NHH

INSTITUTT FOR FORETAKSØKONOMI

DEPARTMENT OF BUSINESS AND MANAGEMENT SCIENCE

FOR 4 2016

ISSN: 1500-4066

February 2016

Discussion paper

Finding the Right Yardstick: Regulation under Heterogeneous Environments

BY

Endre Bjørndal, Mette Bjørndal, Astrid Cullmann AND **Maria Nieswand**

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS

Finding the Right Yardstick: Regulation under Heterogeneous Environments

Endre Bjørndal*, Mette Bjørndal*, Astrid Cullmann†, Maria Nieswand†

February 11, 2016

Abstract

Revenue cap regulation is often combined with systematic benchmarking to reveal the managerial inefficiencies when regulating natural monopolies. One example is the European energy sector, where benchmarking methods are based on actual cost data, which are influenced by managerial inefficiency as well as operational heterogeneity. This paper demonstrates how a conditional nonparametric method, which allows the comparison of firms operating under heterogeneous technologies, can be used to estimate managerial inefficiency. A dataset of 123 distribution firms in Norway is used to show aggregate and firm-specific effects of conditioning. By comparing the unconditional model to our proposed conditional model and the model presently used by the Norwegian regulator, we see that the use of conditional benchmarking methods in revenue cap regulation may effectively distinguish between managerial inefficiency and operational heterogeneity. This distinction leads first to a decrease in aggregate efficient costs and second to a reallocation effect that affects the relative profitability of firms and relative customer prices, thus providing a fairer basis for setting revenue caps.

JEL-Classification: L94, C44, L51

Keywords: Data Envelopment Analysis, Yardstick Regulation, Electricity Distribution

*Norwegian School of Economics, endre.bjorndal@nhh.no / mette.bjorndal@nhh.no

†DIW Berlin, acullmann@diw.de / mnieswand@diw.de

1 Introduction

In general, revenue-cap schemes tend to constrain monopolistic firm behavior by "capping" the revenues regulated firms are allowed to earn. Similar to price cap regulation, revenue caps provide incentives for cost reductions by efficiency improvement since firms are allowed to keep their additional profits from cost savings.

From the perspective of maximizing welfare, the regulator aims to implement a revenue-cap scheme in which the regulated firms sets prices equal to their average costs, i.e. the efficient costs C_i^* (the so-called second-best solution in regulation economics, Laffont and Tirole (1993); Armstrong and Sappington (2007); Shleifer (1985)). The challenge, then, is to determine C_i^* due to information asymmetry.[1] To determine C_i^* and reveal managerial inefficiency, regulators may combine revenue-cap schemes with systematic benchmarking techniques (Agrell and Bogetoft, 2013).

By comparing the performance of the firms via cost functions, benchmarking provides information on the unknown technology, the cost structure and efficient costs and thus is crucial to determining revenue caps.[2] In other words, the more precisely C_i^* are approximated, the closer the revenue caps are to the second-best solution (Laffont and Tirole, 1993).

The precision of efficient cost approximation is incrementally related to the exogenous factors (Shleifer, 1985) describing the operational environment of regulated firms. Exogenous factors cause heterogeneity in terms of technol-

[1] The regulator does not know the efficient cost structures of the regulated firms and they have no incentives to reveal the information.

[2] Generally, benchmarking can be interpreted as creating a hypothetical competition among natural monopolies. Thereby a mechanism of yardstick competition in the spirit of Shleifer (1985) is introduced as the revenue of a particular electricity distribution firm based on the costs of all other distribution firms.

ogy, and hence, efficient costs. Consequently, the level of efficient costs differs between firms according to their environmental conditions.

In this paper, we demonstrate that regulators fail to implement efficient cost levels and revenue caps by not properly accounting for exogenous firm characteristics. We discuss unconditional nonparametric benchmarking applications used in regulation mainly based on data envelopment analysis (DEA), which set revenue caps too high and compensate firms for environmental disadvantages *and* managerial inefficiency. In particular, we analyze the importance of selecting the appropriate reference set for comparing firms to determine efficient costs, C_i^*.

We suggest that conditioning a firm's production process on its operational environment provides more reliable peers. Because Norway has extensive experience with revenue cap regulation since 1997, we use a dataset of 123 Norwegian electricity distribution firms to test and compare three models: 1) an unconditional DEA benchmarking model; 2) a conditional DEA benchmarking model; and 3) the Norwegian regulator's model. We explicitly analyze two effects of conditioning: 1) the effects of a different peer selection; and 2) the effect of compensating for environmental variables under the conditional approach. We also demonstrate the effect of conditioning on the aggregate efficient costs, firms' profits, and consumers' prices for the distribution service. We see, for the Norwegian case, that the use of a conditional approach will lead to a reduction in aggregate efficient cost, but the main effect is a reallocation of revenue between firms, leading to changes in relative profitability and customer prices in the industry.

Even though we use Norway as a case study, our empirical application is relevant for the European regulators using nonparametric benchmarking methods to determine revenue caps. Our analysis is also relevant for general cases in which nonparametric benchmarking is applied to compare decision-making units in different environments.

The remainder of this paper is organized as follows. Section 2 explains the implementation of yardstick regulation via benchmarking and the challenges of separating managerial inefficiency and operational heterogeneity. Section 3 shows that the conditional framework is appropriate when heterogeneous firms are benchmarked against each other. Section 4 presents our empirical strategy and our data. Section 5 summarizes the empirical findings. Section 6 shows the effects of conditioning for regulation. Section 7 concludes.

2 Revenue cap regulation via benchmarking

2.1 Revenue cap regulation

A revenue cap regime can be implemented via the observed costs, C, of the firms (Joskow, 2007) and formalized as

$$R_i = \alpha \cdot C_i^* + (1 - \alpha) \cdot C_i, \tag{1}$$

where the capped revenue R_i for firm i is determined by its actual observed cost C_i and efficient cost C_i^*. The parameter α weights the actual and efficient costs with $0 < \alpha < 1$, according to the strength of the regulatory system.

Obviously, the observed firm-specific costs available to the regulator are influenced by firms' managerial inefficiency and the different operational environments. The major challenge for the regulator is to determine the firm-specific revenue caps, R_i, such that they account for the environmental effects, meaning that firms will only be compensated for environmental disadvantages, not managerial inefficiency.

2.2 Benchmarking and managerial inefficiency

In regulatory practice, the efficient cost, C_i^*, is often determined by means of an *unconditional* nonparametric benchmarking method, where data envelopment analysis (DEA), developed by Charnes et al. (1978), is the most popular variant.[3] DEA estimates the unknown technology or production set Ψ from a given sample of observed vectors of inputs $x_i \in \mathbb{R}_+^p$ and outputs $y_i \in \mathbb{R}_+^q$ used by the firms $i = 1, ..., n$, and where p and q represent the number of input and output factors, respectively. The boundary of Ψ is the frontier and represents the estimated unknown technology (see Appendix A.1 for the standard DEA). Comparing each individual firm to this frontier determines the firm's specific managerial efficiency. Solving a linear program assuming constant returns to scale (CRS) according to Charnes et al. (1978) provides efficiency scores $\hat{\theta} \in \mathbb{R}^n$ such that $0 \leq \hat{\theta}_i \leq 1$ for $i = 1, ..., n$.[4] The efficiency estimate for firm i, i.e., $\hat{\theta}_i$, is a measure relative to the frontier, which is determined by fully efficient observations $j \in \{1, ..., n\}$ with $\hat{\theta}_j = 1$, i.e., the peers.

In the unconditional case, DEA does not consider external factors. Hence, all observations belong to the reference set of the particular observation of

[3] Frontier models are widely applied in performance measurement. The theoretical foundations derive from Koopmans (1957), Debreu (1959) and Farrell (1957). In regulatory practice, nonparametric approaches have outperformed parametric methods due to their easy implementation and interpretation, which is often why regulators favor their use. See Bogetoft and Otto (2011) for a comparison and critical evaluation of the approaches in regulatory practice.

[4] Regulators often assume CRS as proposed by Charnes et al. (1978), because it implies that any observed production plan can be arbitrarily scaled up and down, and it implies that all convex combinations of two observed production plans are assumed to be feasible. The validity of these assumptions, which are beyond the scope of this paper, are not further discussed.

interest and potentially serve as its peers irrespective of their operational environments. However, this involves the risk of evaluating an observation to an infeasible frontier when external factors are disadvantageous or beneficial, respectively, to its production process.

2.3 Managerial inefficiency versus operational heterogeneity

It is worthwhile to emphasize that the external factors can have different channels through which they influence firms' performance. Bădin et al. (2012) point out that environmental factors may influence the technology itself represented by the boundary of the production set Ψ, thus causing a frontier shift, the distribution of inefficiency, or both. The regulator has to compensate for the resulting cost disadvantages only from exogenous frontier shifts due to different environments, and not for managerial inefficiency.

We will assume that the environmental impact can be measured by a set of vectors $z_i \in \mathbb{R}^r$ for $i = 1, ..., n$, where r is the number of environmental factors. In practice, regulators use second-stage regressions (Bjørndal et al., 2010; Agrell et al., 2014) to control for the heterogeneity of operational environments on the frontier. The z-variables are regressed on the DEA estimates to determine the impact of the operational environments on efficiency. The efficiency scores are then adjusted to compensate for the impact of z-variables. One concern, however, is that second-stage regressions are only useful when z-variables fulfill the separability condition (Bădin et al., 2012; Simar and Wilson, 2007).[5] But when separability is given, there is no

[5]Separability is given if z-variable does not influence the attainable set, and thus, the frontier. Only then do second-stage regressions provide meaningful results to explain the differences in the distribution of inefficiency.

frontier shift due to environmental differences and a compensation by means of second-stage regressions will fail to set efficient revenue caps.

In addition, if z-variables cause a frontier shift, and thus are not separable which is usually the case, DEA efficiency scores from the first step lack economic sounding, since they are based on a frontier that production units are unable to reach (Bădin et al., 2012). Using second-stage regressions ignores the potential impact of z-variables on the frontier and their potential impact on the distribution of inefficiency. Therefore, compensation based on the two-stage approach is likely to capture multiple effects and lead to under- or overcompensation of the individual regulated firms, which becomes apparent in a too high or too low C_i^*, and therefore a too high or too low R_i.

3 Conditional benchmarking

Conditional nonparametric benchmarking approaches, which do not require the separability condition, can account for the multiple effects of the operational environment on a firm's performance[6] depending on how the reference sets, i.e. the group of firms use to compare against the firm of interest, are selected. By means of kernel estimation the reference sets are restricted with respect to z-variables, prior to measuring actual performance such that firms are only compared to others with similar environments.[7] Compensa-

[6]To incorporate external factors into the performance evaluation, conditional efficiency estimation was proposed initially by Cazals et al. (2002) in the order-m framework and further developed by Daraio and Simar (2005), Daraio and Simar (2007b), and Daraio and Simar (2007a). The approach aims to compare only units that operate under similar operation environments, i.e. the selection of the reference group for a particular observation is conditional on their z-variables.

[7]Thus, conditioning the performance evaluation on z-variables assumes that the frontier is feasible for the firm to reach, and that the efficiency scores are meaningful (Bădin et al.,

tion for the resulting cost disadvantages is only based on the frontier shift, thus leading to an adequate implementation of the revenue cap scheme.

3.1 Constructing a conditional efficiency estimator

The conditional DEA estimator is based on an attainable production set, $\hat{\Psi}^z$, conditioned on a set of z-variables and implies the estimation of a conditional distribution function where the production process is conditional to a particular level of z (Daraio and Simar, 2007b; Bădin et al., 2010).[8] The latter requires applying a smoothing technique. Therefore, we perform Kernel estimation with an Epanechnikov kernel $K(\cdot)$ as in, e.g., Daraio and Simar (2005) and Daraio and Simar (2007b).[9] The Kernel is defined as

$$K_h = K((z_i - z_k)/h) \tag{2}$$

where z_i and z_k are vectors of z-variables for a unit i and it's reference unit k, respectively, and h is the vector of selected bandwidths. For each of the environmental variables we compute a bandwidth based on least squares cross validation (Hall et al., 2004; Li and Racine, 2007, 2008). Note that the bandwidth selection procedure relies on estimating the conditional probability distribution function of y, given a particular level of z.[10] Hall et al. (2004) emphasize that their proposed method assigns large smoothing parameters to components of z that are irrelevant for estimating the density of y. Therefore, the sizes of the selected bandwidths themselves already contain information about the impact of particular z-variables on the production

2012).

[8]The statistical properties of this estimator are derived in Kneip et al. (2008) and its consistency is established in Jeong et al. (2010).

[9]The authors suggest using kernels with compact support in the framework of conditional boundary estimation.

[10]See Hall et al. (2004) for a detailed presentation of the method.

output y. We then use the obtained bandwidths to estimate the Kernel function in Equation 2 to compute kernel probabilities. Firms closely located to firm i in terms of z thereby receive higher probabilities to be selected into the reference set of the observation of interest, whereas small (or even zero) kernel probabilities are assigned to firms with very different operational environments than firm i.

As shown in Daraio and Simar (2007b), the conditional DEA efficiency estimate for firm i under the assumption of CRS is given by

$$\hat{\theta}_i^c = \min\{\theta \mid \theta x_i \geq \sum\nolimits_{j|z_i-h\leq z_j\leq z_i+h}^{n} \lambda_j x_j,\ y_i \leq \sum\nolimits_{j|z_i-h\leq z_j\leq z_i+h}^{n} \lambda_j y_j, \quad \text{and } \lambda_j \geq 0 \text{ for } j = 1, ..., n\}, \tag{3}$$

where the vector $h \in \mathbb{R}^r$ represents bandwidths of appropriate size. For each observation, the bandwidths determine the range of z in which other observations are considered being similar. Hence, we consider only the observations within this range as potential peers for the unit of interest and select them into the respective reference group. That is, we restrict the reference set of firm i to firms with positive kernel probabilities.

3.2 Bias-correction

DEA efficiency scores are based on finite samples of observations and construct a best-practice frontier, i.e. by construction they are upward-biased Simar and Wilson (1998). We correct for the bias in $\hat{\theta}$ and $\hat{\theta}^c$ by applying the m-bootstrap first proposed by Kneip et al. (2008), and extended by Simar and Wilson (2011).[11] Unlike the naive bootstrap, this approach allows consistent bias-correction by drawing bootstrap subsamples of size $m = n^\kappa$ from

[11] As a variant of the original procedure, we also use the kernel probabilities in order to construct the bootstrap samples for the conditional case, which is consistent with the idea of conditioning the production process and is supposed to give even more precise insights

the given sample of size n with $\kappa \in (0,1)$. Each of the $b = 1, ...B$ bootstrap samples (replications) provides a random subsample of size m which we use to compute the bootstrapped vector of efficiency scores denoted as $\hat{\theta}_{m,b} \in \mathbb{R}^n$. The bias for the unconditional DEA and the conditional DEA is defined as

$$\hat{bias}_B(\hat{\theta}) = \left(\frac{m}{n}\right)^{2/(p+q+1)} \cdot \left(B^{-1}\sum_{b=1}^{B}\hat{\theta}_{m,b} - \hat{\theta}\right) \tag{4}$$

$$\hat{bias}_B(\hat{\theta}^c) = \left(\frac{m}{n}\right)^{2/(p+q+1)} \cdot \left(B^{-1}\sum_{b=1}^{B}\hat{\theta}^c_{m,b} - \hat{\theta}^c\right) \tag{5}$$

where the factor $(m/n)^{2/(p+q+1)}$ controls for the effect of different sample sizes in both the true world and bootstrap world (Simar and Wilson, 2008). Then, we obtain bias-corrected efficiency scores, $\hat{\theta}$ and $\hat{\theta}^c$ by subtracting the bias from $\hat{\theta}$ and $\hat{\theta}^c$ repsectively.

$$\tilde{\theta} = \hat{\theta} - \hat{bias}_B(\hat{\theta}) \tag{6}$$

$$\tilde{\theta}^c = \hat{\theta}^c - \hat{bias}_B(\hat{\theta}^c) \tag{7}$$

Cazals et al. (2002) show that subsampling also overcomes the outlier sensitivity of convex nonparametric frontier models such as DEA models. Although the statistical literature does not precisely define outliers, they can be understood as atypical observations that possibly influence the efficiency estimates of other data points if they distort the frontier Simar (2003). By drawing m out of n observations from the sample, we reduce the influence of potential outliers since they will not always be drawn. Therefore, the computed biases in Equations 4 and 5 are robust toward outliers and helps us tackle the problem. We do not delete outliers from the sample because they give us important information about the heterogeneity of operational environments and the effects on the production frontier.[12]

about the effect of z-variables on the production process.

[12]Regulators also need to include all firms in the analysis.

4 Empirical strategy and data

4.1 Empirical strategy

Our objective is to demonstrate that *conditioning* a firm's production process on its operational environment gives us a better yardstick for determining efficient cost. Therefore, we compare the outcomes of the *unconditional* model used in regulation with our proposed *conditional* model to show the effects of conditioning on efficiency estimates, revenue caps, firms' profits, and prices consumers have to pay. As mentioned, we estimate three different models:

1. The bias-corrected unconditional DEA model ($\tilde{\theta}$).
2. The bias-corrected conditional DEA model ($\tilde{\theta}^c$).
3. The unconditional DEA model with second-stage regression based on the bias-corrected unconditional DEA scores used by the Norwegian regulator ($\tilde{\theta}^{NVE}$).

We run 2,000 replications to obtain the bias-corrected efficiency estimates by the m-bootstrap where drawing the reference sets depends on the respective kernel probabilities. We select $m = 60$ using the *leave-one-out-order-m* algorithm proposed by Daraio and Simar (2007a) and Simar (2003) based on Cazals et al. (2002). Note that $\tilde{\theta}^{NVE}$ are based on $\tilde{\theta}$, and we correct these estimates by the effects of the environmental factors using the second-stage procedure described in Amundsveen et al. (2014).[13] The first effect of conditioning, i.e., different a different peer selections, becomes obvious when comparing $\tilde{\theta}$ to $\tilde{\theta}^c$. The second effect of conditioning i.e. compensating differently for z-variables becomes obvious when comparing $\tilde{\theta}^c$ to $\tilde{\theta}^{NVE}$.

[13]For comparability, we use the bias-corrected unconditional DEA scores as the starting point in the two-stage procedure.

4.2 Data

Our dataset comprise 123 Norwegian electricity distribution firms regulated by the Norwegian Water Resources and Energy Directorate (NVE). Since the deregulation of the Norwegian power market in 1990, the regulatory regime for distribution and regional transmission firms[14] has gone through several phases. In the first years after deregulation, firms were subject to a cost plus (rate of return) regulation with rather week efficiency incentives. In 1997, revenue cap regulation with benchmarking was introduced with five-year regulation periods.

Firm-specific efficiency requirements in each of the first two regulation periods, were based on DEA analysis of historical cost and output data. After 2007, the efficiency incentives were further strengthened with the introduction of a yearly revenue cap implementation and annual updates of the revenue caps based on DEA analyses.[15]

4.2.1 Inputs and outputs

Table 1 lists the input and output variables in our dataset corresponding to those used in the regulatory benchmarking model. The input variable measures total expenditures(TotEx), which comprise five cost elements: the value of lost load (VOLL), thermal power losses, capital depreciation, operation and maintenance expenses, and return on capital.[16] We follow the

[14]The TSO is subject to a separate regulation, which we do not discuss in this paper.

[15]A more detailed discussion of the different regulatory phases is given by Bjørndal et al. (2010) and Amundsveen and Kvile (2015).

[16]Most of the firms also own and operate part of the regional transmission network, and NVE reallocates part of this cost to the (local) distribution activity. For simplicity, we do not include the reallocated cost in our analyses, and our results may differ somewhat from the efficiency measurements published by NVE.

practice of NVE and construct our benchmarks based on average data over a five-year period, having adjusted the data to the price level of a base year.[17] Output variables measure the number of customers, length of the high voltage network, and the number of network stations (transformers).

Table 1: Definitions of variables.

Variable	Name	Sub-variable	Unit
Input			
		O&M costs	1000 NOK
		Value of lost load (VOLL)	1000 NOK
TotEx	x	Thermal power losses	1000 NOK
		Capital depreciation	1000 NOK
		Return on capital	1000 NOK
Outputs			
Customers	y_{Cus}	-	No. of customers
High voltage lines	y_{HV}	-	Kilometers
Network stations	y_{Net}	-	No. of stations
Environment			
Average distance to road	z_{Dis}	-	Meters
HV lines underground	z_{Und}	-	Share of HV network (0-1)
Forest (coniferous)	z_{For}	-	Share of HV network affected (0-1)

[17]We use an industry-specific price index for adjusting operations and maintenance costs and the consumer price index for the VOLL costs. Thermal losses are valued at the average system price for the base year. Capital depreciation is based on reported (nominal) book values, and the return on capital is calculated using the nominal rate of return set by the regulator for the base year.

4.2.2 Environmental variables

We use three variables to measure heterogeneous operational conditions (see Table 1). The first, z_{Dis}, is the average distance between the road network and a firm's network, representing the increased difficulty of maintaining a network that is not easily accessible. The second, z_{Und}, is the share of underground cables that may imply cost disadvantages because of expensive installation as well as cost advantages in terms of fewer outages and lower VOLL. The third is the share of fast-growing forest, z_{For} that may represent a cost disadvantage due to the added cost of forest clearing.[18]

Table 2 lists the summary statistics for the 123 distribution firms. We use average data for the period 2008-2012, and we set 2012 as the base year for adjustment of the cost data. Note that averaging the data does not affect the environmental variables, since their values are constant over time in our dataset.

5 Results

5.1 Selected bandwidths

Table 3 lists the computed bandwidths.[19] Given that they are the smoothing parameters for estimating the conditional density of the multivariate kernel function, they inform us about the impact of the z-variables on the produced

[18] We do not consider the two composite geographical variables used by NVE for this period. With these two variables NVE combines several sub-variables, which involves the use of principal component analysis (PCA). The nature of these variables prevents us from obtaining reasonable bandwidths that remain stable over multiple runs. Nevertheless, excluding both variables does not affect the overall analysis.

[19] The estimates have been scaled so that they are comparable with the corresponding z-values. For the Epanechikov kernel the scaling factor is $\sqrt{5}$.

Table 2: Summary statistics ($n = 123$).

	Input	**Outputs**			**Environment**		
	TotEx	y_{Cus}	y_{HV}	y_{Net}	z_{Dis}	z_{Und}	z_{For}
Mean	95045.6	22669.7	796.2	1006.6	227.8	0.3386	0.1191
Std. dev.	188994.6	59221.2	1328.0	1893.4	208.6	0.1755	0.0994
Minimum	8456.2	1014.2	54.0	59.0	70.4	0.0571	0.0000
Median	35738.9	6386.8	324.8	369.4	142.9	0.3060	0.1191
Maximum	1579179.2	547693.0	8494.6	13491.4	1056.4	0.8641	0.3916

Notes: Average data for the period 2008-2012; 2012 is base year for adjustment of the cost data.

output y. Therefore, they indicate a respective z-variable's importance for selecting the reference set in our conditional efficiency estimation.

For the distance to road variable z_{Dis}, Table 3 shows a very large bandwidth compared to the median value of 142.9 shown in Table 2. Thus, this variable is smoothed out in the kernel density estimation, which implies that the overall impact of z_{Dis} on output y is limited, and therefore, the reference sets are not restricted by z_{Dis}. The magnitudes of bandwidths for the other two variables, 0.29615 for z_{Und} and 0.12891 for z_{For}, are considerably smaller. For interpretation, we again compare the bandwidths to the median values in Table 2, i.e., 0.306 for z_{Und} and 0.1191 for forest z_{For}, respectively. We suggest that the exogenously induced cost disadvantages, which would be measured as inefficiency, if not controlled for, are most likely due to z_{Und} and z_{For}.

Table 3: Estimated bandwidths for z-variables

Variable		Bandwidth
Average distance to road	z_{Dis}	1299522861
HV lines underground	z_{Und}	0.29615
Forest (coniferous)	z_{For}	0.12891

5.2 Reference sets and selected peers

To illustrate how estimated bandwidths affect the selection of the firms comprising the reference set, when individual firms are benchmarked, we start by looking at a firm which we will denote i_0. Figure 1 shows the two-dimensional contour plots of the kernel density function for firm i_0 indicated by the yellow point. Its values for z_{Dis}, z_{Und}, and z_{For} are 162, 0.5236, and 0.1307, respectively. We only select firms with positive kernel density values as members of firm i_0's reference set, i.e., any firm i in the sample for which the values of each z-variable $j \in \{Dis,\ Und,\ For\}$ satisfies $|z_{ij} - z_{i_0 j}| \leq h_j$.

The shaded areas in both panels of Figure 1 represent the respective combinations of two z-variables for which firms obtain positive kernel density values. The left panel shows that the possible peers of firm i_0 will be firms with z_{Und} and z_{For} values within the range of $[0.2274, 0.8197]$ and $[0.0018, 0.2596]$, respectively, i.e., the 70 firms indicated by black dots. Gray dots indicate firms that are too different from firm i_0 to be selected in its reference set. The shaded area in the right panel is not limited in the dimension of distance, which is due to its large bandwidths. Note that this shaded area also includes gray-colored firms because they are similar in terms of z_{Dis} but not in terms of z_{Und}, and therefore, do not belong to the reference set of firm i_0. Given that the kernel centers around firm i_0, the density estimation produces

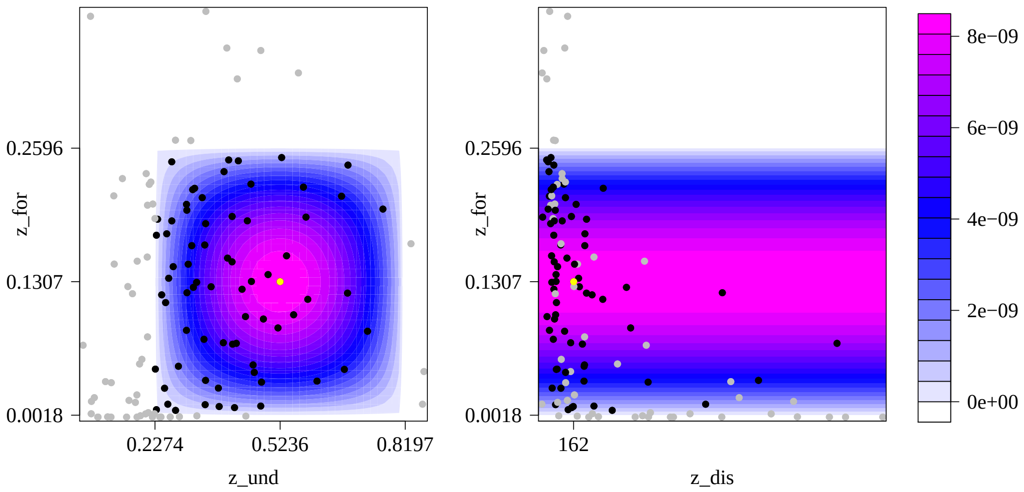

Figure 1: Kernel densities and reference set for firm i_0 .

a reference set specific to each firm. Figure 2 shows that both the size of the reference set and number of peers differ between the unconditional and the conditional approaches. The blue crosses and the black circles in Figure 2 indicate the number of selected peers for each firm by model. The number of selected peers for firm i is the number of firms out of its reference set that serve as peers for this firm in at least one of the bootstrap samples.[20] For almost all firms, note that unconditional DEA selects more peers for efficiency evaluation than conditional DEA. The green line is the number of identical peers within both groups. In extreme cases the overlap of selected peers is very small or even zero, which is particularly true for firms with a small number of selected peers under unconditional DEA. In fact, Figure 2 shows

[20]If λ_{bij}^m is the weight of firm j in the reference set of firm i in bootstrap sample b, we let $\overline{\lambda}_{ij}^m = B^{-1}\sum_{b=1}^{B}\lambda_{bij}^m$ denote the average weight of firm j for efficiency measurement of firm i. The number of selected peers for firm i counts the firms j for which $\overline{\lambda}_{ij}^m > 0$.

that when unconditional DEA is used, each firm in our sample is evaluated against a frontier spanned by peers with dissimilar operational environments (for a detailed list of the sizes of reference sets and identical and different peers in unconditional and conditional models see Appendix 7).

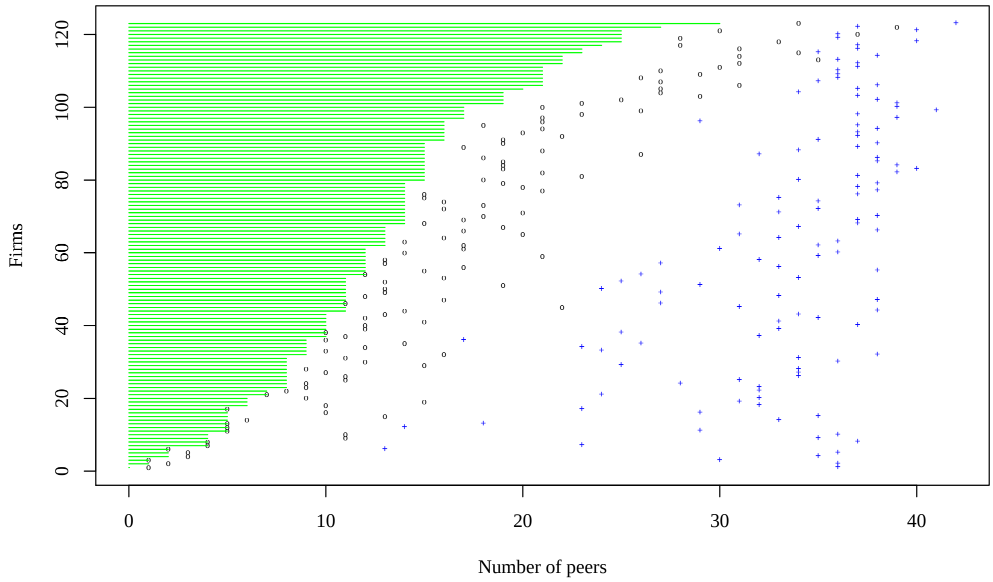

Figure 2: Number of peers.

5.3 Efficiency estimates

We focus on our three efficiency measures ($\tilde{\theta}$, $\tilde{\theta}^c$, $\tilde{\theta}^{NVE}$) to demonstrate the effect of differences with respect to reference sets and peer selection. Table 4 lists the descriptive statistics of the estimated efficiency scores.

As expected, the statistics of $\tilde{\theta}^c$ exceed their unconditional counterparts for $\tilde{\theta}$. For example, the median efficiency score is 0.6977 when exogenous factors are accounted for and greater than 0.6557 where the frontier is estimated without considering heterogeneous operation environments. The difference between the point estimates indicates that exogenous factors indeed produce cost disadvantages for the median firm, shifting its feasible frontier such that its efficiency increases by roughly 4 percent. Hence, if $\tilde{\theta}$ scores are used for any further analysis, the median firm will be evaluated based on a technically infeasible frontier.

Comparing $\tilde{\theta}^c$ with $\tilde{\theta}^{NVE}$ estimates reveals that efficiency is higher when the correction for external factors is made via second-stage regression. We interpret this finding as empirical evidence of biased performance measures and note that this bias results from both the underlying and economically meaningless $\tilde{\theta}$ estimates and the second-stage regression, which assumes that z-variables do not affect the frontier. Based on these results, it is likely that NVE compensates both the frontier shifts due to exogenous variables and the managerial inefficiency because the effects of both on the estimated technology are not appropriately separated.

Table 4: Descriptive statistics of efficiency estimates.

Estimator		Min	Median	Mean	Max*	Std. dev.
1. Unconditional DEA	$\tilde{\theta}$	0.4003	0.6557	0.6744	0.9607	0.1195
2. Conditional DEA	$\tilde{\theta}^c$	0.4277	0.6977	0.7141	1.0000	0.1343
3. NVE DEA	$\tilde{\theta}^{NVE}$	0.4322	0.7246	0.7387	1.0061	0.1232

Note: All estimates are based on bias-corrected efficiency scores. *Maximum values can differ from 1 due to bias-correction.

The rank correlations shown in Table 5 further illustrate the relationship between the estimators. We see that the second-stage adjustment under NVEs method does not affect the firm ranking very much, since the correlation between $\tilde{\theta}$ and $\tilde{\theta}^{NVE}$ is as high as 0.96, while the ranking under conditional

DEA is less correlated with the rankings based on the other two estimators. This could imply that the *relative* effects of adjustment for exogenous cost drivers are greater with the conditional DEA method than with NVEs current procedure.

Table 5: Rank correlations between efficiency estimates.

	$\tilde{\theta}$	$\tilde{\theta}^c$	$\tilde{\theta}^{NVE}$
$\tilde{\theta}$	1.00	0.89	0.96
$\tilde{\theta}^c$		1.00	0.88
$\tilde{\theta}^{NVE}$			1.00

5.4 Impact of the exogenous factors on the frontier

In addition to the bandwidths obtained as mentioned above, we want to analyze the impact of the environmental variables on the production process by using the ratio of conditional to unconditional efficiency scores. Bădin et al. (2012) emphasize that unconditional measures of efficiency are economically meaningless if z-variables impact the frontier since units are compared to infeasible production plans. Knowing that conditional measures control for this, the ratio, therefore, informs us about the local effect of z-variables on the attainable frontier.[21] The ratio ρ is defined as

$$\rho = \frac{\tilde{\theta}^c}{\tilde{\theta}} \tag{8}$$

where we refer to the bias-corrected measures. The ratio ρ takes the value of 1 if both measures are equal, i.e. there is no frontier shift, whereas other values imply that the feasible frontier shifts due to z-variables. A ratio larger

[21] In this paper, this is independent of the inherent efficiency of the firms which the full frontier estimations are supposed to reduce by regulatory incentives.

than 1 implies that the conditional exceeds the unconditional efficiency score indicating cost disadvantages due to the external factors, whereas ratios with values smaller than 1 indicate cost advantages. To analyze the local effect of the exogenous factors on the feasible frontier, we plot the ratio of $\tilde{\theta}^c$ to $\tilde{\theta}$. Both panels in Figure 3 sort firm-specific efficiency ratios in an increasing order of the values of z_{Und} and z_{For}.[22]

Notably, the ratios of the two performance measures differ from 1 for almost all firms; the range is roughly 0.9 to 1.4. Hence, z-variables significantly affect production. Also, in 95 of the 123 firms, the ratio, based on bias-corrected efficiency estimates, is larger than 1. Therefore, the frontier-shifting factors mainly lead to cost disadvantages.

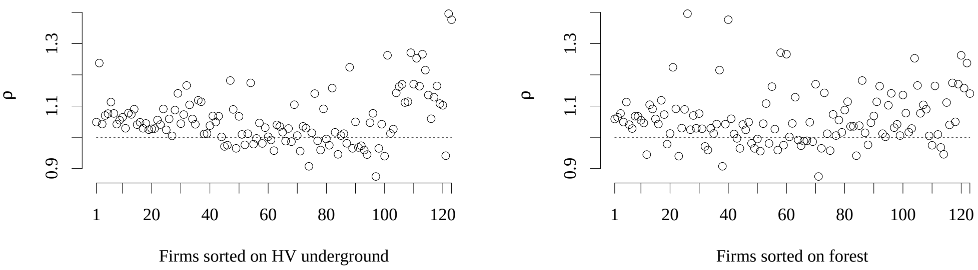

Figure 3: Impact of exogenous factors on the frontier.

6 Effects for regulation

The three models also demonstrate that the choice of benchmarking model, including how to compensate for the effect of exogenous factors, has impor-

[22]From the bandwidth analysis we know that the frontier shift is mainly driven by these two factors.

tant regulatory effects for firms' customers and owners. We start by looking at the impact on the aggregate efficient costs and then look in detail at firm-specific profitability (return on capital) and the average prices custumers pay per kWh. We base our calculations on the current revenue cap model used by the Norwegian regulator (NVE). Similar to Equation 1, as well as the DEA-based mechanism proposed by Bogetoft (1997), the revenue cap for firm i is set by

$$R_i = \alpha \cdot (C_i^* + \Delta_i) + (1 - \alpha) \cdot C_i, \tag{9}$$

where $\alpha = 0.6$. The efficient costs C_i^* are either calculated by $\tilde{\theta}_i^{NVE} \cdot C_i$ or as $\tilde{\theta}_i^c \cdot C_i$. The yardstick formula 9 is applied every year to set the annual revenue caps, but in order to simplify the presentation we will drop the time subscripts in our notation[23].

NVE calibrates the revenue caps, by adding the amount Δ_i to the efficient cost of each firm[24], in order to ensure that revenue equals cost for the industry as a whole, i.e., $\sum R = \sum C$. The rationale for the calibration, as described in Amundsveen and Kvile (2015) and Bjørndal et al. (2010), is to allow the representative firm, with an efficiency equal to the industry (cost-weighted) average, to have a return on its capital equal to the regulated rate of return.

[23] We focus on the most important features of the Norwegian regulation. In practice, there is a two-year time lag in the reporting; the revenue caps in year t must be based on the data available after year $t-2$. We assume that the average of the data for 2008-2012 is representative of a typical year, and we do not consider the timing of the revenue stream. Many firms also own and operate part of the regional transmission network, but we do not consider this part of their revenue caps.

[24] In the present regulation model, the calibration takes the form

$$\Delta_i = \frac{\sum C - \sum C^*}{\sum BV} BV_i,$$

where BV_i is the total book value of capital for firm i. The use of book values in the calibration formula is done to correct for a suspected age bias in the capital costs.

Given the calibration scheme, firms that have above-average efficiency scores will earn more than the regulated rate of return, while firms with below-average efficiency scores will earn less.

6.1 Effect on aggregate efficient costs

Because of the revenue calibration performed by NVE, the compensation scheme for exogenous factors will have no effect on aggregate revenue. In order to study aggregate effects, we will therefore focus on efficient costs. The first row in Table 6 shows the aggregate efficient cost from the unconditional DEA model, i.e., $\sum C^*(\tilde{\theta})$, and the second row shows what the compensation for the exogenous factors will be when it is based on one of the estimators $\tilde{\theta}^c$ or $\tilde{\theta}^{NVE}$. The compensation of 893 MNOK under the current NVE regime is higher compared to the conditional DEA model, where the compensation equals 534 MNOK. The difference of 359 MNOK illustrates the compensation for managerial inefficiency, which the regulator allows under the current regime in addition to the compensation for external factors influencing the production process. Our findings thus show that the currently implemented regime overestimates efficient production costs.

Table 6: Aggregate effects (MNOK).

Efficiency estimator	$v \in \{c, NVE\}$	$\tilde{\theta}^c$	$\tilde{\theta}^{NVE}$
Uncompensated efficient cost	$\sum C^*(\tilde{\theta})$	8,392	8,392
Compensation	$\sum \Big(C^*(\tilde{\theta}^v) - C^*(\tilde{\theta})\Big)$	534	893
Compensated efficient cost	$\sum C^*(\tilde{\theta}^v)$	8,926	9,285
Revenue calibration	$\sum \Delta$	2,765	2,405
Industry revenue	$\sum \Big(C^*(\tilde{\theta}^v) + \Delta\Big)$	11,691	11,691

Under the current regulatory scheme, where the regulator calibrates the final revenue caps by adding Δ_i to the efficient cost of each firm, the overcompensation for exogenous factors will not affect the aggregate revenue. The revenue calibration, i.e., $\sum \Delta$, equals 2,405 MNOK under the current regulatory model, whereas it would be 2,765 MNOK with our alternative model. As shown on the last row of Table 6, the aggregate revenue will be unaffected by the choice of compensation scheme.

6.2 Effect on firm-specific profits and customer prices

Figure 4 illustrates how the choice of compensation ($\tilde{\theta}^c$ versus $\tilde{\theta}^{NVE}$) affects the owners of the network firms. The columns represent the difference in firm-specific revenues between NVE's DEA model and the conditional DEA, $R_i(\tilde{\theta}^{NVE}) - R_i(\tilde{\theta}^c)$. In order to make the example interesting in a more general setting, we also show firm-specific effects without the calibration effect, i.e., with $\Delta_i = 0$, as dots in the figure. We divide the firm-specific revenues by their total book values, BV_i, to see the effects on return on capital. In general, the effects on return on capital vary from -3.8 (-3.2) percent to +5.4 (+6.0) percent, where a positive value reflects a higher return on capital under the NVE model compared to conditional DEA, and where numbers in parenthesis are effects without calibration. The main effect that we see from the figure is a reallocation of revenue due to the change in compensation method, and this is related to the difference in relative efficiency scores under the respective methods, cf., the difference in rank correlations shown in Section 5.3. The difference between the calibrated and the uncalibrated effects is related to the overcompensation of 359 MNOK in the efficient cost level by the NVE two-stage method, as discussed in the previous section. According to formula 9, 60 percent of this overcompensation will be awarded to the firms, and this represents an increase of 0.6 percent in the return on capital for each firm.

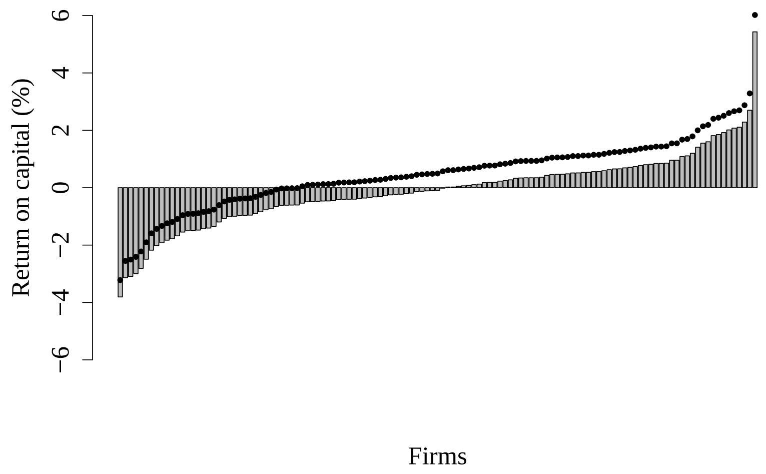

Figure 4: Effect on firms' profits ($r_{NVE2012} = 4.2\%$). Dots = uncalibrated effects.

Figure 5 shows the same differences, $R_i(\tilde{\theta}^{NVE}) - R_i(\tilde{\theta}^c)$, relative to the total quantity of energy delivered by each firm, $Energy_i$, to approximate the average charged prices by each firm. The differences range from -0.030 (-0.021) NOK/KWh to +0.042 (+0.047) NOK/KWh, where a positive value reflects higher prices for customers under the NVE model compared to conditional DEA, and where numbers in parenthesis represent the differences without calibration, i.e., with $\Delta_i = 0$. The average revenue collected from the customers amount to 0.157 NOK/KWh, hence, the reallocation effect of the compensatory scheme is considerable. The overcompensation of 359 MNOK in efficient costs amount to an increase, on average, of only 0.003 NOK/KWh in the customer prices, hence the "overcompensation" effect is small compared to the reallocation effect.

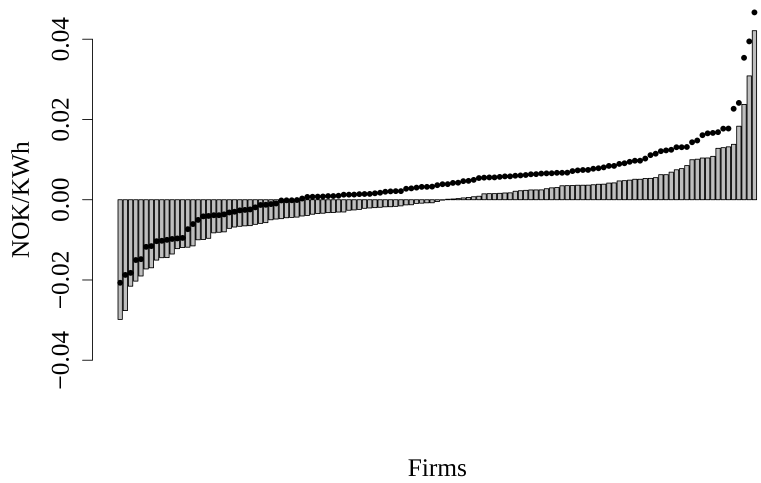

Figure 5: Effect on calibrated revenue caps $\left(\frac{\sum Revenue}{\sum Energy} = 0.157 NOK/KWh\right)$. Dots = uncalibrated effects.

7 Conclusion

In its basic notion, nonparametric benchmarking methods, such as data envelopment analysis that is implemented by several European energy regulators, cannot disentangle the effects of managerial inefficiency and difficult operational environments. Under revenue cap regulation, firms should be compensated for the latter, but not the first. In particular, if environmental factors are suspected to cause non-controllable costs, a two-stage method is often used by energy regulators, including a second stage regression to correct for environmental factors that are not accounted for in the first stage DEA model. In this setting, the two-stage method is likely to capture multiple effects, leading to over- or under-compensation of individual firms. Conditional nonparametric benchmarking approaches have been designed to overcome this challenge, restricting the selection of firms used to compare against the firm of interest, to those with similar environments. This paper proposes a conditional DEA benchmarking model for electricity distribution, and compares it to an unconditional model, as well as the model presently used by the Norwegian regulator. A dataset of 123 Norwegian electricity distribution firms is used to illustrate how managerial inefficiency can be estimated in a meaningful way by comparing among firms operating in comparable environments, i.e., a homogeneous technology. The proposed model is used to compare the effect of conditioning on total efficient costs and revenues, firms' rate of return, and customers' prices. Based on the results, we observe that the use of conditional benchmarking methods in revenue cap regulation may lead not only to a decrease in aggregate efficient costs, but more importantly to a reallocation effect that affects the relative profitability of firms and relative customer prices, and may provide a fairer basis for setting revenue caps. This insight is relevant when using benchmarks for revenue cap regulation, and, more generally, when nonparametric benchmarking is used to compare decision making units in different operational environments.

Acknowledgements

This paper is produced as part of the project Municipal Infrastructure Companies against the Background of Energy Policy and Demographic Change (KOMIED) funded by the Leibniz Commission (SAW-2013-DIW-5), as well as the project Benchmarking for Regulation of Norwegian Electricity Networks (ElBench) at the Centre for Applied Research at NHH (SNF). The paper has been presented at the European Workshop on Productivity and Efficiency Analysis (EWEPA), 2015, and the INFORMS Annual Meeting, 2015. We gratefully ackknowledge Pio Baake, Tomaso Duso and Christian von Hirschhausen for valuable comments and suggestions. We further thank Ann Steward for providing language help and proof reading the article. The usual disclaimer applies.

References

Agrell, P. J. and Bogetoft, P. (2013). Benchmarking and regulation. *CORE Discussion Paper, Louvain-la-Neuve, Belgium*, 2013/8.

Agrell, P. J., Bogetoft, P., Koller, M., and Trinkner, U. (2014). Effizienzvergleich für Verteilernetzbetreiber Strom 2013, Bericht im Auftrag der Bundesnetzagentur (bnetza). Technical report, Swiss Economics and Sumicsid.

Amundsveen, R., Kordahl, O.-P., Kvile, H. M., and Langset, T. (2014). Second stage adjustment for firm heterogeneity in dea: A novel approach used in regulation of norwegian electricity dsos. *Recent Developments in Data Envelopment Analysis and its Applications*, page 334.

Amundsveen, R. and Kvile, H. M. (2015). The development and application of an incentive regulation model - a balancing act. *The ICER Chronicle*, Edition 3.

Armstrong, M. and Sappington, D. E. M. (2007). Recent developments in the theory of regulation. In Armstrong, M. and Porter, R., editors, *Handbook of Industrial Organization*, volume 3, chapter 27, pages 1557–1700. North Holland/ Elsevier: Amsterdam, The Netherlands.

Bădin, L., Daraio, C., and Simar, L. (2012). How to measure the impact of environmental factors in a nonparametric production model. *European Journal of Operational Research*, 223(3):818 – 833.

Bjørndal, E., Bjørndal, M., and Fange, K.-A. (2010). Benchmarking in regulation of electricity networks in Norway: An overview. In Bjørndal, E., Bjørndal, M., Pardalos, P. M., and Rönnqvist, M., editors, *Energy, Natural Resources and Environmental Economics*, Energy Systems, pages 317–342. Springer Berlin Heidelberg.

Bogetoft, P. (1997). Dea-based yardstick competition: the optimality of best practice regulation. *Annals of Operations Research*, 73:277–298.

Bogetoft, P. and Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. Springer-Verlag New York.

Bădin, L., Daraio, C., and Simar, L. (2010). Optimal bandwidth selection for conditional efficiency measures: A data-driven approach. *European Journal of Operational Research*, 201(2):633–640.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1):1 – 25.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444.

Daraio, C. and Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity Analysis*, 24(1):93–121.

Daraio, C. and Simar, L. (2007a). *Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications*. Studies in productivity and efficiency. Springer, New York.

Daraio, C. and Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28(1-2):13–32.

Debreu, G. (1959). *Theory of Value.* Cowles Foundation Monograph. Yale University Press, New Haven, Connecticut, U. S. A.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3):253–290.

Hall, P., Racine, J., and Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association*, 99(468):1015–1026.

Jeong, S.-O., Park, B., and Simar, L. (2010). Nonparametric conditional efficiency measures: asymptotic properties. *Annals of Operations Research*, 173(1):105–122.

Joskow, P. (2007). Regulation of natural monopoly. In A., M. P. and Shavell, S., editors, *Handbook of Law and Economics*, volume 2, chapter 16, pages 1227–1348. Elsevier, 1 edition.

Kneip, A., Simar, L., and Wilson, P. W. (2008). Asymptotics and consistent bootstraps for dea estimators in nonparametric frontier models. *Econometric Theory*, 24:1663–1697.

Koopmans, T. (1957). *Three Essays on the State of Economic Science*, chapter 1. McGraw-Hill, New York, U. S. A.

Laffont, J.-J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation.* The MIT Press.

Li, Q. and Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice.* Princeton University Press, Princeton.

Li, Q. and Racine, J. S. (2008). Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data. *Journal of Business and Economic Statistics*, 26(4):423–434.

Shleifer, A. (1985). A theroy of yardstick competition. *The RAND Journal of Economics*, 16(3):319–327.

Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis*, 20(3):391–424.

Simar, L. and Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1):49–61.

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64.

Simar, L. and Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In Fried, H. O., Lovell, C. K., and Schmidt, S. S., editors, *The measurement of productive efficiency and productivity growth*, pages 421–521. Oxford University Press, Oxford.

Simar, L. and Wilson, P. W. (2011). Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Produ*, 36(1):33–53.

A Appendix

A.1 Unconditional DEA

The main idea is to estimate the unknown technology or production set Ψ from a given sample of observed production units $i = 1, ..., n$. The set Ψ contains all input-output-combinations that are feasible to the production units and is defined as

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

with $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ being the vectors of inputs and outputs. The boundary of Ψ is referred to as frontier and serves as the benchmark against each individual observation is compared to in order to determine its efficiency. The production set assuming constant returns to scale (CRS), as proposed by Charnes et al. (1978), is defined as

$$\hat{\Psi}_{CRS} = \{(x, y) \in \mathbb{R}_+^{p+q} \mid y \leq \sum\nolimits_{i=1}^{n} \lambda_i y_i,\ x \geq \sum\nolimits_{i=1}^{n} \lambda_i x_i, \\ \text{and } \lambda_i \geq 0 \text{ for } i = 1, ..., n\}. \tag{10}$$

Using $\hat{\Psi}_{CRS}$, the level of inefficiency $\hat{\theta}_i$ for firm i can be estimated by solving the following linear program

$$\hat{\theta}_i = \min\{\theta \mid \theta x_i \geq \sum\nolimits_{j=1}^{n} \lambda_j x_j,\ y_i \leq \sum\nolimits_{j=1}^{n} \lambda_j y_j, \\ \text{and } \lambda_j \geq 0 \text{ for } j = 1, ..., n\}. \tag{11}$$

The efficiency estimate $\hat{\theta}_i$ is the Debreu-Farrell input efficiency score and indicates how much the unit of interest can reduce its inputs for a given level of output. If $\hat{\theta}_i = 1$ the observation i is considered fully efficient while $\hat{\theta}_i < 1$ implies inefficiency.

A.2 Reference sets and peers in different models

Table 7: Sizes of reference sets and identical and different peers in unconditional and conditional models.

Firm	Ref. set uncond.	Number of peers uncond.	Ref. set cond.	Number of peers cond.	Number of identical peers	Number of peers in uncond./ not in cond.	Number of peers in cond./ not in uncond.
1	123	35	58	19	16	19	3
2	123	37	61	15	14	23	1
3	123	26	63	12	12	14	0
4	123	38	60	14	11	27	3
5	123	32	46	11	10	22	1
6	123	24	59	7	7	17	0
7	123	23	28	4	4	19	0
8	123	37	84	31	22	15	9
9	123	32	58	9	8	24	1
10	123	34	58	18	15	19	3
11	123	38	60	16	9	29	7
12	123	31	63	22	11	20	11
13	123	35	59	17	13	22	4
14	123	38	81	31	21	17	10
15	123	37	15	4	4	33	0
16	123	37	82	27	20	17	7
17	123	37	73	22	16	21	6
18	123	38	51	15	12	26	3
19	123	37	67	17	14	23	3
20	123	27	53	11	11	16	0
21	123	36	61	14	13	23	1
22	123	33	59	16	13	20	3
23	123	37	73	20	16	21	4
24	123	38	60	16	11	27	5
25	123	29	19	5	5	24	0
26	123	38	53	18	14	24	4
27	123	37	71	23	15	22	8
28	123	36	8	1	0	36	1
29	123	36	7	2	1	35	1
30	123	33	71	20	14	19	6
31	123	30	8	1	1	29	0

Table 7: Sizes of reference sets and identical and different peers in unconditional and conditional models.

Firm	Ref. set uncond.	Number of peers uncond.	Ref. set cond.	Number of peers cond.	Number of identical peers	Number of peers in uncond./ not in cond.	Number of peers in cond./ not in uncond.
32	123	39	67	21	17	22	4
33	123	35	84	27	21	14	6
34	123	28	46	9	8	20	1
35	123	25	50	10	10	15	0
36	123	39	84	23	19	20	4
37	123	37	68	23	17	20	6
38	123	33	58	17	12	21	5
39	123	38	78	25	19	19	6
40	123	35	73	16	14	21	2
41	123	14	51	5	5	9	0
42	123	33	49	12	11	22	1
43	123	24	60	10	9	15	1
44	123	18	47	5	5	13	0
45	123	31	36	11	8	23	3
46	123	31	59	18	14	17	4
47	123	35	63	16	14	21	2
48	123	35	98	34	23	12	11
49	123	36	92	26	21	15	5
50	123	27	61	13	11	16	2
51	123	36	76	29	21	15	8
52	123	33	65	15	14	19	1
53	123	24	60	13	11	13	2
54	123	37	74	15	14	23	1
55	123	35	15	3	2	33	1
56	123	23	47	12	9	14	3
57	123	34	47	11	8	26	3
58	123	38	65	21	16	22	5
59	123	40	81	33	25	15	8
60	123	36	98	28	25	11	3
61	123	27	64	13	12	15	1
62	123	36	106	37	25	11	12
63	123	42	104	34	30	12	4
64	123	37	105	39	27	10	12

Table 7: Sizes of reference sets and identical and different peers in unconditional and conditional models.

Firm	Ref. set uncond.	Number of peers uncond.	Ref. set cond.	Number of peers cond.	Number of identical peers	Number of peers in uncond./ not in cond.	Number of peers in cond./ not in uncond.
65	123	39	63	21	15	24	6
66	123	34	65	10	8	26	2
67	123	38	68	21	14	24	7
68	123	26	57	14	9	17	5
69	123	29	70	19	11	18	8
70	123	32	56	13	12	20	1
71	123	32	39	10	6	26	4
72	123	36	90	35	22	14	13
73	123	40	61	19	15	25	4
74	123	17	53	10	9	8	1
75	123	39	65	19	15	24	4
76	123	36	4	3	2	34	1
77	123	36	79	27	21	15	6
78	123	37	91	31	23	14	8
79	123	25	65	13	11	14	2
80	123	33	54	12	10	23	2
81	123	33	50	6	5	28	1
82	123	41	67	26	17	24	9
83	123	37	46	12	10	27	2
84	123	35	28	11	4	31	7
85	123	31	56	15	6	25	9
86	123	35	69	21	12	23	9
87	123	34	47	9	8	26	1
88	123	37	81	29	19	18	10
89	123	40	93	30	25	15	5
90	123	38	73	19	15	23	4
91	123	38	62	18	15	23	3
92	123	32	56	8	7	25	1
93	123	35	41	13	5	30	8
94	123	36	60	14	12	24	2
95	123	32	88	26	15	17	11
96	123	25	56	15	8	17	7
97	123	34	62	16	11	23	5

Table 7: Sizes of reference sets and identical and different peers in unconditional and conditional models.

Firm	Ref. set uncond.	Number of peers uncond.	Ref. set cond.	Number of peers cond.	Number of identical peers	Number of peers in uncond./ not in cond.	Number of peers in cond./ not in uncond.
98	123	38	88	31	22	16	9
99	123	37	60	18	16	21	2
100	123	37	65	20	14	23	6
101	123	34	76	21	15	19	6
102	123	29	24	10	5	24	5
103	123	36	44	12	8	28	4
104	123	39	66	21	17	22	4
105	123	33	70	15	10	23	5
106	123	31	64	20	13	18	7
107	123	35	44	12	10	25	2
108	123	37	61	17	15	22	2
109	123	34	50	11	8	26	3
110	123	37	81	30	21	16	9
111	123	36	25	11	4	32	7
112	123	37	97	28	24	13	4
113	123	34	43	13	10	24	3
114	123	23	13	5	5	18	0
115	123	38	62	17	13	25	4
116	123	38	61	19	15	23	4
117	123	34	62	19	13	21	6
118	123	38	58	19	14	24	5
119	123	30	59	17	12	18	5
120	123	13	16	2	2	11	0
121	123	32	28	9	6	26	3
122	123	34	91	27	19	15	8
123	123	29	76	21	16	13	5
Average		34	60	17	13	21	4
Min		13	4	1	0	8	0
Max		42	106	39	30	36	13