A new approach to stochastic frontier estimation: DEA+

Dieter Gstach [1]
Working Paper No. 39
Vienna University of Economics

August 1996

[1] University of Economics, VWL 6, Augasse 2-6, 1090 VIENNA, Austria.
phone: +43 (1) 31336-4965, fax: +43 (1) 31336-755, e-mail: gstach@wu-wien.ac.at

Abstract

The outcome of a production process might not only deviate from a theoretical maximum due to inefficiency, but also because of non-controllable influences. This raises the issue of reliability of Data Envelopment Analysis in noisy environments. I propose to assume an i.i.d. data generating process with bounded noise component, so that the following approach is feasible: Use DEA to estimate a pseudo frontier first (nonparametric shape estimation). Next apply a ML-technique to the DEA-estimated efficiencies, to estimate the scalar value by which this pseudo-frontier must be shifted downward to get the true production frontier (location estimation). I prove, that this approach yields consistent estimates of the true frontier.

Keywords: Stochastic DEA, Consistency, Semi-Parametric Frontier Estimation, MLE

JEL-classification: C14, C24, D24

1 Introduction

Frontier estimation based on the early programming approaches ruled out the possibility of noisy data by assumption. So deviations from the frontier could only mean inefficiency. There was no room for uncontrollable good or bad influences to the production process. This might not bother in some circumstances but it definitely does in others.

Early attempts to cope with this unsatisfactory state appeared in the form of chance constraint programming (CCP), as proposed for example in Sengupta (1987), Land, Lovell and Thore (1993), Olesen and Petersen (1993) and Desai, Ratick and Schinnar (1994). This approach makes clear that without one or the other parametric assumption one does not get very far. CCP introduces probability measures on the inequality conditions and modifies the objective function of the programming approach to reflect violations of these inequality conditions appropriately. Furthermore as CCP does not impose parametric restrictions on the type of technology, it has much in common with the DEA+ approach to be proposed below.

Another way of introducing the notion of robustness into DEA is to apply a traditional regression technique to the sample of DEA-efficient observations and to trace out some parametric production (or whatever) frontier. This is the approach taken for example in Thiry and Tulkens (1992), Bardan et.al. (1994) or Sueyoshi (1994). So DEA there is used as a first step screening device. But this does not adress the question of noisy data properly, as the outcome of the second step has no influence on the efficiency estimates from the first step. So no correction takes place.

In a paper by Banker and Maindiratta (1992) a fullscale offence was launched towards the problem at hand. The proposition there is to estimate a monotone and concave frontier by Maximum Likelihood Techniques, imposing parametric structure only on the error and efficiency distributions. No such parametrizing was forced upon technique, because Banker and Maindiratta proposed to estimate each single output observation along with the parameters of the densities involved. The authors showed how to proceed in principle, but mentioned that statistical properties of the estimators are hard to come up with. So this approach has to prove applicability yet.

A very recent attempt that needs to be mentioned comes from Kneip and Simar (1995). They got rid of (almost) all parametrization by assuming the existence of panel data. Assuming in addition time invariance of firm specific efficiencies they went along estimating (nonparametrically) the firm averages and based on them their final DEA frontier estimates. This sort of approach is basically extendable to cover time varying efficiencies too, but of course any such assumption about a chronological structure amounts to replace one parametrization by another.

In section 2 I will propose another approach to deal with the problem of stochastics in DEA, that is semi-parametric but does not require panel data. Before presenting the idea, I shortly have to mention a paper by Banker (1994) that proved consistency of DEA frontier estimates under quite general conditions. Although this proof was formulated for the traditional, deterministic DEA setting, it is highly useful in DEA+ too, as will become clear.

Finally let me draw attention to a paper by Greene (1980), where the problems of Maximum Likelihood Estimation of frontiers in a regression context are discussed at length. In Greenes paper conditions are given under which the typical MLE properties (like consistency and asymptotic efficiency) can be invoked, despite a very particular problem in the frontier estimation context. The problem he adresses there is dependency of the range of the dependent variable (typically output) on the parameters to be estimated (regression coefficients and density parameters). This renders the typical proof of MLE properties impossible. DEA+ has to cope with this problem too, as it uses an ML-technique and claims consistency. DEA+ basically sidesteps the problem by using separate techniques to estimate shape and location of the frontier. So when it comes to MLE, the range of the dependent variable (in this case a rescaled sum of noise and inefficiency) will already be determined.

2 DEA+

The approach I propose basically consists of three steps. These steps of DEA+, as outlined below, ultimately depend on one crucial and unusual assumption, which is the boundedness of noise. In all other respects, my way of modelling stochastics equals the standard procedure in parametric stochastic frontier estimation (=SFE), that is:

$$y = g(x)e^{w}, \quad w \equiv v - u, \qquad v \leq v_{max}, \quad u \geq 0 \tag{1}$$

As with SFE I assume that v and u are i.i.d. random variables with parametrized densities $f_V(v \mid .)$ and $f_U(u \mid .)$. But $f_V(.)$ has a restricted domain $(-\infty, v_{max}]$ here, while u as usual is defined on $[0, \infty)$. The density of w is then of course $f_W(w) = \int_w^{v_{max}} f_V(v) f_U(v - w) dv$. Also unlike in the SFE approach I now employ the DEA assumptions on technology instead of parametrization, i.e.:

$$g(x) \quad \text{is monotonically increasing and concave in} x \tag{2}$$

So clearly (1) is a stochastic frontier model, only slightly different from the ones usually encountered in SFE. But with noise assumed to be bounded from above and i.i.d. error components v and u, there exists a unique deterministic representation of (1), which is

$$y = \tilde{g}(x)\epsilon, \quad \epsilon \leq 1 \tag{3}$$

where the linking equations between (1) and (3) are

$$\epsilon \equiv e^{\tilde{w}}, \quad \tilde{w} \equiv \tilde{v} - u, \quad \tilde{g}(x) \equiv g(x)e^{v_{max}}, \quad \tilde{v} \equiv v - v_{max} \tag{4}$$

Because the deterministic model $\tilde{g}(x)$ is merely a rescaled version of $g(x)$ and the latter satisfies condition (2), so does $\tilde{g}(x)$ which makes model (3) a candidate for DEA. The single-output character of the production process carried along from definition (1) is unduly restrictive from a DEA perspective. But given the assumption of a common disturbance and inefficiency for all output generating processes, the generalization to the multi-output case would be straightforward. For expositional simplicity I will stick to the single-output formulation. [1]

Note further that $f_{\widetilde{W}}(.)$ is just a shifted version of $f_W(.)$ and can be defined as

$$f_{\widetilde{W}}(\widetilde{w}) = \int_{\widetilde{w}}^{0} f_V(v + v_{max}) f_U(v - \widetilde{w}) dv \tag{5}$$

So the natural first step of this DEA+ approach to efficiency measurement in noisy environments can be compactly described as

[1] Actually such an assumption is implicit to all radial variants of efficiency measures employed in the DEA context.

STEP I: Estimate ϵ_i via DEA for model (3).

Assume that this step leads to firm-specific pseudo efficiencies $\{\widehat{\epsilon}_i\}_{i=1}^n$ and corresponding $\widehat{\widetilde{w}}_i = ln(\widehat{\epsilon}_i)$. The next step is to estimate the density $f_{\widetilde{W}}(\widetilde{w}|\theta)$, where the parametrization

$$\theta \equiv \{\theta_V, v_{max}, \theta_U\} \tag{6}$$

is in terms of the parameters of the underlying densities, and it is understood that θ_V are parameters of $f_V(.)$ and θ_U are parameters of $f_U(.)$. The term v_{max} is additionally needed to fully determine $f_V(.)$ and is separately mentioned because of its key role in the concept of DEA+.

The estimation of θ shall be of the ML-variety to take care of the truncated nature of the $\widetilde{w}_i$-distribution and will be based on a sample of $\widehat{\widetilde{w}}_i$ values. I will exclude from this DEA-based sample all observations i with $\widehat{\widetilde{w}}_i = 0$, i.e. all fully efficient observations and I will call the index-set of the remaining observations $C^- \equiv \{j : \widehat{\widetilde{w}}_j < 0\}$ The reason for this screening is discontinuity of the log-likelihood at $\widehat{\widetilde{w}}_i = 0$. Asymptotically this makes no difference of course, as $\lim_{n\to\infty} [\sum_{i=1}^n \mathrm{I}_{i \notin C^-}(i)]/n = 0$ where I(.) is the indicator function.[2] Thus I can formalize this step as

$$\max_\theta \ \ln\Big[\prod_{i\in C^-} f_W(\widehat{\widetilde{w}}_i|\theta)\Big] \tag{7}$$

Strictly speaking (7) is not a maximum-likelihood procedure, because it ignores the interdependencies among the $\widehat{\widetilde{w}}_i$-terms. It nevertheless is asymptotically equivalent to correct MLE (see proof below) and therefore I will simply describe this second step as follows:

STEP II: Estimate θ via MLE based on the sample $\{\widehat{\widetilde{w}}_i\}_{i\in C^-}$.

This yields also an estimate $\hat{v}_{max}$ as part of the estimated parameter vector $\hat{\theta}$. In the next section I will show that this estimate is consistent. So I claim to have consistent point estimates of the frontier $\hat{\tilde{g}}(x_i) = y_i/\hat{\epsilon}_i$ and a consistent estimate $\hat{v}_{max}$ of the distance between the noisy boundary and the true production frontier. A final step links these results together:

STEP III: Estimate $g(x_i)$ via $\hat{g}(x_i) \equiv \hat{\tilde{g}}(x_i)/e^{\hat{v}_{max}}$.

But this Step III estimator is a continuous and real-valued function in arguments that are consistent estimators of some other quantities, and this guarantees consistency of $\hat{g}(x)$ as an estimator of $g(x) \equiv \tilde{g}(x)/e^{v_{max}}$.

So finally we get a semi-parametric estimate of shape and location of a production frontier, that can be applied to noisy data as well, given the assumption of bounded noise is correct. All that remains to be shown is consistency of the estimator given in Step II.

3 Consistency of DEA+

In this section a simple theorem will be formulated that concerns one aspect of asymptotic behavior of the Step II estimator: Consistency. I will not discuss asymptotic normality or efficiency. [3]

[2] This problem is also of little practical importance, when using a bootstrapping variant of DEA, which includes a bias correction. This way very few estimated pseudo-efficiencies will equal one. See Simar and Wilson (1995) or Gstach (1995) for details.

[3] The basic reason being lack of continuous differentiability of the efficiency estimators as functions of the x_i and w_i terms. This continuity is lost during the process of linear programming. It is beyond the scope of this paper to investigate possible alternative conditions to establish the missing asymptotic properties of $\hat{\theta}$.

Basically, it is not hard to imagine that the combination of a consistent estimation procedure (ML) with consistent data estimates ($\widehat{\widetilde{w}}_i$) will yield a consistent estimator in the present context. It should follow from continuity of the underlying densities together with continuity and consistency of the DEA estimators. But neither does consistency hold everywhere strictly nor does continuity, as will become clear. This needs to be taken care of.

In order to prove consistency, I will check the conditions of the following fundamental theorem, which is taken from Bierens (1994):

> **Theorem 1 (Bierens)** *Let $Q_n(\theta)$ be a sequence of random functions on a compact set $\Theta \subset R^m$ such that for a real continuous function $Q(\theta)$ on Θ:*
> *$Q_n(\theta) \longrightarrow Q(\theta)$ in probability pseudo-uniformly on Θ.*
> *Let θ_n be any random vector in Θ satisfying $Q_n(\theta_n) = \sup_{\theta \in \Theta} Q_n(\theta)$*
> *and let θ_0 be a unique point in Θ such that $Q(\theta_0) = \sup_{\theta \in \Theta} Q(\theta)$.*
> *Then: $\theta_n \longrightarrow \theta$ in probability.*

Defining $Q_n(.)$ as an average log-likelihood function it is possible to investigate consistency of an ML-kind of estimator like the one defined in (7). Indexing $Q_n(.)$ with a subscript n emphasizes that we are actually dealing with a random function that depends on some sort of sample information, where the sample size is n. As sample size increases, i.e. asymptotically, this dependence should vanish though, as the convergence condition $Q_n(.) \rightarrow Q(.)$ indicates. In addition to the conditions in Theorem 1 the following requirement has to be checked:

$$\lim_{n \to \infty} \mathrm{E}[Q_n(\theta)] = Q(\theta) \tag{8}$$

Some additional notation for convenience: Define a function $\widehat{Q}_n(.)$ as pseudo average log-likelihood function:

$$\widehat{Q}_n(\theta) \equiv \ln\Big[\textstyle\prod_{i \in C^-} f_W(\widehat{\widetilde{w}}_i|\theta)\Big]/n \tag{9}$$

where the index n refers to a specific sample of size n of observed input-vectors and output-vectors and an unknown n-dimensional vector of composed errors. As required above let $\hat{\theta}_n$ satisfy

$$\widehat{Q}_n(\hat{\theta}_n) \equiv \sup_{\theta \in \Theta} \ln\Big[\textstyle\prod_{i \in C^-} f_W(\widehat{\widetilde{w}}_i|\theta)\Big]/n \tag{10}$$

Next let $\Omega \equiv X \times (-\infty, 0]$ denote the involved sample space with $X \subset R^m$ and typical element x_i. A vector of x_i-observations will be denoted with bold letter $\mathbf{x}_n = \{x_1, x_2, ...x_n\}$. It is clear, that randomness in the context of DEA+ stems not only from inefficiency and noise, but also from the sample design, identified with $\mathbf{x}_n$. Let further

$$0 \leq x_i \leq x_{max} \qquad \forall \quad x_i \in X, \tag{11}$$

so the interior of X, which I designate $\mathcal{I}(X)$, is well defined. Then define a (possibly multivariate) density function $f_X(.)$ on X, with

$$f_X(x_i)\Big|_{x_i \in \mathcal{I}(X)} > 0. \tag{12}$$

The last two requirements do not aim at Theorem 1 but are required to legitimize the use of Banker's consistency result, which is:

> **Theorem 2 (Banker, 1993)** *Let $g(x)$ define a production frontier on a compact input space $X \subset R^m$, with $g(x)$ monotone increasing and concave. Assume that*

observed outputs from this technology might deviate from the frontier due to inefficiency u*, such that* $y \leq g(x)$*, with a density function* $f_U(u)$ *defined on an interval* $W \subseteq (0,1]$ *and satisfying* $f_U(u)|_{u \in W} > 0$*. Assume also, that sampling elements* x_i *from* X *is ruled by a density function* $f_X(x_i)$ *that satisfies condition (11).*
Then the DEA estimator $\hat{g}_n(x)$ *of this frontier, based on a sample of size* n *and evaluated at point* x *is weakly consistent for all* $x \in \mathcal{I}(X)$*.*

Before formulating the next theorem, some additional notation: Let a continuously differentiable density function $f_V(v)$ from a parametric family be defined on a closed set $V \subseteq (-\infty, 0]$ with $\sup\{V\} = 0$. $f_V(v)$ shall be parametrized by $\{\theta_V, v_{max}\}$ where θ_V may be a vector and the scalar v_{max} satisfies $v_{max} = -\mathrm{E}[v]$. It is assumed that

$$f_V(v)\Big|_{\substack{v<0 \\ v \in \mathcal{V}}} > 0 \tag{13}$$

A random variable $v + v_{max}$ will be referred to as noise below. Then let $f_U(u)$ denote a continuously differentiable density function on the domain $[0, \infty)$, parametrized by a vector θ_U. Analogously to (12) assume that

$$f_U(u)\Big|_{u>0} > 0 \tag{14}$$

The u-term will be called inefficiency. Now it is possible to state the following in compact form:

Theorem 3 *Let* $w \equiv v - u$ *define a composed random variable with underlying densities* $f_V(v)$ *and* $f_U(u)$ *as defined above. Assume the maximum likelihood estimator* $\hat{\theta}$ *of the true parameter vector* $\theta_0 \equiv \{\theta_v, v_{max}, \theta_u\}$ *of* $f_{V-U}(w)$ *is well defined.*
Let further $y = g(x)e^{\tilde{v}-u}$ *describe a stochastic frontier model, where the noise component is defined through* $\tilde{v} = v + v_{max}$*, thus satisfying* $E(\tilde{v}) = 0$*.*
Then the DEA+ frontier estimator $\hat{g}(x)$ *as defined in steps 1 to 3 in section 1 is weakly consistent in the sense that* $\hat{g}(x) \to g(x)$ *in probability for all* $X \in \mathcal{I}(X)$*.*

Proof: Define a set $N = \{\{x : x \notin \mathcal{I}(X)\} \times \{0\}\}$. Because of conditions (11) - (13), $\Pr[\{x, w\} \in N] = 0$, i.e. N is a null set. This zero-probability allows to ignore some events, where I encounter continuity problems. I will show first, that $\widehat{Q}_n(\theta)$ as defined in (8) converges in probability uniformly to

$$Q_n(\theta) \equiv \ln[\textstyle\prod_{i \in C^-} f_W(w_i|\theta)] / \sum_{i=1..n} \mathrm{I}_{i \in C^-}(i) \tag{15}$$

by contradiction. So assume instead, this would not hold. Then
$\lim_{n\to\infty} \Pr[\sup_{\theta\in\Theta} |\widehat{Q}_n(\theta) - Q_n(\theta)| \leq \epsilon] < 1$ for some $\epsilon_0 > 0$. But this by continuity of the supremum function would imply existence of a set $X' \times W' \subseteq \Omega \setminus N$ and an infinite dimensional vector $\{\mathbf{x}_\infty, \mathbf{w}_\infty\} \equiv \{x_i, w_i\}_{i=1}^{\infty}$ such that $x_i \in X' \times W' \quad \forall i = 1, 2, ...$ and $\sup_{\theta\in\Theta} |\widehat{Q}_\infty(\theta, \mathbf{x}_\infty, \mathbf{w}_\infty) - Q_\infty(\theta, \mathbf{x}_\infty, \mathbf{w}_\infty)| > \epsilon_0$. But as $n \to \infty$, $\widehat{\widetilde{w}}_i$ approaches $\widetilde{w}_i$ and therefore the difference between $\widehat{Q}_n(\theta)$ and $Q_n(\theta)$ should vanish for all $\theta \in \Theta$ and all $\mathbf{x}_\infty \in \mathcal{I}(X)$. So the supremal absolute difference would eventually get zero and thus $0 > \epsilon_0$. This is a contradiction and I conclude

$$\widehat{Q}_n(\theta) \to Q_n(\theta) \qquad \text{in probability uniformly.} \tag{16}$$

Assuming a well behaved ML-problem when estimating θ with truly observed error values $\{w_i\}_{i=1}^n$ on the other hand implies

$$Q_n(\theta) \to Q(\theta) \quad \text{in probability uniformly.} \tag{17}$$

Now define mappings $\hat{\theta}_1(.) : \Omega \mapsto \Theta$ and $\hat{\theta}_2(.) : \Omega \mapsto \Theta$ to describe the maximands of the supremal absolute differences $\hat{Q}_n(.) - Q_n(.)$ and $Q_n(.) - Q(.)$ respectively and reformulate (16) and (17) in set notation as:

$$\begin{aligned}
\forall\ \epsilon_1 > 0 \qquad & \exists\ \Omega_{1,n} \equiv \Big\{\{\mathbf{x}_n, \mathbf{w}_n\} : \Big|\hat{Q}_n(\hat{\theta}_1(\mathbf{x}_n, \mathbf{w}_n)) - Q_n(\hat{\theta}_1(\mathbf{x}_n, \mathbf{w}_n))\Big| \le \epsilon_1\Big\} \\
& \text{with} \quad \lim_{n\to\infty} \Pr\Big[\{\mathbf{x}_n, \mathbf{w}_n\} \in \Omega_{1,n}\Big] = 1 \qquad (18) \\
\forall\ \epsilon_2 > 0 \qquad & \exists\ \Omega_{2,n} \equiv \Big\{\{\mathbf{x}_n, \mathbf{w}_n\} : \Big|Q_n(\hat{\theta}_2(\mathbf{x}_n, \mathbf{w}_n)) - Q(\hat{\theta}_2(\mathbf{x}_n, \mathbf{w}_n))\Big| \le \epsilon_2\Big\} \\
& \text{with} \quad \lim_{n\to\infty} \Pr\Big[\{\mathbf{x}_n, \mathbf{w}_n\} \in \Omega_{2,n}\Big] = 1 \qquad (19)
\end{aligned}$$

Combining (18) and (19) we have immediately

$$\forall \epsilon_1, \epsilon_2 > 0 \quad \lim_{n\to\infty} \Pr\Big[\{\mathbf{x}_n, \mathbf{w}_n\} \in \Omega_{1,n} \cap\ \Omega_{2,n}\Big] = 1 \qquad (20)$$

Because $\Big|\hat{Q}_n(\theta) - Q_n(\theta)\Big| + \Big|Q_n(\theta) - Q(\theta)\Big| \ge \Big|\hat{Q}_n(\theta) - Q(\theta)\Big|$ for any $\{\mathbf{x}_n, \mathbf{w}_n\}$ we can state

$$\begin{aligned}
\forall\ \epsilon_1, \epsilon_2 > 0 \qquad & \exists\ \Omega_n \equiv \Big\{\{\mathbf{x}_n, \mathbf{w}_n\} : \Big|\hat{Q}_n(\theta, \mathbf{x}_n, \mathbf{w}_n) - Q(\theta, \mathbf{x}_n, \mathbf{w}_n)\Big| \le \epsilon_1 + \epsilon_2\Big\} \\
& \text{with} \quad \lim_{n\to\infty} \Pr\Big[\{\mathbf{x}_n, \mathbf{w}_n\} \in \Omega_n\Big] = 1 \qquad (21)
\end{aligned}$$

Defining a mapping $\theta(\mathbf{x}_n, \mathbf{w}_n)$ by
$\Big|\hat{Q}_n\Big(\theta(.), \mathbf{x}_n, \mathbf{w}_n)\Big) - Q\Big(\theta(.), \mathbf{x}_n, \mathbf{w}_n\Big)\Big| = \sup_{\theta\in\Theta} \Big|\hat{Q}_n(\theta, \mathbf{x}_n, \mathbf{w}_n) - Q(\theta, \mathbf{x}_n, \mathbf{w}_n)\Big|$
and substituting into (20) gives the desired result:

$$\hat{Q}_n(\theta) \to Q(\theta) \quad \text{in probability uniformly.} \qquad (22)$$

The remaining conditions of Theorem 1 are easily verified and thus Theorem 3 is proven.

4 Concluding Remarks

The approach to frontier estimation with noisy data taken here deviates from stochastic frontier estimation (SFE) only by assuming a bounded noise component. The other stochastic assumptions are identical to the ones found in SFE. Under this condition it was shown, that this DEA+ labeled approach gives a consistent, semi-parametric frontier estimation technique, that is easily implemented numerically and does not require panel data. So by its very construction, we can expect DEA+ to share features with the two underlying approaches. Based on preliminary simulation evidence I am convinced that DEA+ is as sensitive to density function parametrization as SFE. But this is exactly the price to be payed for distinguishing uncontrollable influences to production processes from inefficiency in a cross section context. On the other hand DEA+ offers a clear advantage over SFE, as it does not require parametric assumptions about the shape of the production frontier. Also the possibility to investigate technical efficiency of multi-output production adds to these advantages. Large scale simulations are needed now to create evidence as of the comparative performance of DEA+ vs. SFE to delineate the borderline of better suitability of one or the other approach.

References

Aigner, D., C.A.K. Lovell and P. Schmidt. (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models." *Journal of Econometrics 6*, 21-37.

Banker, Rajiv D., A. Charnes and W.W. Cooper. (1984). "Some Models for Estimating Technical and Scale Inefficiency in Data Envelopment Analysis." *Management Science*, Vol. 30, No. 9, 1078-192.

Banker, Rajiv D. and A. Maindiratta. (1992). " Maximum Likelihood Estimation of Monotone and Concave Production Frontiers." *Journal of Productivity Analysis 3*, 401-415.

Banker, Rajiv D. (1993). "Maximum Likelihood, Consistency and Data Envelopment Analysis." *Management Science*, Vol. 39, No. 10, 1265-1273.

Banker, Rajiv D. (1995). "Hypothesis Tests using Data Envelopment Analysis." *Paper presented at INFORMS conference, New Orleans, October 1995*

Bardan, I.R. et.al. (1994). "A Simulation Study of Joint Uses of Data Envelopment Analysis and Statistical Regressions for Production Function Estimation and Efficiency Evaluation." *Paper presented at the Georgia Productivity Workshop, Oct. 21-23, 1994.* Athens, Georgia, USA.

Bierens, Herman J. (1994). *Topics in Advanced Econometrics*, Cambridge University Press, Cambridge, GB.

Desai, Anand, Samuel J. Ratick and Arie P. Schinnar. (1994). "Data Envelopment Analysis with Stochastic Variations in Data." *Workingpaper 94-51 at College of Business, Ohio State University.*

Green, W.H. (1980). "Maximum Likelihood Estimation of Econometric Frontier Functions." *Journal of Econometrics 13*, 27 - 56.

Gstach, D. (1995). " Comparing Structural Efficiency of Unbalanced Subsamples: A Resampling Adaptation of Data Envelopment Analysis." *Empirical Economics 20*, 531 - 542.

Kittelsen, Sverre A.C. (1994). "Monte Carlo simulations of DEA efficiency measures and hypothesis tests." *Paper presented at the Georgia Productivity Workshop, Oct. 21-23, 1994.* Athens, Georgia, USA.

Kneip, Alois and Leopold Simar. (1995). " A General Framework for Frontier Estimation with Panel Data." *Paper presented at the Advanced Research Workshop on Efficiency Measurements, Odense University, 22-24 may 1995.*

Korostelev,A.P., L. Simar and A.B. Tsybakov. (1993). "Efficient Estimation of Monotone Boundaries." *Paper presented at the 3rd European Workshop on Efficiency and Productivity Measurement* held at CORE, Oct 21-23, 1993, Louvain-La-Neuve, Belgium

Land, K., C.A.K. Lovell and S. Thore (1993), "Chance Constrained Data Envelopment Analysis." *Managerial and Decision Economics* 14:6 (November), 541-54

Meeusen, W. and J. Van Den Broeck. (1977). "Efficiency Estimation from Cobb-Douglas Production Functions with composed Error." *International Economic Review 2, Vol. 18*, 435-444.

Olesen, O. and N.C. Petersen (1993). "Chance-Constrained Efficiency Evaluation." *Management Science.*

Sengupta, J.K. (1987). "DEA for Efficiency Measurement in the Stochastic Case." *Comput-*

ers and Operations Research., 14, 1987, 117-129

Simar, Leopold and Paul W. Wilson. (1995). "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models." *CORE Discussion Paper 9543*, Louvain-La-Neuve, Belgium.

Simar, Leopold. (1995). "Aspects of Statistical Analysis in DEA-Type Frontier Models." *CORE Discussionpaper*, Louvain-La-Neuve, Belgium.

Sueyoshi, Toshiyuki. (1994) "Stochastic frontier production analysis: Measuring performance of public telecommunications in 24 OECD countries." *European Journal of Operational Research* 74, 466-478.

Thiry, B. and H. Tulkens (1992), "Allowing for Inefficiency in Parametric Estimation of Production Functions for Urban Transit Firms." *Journal of Productivity Analysis* 3:1/2 (June), 45-65