Pacific Economic Review

Pacific Economic Review, 21: 4 (2016) pp. 483–497
doi: 10.1111/1468-0106.12194

MONTE-CARLO COMPARISON OF CONDITIONAL NONPARAMETRIC METHODS AND TRADITIONAL APPROACHES TO INCLUDE EXOGENOUS VARIABLES

JOSÉ MANUEL CORDERO* *Universidad de Extremadura*
CRISTINA POLO *Universidad de Extremadura*
DANIEL SANTÍN *Universidad Complutense de Madrid*
GABRIELA SICILIA *Universidad Autónoma de Madrid*

Abstract. The aim of this paper is to compare the performance of the conditional nonparametric approach with several traditional nonparametric methods to incorporate the effect of exogenous or environmental variables into the estimation of efficiency measures. To do this, we conduct a Monte Carlo experiment using a translog production function with one output, two discretionary inputs and two exogenous variables to generate simulated data. According to the values of different accuracy measures calculated to evaluate the performance of each method, the conditional data envelopment analysis clearly outperforms all the traditional alternatives.

1. INTRODUCTION

The treatment of so-called exogenous or environmental variables in the evaluation of efficiency using nonparametric methods such as data envelopment analysis (DEA) or free disposal hull (FDH) has been one of the most controversial topics in the research literature on efficiency measurement.[1] Those variables may influence the production process and be responsible for differences in the performances of production units (e.g. type of ownership, weather conditions or quality indicators). Over the past three decades, multiple models have been developed with the aim of providing a correct way of including the effect of such variables in nonparametric production models (see Cordero *et al.* (2008); Fried *et al.* (2008) or Badin *et al.* (2014a) for a review of such methods). Most deliver contradictory results (Muñiz, 2002; Cordero *et al.*, 2009; Yang and Pollitt, 2009). Thus, decision-makers and practitioners face the difficult task of selecting which alternative is the most appropriate (Huguenin, 2015). Unfortunately, there is no agreement on this question as yet.

The aim of this paper is to shed some light on this issue by using a Monte Carlo experimental design to compare the performance of alternative models. This is not an original approach, with earlier studies having used simulated data

**Address for Correspondence*: Department of Economics, University of Extremadura, Av. Elvas s/n, 06006, Badajoz, Spain. E-mail: jmcordero@unex.es. The authors would like to thank two anonymous reviewers for their valuable comments that contributed greatly in improving the quality of the present paper. Jose M. Cordero and Daniel Santín would like to express their gratitude to the Spanish Ministry for Economy and Competitiveness for supporting this research through grant ECO2014-53702-P. Cristina Polo would also like to acknowledge the support and funds provided by the Ramon Areces Foundation in Spain.

[1] There is also another body of literature dealing with those variables in stochastic frontier analysis such as the traditional model developed by Battese and Coelli (1995) or the recent semi-parametric approaches proposed by Johnson and Kuosmanen (2011, 2012).

to evaluate the performance of different alternatives (Ruggiero, 1996, 1998; Yu, 1998; Syrjänen, 2004; Muñiz *et al.*, 2006; Cordero *et al.*, 2009; Estelle *et al.*, 2010; Harrison *et al.*, 2012). However, none of these studies include the conditional approach recently developed by Daraio and Simar (2005, 2007a,2007b). This model can be considered as one of the most promising methodological options proposed in the recent literature, because it avoids the restrictive separability condition between the input–output space and the space of the external factors implicitly assumed by the two-stage approach.

To the best of our knowledge, the accuracy of this conditional approach has only been tested using simulated data by Daraio and Simar (2007b), but it has not been compared with other alternative methods to control for the effect of exogenous variables. In this paper we attempt to fill this gap by testing the performance of both conditional DEA and FDH efficiency measures using two very traditional approaches proposed in the literature on non-discretionary inputs: the one-stage Banker and Morey (1986) and the two-stage Ray (1991) models. Both approaches have been widely applied in empirical studies performed in nonparametric frameworks, even in the latest literature. Thus, the comparison of these methods is still a research topic of interest, providing practitioners with key information regarding the expected performance of each approach in future applications.

To anticipate our main results, they show that the conditional DEA approach outperforms other methods to incorporate the effect of environmental variables, especially in its ability to correctly identify the efficient units. Therefore, we provide empirical evidence to support the use of this method, which is increasingly being applied in empirical studies to overcome some of the limitations identified for the traditional approaches, most of which will be discussed in Section 2.

The article is organized as follows. In Section 2 we briefly explain the methodology, including a brief description of each evaluated method together with their main advantages and limitations. In Section 3 we describe the data generation process used in our Monte Carlo experimental design. In Section 4, we report the main results. Finally, Section 5 outlines the main conclusions.

2. METHODS

We consider a production technology where the activity of the production units is characterized by a set of inputs $x(x \in \mathbb{R}_+^p)$ used to produce a set of outputs $y(y \in \mathbb{R}_+^q)$. In this framework the production set is the set of all feasible input–output combinations (x, y). This can be defined as

$$\psi = \left\{(x, y) \in \mathbb{R}_+^{p+q} \mid x \text{ can produce } y\right\}. \tag{1}$$

The usual assumptions on this set are the free disposability of inputs and outputs, which means that if $(x, y) \in \psi$, then $(x', y') \in \psi$ for all $x' \geq x$ and $y' \leq y$. If we assume an input orientation, the Farrell (1957) measure of technical efficiency for a unit operating at the level (x, y) is given by

$$\theta(x,y) = inf\{\theta > 0 | (\theta x, y) \epsilon\ \psi\}. \tag{2}$$

By construction, $\theta(x,y) \in (0,1)$ for all $(x,y) \in \psi$. This measure gives the feasible, proportionate reduction in input levels at constant output levels for a unit operating at $(x,y) \in \psi$. If $\theta(x,y) = 1$, the unit is said to be technically efficient in the input direction, while if $\theta(x,y) < 1$ the unit is considered technically inefficient.

Given that the production set ψ cannot be observed as well as the efficiency scores, it has to be estimated from a random sample of production units denoted by $\mathcal{S}_{XY,n} = \{(X_i, Y_i) | \ i = 1, \ldots, n\}$. Among the multiple approaches that can be used to achieve this goal, we focus on nonparametric models, which do not impose any *a priori* specification on the functional form of the production technology. The most popular options are based on enveloping estimators such as the FDH developed by Deprins *et al.* (1984):

$$\widehat{\psi}_{FDH} = \left\{(x,y) \epsilon \mathbb{R}_+^{p+q} | y \le Y_{i,}\ x \ge X_i,\ i = 1, \ldots., n\right\}. \tag{3}$$

The convex hull of $\widehat{\psi}_{FDH}$ provides the DEA estimator of ψ, which can be expressed as a linear programming problem (Charnes *et al.*, 1978):

$$\widehat{\psi}_{DEA} = \left\{ \begin{array}{c} (x,y) \in \mathbb{R}_+^{p+q} | \ y \le \sum_{i=1}^{n} \gamma_i Y_i;\ x \ge \sum_{i=1}^{n} \gamma_i X_i for\ \ (\gamma_i, \ldots. \gamma_n \\ st. \sum_{i=1}^{n} \gamma_i = 1;\ \gamma_i > 0; i = 1, \ldots., n \end{array} \right\}. \tag{4}$$

In the following, we explore several approaches developed in the production frontier literature to consider additional information provided by other variables $\left(z \, \epsilon \, \mathbb{R}_+^r\right)$ that can be considered as external factors to, but which may have influence on, the production process.

2.1. *One-stage approach*

The first and simplest alternative for handling efficiency in the presence of exogenous variables is the one-stage model developed by Banker and Morey (1986), although this model was originally designed to include the so-called non-discretionary inputs (i.e. inputs involved in the production process that are beyond the control of managers). This approach includes those factors in the model as an additional restriction in the formulation of the standard DEA programme; thus they are involved in defining the attainable set $\psi \in \mathbb{R}_+^p \times \mathbb{R}_+^q \times \mathbb{R}_+^r$, but without being active in the optimization for the estimation of efficiency scores:

$$\theta(x,y|z) = inf\{\theta | (\theta x, y, z) \epsilon\ \psi\}. \tag{5}$$

The estimator of ψ is straightforward to derive by adding the exogenously fixed variables z in defining the FDH or the DEA enveloping set, with a variable z being considered as an input if it has a favourable effect on efficiency and as an output if the effect is the opposite.

The main advantage of this approach is its non-complexity as it simplifies the estimation of efficiency scores by including all the key variables in a single DEA. As a result, it has been extensively used by researchers and practitioners, because it is included in most DEA software packages (Barr, 2004). Moreover, a statistical DEA bootstrap approach has been recently developed for making inference about the efficiency of estimates (Essid *et al.*, 2010, 2013). However, this model also has some important drawbacks. First, the linear programmes involved in defining the corresponding efficiency scores depend on the assumption of returns to scale made on external variables (Badin *et al.*, 2014a). In particular, this approach implies a shift in the constant returns to scale frontier when non-discretionary factors are treated as such, but no change in the varying returns to scale frontier in relation to the situation where all factors are considered discretionary (Muñiz, 2002). Second, this approach requires some restrictive assumptions like free disposability and convexity of the attainable set. It also calls for the prior specification of the favourable or unfavourable effect of the external factors on the production process, which analysts cannot always foresee (Fried *et al.*, 2002; Badin *et al.*, 2014a). Finally, the efficiency scores estimated with this model can be systematically biased, increasing potential production targets for inefficient decision making units (DMU). This occurs when DMU in better operating environments are included in the reference sets of DMU operating in harsher environments, because the non-discretionary constraints (and the convexity constraint for variable returns to scale) can be satisfied by combinations of multiple DMU where the weights assigned to DMU with favourable environments are relatively small (Ruggiero, 1996).

Despite all these limitations, this model can be considered as the standard approach for including non-discretionary factors in DEA, especially in cases where their influence is recognized but not fully understood (Harrison *et al.*, 2012). Indeed, some authors support its use to adjust for heterogeneity in operating conditions (Haas and Murphy, 2003), because this model has generated more than 200 different publications related to non-discretionary or environmental variables. This suggests that many researchers have found it adequate for their particular contexts. Moreover, according to the results reported by Muñiz *et al.* (2006) or Harrison *et al.* (2012), the variable and the number of units is high.

2.2. *Two-stage approach*

This traditional approach first estimates technical efficiency through DEA using data about inputs and outputs only. The scores are then regressed on exogenous variables using either Tobit regression or OLS in the second stage (for details see Hoff, 2007). This technique also has been extensively applied by practitioners as a simple tool to account for external effects in the estimation of efficiency measures. However, two influential papers by Simar and Wilson (2007, 2011) demonstrate that previous applications of the two-stage approach were invalid due to its failure to account for the bias and serial correlation present among efficiency estimates.

To address this problem, Simar and Wilson (2007) describe a complete data generating process consistent with regression of nonparametric estimates in the second stage and develop two different algorithms based on the use of bootstrap methods to obtain consistent and unbiased estimates for the parameters of the regression. Furthermore, they advocate the use of the truncated regression model that explicitly takes into account the bounded domain of the DEA efficiency estimates. Since then, this method has been mostly used in the literature for identifying the influential determinants of efficiencies in multiple contexts such as education (Alexander *et al.*, 2010; Miningou and Vierstraete, 2013; Santín and Sicilia, 2015), health care (Blank and Valdmanis, 2010; Afonso and St Aubyn, 2011; Nedelea and Fannin, 2013), tourism (Barros *et al.*, 2011) or regional analysis (Demchuk and Zelenyuk, 2009).[2]

However, this approach also has some drawbacks. The most important limitation is that it relies on the separability condition, which states that the support of the inputs and output variables does not depend on the environmental variables in Z. This assumption implies that external factors do not influence the shape of the attainable set ψ, and, thus, can only affect the probability of being more or less efficient.[3] This condition can be very restrictive and may not always being assumed when indeed it is evident that the external variables Z are determining the value of inputs and outputs and, subsequently, are affecting the form of the frontier. Daraio *et al.* (2015) provide a test that allows researchers to check whether their data fulfil the separability condition before applying the two-stage approach.

Finally, it is worth noting that the Simar and Wilson (2007) method was mainly designed to allow for valid inference in the second-stage regression rather than for incorporating the effect of such variables into the estimation of efficiency measures. Thus, this model does not provide a good solution if we are interested in correcting the initial efficiency scores to filter out the influence of the external variables, because all of them have values lower than one (or higher with an output orientation). Therefore, the estimated efficiency scores cannot be interpreted as targets for the units, because there is not a production frontier that can be used as a reference. To overcome this problem and make this method comparable with other alternatives, in this simulation study we estimate corrected efficiency scores following the procedure suggested in Ray (1991). Basically, it consists of regressing the estimated DEA efficiency scores (h) on the external variables Z using a standard least squares model to obtain the predicted value $\hat{h}$, and adding the largest positive residual to the predicted value from the regression[4] to obtain the adjusted efficiency $\left(\hat{\hat{h}}\right)$.[5] The level of

[2] This method can be easily implemented using the rDEA package developed by Simm and Besstremyannaya, available from: https://cran.r-project.org/web/packages/rDEA/.

[3] See Daraio *et al.* (2015) for details.

[4] See Greene (1980) for details.

[5] One problem in the least square regression is that the predicted value $\hat{h}$ may not lie below 1. Following Ray (1991), in these cases we have truncated the adjusted predicted values $\hat{\hat{h}}$ at the value of 1.

inefficiency after discounting the influence of external variables is measured as the shortfall of h from $\hat{\hat{h}}$. Thus, $1-\left(\hat{\hat{h}}-h\right)$ represents the managerial efficiency not caused by external variables, which is the final target of the analysis.

2.3. *Conditional non-parametric approach*

This approach extends the probabilistic formulation of the production process proposed by Cazals *et al.* (2002), where the attainable set is interpreted as the support of some probability measure defined on the input–output space. In terms of the joint probability measure of (X, Y), which represents the probability of a unit operating at level (x, y) being dominated,

$$\begin{aligned} &\theta(x,y) = inf\{\theta | F_X(\theta x|y) > 0\}, \\ &F_X(x|y,z) = Prob(X \le x | Y \ge y,\ Z = z). \end{aligned} \tag{6}$$

This approach allows a natural extension of the model in the presence of environmental factors by conditioning the production process to a given value of $Z=z$. The joint distribution on (X, Y), conditional upon $Z=z$, defines the production process if $Z=z$. This represents the probability of a unit operating at level (x, y) being dominated by other units facing the same environmental conditions z:

$$\begin{aligned} &\theta(x,y|z) = inf\{\theta | F_X(\theta x|y,z) > 0\}, \\ &F_X(x|y,z) = Prob(X \le x | Y \ge y,\ Z = z). \end{aligned} \tag{7}$$

This technique requires the estimation of a nonstandard conditional distribution using a nonparametric kernel function $K(\cdot)$ to select the appropriate reference partners and a smoothing bandwidth parameter h_n for the exogenous variables in z using some bandwidth choice method:

$$\hat{F}_{X,n}(x|y,z) = \frac{\sum_{i=1}^{n} I(x_i \le x, y_i \ge y) K(z - z_i/h_n)}{\sum_{i=1}^{n} I(y_i \ge y) K(z - z_i/h_n)}. \tag{8}$$

The selection of the bandwidth in this complex framework is a key issue because the estimation of the conditional frontier will depend on this parameter. If all the Z variables are continuous, the most common approach is the data-driven selection method suggested by Badin *et al.* (2010) based on the least squares cross-validation (LSCV) approach (Li and Racine, 2004, 2007).

Hence, it is possible to derive a nonparametric estimator of the conditional efficiency $\theta(x, y\,|\,z)$ by plugging in an estimator of $F_X(x\,|\,y, z)$. In particular, the conditional FDH efficiency estimator can be obtained as follows:

$$\hat{\theta}_{FDH}(x,y|z) = inf\left\{\theta|\hat{F}_{X,n}(\theta x|y,z) > 0\right\} = \\ = \min_{\left(i|Y_i \geq y, |Z_i - z| \leq h\right)}\left\{\max_{j=1,\dots,p}\left(\frac{X_i^j}{X^j}\right)\right\}. \tag{9}$$

Therefore, the units selected as benchmarks for each DMU will be those with a higher value of the output than the evaluated unit (in an input orientation) and those that present a value below the chosen bandwidth (h) for the absolute value of the difference between z and the actual value of z for each unit. Moreover, this method allows separating the influential from the irrelevant factors, which are assigned large bandwidth parameters (Badin *et al.*, 2010).

If we assume that the true conditional attainable set is convex,[6] we can also define the conditional DEA efficiency estimator by:[7]

$$\hat{\theta}_{DEA}(x,y|z) = inf\left\{\theta|y \leq \sum_{(i|z-h \leq z_i \leq z+h)}\gamma_i y_i;\ \theta x \geq \sum_{(i|z-h \leq z_i \leq z+h)}\gamma_i x_i,\ \right\} \tag{10}$$
$$s.t.\ \sum_{(i|z-h \leq z_i \leq z+h)}\gamma_i = 1.$$

These conditional approaches have considerable virtues, the most relevant perhaps being the fact that they avoid the aforementioned separability condition because they directly include the environmental variables in the attainable set (eqn 7). In addition, they are defined and estimated nonparametrically, which is a major advantage over semi-parametric methods in terms of flexibility. Moreover, they require no specifications regarding the direction of the influence of environmental variables. Finally, their consistency and asymptotic properties have been proven (Cazals *et al.*, 2002; 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.

As a result, the conditional approach has become very popular in the recent literature on efficiency measurement. Hence, studies using this approach to measure efficiency are available in multiple frameworks (Bonaccorsi *et al.*, 2006; Cherchye *et al.*, 2010; De Witte and Geys, 2011; Halkos and Tzeremes, 2011; Haelermans and De Witte, 2012; Verschelde and Rogge, 2012; De Witte and Kortelainen, 2013; Cordero *et al.*, 2015, 2016). However, we should note that there is not widely available software to conduct conditional DEA so far.[8] Moreover, it has a serious limitation related to the amount of time that

[6] Here, the reference units will also be selected among those with values situated between the assigned bandwidths for each variable.

[7] Note that in this linear program, x_i can be replaced by its projection on the FDH efficient frontier (i.e. by $\hat{\theta}_{FDH}(x_i, y_i)x_i$).

[8] Some works are in progress for the completion of a user friendly toolbox, CONDEFF, that will be available free of charge (see Badin *et al.*, 2014b).

it takes to optimize the bandwidths; thus the estimation of efficiency scores might be a very time-consuming process when using large data sets.[9]

3. DATA GENERATION PROCESS AND EXPERIMENTAL DESIGN

Our Monte Carlo experimental design is based on Cordero *et al.* (2009). However, the production function employed in the data generation process (DGP) reported in that paper had some shortcomings detected by Estelle *et al.* (2010). In this paper we present a corrected DGP to enhance the Monte Carlo experiment and the results.

We first define a flexible and smooth (continuous and easy to derive) translog as the underlying production technology:

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

where the values assigned to the parameters are $\beta_0 = 1$; $\beta_1 = \beta_2 = 0.55$; $\beta_{11} = \beta_{22} = -0.06\beta_{12} = 0.015$; $\beta_{11} = \beta_{22} = -0.06\beta_{12} = 0.015$. With these values the convexity requirement of DEA is not violated and the production function assumes a decreasing return to scale technology. This production technology is closer to economic theory assuming a 'law of diminishing returns' in production processes and to real efficiency measurement problems, because, in practice, most of them are estimated with a DEA assuming variable returns to scale.

Second, we define the observed inefficiency term (u) as a composite term with two components: the effect of exogenous variables (Z_1 and Z_2) and the inefficiency level after discounting the influence of external variables (w) which is the final target of the analysis. Specifically, we define this term using the following expression:

$$u = -\ln\left(\frac{1}{\exp(w)} + z_1 + z_2\right). \tag{12}$$

Third, both inputs are randomly and independently generated from a uniform distribution, $X_1 \sim \mathrm{U}(1, 50)$ and $X_2 \sim \mathrm{U}(1, 50)$, while the two exogenous variables (Z_1 and Z_2) are randomly and independently drawn from a uniform distribution $Z_1 \sim \mathrm{U}(-0.25, 0.25)$ and $Z_2 \sim \mathrm{U}(-0.25, 0.25)$, so their effect on observed inefficiency can be either positive or negative. Fourth, the real inefficiency component (w) is drawn randomly and independently from a half-normal distribution, $W \sim |N(0; 0.3)|$. It is also assumed that 20% of DMU belong to the production frontier in each experiment (i.e. $w = 0$), so

[9] For example, De Witte and Kortelainen (2013) report that the estimation of bandwidths for a sample of 3992 observations took approximately 80 hours on a 2.6-GHz quad-core computer. Cordero *et al.* (2014) indicate that the estimation of bandwidths for a sample of 31 854 observations took more than 10 weeks using a similar computer.

Table 1. Mean efficiency and standard deviation

	50 DMU		400 DMU	
	Mean	SD	Mean	SD
True efficiency (w)	0.8413	0.1423	0.8397	0.1434
One-stage	0.9507	0.1067	0.8862	0.1553
Two-stage corrected (Ray, 1991)	0.8452	0.1806	0.7872	0.1946
Conditional DEA	0.8334	0.1979	0.7997	0.1978
Conditional FDH	0.9677	0.0925	0.9326	0.1348

DEA, data envelopment analysis; DMU, decision making units; FDH, free disposal hull.

these DMU are classified as efficient.[10] Finally, the noise component v is drawn randomly and independently from a normal distribution $v \sim N(0; 0.02)$.

Once the DGP is defined, we use the simulated data to estimate efficiency scores with the four alternative methods (one-stage, two-stage, conditional FDH and conditional DEA). The assessment of the accuracy of the different alternatives is based on the comparison of the estimated scores with the generated real efficiency (i.e. the w term in equation 12). We have considered two alternative sample sizes, 50 and 400 DMU, for the experiment. In this way, we can test how the ratio of the number of observations to the number of variables included in DEA can affect the results for each methodology. Each experiment has been repeated 1000 times to ensure robust results.

4. RESULTS

Table 1 reports the average and standard deviation of the simulated real efficiency and the estimated scores.

According to these values, it is possible to detect that the one-stage approach and the conditional FDH overestimates the mean efficiency, especially in the small sample case. In the case of the one-stage DEA approach, this behaviour may be due to the loss of DEA discriminatory power after including two additional variables. On the other hand, the estimations of the conditional FDH do not meet the convexity assumption, leading to increased mean efficiency estimations. The other two options perform well when a small sample is considered but underestimate average efficiency when sample size is larger. Only the conditional DEA approach appears to maintain mean efficiency levels close to the real ones. This result conforms to the predictions in the literature because average efficiency estimations decrease as sample size increases (Zhang and Bartels, 1998; Perelman and Santín, 2009).

[10] Bardham *et al*. (1998) place 25% of DMU in the production frontier arguing that assurances of full relative efficiency for at least some observations is required to conform with the assumptions normally made in economics. This assumption can be relaxed or made more restrictive depending on the research objectives: e.g. 30% in Holland and Lee (2002), 25% in Perelman and Santín (2009), 12.5% in Ruggiero (1998) and 10% in Muñiz *et al*. (2006). We have chosen 20% as a mid-point between 30% and 10%.

Table 2. Ability to identify the efficient units (percentages)

	50 DMU	400 DMU
One-stage	26.70	36.95
Two-stage corrected (Ray, 1991)	57.70	23.30
Conditional DEA	68.02	72.24
Conditional FDH	37.29	42.51

It is also assumed that 20% of decision making units (DMU) belong to the production frontier in each experiment. DEA, data envelopment analysis; FDH, free disposal hull.

Table 3. Spearman's rank correlation coefficient

	50 DMU	400 DMU
One-stage	0.431***	0.655***
Two-stage corrected (Ray, 1991)	0.469***	0.517***
Conditional DEA	0.555***	0.578***
Conditional FDH	0.424***	0.506***

***Coefficient is significant at 1%. DEA, data envelopment analysis; DMU, decision making units; FDH, free disposal hull.

The second criterion employed to evaluate the performance of the different alternatives is based on their ability to identify the truly efficient units. To do this, we have divided the number of units that are correctly classified as efficient by the total number of efficient units identified by each model (Table 2).[11] According to this criterion, we find that almost all the approaches achieve better results with the larger sample size with the only exception being the two-stage approach. This result is a consequence of the procedure used to correct the efficiency scores detailed in the previous section, which only identifies one DMU as efficient (that with the largest positive residual).

Even more relevant is the fact that the conditional DEA model achieves the best results for both sample sizes. Among the other options, it is noteworthy that the one-stage model is the worst performer when the sample size is small, although its results are similar to the conditional FDH and clearly better than the two-stage model with the larger sample size.

Table 3 reports Spearman's rank correlation coefficient between the true and the estimated efficiency, where a higher value suggests better performance in measuring efficiency. According to this coefficient, the approach with a better performance is again the conditional DEA model, but only for small samples. When the sample is larger (e.g. 400 units), the conditional model is outperformed by the one-stage model, which provides more accurate measures of performance. This finding confirms the evidence shown in previous literature about the good results of this method when there are few exogenous variables compared to

[11] This indicator is not biased by the fact that some approaches are prone to place more units on the production frontier.

the number of observations (Muñiz *et al.*, 2006; Cordero *et al.*, 2009; Harrison *et al.*, 2012). In addition, the rank correlation coefficient increases for all approaches with the number of observations.

Finally, two other accuracy measures have been computed: the mean absolute deviation (MAD) between the true and the estimated efficiency (reported in Table 4) and the percentage of DMU for which the estimated efficiency score deviated less than 10% from real efficiency (reported in Table 5). A lower MAD indicates that the estimated measure is a better fit to the true efficiency, while a higher value for the latter indicator implies better performance.

The results of both accuracy measures indicate that the one-stage model provides the best results for both sample sizes, although there are only minor differences between this model and the DEA conditional approach in the case of the smaller sample. Nevertheless, for large samples, the former clearly outperforms the latter. According to this information and considering the ability of the conditional model to properly identify efficient units, practitioners using this approach should be especially cautious when interpreting measures of performance of inefficient units in terms of production targets, because they could be biased. The two-stage and the conditional FDH approaches perform worse for both indicators, showing a poor ability to identify the true efficiency. It is worth noting that the poor results for small samples achieved by the conditional FDH approach in terms of MAD (Table 4) are due to the selected DGP that assumes production frontier convexity. In the case of the second-stage approach, the deterioration can be explained by two factors. The first factor is our decision of assuming that 20% of DMUs belong to the production frontier in each experiment independently of the sample size. We took this decision with

Table 4. Mean absolute deviation

	50 DMU	400 DMU
One-stage	0.1151	0.0842
Two-stage corrected (Ray, 1991)	0.1209	0.1336
Conditional DEA	0.1207	0.1296
Conditional FDH	0.1387	0.1201

DEA, data envelopment analysis; DMU, decision making units; FDH, free disposal hull.

Table 5. Percentage of units with less than 10% deviation from true efficiency level

	50 DMU	400 DMU
One-Stage	48.56	59.95
Two-stage corrected (Ray, 1991)	43.94	41.26
Conditional DEA	47.20	43.65
Conditional FDH	43.20	46.13

DEA, data envelopment analysis; DMU, decision making units; FDH, free disposal hull.

the aim of facilitating the comparison between the two scenarios, but we are aware that perhaps the proportion of efficient units should be lower when the sample increases, as actually occurs in a standard DEA. Second, as we have discussed above, this result can be a consequence of the procedure used to adjust the efficiency scores.

5. CONCLUDING REMARKS

In this paper we compare the performance of two conditional approaches with two well-known classical nonparametric approaches to include exogenous variables. Our main finding is that the conditional DEA approach is the best option when practitioners are interested in identifying the efficient units, although they should be cautious when interpreting inefficiency measures obtained with this method. Moreover, as we consider different sample sizes in our experiment, we can conclude that the conditional approach seems to provide good results when the sample is small, while the traditional one-stage model seems to provide more accurate measures of performance for larger samples.

In the case of small samples, the FDH conditional approach suffers from the curse of dimensionality problem and does not meet the convexity assumption leading to a problem of discriminatory power. Finally, although in this paper the two-stage approach generally provides the poorest results, this approach might be the best option to explain *ex post* which external factors are affecting the performance of units.

REFERENCES

Afonso A. and M. St. Aubyn (2011) 'Assessing Health Efficiency across Countries with a Two-step and Bootstrap Analysis', *Applied Economics Letters* 18, 1427–30.

Alexander W. R. J., A. A. Haug and M. Jaforullah (2010) 'A Two-stage Double-bootstrap Data Envelopment Analysis of Efficiency Differences of New Zealand Secondary Schools', *Journal of Productivity Analysis* 34, 99–110.

Badin L., C. Daraio and L. Simar (2010) 'Optimal Bandwidth Selection for Conditional Efficiency Measures: A Data-driven Approach', *European Journal of Operational Research* 201, 633–40.

Badin L., C. Daraio and L. Simar (2014a) 'Explaining Inefficiency in Nonparametric Production Models: The State of the Art', *Annals of Operations Research* 214, 5–30.

Badin, L., C. Daraio and L. Simar (2014b) 'Bandwidth Selection Methods for CondEff Toolbox: Introduction and Implementation; VIII North American Productivity Workshop, NAPW 2014, 4–7 June 2014, Ottawa, Canada.

Banker R. D. and R. C. Morey (1986) 'Efficiency Analysis for Exogenously Fixed Inputs and Outputs', *Operations Research* 34, 513–21.

Bardham I. R., W. W. Cooper and S. C. Kumbhakar (1998) 'A Simulation Study of Joint uses of Data Envelopment Analysis and Statistical Regressions for Production Function Estimation and Efficiency Evaluation', *Journal of Productivity Analysis* 9, 249–78.

Barr R. (2004) 'DEA Software Tools and Technology. A State-of-the-art Survey', in W. W. Cooper, L. Seiford and J. Zhu (eds), *Handbook on Data Envelopment Analysis*. Boston: Kluwer.

Barros C. P., L. Botti, N. Peypoch and B. Solonandrasana (2011) 'Managerial Efficiency and Hospitality Industry: The Portuguese Case', *Applied Economics* 43, 2895–905.

Battese G. E. and T. J. Coelli (1995) 'A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data', *Empirical Economics* 20, 325–32.

Blank J. L. and V. G. Valdmanis (2010) 'Environmental Factors and Productivity on Dutch Hospitals: A Semi-Parametric Approach', *Health Care Management Science* 13, 27–34.

Bonaccorsi A., C. Daraio and L. Simar (2006) 'Advanced Indicators of Productivity of Universities, An Application of Robust Nonparametric Methods to Italian Data', *Scientometrics* 66, 389–410.

Cazals C., J. P. Florens and L. Simar (2002) 'Nonparametric Frontier Estimation: A Robust Approach', *Journal of Econometrics* 106, 1–25.

Charnes A., W. W. Cooper and E. Rhodes (1978) 'Measuring The Efficiency of Decision Making Units', *European Journal of Operational Research* 2, 429–44.

Cherchye L., K. De Witte, E. Ooghe and I. Nicaise (2010) 'Efficiency and Equity in Private and Public Education: A Nonparametric Comparison', *European Journal of Operational Research* 202, 563–73.

Cordero J. M., F. Pedraja-Chaparro and J. Salinas-Jiménez (2008) 'Measuring Efficiency in Education: An Analysis of Different Approaches for Incorporating Non-discretionary Inputs', *Applied Economics* 40, 1323–39.

Cordero J. M., F. Pedraja and D. Santin (2009) 'Alternative Approaches to Include Exogenous Variables in DEA Measures: A Comparison Using Monte Carlo', *Computers and Operations Research* 36, 2699–706.

Cordero, J. M., J. Salinas-Jiménez and M. M. Salinas-Jiménez (2014) 'Assessing the Level of Happiness across Countries: A Robust Frontier Approach'. MPRA Working Paper 57784, University Library of Munich. Available from URL: https://mpra.ub.uni-muenchen.de/57784/.

Cordero J. M., E. Alonso-Morán, R. Nuño-Solinis, J. F. Orueta and R. S. Arce (2015) 'Efficiency Assessment of Primary Care Providers: A Conditional Nonparametric Approach', *European Journal of Operational Research* 240, 235–44.

Cordero, J. M., J. Salinas-Jiménez and M. M. Salinas-Jiménez (2016) 'Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis', *European Journal of Operational Research* in press. http://dx.doi.org/10.1016/j.ejor.2016.07.025.

Daraio C. and L. Simar (2005) 'Introducing Environmental Variables in Nonparametric Frontier Models: A Probabilistic Approach', *Journal of Productivity Analysis* 24, 93–121.

Daraio C. and L. Simar (2007a) *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, Methodology and Applications. New York: Springer.

Daraio C. and L. Simar (2007b) 'Conditional Nonparametric Frontier Models for Convex and Non Convex Technologies: A Unifying Approach', *Journal of Productivity Analysis* 28, 13–32.

Daraio, C., L. Simar and P. Wilson (2015) 'Testing the "separability" Condition in Two-stage Nonparametric Models of Production', ISBA Discussion Paper (2015/18), Université Catholique de Louvain.

De Witte K. and B. Geys (2011) 'Evaluating Efficient Public Good Provision: Theory and Evidence from Generalized Conditional Efficiency Model for Public Libraries', *Journal of Urban Economics* 69, 319–27.

De Witte K. and M. Kortelainen (2013) 'What Explains the Performance of Students in a Heterogeneous Environment? Conditional Efficiency Estimation with Continuous and Discrete Environmental Variables', *Applied Economics* 45, 2401–12.

Demchuk P. and V. Zelenyuk (2009) 'Testing Differences in Efficiency of Regions Within a Country: The Case of Ukraine', *Journal of Productivity Analysis* 32, 81–102.

Deprins D., L. Simar and H. Tulkens (1984) 'Measuring Labor Inefficiency in Post Offices', in M. Marchand, P. Pestieau and H. Tulkens (eds), *The Performance of Public Enterprises: Concepts and Measurements*. Amsterdam: North-Holland, 243–67.

Essid H., P. Ouellette and S. Vigeant (2010) 'Measuring Efficiency of Tunisian Schools in the Presence of Quasi-fixed Inputs: A Bootstrap Data Envelopment Analysis Approach', *Economics of Education Review* 29, 589–96.

Essid H., P. Ouellette and S. Vigeant (2013) 'Small is Not That Beautiful After All: Measuring the Scale Efficiency of Tunisian High Schools Using a DEA-Bootstrap Method', *Applied Economics* 45, 1109–20.

Estelle S. M., A. Johnson and J. Ruggiero (2010) 'Three-stage DEA Models for Incorporating Exogenous Inputs', *Computers and Operations Research* 37, 1087–90.

Farrell M. J. (1957) 'The Measurement of Productive Efficiency', *Journal of the Royal Statistical Society. Series A (General)* 120, 253–90.

Fried H. O., C. A. K. Lovell, S. S. Schmidt and S. Yaisawarng (2002) 'Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis', *Journal of Productivity Analysis* 17, 157–74.

Fried H. O., C. K. Lovell and S. S. Schmidt (eds) (2008) *The Measurement of Productive Efficiency and Productivity Growth*. New York: Oxford University Press.
Greene W. (1980) 'Maximum Likelihood Estimation of Econometric Frontier Functions', *Journal of Econometrics* 13, 26–57.
Haas D. A. and F. H. Murphy (2003) 'Compensating for Non-homogeneity in Decision-making Units in Data Envelopment Analysis', *European Journal of Operational Research* 144, 530–44.
Halkos G. and N. Tzeremes (2011) 'A Conditional Nonparametric Analysis for Measuring the Efficiency of Regional Public Healthcare Delivery: An Application to Greek Prefectures', *Health Policy* 103, 73–82.
Haelermans C. and K. De Witte (2012) 'The Role of Innovations in Secondary School Performance: Evidence from a Conditional Efficiency Model', *European Journal of Operational Research* 223, 541–9.
Harrison J., P. Rouse and J. Armstrong (2012) 'Categorical and Continuous Non-discretionary Variables in Data Envelopment Analysis: A Comparison of Two Single-stage Models', *Journal of Productivity Analysis* 37, 261–76.
Hoff A. (2007) 'Second Stage DEA: Comparison of Approaches for Modeling the DEA Score', *European Journal of Operational Research* 181, 425–35.
Holland D. S. and S. T. Lee (2002) 'Impacts of Random Noise and Specification on Estimates of Capacity Derived from Data Envelopment Analysis', *European Journal of Operational Research* 137, 10–21.
Huguenin J. M. (2015) 'Data Envelopment Analysis and Non-discretionary Inputs: How to Select the Most Suitable Model Using Multi-criteria Decision Analysis', *Expert Systems with Applications* 42, 2570–81.
Jeong S. O., B. U. Park and L. Simar (2010) 'Nonparametric Conditional Efficiency Measures: Asymptotic Properties', *Annals of Operations Research* 173, 105–22.
Johnson A. L. and T. Kuosmanen (2011) 'One-stage Estimation of the Effects of Operational Conditions and Practices on Productive Performance: Asymptotically Normal and Efficient, Root-n Consistent StoNEZD Method', *Journal of Productivity Analysis* 36, 219–30.
Johnson A. L. and T. Kuosmanen (2012) 'One-Stage and Two-stage DEA Estimation of the Effects of Contextual Variables', *European Journal of Operational Research* 220, 559–70.
Li Q. and J. S. Racine (2004) 'Cross-Validated Local Linear Nonparametric Regression', *Statistica Sinica* 14, 485–512.
Li Q. and J. S. Racine (2007) *Nonparametric Econometrics: Theory and Practice*. Princeton: Princeton University Press.
Miningou É. W. and V. Vierstraete (2013) 'Households' :Living Situation and the Efficient Provision of Primary Education in Burkina Faso', *Economic Modelling* 35, 910–17.
Muñiz M. (2002) 'Separating Managerial Inefficiency and External Conditions in Data', *European Journal of Operational Research* 143, 625–43.
Muñiz M., J. Paradi, J. Ruggiero and Z. Yang (2006) 'Evaluating Alternative DEA Models Used to Control for Non-Discretionary Inputs', *Computers and Operations Research* 33, 1173–83.
Nedelea I. C. and J. M. Fannin (2013) 'Technical Efficiency of Critical Access Hospitals: An Application of the Two-stage Approach with Double Bootstrap', *Health Care Management Science* 16, 27–36.
Perelman S. and D. Santín (2009) 'How to Generate Regularly Behaved Production Data? A Monte Carlo Experimentation on DEA Scale Efficiency Measurement', *European Journal of Operational Research* 199, 303–10.
Ray S. C. (1991) 'Resource Use Efficiency in Public Schools: A Study of Connecticut data', *Management Science* 37, 1620–28.
Ruggiero J. (1996) 'On the Measurement of Technical Efficiency in the Public Sector', *European Journal of Operational Research* 90, 553–65.
Ruggiero J. (1998) 'Non-discretionary Inputs in Data Envelopment Analysis', *European Journal of Operational Research* 111, 461–9.
Santín D. and G. Sicilia (2015) 'Measuring the Efficiency of Public Schools in Uruguay: Main Drivers and Policy Implications', *Latin American Economic Review* 24, 1–28.
Simar L. and P. W. Wilson (2007) 'Estimation and Inference in Two-Stage, Semi Parametric Models of Production Processes', *Journal of Econometrics* 136, 31–64.
Simar L. and P. W. Wilson (2011) 'Two-stage DEA: Caveat Emptor', *Journal of Productivity Analysis* 36, 205–18.

Syrjänen M. J. (2004) 'Non-discretionary and Discretionary Factors and Scale in Data Envelopment Analysis', *European Journal of Operational Research* 158, 20–33.

Verschelde M. and N. Rogge (2012) 'An Environment-adjusted Evaluation of Citizen Satisfaction with Local Police Effectiveness: Evidence from a Conditional Data Envelopment Analysis Approach', *European Journal of Operational Research* 223, 214–15.

Yang H. and M. Pollitt (2009) 'Incorporating Both Undesirable Outputs and Uncontrollable Variables into DEA: The Performance of Chinese Coal-fired Power Plants', *European Journal of Operational Research* 197, 1095–105.

Yu C. (1998) 'The Effects of Exogenous Variables in Efficiency Measurement. A Monte Carlo Study', *European Journal of Operational Research* 105, 569–80.

Zhang Y. and R. Bartels (1998) 'The Effect of Sample Size on the Mean Efficiency in DEA With An Application to Electricity Distribution in Australia, Sweden and New Zealand', *Journal of Productivity Analysis* 9, 187–204.