ORIGINAL PAPER

CrossMark

Assessment of new methods for incorporating contextual variables into efficiency measures: a Monte Carlo simulation

Jose M. Cordero[1] [ID] · **Cristina Polo[1]** · **Daniel Santín[2]**

Abstract

The treatment of the contextual variables (Z) has been one of the most controversial topics in the literature on efficiency measurement. Over the last three decades of research, different methods have been developed to incorporate the effect of such variables in the estimation of efficiency measures. However, it is unclear which alternative provides more accurate estimations. The aim of this work is to assess the performance of two recently developed estimators, namely the nonparametric conditional DEA method (Daraio and Simar in J Prod Anal 24(1):93–121, 2005; J Prod Anal 28:13–32, 2007a) and the StoNEZD (*Stochastic Non-Smooth Envelopment of Z-variables Data*) approach (Johnson and Kuosmanen in J Prod Anal 36(2):219–230, 2011). To do this, we conduct a Monte Carlo experiment using three different data generation processes to test how each model performs under different circumstances. Our results show that the StoNEZD approach outperforms conditional DEA in all the evaluated scenarios.

Keywords Efficiency · Contextual variables · Conditional model · StoNEZD · Monte Carlo

1 Introduction

Performing an efficiency assessment of a set of decision making units (DMUs) requires accounting for different operating conditions faced by different organizations in order to obtain fairer comparisons (Haas and Murphy 2003). Those operating conditions are referred to different concepts such as external, exogenous or

✉ Jose M. Cordero
   jmcordero@unex.es

1 Department of Economics, University of Extremadura, Av. Elvas s/n, 06006 Badajoz, Spain

2 Department of Applied, Political and Public Economics, Complutense University of Madrid, Campus de Somosaguas, 28223 Pozuelo de Alarcón, Madrid, Spain

contextual variables, commonly known as Z variables in the efficiency literature. Those factors are beyond the control of producers, but their role is crucial because they may influence the production process and be responsible for differences in the performances of DMUs (e.g., type of ownership, location characteristics, weather conditions, legal framework or quality indicators).

Over the last three decades, multiple approaches have been developed with the aim of providing a correct way of accounting for the effect of such variables. In the parametric framework, these factors are usually included as part of the disturbance term,[1] although some methods have also been developed to incorporate the effect of contextual variables in the estimation of the frontier (Zhang et al. 2012). In the nonparametric framework, many different approaches have also been developed in order to account for these contextual variables (see Badin et al. 2014; Huguenin 2015 for a review of such methods), although the most extensively applied in the literature is the two-stage approach. This option is based on a first stage estimation of efficiency scores through DEA using data about inputs and outputs. These scores are then regressed on contextual variables using either a truncated regression (e.g., Simar and Wilson 2007, 2011) or the classic ordinary least squares (Banker and Natarajan 2008) in the second stage. The main concern of this approach is that it relies on the restrictive separability condition that could not be hold in many real situations, since it implies assuming that these factors have no influence on the shape of the attainable set, thus they affect only the probability of being more or less efficient.

Those alternative methods produce divergent results with empirical data (see Muñiz 2002; Yang and Pollitt 2009) and results derived from some previous works comparing their performance using simulated data (e.g., Yu 1998; Ruggiero 1998; Muñiz et al. 2006; Cordero et al. 2009; Estelle et al. 2010; Harrison et al. 2012) are inconclusive. Thus, researchers have not yet agreed on identifying a preferred option for empirical problems (Cordero et al. 2008).

In more recent years, two new appealing approaches have been developed in the literature. On the one hand, we have the nonparametric conditional approach proposed by Daraio and Simar (2005, 2007a, b). This approach extends the probabilistic formulation of the production process proposed by Cazals et al. (2002) in the presence of contextual factors, which might affect both the attainable set and the distribution of the efficiency scores, i.e., it avoids the separability condition. On the other hand, we have the so-called StoNEZD (*Stochastic Non-Smooth Envelopment of Z-variables Data*) approach proposed by Johnson and Kuosmanen (2011) as an extension of the StoNED method (*Stochastic Non-Smooth Envelopment of Data*) developed by Kuosmanen (2006) to accommodate contextual variables. This is a one-stage nonparametric estimator based on convex nonparametric least squares (CNLS), which combines the nonparametric treatment of the frontier with the probabilistic treatment of inefficiency and noise.

The use of these new methods in empirical studies has notably expanded during the last years. For instance, the conditional approach has been applied to measure

---

[1] Kumbhakar et al. (1991) developed this model for cross-sectional data and, subsequently, Battese and Coelli (1995) adapt the model to the existence of panel data.

efficiency in multiple frameworks such as education (Cherchye et al. 2010; Haelermans and De Witte 2012; De Witte and Kortelainen 2013; De Witte et al. 2013; Cordero et al. 2016), health care (Halkos and Tzeremes 2011a; Cordero et al. 2015a, b), banking (Daraio and Simar 2006; Badin et al. 2010; Matousek and Tzeremes 2016), regional analysis (Halkos and Tzeremes 2011b, 2013) or local services (Verschelde and Rogge 2012; De Witte and Geys 2013). The use of the StoNEZD approach has been widely applied as well in recent literature to estimate measures of performance in sectors like electricity distribution (Kuosmanen 2012; Cheng et al. 2014; Saastamoinen and Kuosmanen 2016), energy (Mekaroonreung and Johnson 2014), banking (Eskelinen and Kuosmanen 2013) or agriculture (Vidoli and Ferrara 2015).

The aim of this paper is to examine the performance of these new estimation techniques within the controlled environment of a Monte Carlo simulation. In this sense, it is worth mentioning that the precision of those methods was already tested separately through Monte Carlo simulations in the original papers in which these developments were proposed (Daraio and Simar 2005; Johnson and Kuosmanen 2011). Subsequently, some previous works have compared the performance of each of these novel approaches with some traditional models (Johnson and Kuosmanen 2011; Cordero et al. 2016) finding that both alternatives perform better than traditional options. In this paper, we assess them against each other using the conventional two-stage DEA approach as a baseline reference in the comparison.

Nieswand and Seifert (2018) has also evaluated recently the performance of both techniques through simulated data, although in their definition of the production process they consider the specific characteristics of regulatory datasets, where a small number of large units account for most part of the market. Given that contextual factors may affect the production process in very different ways (Badin et al. 2012), in this paper we use various data generation processes (DGP) previously proposed in the literature representing different situations that may occur in a real production process with the aim of providing more meaningful results.[2] Moreover, the focus of the previous work by Nieswand and Seifert (2018) was placed on the ability of the techniques to identify production frontiers. In contrast, in the present study we are interested in analyzing the ability of these methods to estimate technical efficiency scores because these indicators represent the most frequently used measures for benchmarking DMUs in real-world applications (Andor and Hesse 2014).

The remainder of the paper is organized as follows. Section 2 presents a brief description of the evaluated estimation methods. Section 3 describes the different scenarios considered in our Monte Carlo experimental design and summarizes the main results of the simulations. The final section concludes.

---

[2] A similar strategy was also adopted by Andor and Hesse (2014) to evaluate the performance of several methods for measuring efficiency (DEA, SFA and StoNED), although these authors did not consider the potential influence of contextual variables on efficiency.

2 Methods

In the next lines we introduce the basic notation of the evaluated models considering that the production units use a set of inputs X ($x \in \mathbb{R}_+^p$) to produce a single output Y ($y \in \mathbb{R}_+^q$).[3] We assume that all DMUs share the same production technology Ψ where the frontier is given by the maximum output that can be produced given their inputs. Then, the feasible combinations of inputs and outputs can be defined as the marginal attainable set:

$$\Psi = \left\{ (x, y) \in \mathfrak{R}_+^{p+q} | x \text{ can produce } y \right\} \tag{1}$$

Also, we consider a vector of contextual factors Z ($z \in \mathbb{R}_+^k$) affecting the performance of units. The impact of Z might be diverse, affecting the production frontier, the distribution of the inefficiencies, or both of them. If the separability condition holds, the contextual factors will only have influence on the probability of being more or less efficient, but if this assumption cannot be assumed, z-variables will affect the shape of the attainable set of (x, y) (Badin et al. 2012). In our simulation study we consider different scenarios in which all alternatives are modeled.

2.1 Conditional nonparametric approach

This method was originally developed by Cazals et al. (2002) for partial frontiers and then adapted for full frontiers by Daraio and Simar (2005, 2007a, b). The approach aims to compare only units that operate under similar operational environments, thus the selection of the reference group for each unit is conditional on a given value of the contextual variables ($Z = z$). Given that, the conditional attainable set of inputs and outputs is characterized as

$$\Psi^z = \left\{ (x, y) \in \mathfrak{R}_+^{p+q} | Z = z, x \text{ can produce } y \right\} \tag{2}$$

The conditional model is defined in a completely nonparametric way using a probabilistic formulation of the production process:

$$H_{XY|Z}(x, y|z) = Prob(X \leq x, Y \geq y | Z = z) \tag{3}$$

The joint distribution on (X, Y), conditional upon $Z = z$, defines the production process which represents the probability of a unit operating at level (x, y) being dominated by other units facing the same context. This probability function can be further decomposed as follows:

$$\begin{aligned} H_{XY|Z}(x, y|z) &= Prob(Y \geq y | X \leq x, Z = z)Prob(X \leq x, Z = z) \\ &= S_{Y|X,Z}(Y \geq y | X \leq x, Z = z)F_X(X \leq x, Z = z) \\ &= S_{Y|X,Z}(y|x, z)F_X(x|z) \end{aligned} \tag{4}$$

---

[3] We describe the model for the single-output multiple input case since introducing the multi-output context would involve the use of directional distance functions for the case of StoNEZD method (Kuosmanen and Johnson 2017), so the comparison with the conditional DEA model would be more complex.

where $S_{Y|X,Z}(y|x,z)$ denotes the conditional survivor function of Y and F_X is the cumulative distribution function of X. The conditional output-oriented efficiency measure can be defined as

$$\lambda(x,y|z) = sup\left\{ \lambda \middle| S_{Y|X,Z}(\lambda y|x,z) > 0 \right\} \tag{5}$$

The nonparametric estimators of the conditional frontier $\lambda(x,y|z)$ can be defined by a plug-in rule providing conditional estimators of the full frontier (e.g., FDH or DEA). In this paper we use a DEA estimator of the frontier in order to facilitate the comparison with other models (Daraio and Simar 2007a).

The estimation of $S_{Y|X,Z}(y|x,z)$ requires using some smoothing techniques for the contextual variables in z (due to the equality constraint $Z = z$):

$$\hat{S}_{Y|X,Z,n}(y|x,z) = \frac{\hat{H}_{XY|Z,n}(x,y|z)}{\hat{H}_{XY|Z,n}(x,0|z)} = \frac{\sum_{i=1}^{n} I(x_i \leq x, y_i \geq y)K_h(z - z_i/h_n)}{\sum_{i=1}^{n} I(x_i \leq x)K_h(z - z_i/h_n)} \tag{6}$$

where h is a bandwidth parameter in a kernel function. The selection of the bandwidth in this framework is a relevant issue since the estimation of the conditional frontier will depend on this parameter. If all the Z variables are continuous, the most common option is the data-driven selection approach suggested by Badin et al. (2010), based on the Least Squares Cross Validation procedure (LSCV) developed by Hall et al. (2004) and Li and Racine (2007). This approach has the appealing feature of detecting the irrelevant factors and smoothing them out by providing them with large bandwidth parameters. Therefore, the sizes of the selected bandwidths themselves already contain information about the impact of each contextual variable on the output. In our Monte Carlo simulation, the estimation of bandwidths is made in every loop, i.e., a measure of these parameters is calculated specifically for the corresponding sample of variables for each iteration of the experiment.

This conditional estimator has considerable virtues. First, it is very flexible since it is built in a completely nonparametric context. Second, it allows including the contextual variables in the attainable set, thus it avoids the restrictive separability condition of traditional two-stage approaches. Third, it does not require any a priori specifications regarding the direction of the influence of contextual variables. Finally, its asymptotic properties have been demonstrated by Jeong et al. (2010). This means that the estimators will converge to the true but unknown value that they are supposed to estimate when the sample size increases. Nevertheless, given its deterministic nature, it does not consider the existence of noise in the sample, so all deviations from the frontier are due to inefficiency of producers.

2.2 StoNEZD

The StoNEZD method proposed by Johnson and Kuosmanen (2011) is a nonparametric approach derived from the standard StoNED estimator (Kuosmanen 2006)

to account for contextual variables. This model is based on the regression interpretation of data envelopment analysis (DEA) proposed by Kuosmanen and Johnson (2010) to combine the key characteristics of DEA and SFA into a unified model: the DEA-type nonparametric frontier with the probabilistic treatment of efficiency and noise in stochastic models. One of the main advantages is that the StoNEZD methodology estimates jointly the production frontier and the influence of the contextual variables.

In the production function environment, the actual output (y) may deviate from $f(x)$ due to random noise (v_i), efficiency ($u_i > 0$) and the effect of exogenous factors (z_i). This model can be formally defined as follows

$$
\begin{aligned}
y_i &= f(x_i) \cdot e^{\varepsilon_i} \\
\varepsilon_i &= \delta z_i + v_i - u_i
\end{aligned}
\tag{7}
$$

According to this definition of the composite disturbance term, $\delta z_i - u_i$ can be interpreted as the overall efficiency of a unit, being δz_i the part of technical efficiency explained by the contextual variables being identical for all units while u_i captures the managerial efficiency. Therefore, it is implicitly assumed that the contextual variables only have influence on the distribution of the efficiency scores, but they do not affect the location of the frontier.

The StoNEZD method estimates efficiency in two stages. In the first stage, the shape of the frontier is obtained by applying convex nonparametric least squares (CNLS, Hildreth 1954), which does not assume a priori any assumption about the functional form but it builds upon constraints like monotonicity and convexity. Specifically, the estimator can be solved as the optimal solution to the following quadratic programming problem[4]:

$$
\min_{\alpha,\beta,\delta,\phi,\varepsilon} \sum_{i=1}^{n} \left(\varepsilon_i^{CNLS} \right)^2
$$

subject to
$$
\begin{aligned}
y_i &= \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{x}_i + \boldsymbol{\delta}' \boldsymbol{z}_i + \varepsilon_i^{CNLS} \quad \forall i = 1, \ldots, n \\
\alpha_i + \boldsymbol{\beta}_i' \boldsymbol{x}_i &\leq \alpha_h + \boldsymbol{\beta}_h' \boldsymbol{x}_i \quad \forall h, i = 1, \ldots, n \\
\boldsymbol{\beta}_i &\geq 0 \forall i = 1, \ldots, n
\end{aligned}
\tag{8}
$$

where ε_i^{CNLS} represent the residuals of the regression and the estimated coefficients α_i and $\hat{\boldsymbol{\beta}}_i$ characterize tangent hyperplanes to the unknown production technology

---

[4] Kuosmanen and Johnson (2010) show that this problem is equivalent to the standard (output-oriented, variable returns to scale) DEA model when a sign constraint on residuals is incorporated to the formulation ($\varepsilon_i^{CNLS-} \leq 0 \forall i$) and considering the problem subject to shape constraints (monotonicity and convexity).

$f(\mathbf{x}_i)$ at point $\mathbf{x}_i$. The inequality constraints impose convexity using Afriat inequalities (Afriat 1972) and the last constraint imposes monotonicity.[5]

As pointed out in Johnson and Kuosmanen (2011), the exogenous variables are introduced in the formulation of this method as part of the composite error term, which implies assuming that the exogenous variables affect technical efficiency and not the frontier. Given the residuals of the nonparametric regression (ε_i^{CNLS}), the aim of the next step in the StoNEZD method is to disentangle inefficiency from noise. To do this, and maintaining the assumptions of half-normal inefficiency and normal noise, we use the Method of Moments (Aigner et al. 1977),[6] in which the second and the third central moments can be estimated based on the distribution of the residuals:

$$\hat{M}_2 = \sum_{i=1}^{n} \left(\hat{\varepsilon}_i^{CNLS}\right)^2 / (n-1)$$

$$\hat{M}_3 = \sum_{i=1}^{n} \left(\hat{\varepsilon}_i^{CNLS}\right)^3 / (n-1) \tag{9}$$

The second moment is simply the variance of the distribution and the third is a component of the skewness:

$$M_2 = \left[\frac{\pi - 2}{\pi}\right]\sigma_u^2 + \sigma_v^2$$

$$M_3 = \left(\sqrt{\frac{2}{\pi}}\right)\left[1 - \frac{4}{\pi}\right]\sigma_u^3 \tag{10}$$

As the third moment only depends on the parameter σ_u, it can be written:

$$\hat{\sigma}_u = \sqrt[3]{\frac{\hat{M}_3}{\left(\sqrt{\frac{2}{\pi}}\right)\left[1 - \frac{4}{\pi}\right]}}$$

$$\hat{\sigma}_v = \sqrt{\hat{M}_2 - \left[\frac{\pi - 2}{\pi}\right]\sigma_u^2} \tag{11}$$

---

[5] Several other papers propose using multiplicative error structures when CRS or heteroscedasticity are assumed (Kuosmanen et al. 2015). In the present work we use the additive model because those conditions are not assumed.

[6] Kuosmanen et al. (2015) propose a *pseudo-likelihood* approach (Fan et al. 1996) or nonparametric kernel deconvolution (Hall and Simar 2002) as alternatives to the method of moments. In this study we use the method of moments due to its easier computation and interpretation. Nevertheless, Andor and Hesse (2014) found similar results for the former, while the latter has not been used in any Monte Carlo simulation to evaluate StoNEZD as far as we know.

Finally, parametric techniques can be used to estimate firm-specific inefficiency. The individual technical efficiency can be computed using the point estimator proposed by Battese and Coelli (1988)[7]:

$$\hat{E}\big(\exp(-u_i)|\hat{\varepsilon}_i\big) = \frac{\Phi\big(\hat{\mu}_{*i}/\hat{\sigma}_{*u} - \hat{\sigma}_{*v}\big)}{\Phi\big(\hat{\mu}_{*i}/\hat{\sigma}_{*u}\big)} \cdot exp\Big(\frac{1}{2}\hat{\sigma}_*^2 - \hat{\mu}_{*i}\Big) \qquad (12)$$

where Φ is the cumulative distribution function of the standard normal distribution, with $\mu_* = -\varepsilon_i \sigma_u^2/\sigma^2$ and $\sigma_*^2 = \sigma_u^2 \sigma_v^2/\sigma^2$, and $\hat{\varepsilon}_i = \hat{\varepsilon}_i^{CNLS} - \hat{\sigma}_u\sqrt{2/\pi}$ is the estimator of the composite error term.

The StoNEZD estimator has some desirable properties, since it is shown to be unbiased, asymptotically efficient, asymptotically normally distributed, and converge at the standard parametric rate ($n^{1/2}$) (see Kuosmanen et al. 2015; Yagi et al. 2018 for details). In addition, it is consistent under few assumptions. Therefore, standard techniques from regression analysis like t-tests or confidence intervals can be easily applied for inference. Regarding its limitations, since it is partly based on the SFA methodology, it presents the same peculiarities to be developed in a multi-output context.

2.3 Two-stage DEA

This traditional approach is also included in our evaluation with the aim of having a reference for comparison. In particular, we consider the OLS estimator proposed by Banker and Natarajan (2008), although we estimate corrected efficiency scores following the procedure suggested in Ray (1991) for measuring managerial efficiency as in Cordero et al. (2016). Basically, this procedure consists of re-scaling predicted efficiency values by adding the largest positive residual from the regression to obtain the adjusted efficiency[8] scores. According to this criterion, efficiency is calculated after discounting the effect of Z variables.

3 Monte Carlo simulation

3.1 Design of experiments

In this section we examine the methods presented in the previous section using Monte Carlo simulation techniques. This possibility allows us to evaluate their performance in a controlled environment in which the true shape of the frontier and its characteristics are known. Given that we do not have an a priori preference for either estimator, we attempt to make the fairest comparison possible by considering

---

[7] Kuosmanen et al. (2015) propose to use the point estimator developed by Jondrow et al. (1982), but it is well-known that it is not a consistent estimator of u_i (Andor and Hesse 2014).

[8] See Greene (1980) for details.

different data generation processes (DGPs). In particular, following Badin et al. (2012), we consider two different ways of including the effect of contextual variables. On one hand, we consider that they have influence on the boundary of the attainable set and, on the other hand, we define a production function in which they only affect the distribution of the inefficiencies.

In order to facilitate the comparison of our results with earlier Monte Carlo evidence about the performance of these two models we replicate the main features of three DGPs used in previous published papers analyzing some of them separately. The first one is based on the Cobb–Douglas function defined by Badin et al. (2012), inspired from Simar and Wilson (2011), to evaluate the performance of the conditional approach versus the two-stage model. The second alternative is the translog function proposed by Cordero et al. (2016) to compare the performance of the conditional approach with some traditional alternatives such as the one-stage and the two-stage approach. Finally, we also use the polynomial function originally used by Banker and Natarajan (2008) to evaluate their two-stage approach and subsequently applied in Johnson and Kuosmanen (2011) to examine the performance of StoN-EZD. The specific characteristics of each DGP are explained in detail next.

3.1.1 Cobb–Douglas function (Badin et al. 2012)

This is a well-known functional form of the production function widely used in econometrics. The mathematical form for one output, one input and one contextual variable is determined by:

$$y = [1 - (x-1)^2]^{1/2} \cdot \left(1 + \frac{|z-2|}{2}\right)^{1/2} \exp(-u) \tag{13}$$

The distributions for the input and the z variable are randomly generated from $x \sim U[0,1]$ and $z \sim U[0,4]$, and the random efficiency is given by $u \sim |N(0;0.22)|$. Here the effect of the z variable is included in the production function, thus it affects the shape of the attainable frontier, which implies that the separability condition is violated.

3.1.2 Translog function (Cordero et al. 2016)

The transcendental logarithmic production function is a generalization of the Cobb–Douglas function and belongs to the log-quadratic type. The production technology is defined as the transformation of two inputs (x_1, x_2) to one output (y):

$$\ln(y) = \beta_0 + \beta_1 \ln x_1 + \beta_2 \ln x_2 + \frac{1}{2}\beta_{11}[\ln x_1]^2 + \frac{1}{2}\beta_{22}[\ln x_2]^2 + \beta_{12} \ln x_1 \ln x_2 + v - u \tag{14}$$

The values assigned to the parameters are: $\beta_0 = 1$; $\beta_1 = \beta_2 = 0.55$; $\beta_{11} = \beta_{22} = -0.06$; $\beta_{12} = 0.015$. The random distributions for the inputs and the noise term are built as $x_k \sim U[1,50]$ $(k = 1,2)$ and $v \sim N(0;0.02)$, respectively. The

observed inefficiency term is defined as a composition of two terms: the effect of two contextual variables, z_1 and z_2 (they are obtained as $z_k \sim U[0;0.5]$, $k = 1, 2$), and the efficiency level (w, defined by $w \sim |N(0;0.3)|$) after discounting the influence of the z variables: $u = -\ln(\frac{1}{e^w} + z_1 + z_2)$. It is also assumed that in every loop 20% of units are placed in the production frontier. In this scenario, the contextual variables only have influence on the distribution of inefficiencies, since they are included in the error term. Thus, this DGP implicitly assume the separability condition ($\Psi^z = \Psi$) and Z has effect only on the distribution of the inefficiencies.

3.1.3 Polynomial function (Banker and Natarajan 2008)

Mathematically this is the simplest scenario considered and is built *a la* Battese and Coelli (1995). The output y is defined by a third order polynomial of a single input variable x and the effect of one contextual variable z:

$$y_i = \left(x_i^3 - 12x_i^2 + 48x_i - 37\right) \cdot \exp(\delta z_i - u_i + v_i) \tag{15}$$

The distribution of the input and the z variable are randomly sampled from $x \sim U[1, 4]$ and $z \sim U[0, 1]$, respectively. Similarly, the noise and inefficiency terms are built by $v \sim N(0;0.04)$ and $u \sim |N(0;0.15)|$ respectively. Here, again the contextual variable only influences the distribution of the inefficiencies.

The purpose of selecting different scenarios is to be as fair as possible in the assessment of the alternatives. Hence, we have considered one scenario that may theoretically favor the conditional approach (a), since it does not include noise and introduce the z-variable as part of the production function, so the separability condition is not assumed. In contrast, the scenario (c) should benefit the StoNEZD method, because it includes z in the error term and incorporates a substantial noise variance. Finally, scenario (b) presents the highest variance for inefficiency, because it introduces two contextual variables and two inputs instead of one, so it is not clear which method can be more favored under this framework.

Figure 1a–c included in the Appendix provide a graphical representation of the mathematical specification for all the described scenarios considering one Monte Carlo trial of each DGP for the case of 200 DMUs. From the shape of the graphs it can be derived that the true but unknown production functions belong to the set of continuous, monotonic increasing and globally concave functions, being exactly the same production axioms as standard DEA (Kuosmanen et al. 2015).

Once we have defined our three different DGPs, it is important to note that all of them share some common characteristics such as having only one output and not considering correlation among inputs and contextual variables. All experiments consist of R = 200 Monte Carlo trials; r = 1,…,R. Within each set of experiments, we analyse two sample sizes, 50 and 200 DMUs. In order to assess the performance of the two evaluated methods we consider two statistical indicators frequently used in Monte Carlo simulations: the Pearson correlation coefficient and the Mean Absolute

Deviation (MAD) among the real efficiency and the estimated efficiency score. This can be defined as

$$MAD = \frac{1}{nR} \sum_{r=1}^{R} \sum_{i=1}^{n} \left| \widehat{TE}_{r,i} - TE_{r,i} \right| \qquad (16)$$

where $\widehat{TE}_{r,i}$ is the estimate of technical efficiency of unit i in a given Monte Carlo replication (by a given estimation method) and $TE_{r,i}$ is the true efficiency score determined by the DGP, R is the number of experiment replications and n the number of DMUs. Following Badunenko et al. (2012) we use the median (over the 200 simulations) instead of the average because it eliminates the chance that a few large values may skew the results.[9] Likewise, in order to provide more detailed results, we build three distributions of differences of MADs for each pair of methodologies in each scenario as follows:

$$d1 = MAD_{StoNEZD} - MAD_{CondDEA}$$

$$d2 = MAD_{StoNEZD} - MAD_{DEA+OLS}$$

$$d3 = MAD_{CondDEA} - MAD_{DEA+OLS}$$

From this point, the number of negative differences (<0) is calculated. As we have run each experiment 200 times, the percentage of negative differences indicates that the method in the first position is better than the other one with that percentage of confidence. By doing this, we can check the sample of the errors across simulations for the same DGP instead of examining only at the median values of MAD. This allows us to interpret better in what extent one method performs better than its competitors.

3.2 Results

In this subsection we present the results obtained considering the three different DGPs. In this sense, we need to clarify that, although we estimate output orientation measures for the conditional approach, these scores are reported according to the definition provided by Shephard (1970), i.e., taking values between 0 and 1 in order to simplify the comparison with scores estimated with other alternatives. First of all, we present the median values of the simulated true efficiency calculated in each scenario and the median values of the efficiency scores estimated by the three techniques under evaluation in Table 1. We can observe that the median values of true efficiency are similar for the first two specifications and slightly higher for the polynomial function due to the lower value of the standard deviation for the inefficiency term assumed in the DGP of this case.

---

[9] The mean results are similar and are available from the authors upon request.

Table 1 Median values of true and estimated efficiency scores with different DGPs

DMUs	Cobb–Douglas		Translog		Polynomial	
	50	200	50	200	50	200
True efficiency	0.8427	0.8438	0.8211	0.8195	0.8896	0.8890
Estimated efficiency						
Cond. DEA	0.8217	0.7838	0.8406	0.8068	0.9539	0.9461
StoNEZD	0.8687	0.8445	0.8957	0.8810	0.9150	0.8986
DEA + OLS	0.8212	0.7821	0.7978	0.7089	0.8754	0.8434

Table 2 Correlation and MAD between true and estimated efficiency (Scenario a)

	Correlation coefficient		MAD	
	50 DMUs	200 DMUs	50 DMUs	200 DMUs
Conditional *DEA*	0.7427	0.7606	0.0722	0.0826
StoNEZD	0.7665	0.8022	0.0603	0.0569
DEA + OLS	0.6868	0.7263	0.0720	0.0837

According to these values it can be observed that the StoNEZD method overestimates the true median efficiency in all the specifications, while the conditional approach and the two-stage model underestimate them especially for higher sample sizes. If we compare different scenarios, we can notice that the median StoNEZD estimates are very close to the median real efficiency in the scenario defined by the Cobb–Douglas and the polynomial production functions for both sample sizes, but they differ in a greater extent in the case of the *translog* function. In contrast, the conditional approach provides median estimates closer the true ones in the first two scenarios. Another relevant feature that should be highlighted is that median values of estimated efficiencies with both approaches decrease when sample size is larger in all the scenarios, which for the non-parametric methods conforms to the predictions in the literature (Zhang and Bartels 1998; Perelman and Santín 2009; Cordero et al. 2016).

In order to facilitate the interpretation of the Monte Carlo experiment outcomes, we examine the performance of three methods in each scenario separately.

3.2.1 Cobb–Douglas function

Table 2 reports the correlation coefficients and the MAD among the efficiency scores estimated with the competing techniques and the true efficiency derived from the random distribution using the Cobb–Douglas DGP specification.

The analysis of these results allows us to draw some interesting conclusions. First of all, three methods perform relatively well, although the StoNEZD approach presents slightly better values in both indicators, whereas the two-stage model seems to be dominated by the other two options. The correlation

Table 3 Differences among MADs (Scenario a)

Comparison	50 DMUs		200 DMUs	
	# Negative differences	Confidence (%)	# Negative differences	Confidence (%)
d1	183	92	198	99
d2	167	84	196	98
d3	82	41	124	62

Table 4 Correlation and MAD between true and estimated efficiency (Scenario b)

	Correlation coefficient		MAD	
	50 DMUs	200 DMUs	50 DMUs	200 DMUs
Conditional *DEA*	0.5580	0.5660	0.1246	0.1287
StoNEZD	0.9046	0.9481	0.1028	0.0893
DEA + OLS	0.4792	0.4817	0.1231	0.1481

coefficient is quite similar for both new estimators, although there are more divergences in the case of 200 DMUs. In the same way, values of MAD are lower for the StoNEZD approach, especially when the sample size is greater. Additionally, we observe that the accuracy of the estimates improves (correlation coefficient increases) when the number of units increases from 50 up to 200. This is an expected result that has been proven for the two estimators in some previous works (Simar and Wilson 2015; Johnson and Kuosmanen 2011).

Table 3 shows the differences among MADs for each pair of methods. This information allows us to confirm that StoNEZD performs better than the conditional approach and two-stage DEA with 92 and 84% confidence for 50 DMUs and 99 and 98% for 200 DMUs. In the comparison between the conditional DEA and the two-stage method the results are less clear since one obtain better results for 50 DMUs (two-stage), while the other one performs better when the number of units increases up to 200.

3.2.2 Translog function

According to the values shown in Table 4, the StoNEZD approach presents a significant higher correlation with true efficiency than the conditional DEA in this scenario defined by the translog specification. This result might be explained because this is a noisy scenario and the StoNEZD is able to disentangle noise from inefficiency, while the nonparametric conditional approach attributes all the deviation from the frontier to inefficiency. In addition, it is worth mentioning that

Table 5 Differences among MADs (Scenario b)

Comparison	50 DMUs		200 DMUs	
	# Negative differences	Confidence (%)	# Negative differences	Confidence (%)
d1	121	61	122	61
d2	138	69	196	98
d3	119	60	138	69

Table 6 Correlation and MAD between true and estimated efficiency (Scenario c)

	Correlation coefficient		MAD	
	50 DMUs	200 DMUs	50 DMUs	200 DMUs
Conditional *DEA*	0.3522	0.2577	0.0558	0.0711
StoNEZD	0.8314	0.8853	0.0448	0.0304
DEA + OLS	0.6770	0.7896	0.0531	0.0591

the translog specification includes more variables than other DGPs (one output, two inputs and two exogenous variables *versus* one output, one input and one exogenous variable). This might affect the accuracy of the conditional approach, since its discriminatory power is lower when the dimension of the problem increases due to the *curse of dimensionality* (Daraio and Simar 2007a, b). Finally, traditional two-stage DEA performs worse than the other two methodologies.

The MAD values obtained for this specification are higher than others since here we are considering a higher standard deviation for the inefficiency term. Hence, the StoNEZD approach has better results for both sample sizes, although its accuracy enhances when the number of units in the sample increases. Data showed in Table 5 corroborate that StoNEZD method performs better than other two techniques with confidence bigger than 60% independently of the number of DMUs. Likewise, conditional DEA beats traditional two-stage also with confidence > 60%.

Despite the StoNEZD method presents better results in this scenario, we must note that this approach cannot identify correctly units placed on the production frontier, which in this scenario represent 20% of the sample, since it does not assign full efficiency values to any unit. Therefore, if the main interest of the study relies on identifying efficient units instead of ranking units, the conditional DEA approach should be considered as a good option given that it provides more reliable results with respect to this purpose.[10]

---

[10] Cordero et al. 2016 show that the percentage of accuracy of conditional DEA identifying efficient units is around 68–72%.

Table 7 Differences among MADs (Scenario c)

Comparison	50 DMUs		200 DMUs	
	# Negative differences	Confidence (%)	# Negative differences	Confidence (%)
d1	117	59	198	99
d2	153	77	198	99
d3	122	61	23	12

3.2.3 Polynomial function

Within this framework, the StoNEZD also presents better results both in terms of higher correlation values and lower MAD (see Table 6). However, it is noteworthy that the two-stage approach also shows very good values for both indicators (correlation higher than 0.6 and low MAD values), thus both methods can be considered as valid alternatives under the conditions defined in this scenario. In contrast, the conditional approach presents the worst results.

The differences detected between the values of the correlation coefficient calculated with each methodology might be due to the definition of the production function, which considers the existence of noise and inefficiency. In this scenario, the StoNEZD method performs better because it includes a stochastic part in its mathematical description. Furthermore, the polynomial specification considers the contextual variable as part of the error term, thus only the influence of the contextual factor in the inefficiency is taken into account and not its effect on the frontier. This implies an advantage for the StoNEZD estimator, which also defines its production model in the same way. Due to this specification, conditional DEA model fails to identify units operating in the same environment, which explains the low values of the correlation coefficient found for this method.

In general terms, the MAD values present low values, but it is remarkable that the error increases with the number of DMUs in the conditional DEA model in the three scenarios. This is not surprising in the scenarios that include noise *v*. For the small sample size, the impact of noise and the small sample bias may offset each other, but as the sample size increases, the impact of noise dominates (see Kuosmanen et al. 2013 for further discussion of this point). Another explanationmight be the fact that the three DGPs assume a linear trend among the z variable and the efficiency values or the output, which does not coincide with the formulation of the conditional approach. This can be better illustrated by the differences among MADs reported in Table 7. When we compare the conditional approach with the two-stage DEA estimator, we identify a percentage of 61% for 50 DMUs, but only 12% when the sample size increases. On the other hand, we notice that StoNEZD has advantage over the other two methods, especially for the bigger sample. Specifically, for 50 DMUs

Table 8 Correlation coefficient and MAD in scenario c for different levels of correlation between the input and the contextual variable

	ρ	Estimated efficiency		Correlation coefficient		MAD	
		50 DMUs	200 DMUs	50 DMUs	200 DMUs	50 DMUs	200 DMUs
cDEA	0	0.9539	0.9641	0.3522	0.2577	0.0558	0.0711
	−0.8	0.9472	0.9425	0.4298	0.2609	0.0705	0.0707
	−0.4	0.9464	0.9460	0.3976	0.2207	0.0735	0.0758
	0.4	0.9573	0.9535	0.3515	0.2153	0.0769	0.0789
	0.8	0.9591	0.9584	0.3474	0.2306	0.0765	0.0796
StoNEZD	0	0.9150	0.8986	0.8314	0.8853	0.0448	0.0304
	−0.8	0.9126	0.9005	0.7937	0.8633	0.0411	0.0306
	−0.4	0.9097	0.9012	0.7972	0.8522	0.0406	0.0302
	0.4	0.9172	0.9073	0.7964	0.8411	0.0444	0.0336
	0.8	0.9118	0.8907	0.7940	0.8569	0.0458	0.0355
DEA + OLS	0	0.8754	0.8434	0.6770	0.7896	0.0531	0.0591
	−0.8	0.8892	0.8526	0.7715	0.8018	0.0415	0.0511
	−0.4	0.8757	0.8396	0.6796	0.7548	0.0524	0.0617
	0.4	0.8883	0.8541	0.7458	0.7881	0.0427	0.0512
	0.8	0.8904	0.8603	0.7618	0.8230	0.0402	0.0448

StoNEZD is better than conditional approach with 59% of confidence and with 77% over the traditional two-stage DEA. Moreover, the confidence increases up to 99% when the number of units is 200 for both compared differences.

Finally, following the same formulation proposed by Johnson and Kuosmanen (2011) in a previous work using this polynomial function in a Monte Carlo experiment, we have tested how the existence of correlation between the input and the contextual variable affect the performance of alternative methodologies.[11] Specifically, we have considered different levels of negative and positive correlation represented by the parameter ρ {−0.8, −0.4, 0.4, 0.8}. Table 8 reports the estimated efficiency scores together with correlation coefficients and MADs values. The shaded rows of this table correspond to the baseline DGP in which we assumed that there is no correlation between the input and the Z variable ($\rho = 0$).

We detect that the performance of the StoNEZD method worsens when any level of (positive or negative) correlation is assumed ($\rho \neq 0$), while the opposite is observed for the two-stage approach in most cases. Nonetheless, the former always outperforms the latter in all the scenarios, so the StoNEZD can be considered as the most suitable option. Surprisingly, the values of the correlation coefficients for the conditional approach are very low in all the scenarios, although we identify slightly better results when there is a negative correlation between the input and the contextual variable. Our intuition for explaining these discouraging results is that the effect

---

[11] This additional analysis was included in order to address the suggestion made by an anonymous reviewer.

of the contextual variable is included as part of the error term. Thus, this approach has difficulties to distinguish which units are operating under the same contextual conditions to build reference sets for comparison. Moreover, this method is also penalized by the fact that this scenario does not place units on the efficient frontier (score = 1).

4 Concluding remarks

In this paper we have analyzed the performance of two recent approaches for including the effect of contextual variables in the estimation of efficiency measures, the conditional nonparametric approach and the StoNEZD method, which seem to outperform the traditional ones according to the results obtained in some previous studies using simulated data from Monte Carlo experiments (Cordero et al. 2016; Johnson and Kuosmanen 2011). Our examination of these new methods has also been based on simulated samples of data, but we consider various settings with alternative DGPs representing different ways in which contextual factors can influence the production process. Specifically, we distinguish between a production function in which this effect only relies on the inefficiencies and an alternative option in which this factor influences the boundary of the attainable set.

One of our main findings is that the StoNEZD method provides better results in all the considered scenarios. Moreover, its accuracy in terms of correlation coefficient and MAD consistently increase with sample size. However, we detect that the performance of this method slightly deteriorates when z and x are (negatively or positively) correlated although continues outperforming the other two evaluated methods. This result is in line with other studies using simulated data in different settings. Nevertheless, it is worth mentioning that the StoNEZD approach is computationally intensive, thus, it might be difficult to apply when the sample size is large. In order to solve this limitation, several authors have proposed new algorithms to solve the CNLS formulation (e.g., Lee et al. 2013; Mazumder et al. 2015), but this is still a challenging question. Another relevant issue regarding this approach presents the same peculiarities to be developed in a multi-output context that parametric methods since it is partially based on the SFA methodology. Recently, this issue has been tackled by Kuosmanen et al. (2014) and Kuosmanen and Johnson (2017), who propose that the multi-output technology can be modelled within the directional distance function framework (Chambers et al. 1996) and present illustrative applications using data about Finnish electricity distribution firms. Therefore, it seems that this limitation might be overcome in the near future.

It is also worth mentioning that StoNEZD method does not allow for considering the effect of contextual variables on the shape of the boundary, since it includes the impact of the contextual variables as part of the error term of the production function. Therefore, these factors can only have influence on the distribution of the inefficiencies. This can be considered as the major theoretical shortcoming of this method, since its implementation requires verifying previously whether the separability condition holds using some statistical tools such as those suggested by Daraio et al. (2018).

The conditional DEA performs worse in all the evaluated frameworks, although it can compete with StoNEZD when the production technology assumes that contextual variables have influence on the frontier shape. Actually, this methodological approach should be considered as the most valid option if the purpose of analysis is to identify units placed on the frontier, given that the StoNEZD fails to identify those units. However, we must bear in mind that when the sample size is small and the number of variables included in the model is relatively high, the conditional efficiency measures may be biased upwards due the well-known curse of dimensionality.

As a potential further extension of the present study, we should mention that we could expand our comparison to some new methodological proposals emerged in the most recent literature such as the two-step approach suggested by Florens et al. (2014). This is a nonparametric method for estimating conditional measures based on "whitening" inputs and outputs variables from the effect of the exogenous variables previously to perform the efficiency analysis. This could be an alternative option to mitigate the aforementioned problems related to the dimensions of the model in the conditional approach.
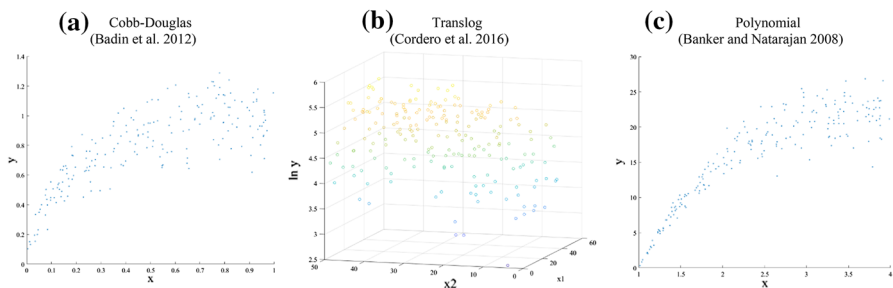
Appendix

See Fig. 1.

Fig. 1 Representation of the three DGPs (one trial with 200 DMUs)

References

Aigner D, Lovell CK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econometrics 6(1):21–37

Andor M, Hesse F (2014) The StoNED age: the departure into a new era of efficiency analysis? A Monte Carlo comparison of StoNED and the "oldies" (SFA and DEA). J Prod Anal 41(1):85–109

Badin L, Daraio C, Simar L (2010) Optimal bandwidth selection for conditional efficiency measures: a data-driven approach. Eur J Oper Res 201:633–640

Badin L, Daraio C, Simar L (2012) How to measure the impact of environmental factors in a nonparametric production model. Eur J Oper Res 223(3):818–833

Badin L, Daraio C, Simar L (2014) Explaining inefficiency in nonparametric production models: the state of the art. Ann Oper Res 214(1):5–30

Badunenko O, Henderson DJ, Kumbhakar SC (2012) When, where and how to perform efficiency estimation. J R Stat Soc A Sta 175(4):863–892

Banker RD, Natarajan R (2008) Evaluating contextual variables affecting productivity using data envelopment analysis. Oper Res 56(1):48–58

Battese GE, Coelli TJ (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. J Econom 38(3):387–399

Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empir Econ 20:325–332

Cazals C, Florens JP, Simar L (2002) Nonparametric frontier estimation: a robust approach. J Econom 106:1–25

Chambers RG, Chung YH, Färe R (1996) Benefit and distance functions. J Econ Theory 70(2):407–419

Cheng X, Bjorndal E, Bjorndal M (2014) Cost efficiency analysis based on the DEA and StoNED models: case of Norwegian electricity distribution companies. In: 2014 11th international conference on the European energy market (EEM). IEEE, pp 1–6

Cherchye L, De Witte K, Ooghe E, Nicaise I (2010) Efficiency and equity in private and public education: a nonparametric comparison. Eur J Oper Res 202:563–573

Cordero JM, Pedraja-Chaparro F, Salinas-Jiménez J (2008) Measuring efficiency in education: an analysis of different approaches for incorporating non-discretionary inputs. Appl Econ 40(10):1323–1339

Cordero JM, Pedraja F, Santín D (2009) Alternative approaches to include exogenous variables in DEA measures: a comparison using Monte Carlo. Comput Oper Res 36:2699–2706

Cordero JM, Alonso-Morán E, Nuño-Solinis R, Orueta JF, Arce RS (2015a) Efficiency assessment of primary care providers: a conditional nonparametric approach. Eur J Oper Res 240(1):235–244

Cordero JM, Santín D, Simancas R (2015b) Assessing European primary school performance through a conditional nonparametric model. J Oper Res Soc 68(4):364–376

Cordero JM, Polo C, Santín D, Sicilia G (2016) Monte-Carlo comparison of conditional nonparametric methods and traditional approaches to include exogenous variables. Pac Econ Rev 21(4):483–497

Daraio C, Simar L (2005) Introducing environmental variables in nonparametric frontier models: a probabilistic approach. J Prod Anal 24(1):93–121

Daraio C, Simar L (2006) A robust nonparametric approach to evaluate and explain the performance of mutual funds. Eur J Oper Res 175(1):516–542

Daraio C, Simar L (2007a) Conditional nonparametric frontier models for convex and non-convex technologies: a unifying approach. J Prod Anal 28:13–32

Daraio C, Simar L (2007b) Advanced robust and nonparametric methods in efficiency analysis. Methodology and applications. Springer, New York

Daraio C, Simar L, Wilson PW (2018) Central limit theorems for conditional efficiency measures and tests of the "separability" condition in nonparametric, two-stage models of production. Econom J. https://doi.org/10.1111/ectj.12103

De Witte K, Geys B (2013) Citizen coproduction and efficient public good provision: theory and evidence from local public libraries. Eur J Oper Res 224(3):592–602

De Witte K, Kortelainen M (2013) What explains the performance of students in a heterogeneous environment? Conditional efficiency estimation with continuous and discrete environmental variables. Appl Econ 45:2401–2412

De Witte K, Rogge N, Cherchye L, Van Puyenbroeck T (2013) Economies of scope in research and teaching: a non-parametric investigation. Omega 41(2):305–314

Eskelinen J, Kuosmanen T (2013) Intertemporal efficiency analysis of sales teams of a bank: stochastic semi-nonparametric approach. J Bank Finance 37(12):5163–5175

Estelle SM, Johnson A, Ruggiero J (2010) Three-stage DEA models for incorporating exogenous inputs. Comput Oper Res 37:1087–1090

Fan Y, Li Q, Weersink A (1996) Semiparametric estimation of stochastic production frontier models. J Bus Econ Stat 14(4):460–468

Florens J, Simar L, van Keilegom I (2014) Frontier estimation in nonparametric location-scale models. J Econom 178:456–470

Greene WH (1980) Maximum likelihood estimation of econometric frontier functions. J Econom 13(1):27–56

Haas DA, Murphy FH (2003) Compensating for non-homogeneity in decision-making units in data envelopment analysis. Eur J Oper Res 144(3):530–544

Haelermans C, De Witte K (2012) The role of innovations in secondary school performance–Evidence from a conditional efficiency model. Eur J Oper Res 223(2):541–549

Halkos G, Tzeremes N (2011a) A conditional nonparametric analysis for measuring the efficiency of regional public healthcare delivery: an application to Greek prefectures. Health Policy 103(1):3–82

Halkos G, Tzeremes N (2011b) Modelling regional welfare efficiency applying conditional full frontiers. Spat Econ Anal 6(4):451–471

Halkos G, Tzeremes N (2013) A conditional directional distance function approach for measuring regional environmental efficiency: evidence from UK regions. Eur J Oper Res 227:182–189

Hall P, Simar L (2002) Estimating a changepoint, boundary, or frontier in the presence of observation error. J Am Stat Assoc 97(458):523–534

Hall P, Racine J, Li Q (2004) Cross-validation and the estimation of conditional probability densities. J Am Stat Assoc 99(468):1015–1026

Harrison J, Rouse P, Armstrong J (2012) Categorical and continuous non-discretionary variables in data envelopment analysis: a comparison of two single-stage models. J Prod Anal 37(3):261–276

Hildreth C (1954) Point estimates of ordinates of concave functions. J Am Stat Assoc 49(267):598–619

Huguenin JM (2015) Data envelopment analysis and non-discretionary inputs: how to select the most suitable model using multi-criteria decision analysis. Expert Syst Appl 42(5):2570–2581

Jeong SO, Park BU, Simar L (2010) Nonparametric conditional efficiency measures: asymptotic properties. Ann Oper Res 173:105–122

Johnson AL, Kuosmanen T (2011) One-stage estimation of the effects of operational conditions and practices on productive performance: asymptotically normal and efficient, root-n consistent StoNEZD method. J Prod Anal 36(2):219–230

Kumbhakar SC, Ghosh S, McGuckin JT (1991) A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms. J Bus Econ Stat 9(3):279–286

Kuosmanen T (2006) Stochastic nonparametric envelopment of data: combining virtues of SFA and DEA in a unified framework. MTT Discussion Paper No. 3/2006. http://dx.doi.org/10.2139/ssrn.905758

Kuosmanen T (2012) Stochastic semi-nonparametric frontier estimation of electricity distribution networks: application of the StoNED method in the Finnish regulatory model. Energy Econ 34:2189–2199

Kuosmanen T, Johnson AL (2010) Data envelopment analysis as nonparametric least-squares regression. Oper Res 58:149–160

Kuosmanen T, Johnson AL (2017) Modeling joint production of multiple outputs in StoNED: directional distance function approach. Eur J Oper Res 262(2):792–801

Kuosmanen T, Saastamoinen A, Sipiläinen T (2013) What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED Methods. Ener Pol 61:740–750

Kuosmanen T, Saastamoinen A, Keshvari A, Johnson A, Parmeter C (2014) Tehostamiskannustin sähkön jakeluverkkoyhtiöiden valvontamallissa, Sigma-Hat Economics Oy (in Finnish)

Kuosmanen T, Johnson AL, Saastamoinen A (2015) Stochastic nonparametric approach to efficiency analysis: a unified framework. In: Zhu J (ed) Data envelopment analysis. A handbook of models and methods. Springer, New York, pp 191–244

Lee CY, Johnson AL, Moreno-Centeno E, Kuosmanen T (2013) A more efficient algorithm for convex nonparametric least squares. Eur J Oper Res 227(2):391–400

Li Q, Racine JS (2007) Nonparametric econometrics: theory and practice. Princeton University Press, Princeton

Matousek R, Tzeremes NG (2016) CEO compensation and bank efficiency: an application of conditional nonparametric frontiers. Eur J Oper Res 251(1):264–273

Mazumder R, Choudhury A, Iyengar G, Sen B (2015) A computational framework for multivariate convex regression and its variants. arXiv preprint arXiv:1509.08165

Mekaroonreung M, Johnson AL (2014) A nonparametric method to estimate a technical change effect on marginal abatement costs of US coal power plants. Energy Econ 46:45–55

Muñiz MA (2002) Separating managerial inefficiency and external conditions in data envelopment analysis. Eur J Oper Res 143(3):625–643

Muñiz M, Paradi J, Ruggiero J, Yang Z (2006) Evaluating alternative DEA models used to control for non-discretionary inputs. Comput Oper Res 33:1173–1183

Nieswand M, Seifert S (2018) Environmental factors in frontier estimation-A Monte Carlo analysis. Eur J Oper Res 265(1):133–148

Perelman S, Santín D (2009) How to generate regularly behaved production data? A Monte Carlo experimentation on DEA scale efficiency measurement. Eur J Oper Res 199(1):303–310

Ray SC (1991) Resource-use efficiency in public schools: a study of connecticut data. Manage Sci 37(12):1620–1628

Ruggiero J (1998) Non-discretionary inputs in data envelopment analysis. Eur J Oper Res 111(3):461–469

Saastamoinen A, Kuosmanen T (2016) Quality frontier of electricity distribution: supply security, best practices, and underground cabling in Finland. Energy Econ 53:281–292

Shephard RW (1970) Theory of cost and production function. Princeton University Press, Princeton

Simar L, Wilson PW (2007) Estimation and inference in two-stage, semi parametric models of production processes. J Econom 136:31–64

Simar L, Wilson PW (2011) Two-stage DEA: caveat emptor. J Prod Anal 36(2):205–218

Simar L, Wilson PW (2015) Statistical approaches for non-parametric frontier models: a guided tour. Int Stat Rev 83(1):77–110

Verschelde M, Rogge N (2012) An environment-adjusted evaluation of citizen satisfaction with local police effectiveness: evidence from a conditional data envelopment analysis approach. Eur J Oper Res 223:214–215

Vidoli F, Ferrara G (2015) Analyzing Italian citrus sector by semi-nonparametric frontier efficiency models. Empir Econ 49(2):641–658

Yagi D, Chen Y, Johnson AL, Kuosmanen T (2018) Shape constrained kernel-weighted least squares: Estimating production functions for Chilean manufacturing industries. J Bus Econ Stat. https://doi.org/10.1080/07350015.2018.1431128

Yang H, Pollitt M (2009) Incorporating both undesirable outputs and uncontrollable variables into DEA: the performance of Chinese coal-fired power plants. Eur J Oper Res 197(3):1095–1105

Yu C (1998) The effects of exogenous variables in efficiency measurement. A Monte Carlo study. Eur J Oper Res 105:569–580

Zhang Y, Bartels R (1998) The effect of sample size on the mean efficiency in DEA with an application to electricity distribution in Australia, Sweden and New Zealand. J Prod Anal 9:187–204

Zhang R, Sun K, Delgado MS, Kumbhakar SC (2012) Productivity in China's high technology industry: regional heterogeneity and R&D. Technol Forecast Soc 79(1):127–141