Journal of
Linear and Topological Algebra
Vol. 02, *No.* 01, 2013, 35- 49

Stochastic Non-Parametric Frontier Analysis

M. Rahmani[a,*] and G. R. Jahanshahloo[a]

[a] *Department of Mathematics, Science and Research Branch, Islamic Azad University, Tehran. Iran.*

Abstract. In this paper we develop an approach that synthesizes the best features of the two main methods in the estimation of production efficiency. Specifically, our approach first allows for statistical noise, similar to Stochastic frontier analysis , and second, it allows modeling multiple-inputs-multiple-outputs technologies without imposing parametric assumptions on production relationship, similar to what is done in non-parametric methods. The methodology is based on the theory of local maximum likelihood estimation and extends recent works of Kumbhakar et al. We will use local-spherical coordinate system to transform multi-input multi-output data to more flexible system which we can use in our approach. We also illustrate the performance of our approach with simulated example.

Keywords: Stochastic frontier, Nonparametric frontier, Local maximum likelihood, DEA, FDH, Directional Distance Function.

1. Introduction

The economic function of a firm is to bid resources away from alternative uses. As a result of such resource transfer, aggregate output may be increased or decreased. If inefficiency exists, an increase in output can be achieved by reallocating resources to more efficient uses. Given the seriousness of the issue (viz., the economic, political and social implications of inefficiency), it is essential that the measurement of efficiency/performance should be theoretically valid and subject to unambiguous interpretations.

Econometric measurement of efficiency of firms goes back to [1]Aigner, Lovell and Schmidt (1977) and [25]Meeusen and van den Broeck (1977). When analyzing the performance of a firm that is observed to produce an output level $y \in \mathbb{R}$ using input quantities $x \in \mathbb{R}^d_+$, one typically compares the observed output level with the maximum possible output that can be obtained from the production frontier (defined as

*Corresponding author.
E-mail addresses: smrds2003@yahoo.com (M. Rahmani).

ISSN: 2252−0201
http://jlta.iauctb.ac.ir

$f(x) = \max y : y \in P(x)$ where $P(x)$ describes the set of outputs that are feasible for each input vector $x \in \mathbb{R}^d_+$). The estimation of the production frontier is obtained from a random sample of observed firms $(X_i, Y_i) : i = 1, ..., n$. Then an efficiency score for a given production plan (x, y) is derived from the distance of this point to the estimated production frontier. The same approach can also be applied when a cost frontier is analyzed. In the latter we seek the minimal cost achievable for a given level of output(s).

Since the publication of the seminal papers by [1]Aigner et al. (1977) and [25]Meeusen and van den Broeck (1977), econometric estimation of parametric stochastic frontier (SF) functions has become a standard practice in efficiency measurement studies. However, in this approach, the estimation relies heavily on the particular choices of the functional form of the production/cost frontier (CobbDouglas, Translog, etc.) as well as the specific distributional assumptions on the error term (a convolution of a one-sided inefficiency term and a two-sided noise term). Typically, these parametric models are written as

$$Y_i = \beta_0 + \beta^T X_i - \mu_i + \nu_i, \quad i = 1, ..., n. \tag{1}$$

where $\mu_i > 0$ is the inefficiency term and $\nu_i \in \mathbb{R}$ represents random noise. The estimation technique is straightforward and is primarily based on the maximum likelihood principle.Of course, in practice we cannot be confident about the validity of these parametric assumptions that are used to estimate the model. The parametric form of the frontier function might be wrong due to several reasons. For example, the parametric functional form might be wrong, the stochastic specifications of the error components (particularly for the inefficiency component) might be wrong, among others.

An alternative to the parametric SF is the deterministic nonparametric approach where no specific parametric assumptions are made on the model. The frontier is defined as the upper boundary of the attainable set, say $\Psi = (x, y) : x \textit{ can produce } y$. In these nonparametric approaches the statistical properties of envelopment estimators like DEA and FDH ([12]Farrell, 1957; [4]Charnes et al., 1978; [10]Deprins et al., 1984), rely on the so-called deterministic assumption, viz.,

$$prob((X_i, Y_i) \in \Psi) = 1. \tag{2}$$

This assumption implies that no noise is allowed in these deterministic frontier (DF) models. The introduction of noise in a full nonparametric setup is problematic due to identification problems (see [17]Hall and Simar, 2002). Statistical inference is now available in these nonparametric DF models (see [32]Simar and Wilson, 2000, for a recent survey) but assumption $prob((X_i, Y_i) \in \Psi) = 1$ is too strong in many practical situations where we might expect measurement error, random shocks, etc. Recently [3]Cazals et al. (2002), and [6]Daouia and Simar (2007) have proposed robust versions of the FDH estimator, robust to extremes values and/or outliers since they do not envelop all the data. But these approaches still rely heavily on the deterministic assumption, where no noise is allowed.

In the presence of panel data, Park and Simar in a series of papers, consider the semi-parametric estimation of SF panel models under various assumptions on the joint distribution of the random firm effects and the regressors and on various dynamic specifications. The nonparametric part of these models concerns the distribution of the inefficiency terms. However, the estimators in these panel models are based on the linearity of the efficient frontier.

Fan et al. propose a two-step pseudo-likelihood estimator in a semi-parametric model where the production frontier is not specified, but distributional assumptions are imposed on the stochastic components as in [1]Aigner et al. (1977). An average production frontier is then estimated through standard kernel methods, the shift for the frontier is obtained through a moment condition, as in the MOLS approach and the remaining parameters of the stochastic components are estimated by maximizing a pseudo-likelihood function.

Our purpose in this paper is to propose a new approach to handle nonparametric Stochastic Frontier models by introducing noise in models. We develop an approach that synthesizes the best features of the two main methods in the estimation of production efficiency. Specifically, our approach first allows for statistical noise, similar to Stochastic frontier analysis (even in a more flexible way), and second, it allows modeling multiple-inputs-multiple-outputs technologies without imposing parametric assumptions on production relationship, similar to what is done in non-parametric methods, like Data Envelopment Analysis (DEA), Free Disposal Hull (FDH), etc. The method is based on the local maximum likelihood principle, which is nonparametric in the sense that the parameters of a given local polynomial model are localized with respect to the covariates of the model. Localizing can be viewed as a way of non-parametrically encompassing a parametric anchorage model. The idea to use local likelihood method for SF models was first suggested by Kumbhakar et al. for a particular case of the model proposed here. The method also improves DEA/FDH estimators, by allowing them to be quite robust to statistical noise and especially to outliers, which were the main problems of the original DEA/FDH estimators. The procedure shows great performance for various simulated cases and is also illustrated with simulated data.

2. Efficient frontier and efficiency scores

2.1 *Economic Theory*

In productivity analysis, the performance of production units is measured through efficiency scores. These efficiency measures are typically given by the distance of the individual decision making units (DMU) to the technology which is defined as the frontier of the production set. In the classical setting of productivity analysis and technical efficiency study, we consider a set of p inputs and q outputs used in the production process. The production set is the set of technically feasible combinations of inputs and outputs. It is defined as

$$\Psi = (x, y) \in \mathbb{R}_+^{p+q} : x \ can \ produce \ y \tag{3}$$

and its efficient frontier (the technology) is defined by

$$\Psi_\vartheta = (x, y) \in \Psi : (\gamma^{-1}x, \gamma y) \notin \Psi for \ all \ \gamma > 1. \tag{4}$$

The Farrell-Debreu input distance (see [9]Debreu, 1951, [12]Farrell, 1957) for a firm operating at the level (x, y) is then determined by the radial distance from (x, y) to the efficient frontier. It indicates how much all input quantities can be proportionately reduced so that the output levels y can still be produced. Formally, the Farrell input distance for a firm at a point (x, y) is given by

$$\theta(x, y) = \inf \theta > 0 : (\theta x, y) \in \Psi. \tag{5}$$

In the same spirit, the Farrell output distance of (x; y)

$$\lambda(x, y) = \sup \lambda > 0 : (x, \lambda y) \in \Psi. \tag{6}$$

is the maximum feasible equi-proportionate expansion of all outputs attainable with the input level x. Note that the [34]Shephard (1970) distances functions are the reciprocals of the Farrell distance functions.

The hyperbolic graph efficiency measure, proposed by [13]Färe et al. (1985), [14]Färe and Grosskopf (2000), provides an alternative measure of the distance of a firm (x, y) to the efficient frontier. It is given by

$$\gamma(x, y) = \inf \gamma > 0 : (\gamma^{-1}x, \gamma y) \in \Psi. \tag{7}$$

A use of hyperbolic path allows to avoid some of the ambiguity in choosing output or input orientation. In this case, input and output levels are adjusted simultaneously.

All the above methods rely on multiplicative measures of the distance and so require to deal with strictly positive inputs and outputs. This can be critical when the data contain zero or negative values as in financial data bases with the measure of funds performances.

A natural idea suggested by several authors is to translate the data to avoid negative values, but as pointed e.g. by [24]Lovell and Pastor (1995), if multiplicative efficiency measures satisfy the unit invariant property they are not invariant from an affine translation of inputs or outputs, and only additive models can satisfy the translation invariant property. As these authors note, some specific DEA estimator can satisfy translation invariance for inputs or outputs, but not for both. Such a restriction may strongly constraint the choice of inputs and outputs.

Recently directional distance functions have been introduced (see [5]Chambers, Chung and [14]Fare et al., 2000) that generalize both input and output distance functions. The directional distance function projects the input-output vector (x, y) onto the technology frontier in a direction given by vector $(-g_x, g_y)$ where $g = (-g_x, g_y) \in \mathbb{R}_+^{p+q}$. It is defined as

$$D(x, y; g_x, g_y) = \sup \beta > 0 : (x - \beta g_x, y + \beta g_y) \in \Psi. \tag{8}$$

It encompasses input or output oriented radial distance measures as special cases: it can indeed be seen that if $g = (x, 0)$ (resp. $g = (0, y)$), the input (resp. output) radial distance introduced above can be recovered. The choice of the direction vector g is to be done by the researcher the only restriction is that it has to be positive and in the units chosen for the inputs and the outputs.This direction can be different for each point to be evaluated but it can also be the same for all the points (x, y) which would be like assuming all the firms face the same prices.

This directional distance function can indeed be viewed as an additive definition of efficiency since to reach the frontier we subtract portions of g_x from the input x and add g_y to the outputs y. So, the distance can be defined with negative inputs or outputs. We can see that $D(x, y; g_x, g_y) \geqslant 0$ if and only if $(x, y) \in \Psi$ and a point on the frontier is characterized by $D(x, y; g_x, g_y) = 0$. It benefits from translation property and is also independent of unit of measurement (such as radial distances). The translation property can be written as

$$D(x - \eta g_x, y + \eta g_y; g_x, g_y) = D(x, y; g_x, g_y) - \eta, \forall \eta \in \mathbb{R}, \tag{9}$$

and the independence of unit of measurement as

$$D(\theta.*x, \lambda.*y; \theta.*g_x, \lambda.*g_y) = D(x, y; g_x, g_y) where \theta \in \mathbb{R}^p_+, and \lambda \in \mathbb{R}^q_+, \quad (10)$$

where .∗ refers to componentwise multiplication of vectors.

For presentation in follow let $g = (-g_x, g_y)$ be arbitrary but fixed and let $\beta(x, y) = D(x, y; g_x, g_y)$ so

$$\beta(x, y) = \sup \beta > 0 : (x - \beta g_x, y + \beta g_y) \in \Psi. \quad (11)$$

A value $\beta(x, y) = 0$ means that the producer operating at the level (x, y) is efficient, while a value $\beta(x, y) > 0$ indicates inefficiency, in the sense that it is possible to have the directional increase in all the inputs and outputs in order to reach the efficient frontier. For a given level of input and an output mix, the efficient level of input-output mix is then given by:

$$\partial(x, y) = (x, y) + \beta(x, y)g \quad (12)$$

Note that under free disposability, the directional efficiency can also be defined as:

$$\beta(x, y) = \sup \beta > 0 : \mathrm{H}_{X,Y}(x - \beta g_x, y + \beta g_y) > 0, \quad (13)$$

where $\mathrm{H}_{X,Y}(x, y) = Prob(X \leqslant x, Y \geqslant y)$ is the probability of the firm (x, y) to be dominated (see [3]Cazals et al. 2002, [8]Daouia et al. 2013, and [30]Simar et al. 2012 for details)

Since the attainable set Ψ is unknown, so is its frontier $\partial(x, y)$ and the directional efficiency score $\beta(x, y)$. The best we can do is to estimate these quantities from a sample of i.i.d. observations $\chi_n = (X_i, Y_i) : i = 1, ..., n$ generated according to the joint probability measure of (X, Y).

In order to make inference, we need to define the Data Generating Process (DGP), i.e., the statistical model providing the assumptions under which the random sample (X_i, Y_i) has been generated. Nonparametric models avoid restrictive parametric assumptions on the various components of the DGP, in order to provide inference in a very flexible and general framework.

2.2 *Non-Noisy Models and Envelopment estimators*

Most of the existing literature on nonparametric frontier has been first developed in the framework of non-noisy frontier models, where it is assumed that $Prob((X_i, Y_i) \in \Psi = 1$. In this framework the FDH estimator has been proposed by [10]Deprins, Simar and Tulkens (1984) when only free disposability of inputs and outputs are assumed and the DEA estimator, initiated by [12]Farrell (1957), assumes in addition the convexity of the attainable set.

The FDH is the most natural nonparametric estimator obtained by replacing the unknown $H_{X,Y}(x, y)$ by its empirical version, the proportion of points in the sample dominating (x,y) i.e. $\hat{H}_{X,Y,n}(x, y) = (1/n)\sum_{i=1}^{n} \Upsilon(X_i \leqslant x, Y_i \geqslant y)$ where $\Upsilon(.,.)$ is the usual indicator function. It turns out ([3]Cazals et al. 2002) that the corresponding estimator of the attainable set is given by the free disposal hull of the sample points:

$$\hat{\Psi}_{FDH,n} = (x, y) \in \mathbb{R}+^{p+q} : X_i \leqslant x, Y_i \geqslant y, forall(X_i, Y_i) \in \chi. \quad (14)$$

The DEA estimator is the convex hull of the FDH estimator, it can be written as:

$$\hat{\Psi}_{DEA,n} = (x,y) \in \mathbb{R}+^{p+q} : \begin{array}{l} \Sigma_{i=1}^{n}\gamma_i X_i \leqslant x, \\ \Sigma_{i=1}^{n}\gamma_i Y_i \geqslant y, \\ \Sigma_{i=1}^{n}\gamma_i = 1, \end{array} \quad \gamma_i \geqslant 0, \;\; i = 1, ..., n. \tag{15}$$

The resulting estimators of the efficiency scores can be obtained by plugging the estimators of Ψ.

Deterministic frontier models have the drawbacks of (i) not allowing random noise in the DGP and, as a result, (ii) being very sensitive to extreme data points and outliers. The latter issue is addressed by extensions of the DEA/ FDH estimators to partial orders frontiers, as the order-m frontiers ([3]Cazals et al. 2002) and the order-α quantile frontiers ([6]Daouia and Simar 2007), and can be extended to directional efficiency approach, but in essence, they remain non-noisy models since they do not allow for the presence of noise in the DGP. [31]Simar and Wilson (2008) propose a survey on the properties of these nonparametric estimators and their variations and they provide details on how to make inference in these models.

2.3 *Introducing Noise to Model*

When introducing noise in the model, we are not sure that all the observations are really in the attainable set and envelopment estimators are no more appropriate since they would also envelop noisy data points. This complicates the analysis and the estimation of the frontier because we have to identify inefficiency from noise when analyzing the distance from a data point to the frontier. This family of models allowing noise are known as stochastic frontier models and, due to the complexity of the models, the literature was first developed in the simpler parametric setup.

The pioneering work in this framework is [1]Aigner et al. (1977), and [25]Meeusen and van den Broek (1977). In their approaches (and all the existing variants), specific parametric analytical forms are needed for the shape of the boundary of Ψ, and for the probability structure of the noise and of the efficiency distributions. Typically, the models take the (log-) linear form

$$Y_i = \beta_0 + \beta' X_i - u_i + v_i, i = 1, ..., n \tag{16}$$

where, the inputs/outputs are usually in the log-scale and the random inefficiency term is, for instance,$u_i \sim Exp(\lambda)$ or $u_i \sim N^+(0, \sigma_u^2)$ and the noise term is $v_i \sim N(0, \sigma_v^2)$.

Generally, it is supposed that v is independent of (u, X) and that also u is independent of X. These approaches work well and have well established properties but they are limited by all these restrictive parametric hypotheses.

Recently new attempts have been developed to try to handle noise in multivariate nonparametric frontier models. [29]Simar (2007) offers a recent survey on the topic explaining the limitations of the different approaches proposed so far (this includes, e.g., [2]Atkinson and Primont 2002; [16]Gstach 1998; [23]Land et al. 1993; [26]Olesen and Petersen 1995) and suggests, as a first attempt, a DGP allowing to introduce noise in a fully multivariate nonparametric setup. It is noticed there, using results from [17]Hall and Simar (2002), that a full nonparametric specification of the DGP leads to non-identifiable models. However, if the noise is not too big in terms of noise to signal ratio, [29]Simar (2007) proposes stochastic versions of DEA/FDH estimators that perform better than the envelopment estimators in the presence of noise and perform as well when in fact there is no noise in the DGP. It appears also that these estimators are much more robust to extreme values and outliers than the original DEA/FDH estimators. Still the approach is limited to situations where the noise has a small size.

Recently, [22]Kumbhakar et al. (2007) have proposed a local maximum likelihood approach for nonparametric stochastic frontier models. They analyze local polynomial models which encompass a particular anchoring parametric model like. The local model is a parametric one and so is well identified, but by localizing the polynomial approximations, the model shares the flexibility of nonparametric models. This is a very promising approach, but so far, it was limited to one-dimensional response variable (input or output).

The idea of this paper is to extend the [22]Kumbhakar et al. (2007) ideas to a full multivariate setup by using the full multivariate DGP. This allows to handle noise in a quite general multivariate nonparametric stochastic frontier setup. In addition, the method provides attractive stochastic versions of the envelopment estimators that circumvents the main weakness of these methodspresence of noise and especially outliers.

2.4 *The DGP Framework*

In parametric models, we have, in the log-scale, a parametric frontier model with additive inefficiency u_i and noise v_i. In multivariate setups, since the efficiency measures are radial measures, the noise can be introduced, as it is done for the inefficiency, in the appropriate radial direction, by using local-spherical coordinates (see e.g., [19]Kneip et al. 1998, [20]Kneip et al. 1996, and [21]Kneip et al. 2008).

We consider first the DGP, generating points inside Ψ and then we will introduce the noise. The data (X_i, Y_i) are iid random variables generated according to a density $f_{X,Y}(x, y)$ having support Ψ. We can formulate the joint density in terms of the local-spherical coordinates (ρ, η), so that $(x, y) \Leftrightarrow (\rho, \eta)$, where $\rho \in \mathbb{R}_+$ is the modulus and $\eta_i \in [0, 2\pi]^{p+q-2}, i = 1, .., p + q - 2.$ and $\eta_{p+q-1} \in [0, \pi]$ is the amplitude (angle) of the vector (x,y). To be explicit, by $z = (x, y)$ we have : $\rho = \rho(z) = \sqrt{(z_1^2 + ... + z_p^2 + z_1^2 + ..., z_q^2)}$ and $\eta_i = \arccos(z_i/\rho), i = 1, ..., p + q - 2$, $\eta_{p+q-1} = \arccos(z_{p+q-1}/\sqrt{z_{p+q}^2 + z_{p+q-1}^2})$. in what will be flow we will transform origin to $(\min_i x_i, \max_i y_i)$ and find ρ and η for (x, y) reference to this new origin, so from now we call this new modification of spherical coordinate as local-spherical coordinate(this will help us for do efficiency analysis in right direction).

The joint density $f_{X,Y}(x, y)$ induces a density $f_{\rho,\eta}(\rho, \eta)$ on the local-spherical coordinates and we decompose this joint density as follows

$$f_{\rho,\eta}(\rho, \eta) = f_\rho(\rho \mid \eta) f_\eta(\eta), \tag{17}$$

where we assume all the densities exist. For a given (ρ, η) the frontier point $\partial(x, y)$ has a modulus which can be described through the upper boundary of the support of the density $f_\rho(\rho \mid \eta)$:

$$\rho(\partial(x, y)) = \sup \rho \in \mathbb{R}^+ : f_\rho(\rho \mid \eta) > 0 \tag{18}$$

Note that the directional efficiency using $g = (\min_i x_i, \max_i y_i) - (x, y)$ as direction is $\beta(x, y) = \sup \beta : (x, y) + \beta g \in \Psi$ and score can be redefined as

$$\beta(x, y) = \rho(x, y) - \rho(\partial(x, y)), \tag{19}$$

In a non-noisy frontier framework, in order to achieve consistency of DEA/FDH estimators, we need the free disposability of Ψ assumption when FDH estimators are used and in addition, the convexity of Ψ when DEA estimators are used. [19]Kneip et al. (1998) and [27]Park, Simar and Weiner (2000) also introduce two additional regularity conditions:

- The function $\beta(x, y)$ is differentiable in both arguments (smoothness of the frontier);
- For all η, $f_\rho(\rho \mid \eta) > 0$ (positive density on the efficient frontier).

All what precedes defines a DGP that generates data points inside the deterministic frontier of Ψ.

The radial-noise can now be introduced through the univariate modulus ρ, conditionally on angular mix η (this is the multivariate analog of the parametric stochastic model). We suppose here that the observations are made on noisy data in the radial direction: so we observe n i.i.d. random variables $(X_i, Y_i) : 1 = 1, ..., n$, with local-spherical coordinates (ρ_i, η_i) where

$$\rho_i = \rho(\partial(X_i, Y_i)) + u_i + v_i, \tag{20}$$

where$\rho(\partial(X_i, Y_i))$ is the frontier level. It is only a function of η_i. Finally, the model could be written as

$$\rho_i = r(\eta_i) + u_i + v_i, \tag{21}$$

with $u_i > 0$ and $E(v_i \mid \eta_i) = 0$. Hence $r(\eta)$ is an unknown function determining the frontier level for a given angular mix η. Conditionally on η, the random elements u and v are assumed to be independent of each other, but in our approach, we do not need the usual homoskedastic assumption: u and v can be stochastically dependent of η. Assumptions above implies that we consider conditional densities for u such that $f_u(0 \mid \eta) > 0$.

3. Local Maximum Likelihood Estimation

For a given point of interest (x, y) with local-spherical coordinates (ρ, η), the problem is back to an univariate nonparametric stochastic frontier problem: given η, we search for the estimation of $r(\eta) = \rho(\partial(x, y))$, from a random sample $(\rho_i, \eta_i), i = 1, ..., n$. For the estimation technique, we may adapt the nonparametric approach developed in [22]Kumbhakar et al. (2007) to our particular setup. It is based on local polynomial maximum likelihood estimation techniques. The idea is to choose a local polynomial approximation for the unknown function $r(\eta)$ and to choose a localized parametric model for the shape parameters of the distribution of the random convolution error $\varepsilon = u + v$, where other local polynomial approximations will be used. It is known from the literature on local maximum likelihood (see e.g. [11]Fan and Gijbels 1996) that the local parametric specification can be viewed as an anchorage model but its localization provides models sharing the flexibility of nonparametric models.

In summary, the nonparametric localized model can be written as

$$\rho = r(\eta) + u + v, \tag{22}$$

where $(u \mid \eta) \sim\mid N(0, \sigma_u^2(\eta) \mid$ and $(v \mid \eta) \sim N(0, \sigma_v^2(\eta)$, u and v being independent conditionally on η. All functional parameters $r, \sigma_u^2, \sigma_v^2$ are unknown functions and will be approximated by local polynomials. The procedure provides for any value of η estimates $\hat{r}(\eta), \hat{\sigma}_u^2(\eta), \hat{\sigma}_v^2(\eta)$ and if needed estimates of their partial derivatives with respect to the elements of η.

3.1 *local MLE in use : Local polynomial estimation*

This subsection summarizes the main aspects of local maximum likelihood estimation using local polynomial approximation for the functional parameters of a model. The presentation of these techniques in a multiple regression framework with high order polynomials can be notationally very complex. In order to simplify the notation, in this section, we will use a general notation, independent of the preceding one, but we will give the correspondence when necessary.

We consider indeed a general nonparametric stochastic frontier model which can be viewed as a nonparametric regression of a univariate dependent variable Y on a d-dimensional variable Z, where locally (for $Z = z$), the error term is the convolution of a positive random variable u (inefficiency) with a known distribution and a real random variable v having also a known distribution with zero mean, the parameters of these distributions being unknown function of z. In its most general version the model could be presented as follows.

We observe a set of i.i.d. random variables (Z_i, Y_i) for $i = 1, ..., n$ with $Z_i \in \mathbb{R}^d$ and Y_i where

$$Y_i = f(Z_i) + u_i + v_i \tag{23}$$

for some unknown function f, where, conditionally on $Z = z$, u and v independent and $\varepsilon = u + v$ has a known conditional continuous distribution $G(., \tau)$, where τ is some k-valued unknown function. The unknown functions f and τ are called the functional parameters of the model. In other words, the conditional density of Y given $Z = z$ equals:

$$\phi(y, f(z), \tau) \equiv g_\varepsilon(y - f(z), \tau), \tag{24}$$

where $g_\varepsilon(\varepsilon, \tau) = \partial G(\varepsilon, \tau)/\partial\varepsilon$. Our main interest here is estimation of the function f and eventually its derivatives. We now describe the local polynomial estimation of f in a general setting of multivariate Z.

Define $l = log\phi$. Then, the conditional log-likelihood equals $\sum_i^n l(Y_i, f(Z_i), \tau)$. Let z be a point at which one wants to estimate the values of the function f and its derivatives. A local conditional log-likelihood is obtained by replacing f in the conditional log-likelihood by its mth order polynomial approximation in a neighborhood of z and putting the weight $K_h(Z_i - z)$ for each observation (Z_i, Y_i), where $K_h(u) = h^{-d}K(h^{-1}u)$, K is a d-variate kernel function, typically a symmetric density function defined on $\mathbb{R}^d$, and h is a positive scalar, called the bandwidth. Precisely, it is given by

$$L_n(\theta_0, \theta_1, ..., \theta_{r(m)-1}, \tau; z) = \sum_i^n l(Y_i, \theta_0 + \theta_1(Z_{i1} - z_1) + ...+$$
$$\theta_{r(m)-1}(Z_{id} - z_d)^{r(m)-1}, \tau)K_h(Z_i - z)$$

where $r(m)-1$ is the total number of partial derivatives up to order m. and

$$Z_i \equiv (Z_{i1},...,Z_{id})^T, \;\; z_i \equiv (z_i,...,z_{id})^T.$$

The m-th order local polynomial estimators of f and its derivatives at z are obtained by maximizing $L_n(\theta_0,\theta_1,...,\theta_{r(m)-1},\tau;z)$. For example,$\hat{f}(z)=\hat{\theta}_0(z)$, where

$$(\hat{\theta}_0(z),\hat{\theta}_1(z),...,\hat{\theta}_{r(m)-1}(z),\hat{\tau}(z) = arg \max_{\theta(0),...,\theta(r(m)-1),\tau} L_n(\theta_0,\theta_1,...,\theta_{r(m)-1},\tau;z). \quad (25)$$

By this introduction now we can formulate more applicable version in out context. we first have to select the distribution for the convolution term. As is well known in local MLE procedure, this can be viewed as selecting an anchorage parametric model for the convolution term; the localized version will encompass this anchorage model, ensuring much more flexibility, as shown e.g. in [22]Kumbhakar et al. (2007), in [28]Park et al. (2008).

We will focus here on the two-parameters (k=2) distribution $G(.,\tau)$ obtained from the Normal/Half-Normal convolution case as proposed in [1]Aigner et al. (1977). To be explicit, we choose a local normal $N(0,\sigma_v^2(z))$ for the noise v and a local half-normal $N^+(0,\sigma_v^2(z))$ for the inefficiency u. An alternative could be the Normal/Exponential convolution also analyzed by [1]Aigner et al. (1977), but to save place we will not present this case. Although theoretically valid, we also avoid one sided distributions for the inefficiency term u that would be characterized by two free parameters (scale and shape) as the Gamma ([15]Greene 1990) or the truncated normal ([33]Stevenson 1980). The local version of our model showed to be flexible enough to handle many different situations.

So our nonparametric localized model can be written, as :

$$Y = f(Z) + u + v \quad (26)$$

where $(u \mid Z=z) \equiv N^+(0,\sigma_u^2(z))$ and $(v \mid Z=z) \equiv N^+(0,\sigma_v^2(z))$, u and v being independent conditionally on Z. The conditional probability density function of $\varepsilon = u+v$ is given by

$$g_\varepsilon(\varepsilon \mid Z=z) = \frac{2}{\sqrt{\sigma_u^2(z)+\sigma_v^2(z)}}\phi(\frac{\varepsilon}{\sqrt{\sigma_u^2+\sigma_v^2}})\Phi(\varepsilon\sqrt{\frac{\sigma_u^2(z)}{\sqrt{\sigma_v^2(z)(\sigma_u^2(z)+\sigma_v^2(z))}}}) \quad (27)$$

where $\phi(.)$ and $\Phi(.)$ are the pdf and cdf of a standard normal variable. Note that we keep the parametrization in terms of the functionals σ_v^2 and σ_u^2 because we could be interested in estimating their derivatives, but any other standard parametrization could be used.

As suggested in [22]Kumbhakar et al. (2007), in order to avoid non negativity restrictions on the variance functions in the local polynomial approximations, we rather choose the following coordinate system for the shape parameters: $\tau(z) = (\hat{\sigma}_u^2,\hat{\sigma}_v^2)^T$ where $\hat{\sigma}_u^2 = log(\sigma_u^2)$ and $\hat{\sigma}_v^2 = log(\sigma_v^2)$.

The conditional pdf of Y given Z is thus given by:

$$\phi(y, f(z), \tau(z)) = \frac{2}{\sqrt{\exp(\hat{\sigma}_u^2(z)) + \exp(\hat{\sigma}_v^2(z))}} \times \phi(\frac{(y - f(z))}{\sqrt{\exp(\hat{\sigma}_u^2) + \exp(\hat{\sigma}_v^2)}}) \times \Phi((y - f(z)) \frac{\exp(\hat{\sigma}_u^2(z)/2 - \hat{\sigma}_v^2(z)/2)}{\sqrt{\exp(\hat{\sigma}_u^2(z)) + \exp(\hat{\sigma}_v^2(z))}}) \tag{28}$$

After some analytical manipulation, it is found that

$$l((y, f(z), \tau(z))) = \log(\phi(y, f(z), \tau(z))) \propto (-1/2) \log(\exp(\hat{\sigma}_u^2(z)) + \exp(\hat{\sigma}_v^2(z))) \times - (1/2) \frac{(y - f(z))^2}{\exp(\hat{\sigma}_u^2) + \exp(\hat{\sigma}_v^2)} \times + \log \Phi((y - f(z)) \frac{\exp(\hat{\sigma}_u^2(z)/2 - \hat{\sigma}_v^2(z)/2)}{\sqrt{\exp(\hat{\sigma}_u^2(z)) + \exp(\hat{\sigma}_v^2(z))}}) \tag{29}$$

where the constants have been eliminated. The conditional local log-likelihood is

$$L(f, \tau) \propto \sum_{i=1}^{n} [(-1/2) \log(\exp(\hat{\sigma}_u^2(Z_i)) + \exp(\hat{\sigma}_v^2(Z_i)))] \times -(1/2) \frac{(Y_i - f(Z_i))^2}{\exp(\hat{\sigma}_u^2(Z_i)) + \exp(\hat{\sigma}_v^2(Z_i))} \times + \log \Phi((Y_i - f(Z_i)) \frac{\exp(\hat{\sigma}_u^2(Z_i)/2 - \hat{\sigma}_v^2(Z_i)/2)}{\sqrt{\exp(\hat{\sigma}_u^2(Z_i)) + \exp(\hat{\sigma}_v^2(Z_i))}}) \tag{30}$$

By considering the local polynomial approximations for the functional parameters f (linear approximation), $\hat{\sigma}_u^2$ (quadratic approximation) and $\hat{\sigma}_v^2$ (quadratic approximation) the desire model for estimation is in hand. we use that model for our calculation in what follows (For more details see [22]Kumbhakar et al. (2007)).

3.2 *Estimation of the individual efficiencies*

Once the model has been estimated, the estimation of the individual efficiency score for a particular fixed point $(x, y) \Leftrightarrow (\rho, \eta)$ is easy: we have indeed under our parametrization

$$E(u \mid \eta) = \sqrt{2/\pi} \sigma_u(\eta), \tag{31}$$

where $\sigma_u(\eta)$ can be estimated by $\sqrt{\hat{\sigma}_u^2(\eta)}$. This can be done for any value of η and a sensitivity analysis could be performed.

When individual efficiencies are wanted for data points in the original sample, the problem is more complicated since these data points are perturbed by noise. This is a well-known identification problem in stochastic frontier models but we can use, as suggested by [22]Kumbhakar et al. (2007), the [18]Jondrow et al. (1982) procedure to predict the individual efficiency score of a particular data point from the predicted convoluted error $\varepsilon = \rho - r(\eta)$. The procedure is as follows: again denoting η by z, it can be shown that

$$(u \mid \varepsilon, z) \sim N^{+}(\mu^{*}(z), \sigma^{2*}(z)) \tag{32}$$

i.e., a truncated (positive values) normal, where

$$\mu^{*}(z) = \frac{-\varepsilon \sigma_u^2(z)}{\sigma_u^2(z) + \sigma_v^2(z)} \tag{33}$$

$$\sigma^{2*}(z) = \frac{\sigma_u^2(z)\sigma_v^2(z)}{\sigma_u^2(z) + \sigma_v^2(z)}, \tag{34}$$

In particular, we can compute $E(u \mid \varepsilon, z)$ (see [18]Jondrow et al. (1982)). So a point predictor of the individual efficiency score for a data point (ρ_i, η_i) generated by the DGP, can be obtained.

We know of course, see [32]Simar and Wilson (2010), as in the full parametric setup, that this procedure provides point predictor with uncertain statistical properties since they are based on the observation of one residual.

4. Stochastic Version of Envelopment estimators

Thus, the above method provides, for any (x, y), an estimate of the frontier. Of course this estimator will not show the usual properties of usual envelopment estimator (DEA/FDH) (e.g. free disposability and/or convexity). Our estimator $\hat{r}(\eta)$ is consistent even under these additional assumptions, but if such properties are desired, we suggest to modify our estimator by using the appropriate FDH or DEA program on the projections of grid of points. The resulting estimated frontier is the free disposal hull (if FDH is used) or the convex free disposal hull (if DEA is used) of our nonparametric estimator derived above.

The whole procedure provides a stochastic version of the DEA/FDH frontier. It may be summarized as follows :

- Transform all the grid points (x_i, y_i) into local-spherical coordinates (ρ_i, η_i).
- Compute for each point (x_i, y_i), the estimates $\hat{r}(\eta(i))$, using the original sample χ_n
- Project the grid points on the estimated frontier to obtain $\partial(x_i, y_i)$, where $\partial(x_i, y_i) = (x_i, y_i) + (\rho_i - \hat{r}(\eta_i))g$
- For any given fixed value of interest (x, y), run a FDH or DEA program with reference set $\chi^{*} = \partial(x_i, y_i) : i = 1, ..., n$to compute an efficiency estimator.

We could also compute efficiency for $(X_i, Y_i), i = 1, ..., n$, by running n DEA or FDH programs for each of the data points generated by the DGP, with the reference set χ^{*}. Of course, due to the presence of noise in the data some of the resulting values might be smaller than 0. Due to the lack of information on the noise structure, we are unable to identify in efficiency estimator, the part which is due to noise from the part due to inefficiency.

Figure 1. Simulated Example : true frontier in black, noisy sample data in red and estimated points in blue

4.1 *Simulated Example*

We used in this section simulated data set to illustrate the flexibility of our approach. In order to evaluate the relative performances of our estimator we compute the Mean Integrated Squared Error (MISE) between the true frontier level and the nonparametric estimated one. This MISE is defined as follows :

$$MISE = \int_{\Psi} \|\partial(x,y) - \hat{\partial}(x,y)\|^2 f(x,y)dxdy, \tag{35}$$

where $\partial(x,y)$ is the true frontier and $\hat{\partial}(x,y)$ is the selected estimator of the frontier level, $f(x,y)$ is the density of (x,y) defined over Ψ and $\|.\|$ stands for the Euclidean norm. In practice, we will estimate this quantity for a given simulated example by its empirical counterpart:

$$M\hat{I}SE = 1/n \sum_{i=1}^{n} \|\partial(x,y) - \hat{\partial}(x,y)\|^2, \tag{36}$$

In what follows we will compute this $M\hat{I}SE$ for the DEA estimator and for their stochastic versions.

For a simulated sample we compute the optimal bandwidth h_{opt} for our non-parametric stochastic estimator by using Leave-One-Out Least- Squares cross validation techniques.

We simulate a sample of size 100 according the Cobb- Douglas model:

$$Y_i = X_i^{.5} \tag{37}$$

For this we use $X_i \sim \mathrm{U}(0,1)$, and draw the sample and calculate Y_i for our sample. After this we have data on frontier so for next step we transform sample data (X_i, Y_i) to local-spherical coordinate same as method described in above, as result we have (ρ_i, η_i). Now we include some inefficiency for data in radial direction with $U \sim \mathrm{N}^+(0, \sigma_u^2)$ where $\sigma_u = 0.1$ and then introduce noise in radial direction by $V \sim \mathrm{N}(0, \sigma_v^2)$ where $\sigma_v = 0.1$. below Figures illustrate inefficiency, noise, and their related frontiers.

Looking to Figure, we can see that DEA envelopment on noisy data for estimating the frontier suffers from a serious drawback in this scenario, because of its deterministic foundation, on the other hand our estimated frontier(by using DEA as envelopment tool) share the better result, because of its stochastic nature. Translating this visual explanation to number by some calculation we have $M\hat{I}SE_{DEA} = 0.02589437$ for DEA frontier by noisy data and for our proposed model we have $M\hat{I}SE_{SNFA} = 0.004846924$. as expected this show the good result for proposed model. If some one interested to evaluate the effecting of some activity (e.g. DEA or FDH estimates) this will be done by considering new frontier as reference.(refer to Figure 1)

5. Conclusion

This paper provides one promising way to handle noise in a multivariate nonparametric stochastic frontier setup. The approach also allows the researcher to shed lights on the effect of the inputs on the frontier level but also on the level of the efficiency. The procedure can handle multi-inputs and multi-outputs situations. as expected, in the case of the presence of noise in the DGP, the proposed model outperform the standard envelopment models (e.g. DEA/FDH estimators). The numerical burden in model is linked to the bandwidth selection (leave-one-out least-squares cross validation).Many of the statistical issues are still open. The statistical theory of the local polynomial MLE in this frontier setup has been done (see [22]Kumbhakar et al. 2007). We hope to be able in future work to derive the full properties of our stochastic versions of the envelopment estimators, but we know this is a serious challenge for further research, but the obtained results certainly encourage such a research.

References

[1] Aigner DJ., Lovell CAK., Schmidt P., 1997. Formulation and estimation of stochastic frontier models. Journal of Econometrics 6, 21–37.
[2] Atkinson SE., Primont D., 2002. Stochastic estimation of firm technology, and productivity growth using shadow cost and distance function. Journal of Econometrics 108, 203-225.
[3] Cazals C., Florens JP., Simar L., 2002. Nonparametric frontier estimation: a Robust approach. Journal of Econometrics 106, 1-25.
[4] Charnes A., Cooper WW., Rhodes E., 1978. Measuring the inefficiency of decision making units. European Journal of Operational Research 2, 429-444.
[5] Chambers, R.G., Y.H. Chung, Färe R., 1998. Proft, Directional Distance Functions and Nerlovian Efficiency. Journal of Optimization Theory and Applications 98, 351–364.
[6] Daouia A, Simar L., 2007. Nonparametric efficiency analysis: a multivariate conditional quantile approach. Journal of Econometrics 140, 375-400.
[7] Daraio C, Simar L., 2007. Advanced Robust and nonparametric methods in efficiency analysis: methodology and applications. Springer, New York.
[8] Daraio C, Simar L., Wilson PW., 2013. Measuring Firm Performance using Nonparametric Quantile-type Distances. TSE Working Papers 13-412, Toulouse School of Economics (TSE).
[9] Debreu G., 1951. The coefficient of resource utilization. Econometrica 19, 273-292
[10] Deprins D, Simar L, Tulkens H., 1984. Measuring labor inefficiency in post offices. In: Marchand M, Pestieau P, Tulkens H (eds) The performance of public enterprises: concepts and measurements.Amsterdam. North-Holland, pp 243-267.
[11] Fan J., Gijbels I., 1996. Local polynomial modelling and itsapplications. Chapman and Hall, London.
[12] Farrell MJ., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society 120, 253-281.
[13] Färe, R., Grosskopf, S. and Lovell C.A.K., 1985. The Measurement of Effciency of Production. Boston, Kluwer-Nijho Publishing.
[14] Färe, R., and Grosskopf S., 2000. Theory and application of dierectional distance functions. Journal of Productivity Analysis 13, 93–103.
[15] Greene WH., 1990. A gamma-distributed stochastic frontier model. Journal of Econometrics 46, 141-163.
[16] Gstach D., 1998. Another approach to data envelopment analysis in noisy environments: DEA +. Journal of Productivity Analysis 9, 161-176.
[17] Hall P., Simar L., 2002. Estimating a changepoint, boundary or frontier in the presence of observation error. Journal of the American Statistical Association 97, 523-534.
[18] Jondrow J., Lovell CAK., Materov IS., Schmidt P., 1982. On the estimation of technical inefficiency in stochastic frontier production models. Journal of Econometrics 19, 233-238.
[19] Kneip A., Park BU., Simar L., 1998. A note on the convergence of nonparametric DEA estimators for production efficiency scores. Econometric theory 14, 783-793.
[20] Kneip A., Simar L., 1996. A general framework for frontier estimation with panel data. Journal of Productivity Analysis 7, 187-212
[21] Kneip A., Simar L., Wilson PW., 2008. Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models. Econometric Theory 24, 1663-1697
[22] Kumbhakar SC., Park BU., Simar L., Tsionas EG., 2007. Nonparametric stochastic frontiers: a local likelihood approach. Journal of Econometrics 137, 1-27.
[23] Land KC., Lovell CAK., Thore S., 1993. Chance-constrained data envelopment analysis. Managerial and Decision Economics 14, 541-554.
[24] Lovell, K.C.A., Pastor J., 1995. Units invariant and translation invariant DEA models. Operations Research Letters 18, 147–151.
[25] Meeusen W., van den Broek J., 1977. Efficiency estimation from Cobb-Douglas production function with composed error. International economic review 8, 435-444.
[26] Olesen OB., Petersen NC., 1995. Chance-constrained efficiency evaluation. Management Science 41, 442-457.
[27] Park B., Simar L., Weiner Ch., 2000. The FDH estimator for productivity efficiency scores: asymptotic properties. Econometric Theory 16, 855-877.

[28] Park B., Simar L., Zelenyuk V. 2008. Local likelihood estimation of truncated regression and its partial derivatives: theory and application. Journal of Econometrics 146, 185-198.
[29] Simar L., 2007. How to improve the performances of DEA/FDH estimators in the presence of noise. Journal of Productivity Analysis 28, 183-201.
[30] Simar L., Vanhems A., Wilson PW. 2012. Statistical inference for DEA estimators of directional distances. European Journal of Operational Research 230, 853–864.
[31] Simar L., Wilson PW., 2008. Statistical inference in nonparametric frontier models: recent developments and perspectives. In: Fried H, Lovell CAK, Schmidt S (eds) The measurement of productive efficiency, 2nd edn. Oxford University Press, Oxford.
[32] Simar L., Wilson PW., 2010. Inference from cross-sectional, stochastic frontier models. Econometric Reviews 29, 62-98.
[33] Stevenson RE., 1980. Likelihood functions for generalized stochastic frontier estimation. Journal of Econometrics 13, 57-66.
[34] Shephard, R.W., 1970. Theory of Cost and Production Function. Princeton University Press, Princeton, New-Jersey.