A Note on the Convergence of Nonparametric DEA Efficiency Measures

by

Alois Kneip*, Byeong U. Park and Léopold Simar***

April 1996

Abstract

Efficiency scores of production units are measured by their distance to an estimated production frontier. Nonparametric DEA estimators are based on a finite sample of observed production units and radial distances are consisered. We investigate the consistency and the speed of convergence of these estimated efficiency scores (or of the radial distances) in the very general setup of multi-output and multi-input case. It is shown that the speed of convergence relies on the smoothness of the unknown frontier and on the number of inputs and of outputs. Furthermore, one has to distinguish between the output- and the input-oriented case.

* Institut de Statistique and CORE, Université Catholique de Louvain, Voie du Roman Pays, 34, Louvain-la-Neuve, Belgium.

** Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

Research support from the contract "Projet d'Actions de Recherche Concertées" (PARC No. 93/98-164) of the Belgian Government is gratefully acknowledged.

1 Introduction

The idea of estimating production efficiency scores in a nonparametric framework dates to the original work of Farrell (1957). The efficiency of a production unit is measured by its distance to some production frontier. This frontier has to be estimated from a set of n observed production units.

The Data Envelopment Analysis (DEA) approach is nonparametric, in the sense that no particular functional assumptions are made on the production frontier. It follows Farrell's ideas and is based on linear programming techniques. When allowing for variable returns to scale, the frontier is taken as the boundary of the convex hull of the set of observations in the input/output space. This approach relies on convexity assumptions of the attainable set of productions (see e.g.: Charnes *et al.* (1978) or Färe *et al.* (1985)).

Although the literature typically refers to DEA as being deterministic, the efficiency is measured relative to an **estimate** of the true (but unknown) production frontier. Since statistical estimators of the frontier are obtained from finite samples, the corresponding measures of efficiency are sensitive to the sampling variations of the obtained frontier. Banker (1993) and Korostelev, Simar and Tsybakov (1995a, 1995b) have given some statistical foundations to the DEA type estimators, and Simar and Wilson (1995) propose a bootstrap strategy for analyzing this sensitivity to sampling variations (see *e.g.* Grosskopf (1995) for a survey of statistical inference in DEA models).

Banker (1993) proves the consistency of the *output*-oriented efficiency scores in the case of a *single output* and no indication is given on the achieved rate of convergence. Korostelev *et al.* (1995a, 1995b) analyze the same particular problem and derive the speed of convergence not for the estimated measures of efficiency but rather for the estimated attainable production set (using either the Hausdorff metric or the Lebesgue measure of symmetric differences between the true and the estimated production sets).

Therefore, the theory of statistical inference in DEA models has still to be extended to the general multi-input *and* multi-output case for both input- and output-oriented efficiency measures. First a coherent statistical model has to be defined and then the convergence of the DEA estimators of the quantities of interest (the Farrell's efficiency scores) has to be investigated. We are unaware of any such asymtotic result in this general setup.

Section 2 introduces the basic concepts and notations in frontier models along with a brief introduction to DEA estimators. Section 3 analyses the underlying statistical model

and defines the Data Generating Process (DGP), it resumes all the needed assumptions. The basic result is in section 4 where a convergence theorem is proposed and proved. Section 5 concludes.

2 Frontier Analysis and DEA Estimators

2.1 The Frontier Model

Given a list of p inputs ($x \in I\!R^p_+$) and of q outputs ($y \in I\!R^q_+$), it is common practice in economic analysis to describe the activity of a productive organization by means of the production set Ψ of physically attainable points (x, y):

$$\Psi = \{(x, y) \in I\!R^{p+q}_+ \mid x \text{ can produce } y\}. \tag{2.1}$$

This set can be described by its sections, either an input requirement set defined $\forall\, y \in \Psi$:

$$X(y) = \{x \in I\!R^p_+ \mid (x, y) \in \Psi\} \tag{2.2}$$

or an output correspondence set defined $\forall\, x \in \Psi$:

$$Y(x) = \{y \in I\!R^q_+ \mid (x, y) \in \Psi\}. \tag{2.3}$$

clearly

$$x \in X(y) \Leftrightarrow y \in Y(x) \tag{2.4}$$

The relations between the two sets, along with standard assumptions one may reasonably make on them, are discussed in section 9.1 of Shephard (1970). Convexity of $X(y)$ for all y (and of $Y(x)$ for all x) and disposability of inputs and outputs are the most usual. The disposability assumptions correspond to some monotonicity of the frontier :

$$y_1 \leq y_2 \Rightarrow X(y_2) \subseteq X(y_1) \tag{2.5}$$

$$x_1 \leq x_2 \Rightarrow Y(x_1) \subseteq Y(x_2). \tag{2.6}$$

where the inequality between the vectors is to be understood componentwise.

The Farrell efficiency boundaries are subsets of $X(y)$ (and $Y(x)$ respectively) denoted by $\partial X(y)$ (resp. $\partial Y(x)$):

$$\partial X(y) = \{x \mid x \in X(y), \theta x \notin X(y) \quad \forall\, 0 < \theta < 1\} \tag{2.7}$$

$$\partial Y(x) = \{y \mid y \in Y(x), \beta y \notin X(y) \quad \forall\, \beta > 1\}. \tag{2.8}$$

These may be used to define the Farrell input and output measures of efficiency (respectively) for a given point (x, y):

$$\theta(x, y) = \inf\{\theta \mid \theta x \in X(y)\} \tag{2.9}$$

$$\lambda(x, y) = \sup\{\lambda \mid \lambda y \in Y(x)\}. \tag{2.10}$$

We will make the presentation for the input-oriented case. The output-oriented case largely involves a straightforward translation of the notation in what follows [1] .

The input efficiency score θ measures, for a fixed level of output y and in a radial way (keeping fixed the direction of input vector), the distance from a point (x, y) to the frontier $\partial X(y)$: it is the feasible proportionate reduction of input for a firm working at the level (x, y) for being efficient. Clearly, $\theta(x, y) \leq 1$ if and only if $x \in X(y)$. It will also be useful for later development to denote the efficient level of input, corresponding to the output level y and to the direction of input vector determined by x, as

$$x^{\partial}(y) = \theta(x, y)x \tag{2.11}$$

Note that $x^{\partial}(y)$ is the intersection of $\partial X(y)$ and the ray θx.

Since radial distance are considered, we will often refer to the polar coordinates of x defined by its modulus $\omega = \omega(x) \in I\!R^+$ and its angle $\eta = \eta(x) \in [o, \frac{\pi}{2}]^{p-1}$.

Typically, Ψ, $X(y)$, and $\partial X(y)$ are unknown; hence, for a given unit (x, y), $\theta(x, y)$ is also unknown : in this paper we are interested in finding a consistent estimator of $\theta(x, y)$ from a random sample $\mathcal{X} = \{(x_i, y_i) | i = 1, \ldots, n\}$, and we will derive that the DEA technique provides such an estimator.

2.2 The DEA Approach

The DEA approach is based on Farrell's (1957) ideas, it involves measurement of efficiency for a given unit (x, y) relative to the boundary of the convex hull of $\mathcal{X} = \{(x_i, y_i),\ i = 1, \ldots, n\}$.

[1] The Farrell's input efficiency is sometimes also described in terms of the Shephard input distance function

$$\delta(x, y) = (\theta(x, y))^{-1} = \sup\{\delta \mid \frac{x}{\delta} \in X(y)\}$$

We choose here the presentation in terms of $\theta(x, y)$, but of course everything can be translated in terms of the distance function.

More precisely, we have [2]:

$$\widehat{\Psi}_{DEA} = \{(x,y) \in I\!R^{p+q} \quad | \quad y \leq \sum_{i=1}^{n} \gamma_i y_i; x \geq \sum_{i=1}^{n} \gamma_i x_i \text{ for some } (\gamma_1, \ldots, \gamma_n) \text{ such that } \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, \ldots, n\}. \tag{2.12}$$

From (2.2) and (2.7) the input requirement set and the input-efficient boundary for the output level y is

$$\widehat{X(y)} = \{x \in I\!R^p | (x,y) \in \widehat{\Psi}_{DEA}\} \tag{2.13}$$

$$\partial\widehat{X(y)} = \{x \in I\!R^p | x \in \widehat{X(y)}, \theta x \notin \widehat{X(y)} \quad \forall 0 < \theta < 1\}. \tag{2.14}$$

Finally, for any given point (x,y), $\hat{\theta}(x,y)$ is obtained from (2.9) by solving the following linear program :

$$\hat{\theta}(x,y) = \min\{\theta > 0 \quad | \quad y \leq \sum_{i=1}^{n} \gamma_i y_i; \theta x \geq \sum_{i=1}^{n} \gamma_i x_i \text{ for some } (\gamma_1, \ldots, \gamma_n) \text{ such that } \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, \ldots, n\}. \tag{2.15}$$

Consistently with (2.11) we have :

$$\hat{x}^{\partial}(y) = \hat{\theta}(x,y)x. \tag{2.16}$$

Note that $\widehat{\Psi}_{DEA} \subseteq \Psi$, and so $\partial\widehat{X(y)}$ is an upward-biased estimator of $\partial X(y)$ and $\hat{\theta}(x,y)$ is an upward-biased estimator of $\theta(x,y)$. We have also

$$\forall (x,y) \in \Psi, \quad \theta(x,y) \leq \hat{\theta}(x,y). \tag{2.17}$$

and

$$\forall (x,y) \in \widehat{\Psi}_{DEA}, \quad 0 < \hat{\theta}(x,y) \leq 1. \tag{2.18}$$

In particular, for the sample points, we have

$$0 < \hat{\theta}(x_i, y_i) \leq 1 \text{ for } i = 1, \ldots, n. \tag{2.19}$$

We want to analyse the behavior of $(\hat{\theta}(x,y) - \theta(x,y))$ for a given (x,y) when the sample size n increases.

[2] This is the definition of $\hat{\Psi}_{DEA}$ with varying returns to scale. If the equality constraint in (2.12) is replaced by the inequality $\sum_{i=1}^{n} \gamma_i \leq 1$, this adds the origin in the feasible set and implies decreasing returns to scale. For constant returns to scale (conical hull of $\mathcal{X}$) this constraint is suppressed. Since we do not know the true production set Ψ, we use the less constraint $\hat{\Psi}_{DEA}$ given by (2.12) in what follows.

3 The Statistical Model

In order to evaluate the statistical performances of the estimator $\hat{\theta}(x,y)$, we must first define how the data $\mathcal{X}$ are generated. The model we propose is in the same lines of that presented in Simar (1995) or in Simar and Wilson (1995). In order to derive the speed of convergence, it will also be necessary to particularize the degree of smoothness of the true frontier.

Assumption A1. $\{(x_i,y_i),\ i=1,\ldots,n\}$ *are i.i.d. random variables on the convex attainable set* Ψ.

In the input-oriented case presented here, y is considered as being exogenous and for a given y, inefficiency generates x away from the efficient frontier $\partial X(y)$ along rays.

Assumption A2. y *possess a density* $f(.)$ *whose support* $\mathcal{D} \subseteq I\!R^q_+$ *is compact.*

Due to the radial nature of the inefficiency, the conditional pdf of x for a given y is more naturally introduced through the polar coordinates of x.

Assumption A3. For all $y \in \mathcal{D}$, $\eta = (\eta_1,\ldots,\eta_{p-1})$ *has a conditional pdf* $f(\eta \mid y)$ *on* $[o,\frac{\pi}{2}]^{p-1}$ *and conditional on* (y,η), *the modulus* ω *has a density* $f(\omega \mid y,\eta)$.

Note that, for a given (y,η) the frontier point $x^\partial(y)$ defined by (2.11) has modulus

$$\omega(x^\partial(y)) = inf\{\omega \in I\!R^+ \mid f(\omega \mid y,\eta) > 0\} \tag{3.1}$$

The relation between $\omega(x)$ and the input distance function $\theta(x,y)$ is given by [3]

$$\theta(x,y) = \frac{\omega(x^\partial(y))}{\omega(x)} \tag{3.2}$$

In order to achieve consistency, in any nonparametric estimation of boundaries, the DGP must insure that points will be observed near the frontier when n is sufficiently large. This imposes in our case, an assumption on the conditional pdf of ω given (y,η) :

Assumption A4. For all $y \in \mathcal{D}$ *and all* $\eta \in [o,\frac{\pi}{2}]^{p-1}$, *there exist constants* $\epsilon_1 > 0$ *and* $\epsilon_2 > 0$ *such that* $\forall \omega \in [\omega(x^\partial(y)), \omega(x^\partial(y)) + \epsilon_2]$, $f(\omega \mid y,\eta) \geq \epsilon_1$.

[3] The assumption $A3$ induces, by (3.2), a conditional pdf for $\theta(x,y)$ given (y,η) : $f(\theta \mid y,\eta)$ with support $(0,1]$. Even with this non restrictive assumption, consistency of $\hat{\theta}(x,y)$ for a given (x,y) is achieved, but if an estimation of the density of $\theta(x,y)$ is wanted, some restrictions (like homogeneity of inefficiency distributions) will facilitate the job. This was in fact done in Simar and Wilson (1995), where, in order to obtain a consistent bootstrap process, it is supposed that $f(\theta \mid y,\eta) = f(\theta)$.

Finally, the statistical performance of the estimator will also depend on the smoothness of the frontier. For simplicity, we express this smoothness in terms of the input-efficiency function $\theta(x,y)$ itself [4].

Assumption A5 : For some β, $0 < \beta \leq 1$, *some integer* $j \geq 0$ *and some finite* M_1 *we have*

$$||\theta^{(j)}(x,y_1) - \theta^{(j)}(x,y_2)|| \leq M_1||y_1 - y_2||^{\beta}$$

for all y_1 *and* y_2 *in* $\mathcal{D}$, *and for all* $x \in X(y_1) \cap X(y_2)$, *where* $\theta^{(j)}$ *denotes the* j^{th} *derivative w.r.t.* y. *Call* $\gamma = \beta + j$ *the smoothness of* $\theta(x,y)$ *w.r.t.* y.

Assumption A6 : For some α, $0 < \alpha \leq 1$, *some integer* $k \geq 0$ *and some finite* M_2 *we have*

$$||\theta^{<k>}(x_1,y) - \theta^{<k>}(x_2,y)|| \leq M_2||x_1 - x_2||^{\alpha}$$

for all $y \in \mathcal{D}$, *and for all* x_1 *and* x_2 *in* $X(y)$, *where* $\theta^{<k>}$ *denotes the* k^{th} *derivative w.r.t.* x. *Call* $\nu = \alpha + k$ *the smoothness of* $\theta(x,y)$ *w.r.t.* x.

In many situations, we will have γ and ν both greater ($>$) than 1 : this is the case of continuous and differentiable frontiers in both directions. But if we take for instance a piecewise linear frontier (like the facets of the convex polyhedra constructed by the DEA approach), we have $\beta = \gamma = 1$ and $\alpha = \nu = 1$. In fact, we derive the consistency and the speed of convergence for true frontiers which are even less smooth (γ and/or ν less than 1), but this will slow down the speed of convergence, as stated in the next theorem.

4 The Convergence Theorem

We now state the theorem which proves the consistency of $\hat{\theta}(x,y)$ [5] for a given (x,y) and gives the rates of convergence to the true but unknown $\theta(x,y)$ [6] .

Theorem 1 *Under Assumptions A1 to A6, we have for all* (x,y) *in the interior of* Ψ :

[4]The assumptions $A5$ and $A6$ could also equivalently be written in terms of the Shephard input-distance function $\delta(x,y)$.

[5]We could have the equivalent theorem for $\hat{\delta}(x,y) = (\hat{\theta}(x,y))^{-1}$ the DEA estimator of the Shephard distance function $\delta(x,y)$. We could also obtain the same results for the output-efficiency scores $\lambda(x,y)$ defined in (2.10), but the role of p and q has to be reversed !

[6]Note the particular role of p, the dimension in which we are defining the frontier (the input space in our case). Note also, that in most applications, γ and ν may be both larger than 1 : in that case, the theorem simplifies since $min(1,\nu) = min(1,\nu,\gamma) = 1$.

(1) If $0 < \gamma \leq 2$,

(a) If $p = 1$ *:* $\hat{\theta}(x,y) - \theta(x,y) = \mathcal{O}_P(n^{-\frac{\gamma}{q+\gamma}})$

(b) If $p > 1$ *:* $\hat{\theta}(x,y) - \theta(x,y) = \mathcal{O}_P(n^{-\frac{min(1,\nu,\gamma)}{p+q-1+min(1,\nu,\gamma)}})$

(2) If $\gamma > 2$

(a) If $p = 1$ *:* $\hat{\theta}(x,y) - \theta(x,y) = \mathcal{O}_P(n^{-\frac{2}{q+2}})$

(b) If $p > 1$ *:* $\hat{\theta}(x,y) - \theta(x,y) = \mathcal{O}_P(n^{-\frac{min(1,\nu)}{p+q-1+min(1,\nu)}})$

Proof : Let (x,y) be in the interior of Ψ, so y is some element in the interior of $Y(x) \subseteq \mathcal{D}$ and x is in the interior of $X(y)$. For some $h > 0$, and for $i = 1, \ldots, q$ let $y^{(2i-1)} := y + he_i$ and $y^{(2i)} := y - he_i$, where e_i is the i^{th} column of the identity matrix of order q.

For $s = 1, \ldots, 2q$, let B_s denote the $q-$dimensional ball with radius $\frac{h}{2}$ around $y^{(s)}$. If h is sufficiently small, $B_s \subset Y(x) \subseteq \mathcal{D}$. By the convexity of $Y(x)$, we have

(i) For any combination of $z_1, \ldots, z_{2q}, z_s \in B_s, s = 1, \ldots, 2q$, there exist $\alpha_1, \ldots, \alpha_{2q} \geq 0, \sum_{s=1}^{2q} \alpha_s = 1$ such that $y = \sum_{s=1}^{2q} \alpha_s z_s$.

(ii) We also have

$$\theta(x,y) = \theta(x, \sum_{s=1}^{2q} \alpha_s z_s) \leq \sum_{s=1}^{2q} \alpha_s \theta(x, z_s) \tag{4.1}$$

For the last inequality of (4.1), observe that $z_s \in Y(x)$ so that $x \in X(z_s)$ and $x^{\partial}(z_s) = \theta(x,z_s)x$. Now define

$$x^* = \sum_{s=1}^{2q} \alpha_s x^{\partial}(z_s) = (\sum_{s=1}^{2q} \alpha_s \theta(x,z_s))x.$$

Then, by convexity of Ψ, $(x^*, y) \in \Psi$ and so $x^* \in X(y)$ which implies by definition of $\theta(x,y)$ the inequality (4.1). Note that the same inequality holds for $\hat{\theta}(x,y)$ by the same reasons.

Consider also $\mathcal{C}_h(x)$, the conical hull of a ball of radius h around x in $I\!R_+^p$, *i.e.*

$$\mathcal{C}_h(x) = \{x' \in I\!R_+^p \mid \exists \mu > 0 \text{ s.t. } ||\mu x' - x|| \leq h\} \tag{4.2}$$

In fact, $\mathcal{C}_h(x)$ is, in polar coordinates, a $h-$neighbourhood of η, the angle of x.

For the given (x,y) consider now a particular choice of $(x_s, z_s) \in \mathcal{C}_h(x) \times B_s, s = 1, \ldots, 2q$, let

$$(x_s, z_s) := arg\ max_{(x_i,y_i) \in \mathcal{C}_h(x) \times B_s} (\theta(x_i, y_i)), \tag{4.3}$$

so that (x_s, z_s) equals the particular observation (x_i, y_i) such that $\theta(x_i, y_i)$ is maximal with respect to all the observations falling into $\mathcal{C}_h(x) \times B_s$.

Since B_s has radius $\frac{h}{2}$ and $\mathcal{C}_h(x)$ is a $h-$neighbourhood of $\eta = (\eta_1, \ldots, \eta_{p-1})$, the number of observations falling into $\mathcal{C}_h(x) \times B_s$ is proportional to nh^{q+p-1}. Hence, due to $A4$,

$$\theta(x_s, z_s) = 1 - \mathcal{O}_P(n^{-1}h^{-(q+p-1)}). \tag{4.4}$$

Since $x_s \in \mathcal{C}_h(x)$, there exists a $\mu_s > 0$ such that $||\mu_s x_s - x|| = \mathcal{O}_P^{*p}(h)$ where [7]

$$\mathcal{O}_P^{*p}(h) = 0 \quad \text{if } p = 1 \tag{4.5}$$

$$\mathcal{O}_P^{*p}(h) = \mathcal{O}_P(h) \text{ if } p > 1. \tag{4.6}$$

Therefore, we have by $A6$

$$|\theta(\mu_s x_s, z_s) - \theta(x, z_s)| = \mathcal{O}_P^{*p}(h^{min(1,\nu)}) \tag{4.7}$$

and since the DEA frontier is piecewize linear we have :

$$|\hat{\theta}(\mu_s x_s, z_s) - \hat{\theta}(x, z_s)| = \mathcal{O}_P^{*p}(h) \tag{4.8}$$

Now, due to the radial nature of the distance functions, we have

$$\forall (x, y) \in \Psi \text{ and } \forall \mu > 0, \frac{\hat{\theta}(x, y)}{\theta(x, y)} = \frac{\hat{\theta}(\mu x, y)}{\theta(\mu x, y)} \tag{4.9}$$

Finally, we obtain from (4.7) to (4.9)

$$\frac{\hat{\theta}(x_s, z_s)}{\theta(x_s, z_s)} = \frac{\hat{\theta}(\mu_s x_s, z_s)}{\theta(\mu_s x_s, z_s)} = \frac{\hat{\theta}(x, z_s)}{\theta(x, z_s)} + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) \tag{4.10}$$

or equivalently

$$\frac{\hat{\theta}(x, z_s)}{\theta(x, z_s)} = \frac{\hat{\theta}(x_s, z_s)}{\theta(x_s, z_s)} + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) \tag{4.11}$$

$$\leq \frac{1}{\theta(x_s, z_s)} + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) \tag{4.12}$$

where the last inequality follows from (2.19). Now we have

$$\hat{\theta}(x, z_s) \leq \frac{\theta(x, z_s)}{\theta(x_s, z_s)} + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) \tag{4.13}$$

[7] There is indeed a qualitative step when passing from $p = 1$ to $p > 1$, this is due to the fact that when $p = 1$ the ray of x is trivial (η is undefined) : for all x_s there exists a μ_s such that $\mu_s x_s = x$.

Applying (4.4), we obtain

$$\hat{\theta}(x, z_s) \leq \theta(x, z_s) + \mathcal{O}_P(n^{-1}h^{-(q+p-1)}) + \mathcal{O}_P^{*p}(h^{min(1,\nu)}). \tag{4.14}$$

Combining the last inequality with (4.1) and (2.17) we arrive at the following basic inequality

$$\theta(x, y) \leq \hat{\theta}(x, y) \leq \sum_{s=1}^{2q} \alpha_s \hat{\theta}(x, z_s) \leq \sum_{s=1}^{2q} \alpha_s \theta(x, z_s) + \mathcal{O}_P(n^{-1}h^{-(q+p-1)}) + \mathcal{O}_P^{*p}(h^{min(1,\nu)}). \tag{4.15}$$

Now the theorem follows from the particularization of (4.15) to the smoothness of the frontier.

(i) If $0 < \gamma \leq 1$

$$\sum_{s=1}^{2q} \alpha_s \theta(x, z_s) = \sum_{s=1}^{2q} \alpha_s \theta(x, y) + \mathcal{O}_P(h^{\gamma}) \tag{4.16}$$

$$= \theta(x, y) + \mathcal{O}_P(h^{\gamma}). \tag{4.17}$$

So we have

$$\theta(x, y) \leq \hat{\theta}(x, y) \leq \theta(x, y) + \mathcal{O}_P(n^{-1}h^{-(q+p-1)}) + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) + \mathcal{O}_P(h^{\gamma}). \tag{4.18}$$

(ii) If $1 < \gamma \leq 2$

$$\sum_{s=1}^{2q} \alpha_s \theta(x, z_s) = \sum_{s=1}^{2q} \alpha_s \theta(x, y) + \sum_{s=1}^{2q} \alpha_s \theta^{(1)'}(x, y)(z_s - y) + \mathcal{O}_P(h^{\beta+1}) \tag{4.19}$$

$$= \theta(x, y) + 0 + \mathcal{O}_P(h^{\gamma}). \tag{4.20}$$

So we have again

$$\theta(x, y) \leq \hat{\theta}(x, y) \leq \theta(x, y) + \mathcal{O}_P(n^{-1}h^{-(q+p-1)}) + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) + \mathcal{O}_P(h^{\gamma}). \tag{4.21}$$

(iii) If $\gamma > 2$

$$\sum_{s=1}^{2q} \alpha_s \theta(x, z_s) = \sum_{s=1}^{2q} \alpha_s \theta(x, y) + \sum_{s=1}^{2q} \alpha_s \theta^{(1)'}(x, y)(z_s - y) + \mathcal{O}_P(h^2) \tag{4.22}$$

$$= \theta(x, y) + 0 + \mathcal{O}_P(h^2). \tag{4.23}$$

Here we obtain

$$\theta(x, y) \leq \hat{\theta}(x, y) \leq \theta(x, y) + \mathcal{O}_P(n^{-1}h^{-(q+p-1)}) + \mathcal{O}_P^{*p}(h^{min(1,\nu)}) + \mathcal{O}_P(h^2). \tag{4.24}$$

Choosing for some constant $c > 0$

$$h = cn^{-\frac{1}{q+min(\gamma,2)}} \text{ if } p = 1 \tag{4.25}$$

$$= cn^{-\frac{1}{p+q-1+min(1,\nu,\gamma)}} \text{ if } p > 1, \tag{4.26}$$

completes the proof of the theorem.

5 Conclusions

In this paper, we have proven the consistency of the DEA nonparametric measures of efficiency in the very general setup of a multi-output and multi-input case. A statistical model has been proposed which allows to define the data generating process. The speed of convergence has also been investigated.

Convergence rates deteriorate for high dimensional problems (which is known as the curse of dimensionality in most of the nonparametric problems). We pointed the particular role of the number of inputs for the estimation of the input-oriented measure : there is a qualitative step going from $p = 1$ to $p > 1$ (we would have obtained the same result for the number of outputs in the output-oriented framework).

Finally, as usual in nonparametric curves estimation, the degree of smoothness of the true frontier influences the speed of convergence of the DEA estimator.

References

[1] Banker, R.D. (1993), Maximum Likelihood, consistency and data envelopment analysis : a statistical foundation, *Management Science*, 39,10,1265-1273.

[2] Charnes, A., Cooper W.W. and E. Rhodes (1978), Measuring the inefficiency of decision making units, *European Journal of Operational Research* 2, 429-444.

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

[4] Färe, R., Grosskopf, S. and C.A.K. Lovell (1985), *The Measurement of Efficiency of Production*. Boston, Kluwer-Nijhoff Publishing.

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

[6] Grosskopf,S. (1995), Statistical inference and nonparametric efficiency : a selective survey, manuscript to appear in *Journal of Productivity Analysis*.

[7] Korostelev, A., Simar, L. and A. Tsybakov (1995a), Efficient estimation of monotone boundaries, *The Annals of Statistics*, Vol. 23, (2), 476-489.

[8] Korostelev, A., Simar, L. and A. Tsybakov (1995b), On estimation of monotone and convex boundaries, *Pub. Inst. Stat. Univ. Paris*, XXXIX, 1, 3-18.

[9] Shephard, R.W. (1970), *Theory of Cost and Production Function*. Princeton University Press, Princeton, New-Jersey.

[10] Simar, L. (1995), Aspects of Statistical Analysis in DEA-Type Frontier Models, Discussion paper 9512, Institut de Statistique, UCL, Louvain-la-Neuve, Belgium, to appear in *Journal of Productivity Analysis*.

[11] Simar, L. and P. Wilson (1995), Sensitivity analysis of efficiency scores : how to bootstrap in nonparametric frontier models, Discussion paper 9503, Institut de Statistique, UCL, Louvain-la-Neuve, Belgium, to appear in *Management Science*.