QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 317 – 329

A MULTI-PRODUCT VERSION OF THE DEA+ METHOD [1]

Artur Prędki
Department of Econometrics and Operational Research
Cracow University of Economics
e-mail: predkia@uek.krakow.pl

Abstract: The paper presents the DEA+ method as a tool for estimating the production function and the measure of technical efficiency in data points. A multi-product case is considered. Presentation of the underlying semiparametric frontier model is followed by demonstration of the very algorithm of DEA+ and a discussion of its validity. Finally, the method is illustrated with an empirical example with selected model distributions for each random variable constituting the composed error.

Keywords: DEA+, semiparametric frontier model, production function, technical efficiency

INTRODUCTION

DEA+ is a two-stage procedure of point estimation of the production function (transformation) and the measure of technical efficiency of a production unit within the semiparametric frontier model. It was first presented by Gstach [Gstach 1998, 1999] but did not gain popularity. However, chronologically, it is the first method in which DEA (*Data Envelopment Analysis*) is connected with the composed error term. Construction of the model and the method is based on SFA (*Stochastic Frontier Analysis*) – see, e.g., [Kumbhakar, Lovell 2000]. It is thus a way of linking DEA with the methods of production process analysis based on parametric models. Additionally, it can be considered a predecessor of now commonly used StoNED (*Stochastic Non-smooth Envelopment of Data*) – see [Kuosmanen, Kortelainen 2012] or a paper in Polish [Prędki 2012].

[1] The study conducted with financial support from the Faculty of Management, Cracow University of Economics.

The paper presents a multiproduct version of the method briefly described in the source paper [Gstach 1999]. Originality of this paper can be seen in, firstly, organizing and describing the assumptions of a corresponding semiparametric model, which make the considerations that follow clearer. Certain assumptions are not explicitly mentioned by Gstach, but their introduction results from the context and sparse hints, while others are slightly changed in comparison to their original form. Secondly, critical comments included in this paper can provide an explanation to the reasons behind the lack of popularity of this method. The empirical example in which the method is used is also original, as it assumes a different distribution of one of the components of composed error term than in the source paper.

STATISTICAL MODEL WITH A DISCUSSION OF ASSUMPTIONS

Let us begin with defining the semiparametric model mentioned above with a set of assumptions. The idea was, to a great extent, borrowed from the theory of parametric frontier models, introduced in the late 1970s – see [Aigner et al. 1977] and [Meeusen, Van den Broeck 1977].

Assumption 1. Economic units produce *s* sorts of outputs out of *m* sorts of inputs and use the same technology, represented by T – a compact and convex production possibility set, satisfying the inefficiency condition.

Assumption 2. The quantity of inputs and outputs is given for *n* production units by sample $\mathcal{X}_n = ((\mathbf{x}_j, \mathbf{y}_j) \in T, j = 1, \ldots, n)$. Vector $\mathbf{x}_j$ is characterized by the density function h_{Xj}:

$$\forall \mathbf{x} \in (\mathbf{0}, \underline{\mathbf{x}}): h_{Xj}(\mathbf{x}) > 0, \quad (1)$$

where

$\mathbf{x}_j = [x_{1j}, \ldots, x_{mj}]$ – the vector of the quantity of inputs of j^{th} production unit,

$\mathbf{y}_j = [y_{1j}, \ldots, y_{sj}]$ – the vector of the quantity of outputs of j^{th} production unit.

The vector notation of belongingness of vector $\mathbf{x}$ in formula (1) should be understood ”by coordinates”. Most probably, the identity of distribution of $\mathbf{x}_j$, j = 1, …, *n* is assumed here as well. The support of all densities is the same, although the density is indexed by j.

The description of generating vector $\mathbf{y}_j$ is performed separately. Using set T, we first define the so called *Farrell output measure of technical efficiency* for feasible production plan $(\mathbf{x}_o, \mathbf{y}_o)$:

$$\theta_P(\mathbf{x}_o, \mathbf{y}_o) = \max\{\theta \in R: (\mathbf{x}_o, \theta\mathbf{y}_o) \in T\}. \quad (2)$$

What is important here is the fact that its value is greater than or equal to unity and equal to unity for objects technically efficient. Because the form of T is unknown, we also do not know the value of the measure in point $(\mathbf{x}_o, \mathbf{y}_o)$. So, it will be estimated in the model.

Next, we define the set:

$$F = \{(\mathbf{x}, \mathbf{y}) \in T: \theta_P(\mathbf{x}, \mathbf{y}) = 1\}, \quad (3)$$

called *the production possibility frontier*. The Farrell technical efficiency measure here plays the role of a transformation function in its implicit form[2]. Introducing F, leads to formulation of the next model assumption.

Assumption 3.

$$\mathbf{y}_j = \mathbf{y}_{jF} \cdot e^{v_j - u_j} = \mathbf{y}_{jF} \cdot e^{w_j}, \text{ for } (\mathbf{x}_j, \mathbf{y}_{jF}) \in F, \quad (4)$$

where $w_j = v_j - u_j$ is so called the composed error term.

To obtain a full description of DGP for $\mathbf{y}_j$, it is necessary to provide a way in which quantities on the right side of equation (4) are generated. Let us start with the assumption regarding component v_j

Assumption 4a. Noise components v_j, j = 1, …, *n*, have independent and identical parametric distributions set by density function f_V which depends on parameters (θ_V, v_{max}). Additionally:

$$E(v) = 0 \text{ oraz } f_V(v) = 0, \text{ dla } v > v_{max}. \quad (5)$$

Bounding the support of noise by parameter v_{max} was Gstach's idea. Its justification will be provided in the next section. In his earlier paper [Gstach 1998], he presented a slightly different version of this postulate.

Assumption 4b. Noise components v_j, j = 1, …, *n*, have independent and identical parametric distributions set by symmetric density function f_V with the support ($-v_{max}$, v_{max}) dependent on parameters (θ_V, v_{max}).

It should be noticed that a more general version of 4a can be derived from Assumption 4b. This change was probably caused by two reasons. Firstly, in order to prove properties of estimators obtained by using the DEA+ method, it is enough to bound the noise up. Secondly, a lower bound is problematic while introducing the joint density function of the composed error term, which is not mentioned by the author. The details can be found in the paper [Prędki 2014]. The author of this paper supports version 4b as the more practical one. There are numerous well-known and widely-used distributions which satisfy Assumption 4b, yet it is difficult to construct a useful distribution satisfying more general Assumption 4a and, at the same time, not satisfying Assumption 4b. Besides, bounding random noise arbitrarily only up contradicts the idea of disturbances which are to be modelled by this component. That is why in practical applications it is not possible to avoid introducing lower bound of the support of noise and difficulties involved in it.

Both source papers mentioned above offer the same assumption regarding component u_j.

Assumption 5. Components u_j connected with modelling inefficiency have independent and identical distributions set by density function f_U with support R_+ dependent on parameters θ_U.

[2] It means that we have the equation linking inputs and maximal outputs, but it is usually not possible to use it for deriving an analytical formula for the maximal quantity of outputs depending on the quantity of inputs.

Theoretically, θ_V, θ_U can be the vectors of parameters, most often, however, they are just single parameters.

In the DEA+ method, the maximum likelihood method (ML) is used for estimating parameters, and to do so, the likelihood function for composed errors is introduced. This poses a question regarding independence of these components as well as independence of components u_j and v_j. In order to estimate the production frontier, it is necessary to know whether factor $\mathbf{y}_{jF}$ is not dependent[3] on a multiplicative version of the composed error e^{w_j} – see factors in equation (4). The source papers do not address this problem, and that is why it is necessary to introduce an additional assumption.

Assumption 6. Components u_j, v_j are not dependent on each other nor on vector $\mathbf{y}_{jF}$.

The last model assumption is connected with a way of generating $\mathbf{y}_{jF}$ values.

Assumption 7. $\mathbf{y}_{jF}$ value is generated from conditional distribution $Y_F | X_j$ set by density $g_{Y_F|X_j}$ with the support Int(Isoq($\mathbf{x}_j$)).

In the paper [Gstach 1999], the support of density was called *the interior of an input-isoquant* and was not precisely defined. Using the context, it can be deduced that it probably refers to the topological interior of set:

$$\text{Isoq}(\mathbf{x}_j) = \{\mathbf{y}_{jF}: (\mathbf{x}_j, \mathbf{y}_{jF}) \in F\}, \qquad (6)$$

where the quantity of inputs $\mathbf{x}_j$ is fixed.

However, the author of this paper has certain doubts regarding correctness of this definition. It is true that there might be numerous combinations of outputs, for which $\theta_P(\mathbf{x}_j, \mathbf{y}_{jF}) = 1$, yet this equation suggests that this set can be a measure-zero set in the space of outputs, so its topological interior might be an empty set as well.

We take the logarithms of both sides of equation (4), itemise it for components according to sorts of output , and then we obtain:

$$\forall j = 1, \ldots, n \ \forall r = 1, \ldots, s: w_j = \ln(y_{rj}) - \ln(y_{rjF}). \qquad (7)$$

Let us pay attention to the fact that the composed error term is not dependent on the sort of output. In [Gstach 1998, p. 163] it is called the equi-proportional impact of error terms on particular outputs. So, it is a stochastic equivalent of radial property of the Farrell technical efficiency measure, which means a proportional change of all outputs, thus not dependent on the sort of output.

DEA+ METHOD

Let us start with introducing certain auxiliary terms. The set:

$$\tilde{F} = \{(\mathbf{x}, \mathbf{y}) \in T: (\mathbf{x}, \mathbf{y}\, e^{-v_{max}}) \in F\}. \qquad (8)$$

[3] In parametric models it is usually assumed that components u_j, v_j are not dependent on $\mathbf{x}_j$, which is caused by the analytical form of the production frontier dependent explicitly on the elements of vector $\mathbf{x}_j$.

is called a production pseudo-frontier.
The error term:

$$\tilde{w}_j = v_j - v_{max} - u_j = w_j - v_{max} \leq 0, \tag{9}$$

is called a pseudo-efficiency of j[th] object.
Let us notice that :

$$\mathbf{y}_j = \mathbf{y}_{jF} \cdot e^{w_j} = \mathbf{y}_{jF} \cdot e^{w_j - v_{max} + v_{max}} = \tilde{\mathbf{y}}_{jF}\, e^{\tilde{w}_j}, \tag{10}$$

where $(\mathbf{x}_j, \tilde{\mathbf{y}}_{jF}) \in \tilde{F}$.

The sign of pseudo-efficiency and the sequence of equalities (10) follows directly from Assumptions 3-5. This means that an observed quantity of outputs $\mathbf{y}_j$ can be looked at from two perspectives. In fact, we can observe optimal quantities of outputs connected with the production frontier disturbed by two types of shocks (exogenous and inefficiency). On the other hand, it can be assumed that it is the optimal quantity of outputs connected with the pseudo-frontier, disturbed by a shock called pseudo-efficiency. This second concept will be used during Stage I of the DEA+ method.

As it directly follows from definition (8), the shape of a pseudo-frontier is identical with the shape of a real frontier, only shifted by the quantity of $e^{v_{max}}$. Similarly, the same happens to the distribution of deviations $\tilde{w}_j$, j = 1, …, *n*. From Assumption 4 it follows that pseudo-efficiencies form i.i.d. sequence, and their distribution is the distribution of the composed error term w_j shifted downwards by the quantity of v_{max}.

As far as the DEA+ procedure itself is concerned, during Stage I we calculate the estimate of the technical efficiency measure for all objects in the sample:

$$\hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j) = \max\{\theta \in R: (\mathbf{x}_j, \theta\mathbf{y}_j) \in \hat{T}\}, \tag{11}$$

where

$$\hat{T} = \{(\mathbf{x}, \mathbf{y}) \in R_{0+}^{m+s}: \exists\lambda_j \geq 0: \sum_{j=1}^{n} \lambda_j = 1,\ \mathbf{x} \geq \sum_{j=1}^{n} \lambda_j \mathbf{x}_j,\ \mathbf{y} \leq \sum_{j=1}^{n} \lambda_j \mathbf{y}_j\}. \tag{12}$$

A deterministic version of the DEA is used to estimate the value of the technical efficiency measure, precisely, the envelopment form of the BCC model – see [Banker et al. 1984]. Here a set $\hat{T}$ is an approximation of an unknown production possibility set T.

The estimate of the multidimensional equivalent of a production pseudo-frontier in point $\mathbf{x}_j$ is then:

$$\hat{\tilde{\mathbf{y}}}_{jF} = \hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j) \cdot \mathbf{y}_j. \tag{13}$$

Consequently, the estimate of pseudo-efficiency is expressed by:

$$\forall r = 1, \ldots, s: \hat{\tilde{w}}_j = \ln(y_{rj}) - \ln(\hat{\tilde{\mathbf{y}}}_{jF}) = -\ln[\hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j)]. \tag{14}$$

Equation (14) indicates that the estimates of pseudo-efficiency are also not dependent on the sort of output. Besides, they are closely connected with the estimate of the Farrell technical efficiency measure obtained by applying the DEA method. The minus in equation (14) indicates that the quantity $\hat{\tilde{w}}_j$ should actually be called an estimate of *pseudo-inefficiency*, according to the convention adopted in parametric models. Let us stress once more that radiality of the Farrell technical efficiency measure is desirable here, because it corresponds with Assumption 3 regarding equi-proportionality of the composed error term.

Stage I of the DEA+ method is justified by the following theorem:

Theorem 1. Under Assumptions 1-7, the asymptotic distribution of the estimator $\hat{\tilde{w}}_j$ is identical to deviation $\tilde{w}_j$ in the interior of set F.

Proof: According to the author of the method, the assumptions of Theorems 5 and 6 from [Banker 1993], which implicate the above theorem, are satisfied.

However, using set intF evokes the same doubts as using set Int[Isoq($\mathbf{x}_j$)] previously.

During Stage II the estimates of parameters $\theta = (\theta_U, \theta_V, v_{max})$ are calculated by the ML method, on the basis of the estimates of pseudo-efficiency $\hat{\tilde{w}}_j$ obtained during Stage I. This means that:

$$(\hat{\theta}_U, \hat{\theta}_V, \hat{v}_{max}) = \operatorname{argmax}_\theta \ln\left[\prod_{j\in J} f_{\tilde{W}}\left(\hat{\tilde{w}}_j|\theta\right)\right], \tag{15}$$

where

$$f_{\tilde{w}}\left(\hat{\tilde{w}}\right) = \int_{\hat{\tilde{w}}}^{0} f_V\left(v + v_{max}\right) f_U\left(v - \hat{\tilde{w}}\right) dv \text{ oraz } J = \{j: \hat{\tilde{w}}_j < 0\}. \tag{16}$$

Introducing set J means that in the estimation we do not take into account pseudo-efficient objects ($\hat{\tilde{w}}_j = 0$). For such objects it may happen that $\mathbf{x_j} \notin$ intX, and then consistency of Stage I of the DEA+ method is not guaranteed. Secondly, the next stage of DEA+ is infeasible due to degeneration of limits of integration in formula (16) for joint density function. Gstach claims that, asymptotically, a fraction of objects outside set J is neglected, that is:

$$\lim_{n\to\infty} \frac{\left[\sum_{j=1}^{n} I_{j\notin J}(j)\right]}{n} = 0, \tag{17}$$

takes place, where I(·) – is the indicator function. So, in an asymptotic sense, it will not matter whether the procedure is performed on all observations or only on the pseudo-inefficient ones from set J.

Once we have the estimates of all parameters characterizing distributions of both components of the composed error, we can, firstly, estimate the actual production frontier in data points:

$$\hat{\mathbf{y}}_{jF} = \hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j)\cdot\mathbf{y}_j\cdot e^{-\hat{v}_{max}} . \quad (18)$$

Secondly, using the same methods as the ones used in the SFA approach, we can also obtain the estimate of the efficiency measure of j^{th} object - see, e.g., [Kumbhakar, Lovell 2000, p. 78]. The author of the method only mentions such a possibility in the paper [Gstach 1999, p. 102], not performing it[4]. It should be remembered that, if we obtain the estimates of variances of both components of the composed error, it is not necessary to apply the method of moments nor the pseudo-likelihood method. The first one, although simple, is not free from drawbacks and limitations - see [Kumbhakar, Lovell 2000, p. 92]. The second one, however, requires laborious computations – see, e.g., [Kuosmanen, Kortelainen 2012, p. 18].

According to Gstach, consistency of the estimator used during Stage I is conditioned by introducing the bounded support of noise. Parameter v_{max} assures one-sidedness of the error term $\tilde{w}_j$, which is consistent with the nature of the DEA estimator, which is also one-sided. The author of this paper would like to highlight another role of parameter v_{max}. It is a necessary component of the correction of the initial production pseudo-frontier. It is thus a similar procedure as in case of the COLS method or the MOLS, where the initial estimator of the production frontier is also corrected, by the largest residual or by $E(u_j)$ characteristic, respectively.

In his source paper [Gstach 1998, pp. 165-167], he attempts to prove consistency of the whole DEA+ procedure, although only its single-product version. Particularly, on the basis of Theorem 1 (consistency of Stage I) and the theorem from paper [Bierens 1994], Gstach proves the theorem regarding consistency of the estimator of the production frontier in data points, obtained as a result of using the DEA+. In paper [Gstach 1999, p. 102], the author of the method claims that this result can be transferred to a multi-product case. Let us write down a corresponding theorem.

Theorem 2. Under Assumptions 1-7, the DEA+ method provides a consistent estimator $(\mathbf{x}_j, \hat{\mathbf{y}}_{jF})$ of point $(\mathbf{x}_j, \mathbf{y}_{jF})$ in the interior of set F.

It should be noted, however, that consistency of the DEA+ procedure is questionable, because of the reason mentioned by its author – see [Gstach 1998, p. 165]. In the ML procedure, in the notation of the likelihood function (15), we obtain the product of densities of random variables $\hat{\tilde{w}}_j$, even though they do not have to be independent. Further elaborations on this issue can be found in paper [Prędki 2014].

[4] Only the characteristic of $E(U|\hat{\theta}_U)$, called average inefficiency, is computed, while the main objective of the paper is a comparison of the SFA and the DEA+ methods.

EMPIRICAL STUDY

The study is based on the data from 2000 gathered by the paper's authors [Osiewalski, Osiewalska 2006], describing 240 county, urban and municipal public libraries in Poland[5]. These libraries are non-profit institutions, which can be treated as production (service) units using certain production factors to manufacture specific products (library services). Suggestions included in the source paper were used to select a set of inputs and outputs.

The following factors are taken as inputs:

x_1 – the number of job positions,
x_2 – the number of books,
x_3 - the number of magazine titles,
x_4 – the usable area of a library,
x_5 – the number of seats in reading rooms

Outputs of libraries include the following quantities, that is the ones which could generate profits if a library were a private firm – following the suggestions from paper [DeBoer 1992].

y_1 – the number of library members registered in a library,
y_2 – the number of books borrowed,
y_3 – the number of visits in reading rooms and reading corners.

Due to the numerousness of data, they are presented as selected empirical characteristics – see Table 1.

Table 1. Selected empirical statistics of data

	median	mean	deviation	min.value	max.value
x1	3,49	8,39	13,27	0,88	93
x2	35779	66886,81	88461,84	345	525441
x3	22	49,64	77,82	1	559
x4	211	524,84	949,04	40	10545
x5	30	53,23	61,94	2	441
y1	1364	4395,34	8135,12	263	74003
y2	29797	85932,70	162064,67	698	1643662
y3	2836	10567,74	20841,96	47	232300

Source: own elaboration

The first stage of the DEA+ method was performed by computing the technical efficiency measures $\hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j)$, j = 1,…,*n* using formula (11) and formula

[5] Due to the multidimensionality of the model and asymptotic properties, numerous multiproduct data were selected. The author of the paper would like to thank here Prof. Jacek Osiewalski and Dr. Anna Osiewalska for granting him access to their data.

(12). Then their logarithms were taken and pseudo-efficiencies $\hat{\tilde{w}}_j$, j = 1,…,*n* were obtained, which allowed us to move to Stage II. Following the source paper, [Gstach 1998], half-normal distribution $N^+(0, \sigma_u^2)$ was assumed for component u modelling inefficiency. Normal distribution $N(0, \sigma_v^2)$ was assumed for noise *v*, truncated to interval $(-v_{max}, v_{max})$, unlike in the source paper, where truncated symmetric beta distribution was used.

As it was mentioned above, the lower bound of the support of noise causes certain problems while introducing joint density of the random variable $\tilde{W}$. That is why it was necessary to modify the formula suggested by Gstach to the following form:

$$f_{\tilde{w}}(\tilde{w}) = \begin{cases} \int_{\tilde{w}}^{0} f_V(v + v_{max}) f_U(v - \tilde{w})dv, & \tilde{w} > -2v_{max} \\ \int_{-2v_{max}}^{0} f_V(v + v_{max}) f_U(v - \tilde{w})dv, & \tilde{w} < -2v_{max} \end{cases}. \quad (19)$$

After a number of arduous but simple transformations, the following form of density was obtained:

$$f_{\tilde{w}}(\tilde{w}) = \frac{\sqrt{\frac{2}{\pi(\sigma_u^2 + \sigma_v^2)}} \exp\left\{\frac{-1}{2}\left[\frac{(v_{max} + \tilde{w})^2}{\sigma_u^2 + \sigma_v^2}\right]\right\}\left[\Phi\left(\frac{-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]}{2\Phi\left(\frac{v_{max}}{\sigma_v}\right) - 1}, \quad (20)$$

where $\Phi(\cdot)$ is cumulative distribution function of standard normal, and:

$$\mu = \frac{-v_{max}\sigma_u^2 + \tilde{w}\sigma_v^2}{\sigma_u^2 + \sigma_v^2}, \ \sigma = \frac{\sigma_u \sigma_v}{\sqrt{\sigma_u^2 + \sigma_v^2}}, \ a = \begin{cases} \tilde{w}, & \tilde{w} > -2v_{max} \\ -2v_{max}, & \tilde{w} < -2v_{max} \end{cases}. \quad (21)$$

Next, reparametrization was performed:

$$\tilde{\lambda} = \frac{\sigma_u}{\sigma_v}, \ \tilde{\sigma} = \sqrt{\sigma_u^2 + \sigma_v^2}, \quad (22)$$

and, using the residuals $\hat{\tilde{w}}_j$ of pseudo-inefficient units, the formula for the log-likelihood function was derived:

$$\ln\left[\prod_{j \in J} f_{\tilde{W}}(\hat{\tilde{w}}_j | \theta)\right] = \quad (23)$$

$$\sum_{j \in J}\left\{\ln\sqrt{\frac{2}{\pi}} - \ln\tilde{\sigma} - \frac{1}{2}\left(\frac{v_{max} + \hat{\tilde{w}}_j}{\tilde{\sigma}}\right)^2 + \ln\left[\Phi\left(\frac{v_{max}\tilde{\lambda}^2 - \hat{\tilde{w}}_j}{\tilde{\sigma}\tilde{\lambda}}\right) - \Phi(b_j)\right] - \ln\left[2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2 + 1}}{\tilde{\sigma}}\right) - 1\right]\right\},$$

where

$$b_j = \begin{cases} \frac{(\hat{\tilde{w}}_j + v_{max})\tilde{\lambda}}{\tilde{\sigma}}, & \hat{\tilde{w}}_j > -2v_{max} \\ \frac{-\tilde{\lambda}^2 v_{max} - 2v_{max} - \hat{\tilde{w}}_j}{\tilde{\lambda}\tilde{\sigma}}, & \hat{\tilde{w}}_j < -2v_{max} \end{cases}. \qquad (24)$$

In order to obtain convergence of the ML procedure, the following restrictions on parameters in the log-likelihood function were imposed:

1. The arguments of cumulative distribution functions of standard normal from formula (23) were bounded to interval [-5;5].
2. The arguments of natural logarithms from formula (23) and parameters $\tilde{\sigma}$, $\tilde{\lambda}$, v_{max} are not lower than 10^{-4}.
3. Additionally, more substantial restrictions[6] were assumed:

$$2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2+1}}{\tilde{\sigma}}\right)-1 \geq 0{,}955,\ v_{max} \geq 2\sigma_v. \qquad (25)$$

After numerous tests, the following starting points were selected:

$$[\tilde{\sigma}_o, \tilde{\lambda}_o, v_{max,o}] = [0{,}83;\ 1{,}65889674;\ 0{,}85956661], \qquad (26)$$

whose choice was not purely accidental. The starting value $\tilde{\lambda}_o$ results from assuming equal dispersion of two distributions at the start: the normal distribution of nontruncated noise v and half normal distribution of component u. The starting value $v_{max,o}$ is linked with residuals by:

$$\min_{j \in J} \hat{\tilde{w}}_j = -2v_{max,o}, \qquad (27)$$

and this relation is strongly connected with the author's correction of joint density of the random variable $\tilde{W}$ given by formula (19). The starting value $\tilde{\sigma}_o$ is close to the local maximum of the log-likelihood function, with starting values of other parameters obtained earlier.

As a result of ML procedure in selected starting points, convergence was obtained reaching the values:

$$[\hat{\tilde{\sigma}}, \hat{\tilde{\lambda}}, \hat{v}_{max}] \approx [0{,}37889;\ 1{,}46353;\ 0{,}42850], \qquad (28)$$

which were assumed as the final estimates of unknown parameters of corresponding distributions. Only the restriction $2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2+1}}{\tilde{\sigma}}\right)-1 \geq 0{,}955$ turned out to be valid.

Using $\hat{v}_{max}$, the production frontier in data points was estimated from formula (18). Next, the estimates of parameters were used to compute the value of

[6] Unfortunately, they turned out necessary to obtain convergency of the procedure. It should be noticed that truncating the support of the noise by v_{max} occurs deeply in tails of corresponding distribution, and does not seem particularly limiting.

the efficiency measure for particular objects following the scheme described in the paper [Kumbhakar, Lovell 2000, pp. 78-82]. This measure is given by[7]:

$$TE_j = \exp(-\hat{E}(u_j|\tilde{w}_j)),\ j = 1,\dots,n. \tag{29}$$

To use the above formula, first of all, conditional distribution $U|\tilde{w}$ is needed. Using known forms of the densities of random variables U and V, after simple transformations, we obtain:

$$U|\tilde{w} \sim D(\tilde{w})\cdot N^{+}(\mu-\tilde{w},\sigma^2), \tag{30}$$

where

$$D(\tilde{w}) = \frac{1-\Phi\left(-\frac{\mu-\tilde{w}}{\sigma}\right)}{\Phi\left(-\frac{\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}. \tag{31}$$

To compute the expected value $E(u|\tilde{w})$, formulae from the paper [Kumbhakar, Lovell 2000, p. 78] were used, and a missing $D(\tilde{w})$ was added:

$$E(u|\tilde{w}) = D(\tilde{w})[\mu_* + \sigma\frac{\varphi(-\mu_*/\sigma)}{\Phi(\mu_*/\sigma)}], \tag{32}$$

where $\mu_* = \mu - \tilde{w}$ and $\varphi(\cdot)$ is the density function of standardized normal distribution. Formula (32) was realized for particular observations on residuals $\hat{\tilde{w}}_j$ and on the estimates of parameters obtained from the DEA+ method, which yielded the estimate of the appropriate expected value, and, consequently, the value of the efficiency measure TE_j.

The results obtained are presented in Table 2, and, again, selected empirical characteristics are provided. Pseudo-efficient objects, for which the DEA+ method degenerates, (as it was mentioned above) were omitted .

Table 2. Selected empirical statistics of the maximal products and the technical efficiency measure

	median	mean	deviation	min.value	max.value
y_{1F}	2072,01	3846,242	5255,20266	293,6984	33269,87
y_{2F}	38936,2	73289,39	91931,9138	1227,143	504389,4
y_{3F}	3724,3	8739,437	11864,0055	68,62053	71196,99
TE	0,82382	0,781215	0,13075852	0,085188	0,896236

Source: own elaboration

Lines 2-4 in Table 2 refer to estimated maximal values of respective outputs – formula (18). Line 5 refers to the value of the technical efficiency measure calculated using formula (29). Due to a methodological nature of the paper and

[7] It is one of the ways of measuring efficiency presented in the source paper.

limited space available, the author did not include a broader economic interpretation of the results.

CONCLUSIONS

The results of the author's methodological research (see also [Prędki 2014]) indicate two major reasons behind only scarce popularity of DEA+. Firstly, ordinary, deterministic DEA also yields technical efficiency measure as well as the value of production frontier at the data points. Hence, according to Ockham's razor, there appears to be no point in turning to a more complex approach to obtain (however different) estimates of the above quantities. Presumably, what Gstach intended was to render DEA a method of estimation of these quantities within a semiparametric statistical model comparable to a more common approach based on stochastic frontier analysis. However, it still appears to lack validity, for such a model does not allow one to obtain dispersion measures of the new estimator of efficiency measure. What is more, it is not possible to make statistical inference about either the model assumptions or the production process.

Secondly, as pointed by the author of the current paper, the efficiency measure estimator resulting from the DEA+ method has not been proven to be consistent so far. Moreover, the algorithm itself is rather cumbersome, entailing numerous methodological and numerical obstacles. In addition, the statistical model itself, as formulated by its original author, seems to be lacking in its underlying assumptions – one of the relevant assumptions is clearly missing, whereas the formulation of the other is disputable. Therefore, the author of the current article would not recommend employing DEA+ in practice. Simultaneously, it should be noted that such apparently negative conclusions are of genuine scientific merit, and the methodological objectives of the study (presented in the Introduction), constituting major contribution of the article, have been successfully achieved.

REFERENCES

Aigner D., Lovell C. A. K., Schmidt P. (1977) Formulation and estimation of stochastic frontier models, Journal of Econometrics, Vol. 6, pp. 21–37.

Banker R.D. (1993) Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation, Management Science, Vol. 39 (10), pp. 1265-1273.

Banker R., Charnes A., Cooper W. 1984 Some models for estimating technical and scale inefficiencies in DEA, Management Science, Vol. 30 (9), pp. 1078-1091.

Bierens H. J. (1994) Topics in advanced econometrics, Cambridge University Press, Cambridge.

DeBoer L. (1992) Economies of scale and input substitution in public libraries, Journal of Urban Economics, Vol. 32, pp. 257-268.

Gstach D. (1998) Another approach to Data Envelopment Analysis in noisy environments: DEA+, Journal of Productivity Analysis, Vol. 9, pp. 161-176.

Gstach D. (1999) Technical efficiency in noisy multi-output settings, CEJOR, Vol. 7, pp. 93-110.

Kumbhakar S. C., Lovell C. A. K. (2000) Stochastic frontier analysis, Cambridge University Press, Cambridge.

Kuosmanen T., Kortelainen M. (2012) Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints, Journal of Productivity Analysis, Vol. 38, pp. 11-28.

Meeusen W., Van den Broeck J. (1977) Efficiency estimation from Cobb-Douglas production functions with composed error, International Economic Review, Vol. 18 (2), pp. 435-444.

Osiewalski J., Osiewalska A. (2006) Stochastic cost frontier function for Polish public libraries, in “Space-time modelling and forecasting of economic phenomena” (ed. Zeliaś), pp. 179-193, Cracow Academy of Economics Publishing House, Cracow (in Polish).

Prędki A. (2012) Selected estimation methods in the semiparametric frontier model, Statistical Review, Vol. 59 (3), pp. 215-232 (in Polish).

Prędki A. (2014) Single-product version of the DEA+ method, Statistical Review, Vol. 61 (2), pp. 115-129 (in Polish).