limited to pictures in three dimensions.

6. Conclusions

This paper develops the conditional nonparametric methodology to investigate the role of environmental variables Z by introducing these external factors in a non restrictive way.

The paper clarifies what can be learned by analyzing conditional efficiency measures and proposes a general approach to measure and draw inference about the impact of these factors on the production process.

By using conditional efficiency measures we can indeed measure the impact of external factors on the attainable set in the input–output space, and/or we can investigate the impact of the external factors on the distribution of inefficiency scores. We extend existing methodological tools to explore these interrelationships, both from individual and global perspectives.

Further, we propose a two-stage type approach of the conditional efficiencies on the factors that do not suffer from the limitations of traditional two-stage approaches. We suggest a flexible model to eliminate the location and the scale effect of Z on the efficiencies. The analysis of the obtained residuals provides a measure of efficiency whitened from the main effects of the environmental factors. This allows us to rank the firms according to their managerial efficiency, even when facing heterogeneous environmental conditions.

The procedure is illustrated through simulated samples and with a real data set on US commercial banks.

Future research should address the following issues: testing the partial separability condition to simplify the analysis; testing the independence between the error term and the environmental factors in the second-stage regression; and, finally establishing the asymptotic properties of the second-stage regression. The latter two should be interrelated.

Acknowledgements

Financial support from the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, Project number PN-II-ID-PCE-2011-3-0893, from the Inter-university Attraction Pole, Phase VI (No. P6/03) of the Belgian Government (Belgian Science Policy) and from the INRA-GREMAQ, Toulouse, France are gratefully acknowledged. This work was completed during an academic visit in 2012 of L. Simar at the DIAG, Sapienza University of Rome, Italy.

Appendix A. Effect of Z on the frontier levels

We illustrate the basic ideas for the input orientation in a simple scenario where $p = q = r = 1$. In this case, we can describe the (input) efficient frontier and its conditional version by the functions $\phi(y)$ and $\phi(y|z)$. Suppose that when $Y = y_1$, Z is favorable to the production process (it acts like a free disposal input), the frontier $\phi(y_1|z)$ is displayed in dashed line in the top panel of Fig. 9, we see also the marginal input-frontier $\phi(y_1)$. Suppose that when $Y = y_2$, Z is favorable until level z_a and then unfavorable (acting like an undesired output), the conditional frontier $\phi(y_2|z)$ is represented by the solid line in the figure. Finally, suppose that when $Y = y_3$, Z is unfavorable, the conditional frontier $\phi(y_3|z)$ is then displayed as the dotted line in the figure. We see that for all cases, $R_I(x,y|z) \geqslant 1$, but the shape of the ratios as a function of z can be different according to the values of Y (see the bottom panels for the three different levels of Y). In the cases illustrated here, the analysis of these ratios $R_I(x,y|z)$ as a function of z only, would be problematic, so in general, without additional assumptions, it is better to carry the analysis for fixed levels of the outputs Y.

Of course, the example illustrated in Fig. 9 is rather extreme, and in many situations, the interactions between Z and Y on the frontier levels will be less complicated. In particular, if Y is independent of Z, or in a less restrictive way, under the assumption that the shape of the boundaries of $\mathcal{P}$ in the sections $Y = y$ (in the (X,Z) space) would not change with the level y, the conditional frontiers in the top panel of Fig. 9 would be parallel, so that the ratios $R_I(x,y|z)$ would have the same shape when considered as a function of z for all values of Y. For instance, this would be the case if Z would act as a free disposal input for all the values of Y (Panel I in the bottom panels of Fig. 9). In the lines of Simar and Wilson (2007), this corresponds to an assumption of partial separability, which was implicitly assumed and illustrated, in Daraio and Simar (2005, 2007a). In this case, the analysis of the effect of Z on the efficient frontier is largely simplified. Testing this partial separability assumption remains an open issue for future work, in the numerical illustrations provided in the paper we give some descriptive tools to investigate this issue.

Appendix B. Complementarity of full-frontier and partial-frontier measures

In Fig. 10 we illustrate the basic ideas of Section 3.3 for the output orientation, for the particular case of a univariate output, for a fixed level of input x_0, and for a fixed value z_0. The figure displays the conditional distributions $F(y|X \leqslant x_0)$ and $F(y|X \leqslant x_0, Z = z_0)$, for various scenarios, along with the upper boundary of their support and their α-quantiles. Remember that in this univariate output case, $R_O(x_0,y_0|z_0) = \varphi(x_0|z_0)/\varphi(x_0)$, with a similar expression for $R_{O,\alpha}(x_0,y_0|z_0)$. In the left panels, where the separability condition is verified (see panels II and III), the conditional and unconditional distributions of the inefficiencies are different, but they share the same support; this results in ratios $R_O(x_0,y_0|z_0) = 1$ (we see indeed that $\varphi(x_0|z_0) \equiv \varphi(x_0)$).

Second, the information carried by the ratios $R_{O,\alpha}(x,y|z)$, when defined relative to the partial frontiers, is multiple. Suppose that $\Psi^z = \Psi$ and so $R_O(x,y|z) = 1$ for all points (x,y,z) (left panels in Fig. 10). Then, if the distribution of the inefficiencies is affected by Z, the quantiles of $S_{Y|X,Z}$ will be different from those of $S_{Y|X}$. Therefore, for all $(x,y) \in \Psi^z$, the ratios $R_{O,\alpha}(x,y|z)$ will be affected. Note that in this case ($\Psi^z = \Psi$), the changes can go in two directions for the partial parameter: if the distribution of the

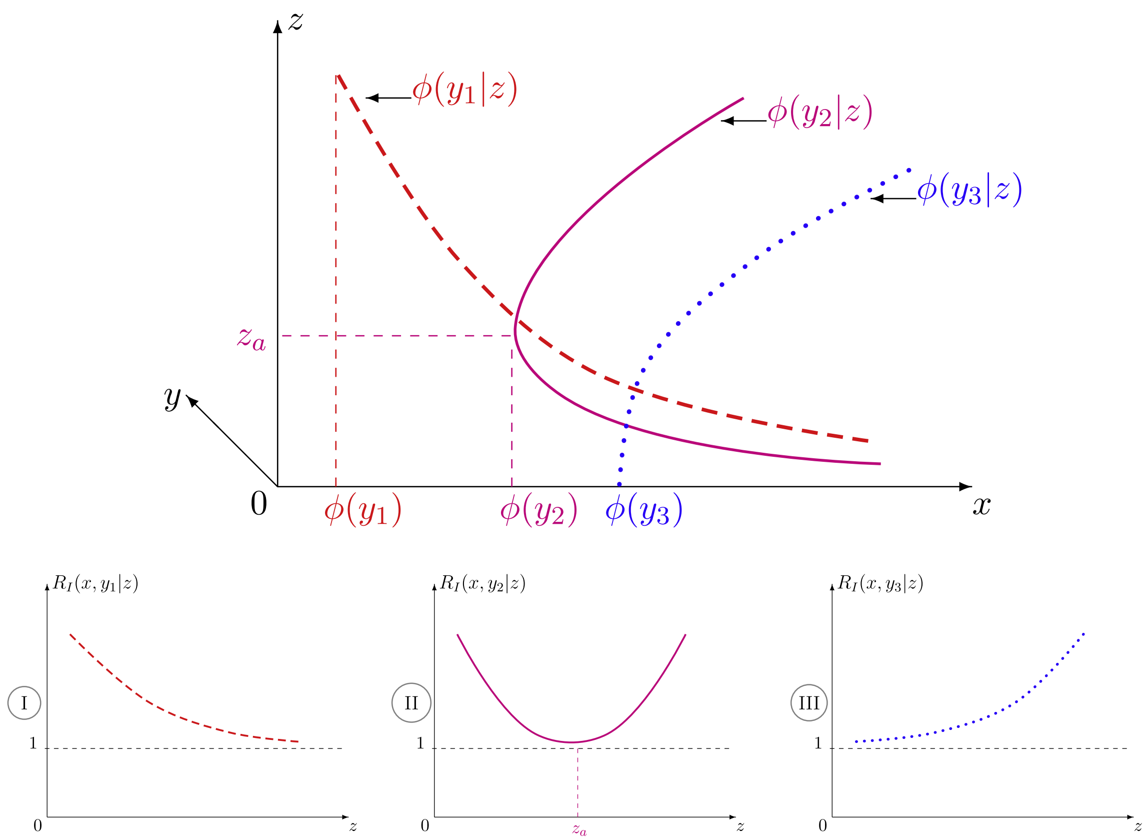

Fig. 9. Various scenarios for interpreting the effect of Z. Top panel, the minimal input frontiers, in the coordinates (x,z) for different levels of Y. Bottom panels, the corresponding ratios $R_I(x,y|z)$ as a function of z.

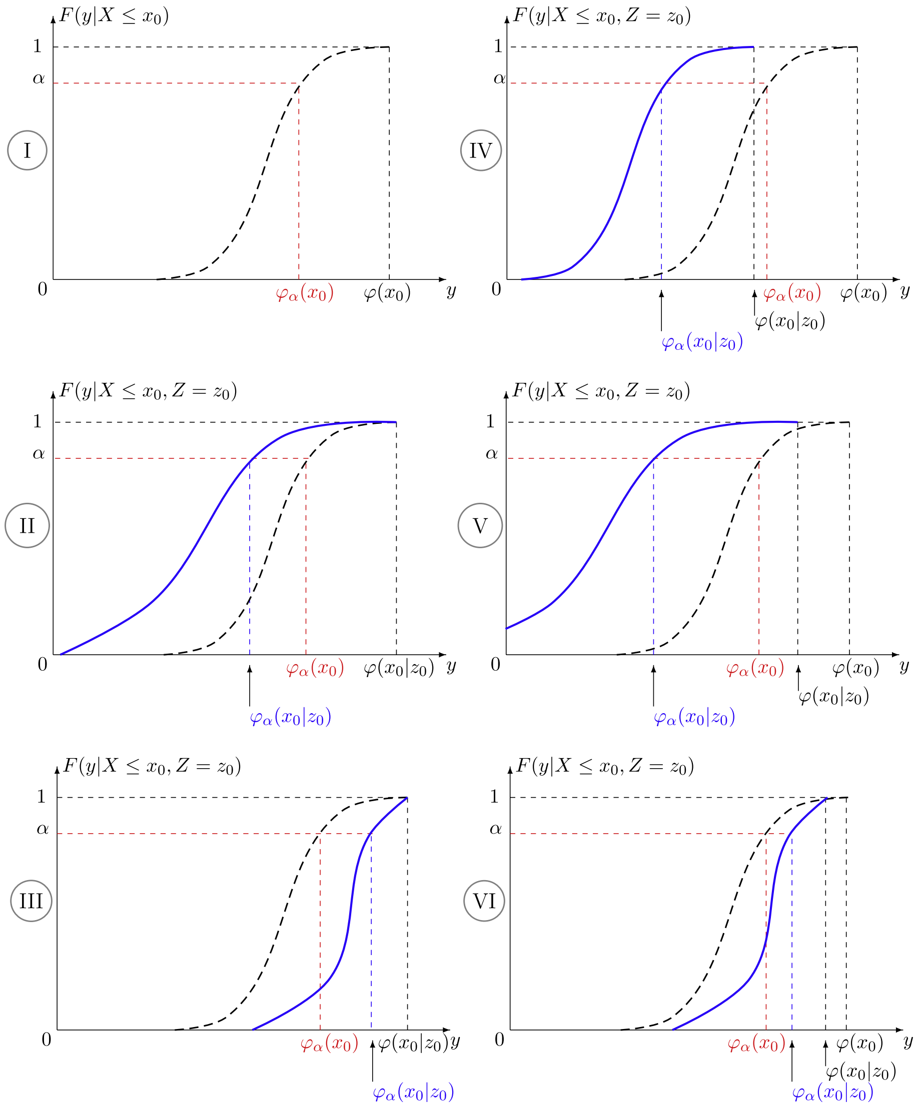

Fig. 10. Various scenarios for $F(y|X \leqslant x_0)$ and $F(y|X \leqslant x_0, Z = z_0)$. In the left panels the separability condition is verified at (x_0,z_0), while on the right panels, this condition is violated. In all six panels, the dashed black line represents $F(y|X \leqslant x_0)$, with upper boundary of support $\varphi(x_0)$, and the solid blue line is $F(y|X \leqslant x_0, Z = z_0)$, with upper boundary of support $\varphi(x_0|z_0)$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inefficiency is more spread in the direction of less efficient behavior (as in panel II), we observe $\varphi_\alpha(x_0|z_0) < \varphi\alpha(x_0)$ giving $R_{O,\alpha}(x_0,y_0|z_0) < 1$. On the contrary, if z_0 provides a favorable environment to efficient behavior of the firms (without affecting the upper boundary), the distribution of Y will be more concentrated near the efficient boundary when $Z = z_0$ (as in panel III), we have $\varphi_\alpha(x_0|z_0) > \varphi_\alpha(x_0)$ giving $R_{O,\alpha}(x_0,y_0|z_0) > 1$. This is, of course, less probable when $\alpha \to 1$. That is the reason why the global test of separability of Daraio et al. (2010) uses test statistics based only on the full measures of efficiency and not on the partial efficiency scores, unless α is not far from one.

Third, if there is a shift on the frontier $\Psi^z \subset \Psi$, it is much more difficult to interpret the ratios $R_{O,\alpha}(x,y|z)$. It is clear that a shift of the boundary will be transferred to the partial frontier, at least for large values of α, near 1, but this effect can either be increased or compensated by a simultaneous change of the distribution of the inefficiencies from the unconditional to the conditional. So, in the case of a shift of the boundary (see the right panels of Fig. 10), we could observe $R_{O,\alpha}(x_0,y_0|z_0)$ less, equal or greater than 1. We illustrate three cases in Fig. 10. We see that in panel IV, the shift of $\varphi_\alpha(x_0|z_0)$ with respect to $\varphi_\alpha(x_0)$ is the same as the shift of $\varphi(x_0|z_0)$ with respect to $\varphi(x_0)$, giving here $R_{O,\alpha}(x_0,y_0|z_0) < R_O(x_0,y_0|z_0) < 1$. In panel V, we have more spread toward inefficiencies when conditioning on z_0, the shift of the quantile of the conditional distribution is much more important so $R_{O,\alpha}(x_0,y_0|z_0) \ll R_O(x_0,y_0|z_0) < 1$. But we could observe, as in panel VI, a different behavior when for a given z_0 it is more probable to reach the efficient frontier $\varphi(x_0|z_0)$ implying that we could obtain for some quantiles $R_{O,\alpha}(x_0,y_0|z_0) > R_O(x_0,y_0|z_0)$. So, even if $R_O(x_0,y_0|z_0) < 1$, we could have in extreme cases $R_{O,\alpha}(x_0,y_0|z_0) \geqslant 1$ (as in panel VI of Fig. 10).

To summarize the second and third points if $\Psi^z = \Psi$, the ratios $R_{O,\alpha}(x,y|z)$ are useful to shed light on the local impact of Z on the shape of the distribution of the inefficiencies. But it does not allow us to detect, when considered alone, a local shift of the boundary of the support of (X,Y). Unless $\alpha \to 1$, because in this case, the partial frontier can serve as a robust estimator of the full frontier.

In any case, these partial measures bring useful complementary information of the relative position of the quantiles of $S_{Y|X,Z}$ with respect to those of $S_{Y|X}$. It will therefore be useful to provide some detailed analysis of the ratios $R_O(x,y|z),\ R_{O,\alpha_1}(x,y|z),\ldots,R_{O,\alpha_k}(x,y|z)$, as described in Section 3.1, for a grid of selected values for α like, 0.99, 0.95, 0.90…,0.50. The latter case $\alpha = 0.50$ provides for instance, a picture on the impact of z on the median of the inefficiency distribution.

The same would be true for the order-m partial ratios $R_{O,m}(x,y|z)$ where the particular case $m = 1$ would allow us to investigate the effect of Z on the average frontier. Here, the choice of large values of m would provide the same information as for the full-frontier case.

References

Aly, H.Y., Grabowski, C.P.R.G., Rangan, N., 1990. Technical, scale, and allocative efficiencies in U.S. banking: an empirical investigation. Review of Economics and Statistics 72, 211–218.

Banker, R.D., Morey, R.C., 1986a. Efficiency analysis for exogenously fixed inputs and outputs. Operations Research 34 (4), 513–521.

Banker, R.D., Morey, R.C., 1986b. The use of categorical variables in data envelopment analysis. Management Science 32 (12), 1613–1627.

Banker, R.D., Natarajan, R., 2008. Evaluating contextual variables affecting productivity using data envelopment analysis. Operations Research 56 (1), 48–58.

Bădin, L., Daraio, C., 2011. Explaining efficiency in nonparametric frontier models. Recent developments in statistical inference. In: Van Keilegom, I., Wilson, P.W. (Eds.), Exploring Research Frontiers in Contemporary Statistics and Econometrics. Springer-Verlag, Berlin Heidelberg.

Bădin, L., Daraio, C., Simar, L., 2010. Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. European Journal of Operational Research 201 (2), 633–640.

Cazals, C., Florens, J.P., Simar, L., 2002. Nonparametric frontier estimation: a robust approach. Journal of Econometrics 106, 1–25.

Chen, Y., Cook, W.D., Li, N., Zhu, J., 2009a. Additive efficiency decomposition in two stage DEA. European Journal of Operational Research 196 (3), 1170–1176.

Chen, Y., Liang, L., Zhu, J., 2009b. Equivalence in two-stage DEA approaches. European Journal of Operational Research 193, 600–604.

Daouia, A., Gijbels, I., 2011a. Robustness and inference in nonparametric partial-frontier modeling. Journal of Econometrics 161, 147–165.

Daouia, A., Gijbels, I., 2011b. Estimating frontier cost models using extremiles. In: Van Keilegom, I., Wilson, P.W. (Eds.), Exploring Research Frontiers in Contemporary Statistics and Econometrics. Springer-Verlag, Berlin Heidelberg.

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

Daraio, C., Simar, L., 2005. Introducing environmental variables in nonparametric frontier models: a probabilistic approach. Journal of Productivity Analysis 24, 93–121.

Daraio, C., Simar, L., 2006. A robust nonparametric approach to evaluate and explain the performance of mutual funds. European Journal of Operational Research 175 (1), 516–542.

Daraio, C., Simar, L., 2007a. Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications. Springer, New York.

Daraio, C., Simar, L., 2007b. Conditional nonparametric Frontier models for convex and non convex technologies: a unifying approach. Journal of Productivity Analysis 28, 13–32.

Daraio, C., Simar, L., Wilson, P., 2010. Testing whether two-stage estimation is meaningful in nonparametric models of production. Discussion Paper #1031. Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Debreu, G., 1951. The coefficient of resource utilization. Econometrica 19 (3), 273–292.

Fan, J., Gijbels, I., 1996. Local Polynomial Modelling and Its Applications. Chapman and Hall.

Fan, J., Yao, Q., 1998. Efficient estimation of conditional variance functions in stochastic regression. Biometrika 85, 645–660.

Farrell, M.J., 1957. The measurement of the productive efficiency. Journal of the Royal Statistical Society, Series A CXX (Part 3), 253–290.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1985. The Measurement of Efficiency of Production. Kluwer-Nijhoff Publishing, Boston.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. Production Frontiers. Cambridge University Press, Cambridge.

Färe, R., Grosskopf, S., Lovell, C.A.K., Pasurka, C., 1989. Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. Review of Economics and Statistics 71 (1), 90–98.

Jeong, S.O., Park, B.U., Simar, L., 2010. Nonparametric conditional efficiency measures: asymptotic properties. Annals of Operations Research 173, 105–122.

Jeong, S.O., Simar, L., 2006. Linearly interpolated FDH efficiency score for nonconvex frontiers. Journal of Multivariate Analysis 97, 2141–2161.

Kao, C., Hwang, S.N., 2008. Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. European Journal of Operational Research 185 (1), 418–429.

Kneip, A., Simar, L., Wilson, P.W., 2012. Central Limit Theorems for DEA Estimators: When Bias Can Kill Variance, manuscript.

Leibenstein, H., 1966. Allocative efficiency vs "X-efficiency". American Economic Review 56, 392–415.

Leibenstein, H., 1979. A branch of economics is missing: micro–micro theory. Journal of Economic Literature XVII, 477–502.

Leibenstein, H., Maital, S., 1992. Empirical estimation and partitioning of X-inefficiency: a data-envelopment approach. AEA Papers and Proceedings 82 (2), 428–433.

Li, Q., Racine, J., 2007. Nonparametric Econometrics: Theory and Practice. Princeton University Press.

Pagan, A., Ullah, A., 1999. Nonparametric Econometrics. Cambridge University Press.

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

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

Simar, L., 2003. Detecting outliers in frontiers models: a simple approach. Journal of Productivity Analysis 20, 391–424.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 136 (1), 31–64.

Simar, L., Wilson, P.W., 2008. Statistical inference in nonparametric frontier models: recent developments and perspectives. In: Harold, Fried, KnoxLovell, C.A., Schmidt, Shelton. (Eds.), The Measurement of Productive Efficiency, 2nd ed. Oxford University Press.

Simar, L., Wilson, P.W., 2011a. Inference by the m out of n bootstrap in nonparametric frontier models. Journal of Productivity Analysis 36, 33–53.

Simar, L., Wilson, P.W., 2011b. Two-stage DEA: Caveat Emptor. Journal of Productivity Analysis 36, 205–218.

Zha, Y., Liang, L., 2010. Two-stage cooperation model with input freely distributed among the stages. European Journal of Operational Research 205 (2), 332–338.

Ziegelmann, F.A., 2002. nonparametric estimation of volatility functions: the local exponential estimator. Econometric Theory 18 (4), 985–991.