UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Benchmarking Heterogeneous Distribution System Operators: Evidence from Norway

George Elias

16-06

March 2016

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Benchmarking Heterogeneous Distribution System Operators: Evidence from Norway

George Elias*

Abstract

Regulatory authorities in the European electricity sector use benchmarking techniques to determine the cost-efficient production level for an incentive regulation of distribution system operators (DSOs). With nearly 900 DSOs operating in the German electricity sector, of which 200 subject to incentive regulation, the issue of heterogeneity of DSOs has to be addressed. Using publicly available data of 133 Norwegian DSOs and replicating the model employed by the German regulator (who refuses access to the data), I show its assumption of homogeneous technology cannot be maintained. Quantile regressions (QR) across the cost distribution reveal heterogeneity in the coefficients of the explanatory variables, resulting in biased efficiency scores derived from stochastic frontier analysis. To correct for this heterogeneity in coefficients, I propose a Bayesian estimation of a more flexible SFA with latent classes for selected parameters that reflect variation in technologies. This estimation has better goodness of fit, reduced variance of all coefficients, and higher efficiency scores for nearly all DSOs, compared to the conventional alternative.

Keywords: Efficiency measurement, cost function, incentive regulation, electricity sector
JEL: C11, C21, D24, L94

* Corresponding author: George Elias, Phone: +41 (0)78 873 83 90, E-mail: george.elias@polynomics.ch

1 Introduction

Regulatory authorities in the electricity sector across many European countries employ benchmarking techniques to determine the cost-efficient production levels of distribution system operators (DSOs) in a quest to increase efficiency. The ensuing cost or profit reduction targets have direct and often severe consequences for the regulated DSOs. In Germany, benchmarking of DSOs for both electricity and gas started in 2008, using revenue caps and individual efficiency scores. The cost-weighted average efficiency score of the DSOs in the first regulatory period amount to 96 (electricity) and 90 (gas) percent, respectively, suggesting excess network costs of hundreds of million of Euros to be cut within two regulatory periods (ten and nine years, respectively). The magnitude of this regulatory impact calls for an inquiry into the benchmarking methods used to determine individual efficiency scores. In fact, stochastic frontier analysis (SFA) used by the German authorities requires strong assumptions with respect to stochastic influences as well as technological homogeneity.

Manski (2007) proposes the law of decreasing credibility stating that the credibility of inference decreases with the restrictiveness of the assumptions on which is is based. According to him, the problem is researchers' fixation on the identification of parameters for point estimation, causing them to impose assumptions that suffice to obtain an exact value. He further notes that conventional parametric methods are strong on identifying power but weak on substantive foundations, especially when it comes to missing data.

In benchmarking, the problem usually is not that of identification because of missing data, since the regulatory authority requests all firms to supply complete data sets. This results in a full sample of national firms that are comparable with respect to technological and legal characteristics, permitting to focus on the methodological assumptions made rather than on coping with missing data. While this paper is motivated by the German regulation of electricity DSOs, it analyses publicly available Norwegian data. The German regulatory authority («Bundesnetzagentur», BNetzA) refuses to provide firm-specific data on the ground of privacy concerns. As according to energy experts of PricewaterhouseCoopers, not even aggregate cost statistics of the regulated German DSOs are available, it is not directly possible to compare Norwegian and German DSOs w.r.t. cost ratios. However, evidence from Norway should be relevant for Germany for at least two reasons. First, the only statistic published by the BNetzA, electricity prices for household customers, are very similar for Germany and Norway at around 0.15 US$ per kWh at 2005 exchange rates[1]. Second, the definitions of the relevant cost and structural parameters to be used in the benchmarking process are comparable as stipulated by the Norwegian and German regulatory authority.

Quantile regressions (QR) are used to check for technological heterogeneity of Norwegian DSOs. A Bayesian SFA framework with parameter-specific latent classes of DSOs is then proposed that allows for variation in coefficients.

Section 2 explains the problem of heterogeneity in benchmarking methods. Section 3 presents the data and the model specification used in the analysis. Section 4 expands conventional latent class SFA, and quantile regression. In Section 5, results are presented and discussed. Section 6 concludes.

2 The problem of heterogeneity in benchmarking electricity and gas utilities

Non-observable heterogeneity in the data can influence efficiency scores in a benchmarking in two ways. First, there may be separable heterogeneity. This does not bias estimated coefficients of the minimum cost function and can be captured through the specification of a stochastic component. Second, non-separable heterogeneity does affect the coefficients (see e.g. Cullmann et al. (2009)). Heterogeneity therefore calls for care in the design and implementation of benchmarking if DSOs are numerous and variable in size, as is the case in Norway (see Table 1) and in Germany. The problem is exacerbated by the fact that there is no general agreement as to the choice of method in the estimation of the efficient frontier. Accordingly, regulatory authorities use different approaches ranging from data envelopment analysis (DEA) and variations of ordinary least squares (corrected, COLS; modified, MOLS) to stochastic frontier analysis (SFA). Additionally, the variables potentially characterized by heterogeneity need to be selected carefully in order to separate inefficiency from exogenous heterogeneity. By assumption the DSOs can influence cost for network construction and maintenance. Given technology and factor prices, failure to minimize cost reflects inefficiency. By way of contrast, the technology used by the DSO can be (at least in the short run), be regarded as exogenous, a view adopted e.g. in the German ordinance on incentive regulation for energy distribution (Bundesregierung (2012) § 13). That

[1] Eurostat (2013)

is, the parameters of the cost function reflect the structure of the network operated by the DSO, subject to long-term path dependence. Failure to account for technological heterogeneity may therefore cause misrepresentation of a DSO's cost efficiency. However, coefficient heterogeneity might still reflect inefficiency because a DSO is slow to adopt a least-cost technology. Therefore, certain parameters could be fixed on values reflecting efficient provision. Finally, heterogeneity in both energy output and input variables may be the consequence of inefficiency. On the output side, some DSOs may be lagging behind in their implementation of structural change that is desirable from the regulator's perspective (e.g., a shift from peak load to base load provision). On the input side, the regulator might prefer an increased use of renewables such as wind or solar energy. Other input or control variables such as the population in the service area arguably amounts to exogenous heterogeneity; for them, there is no benchmark value associated with cost efficiency.

There is a vast empirical literature on the technological heterogeneity in DSOs (see e.g. Greene (2005)). However, it presupposes the availability of panel data which permit tracking a DSO over several periods of time. In the case of Norway, and several other regulated countries, panel data for longer time periods are simply not available. Moreover, the structure of the DSO tends to undergo changes over the years through network mergers or remunicipalisations. Due to these problems with panel data, benchmarking methods based on panel data are currently to my knowledge not applied by European regulatory authorities. Finally, cost drivers often cannot be determined by statistical methods but do have a political or ecological dimension.[2] Therefore, the analysis needs to detect exogenous heterogeneity even if only cross sectional data is available.

3 Model specification and data

Starting in late 2013, the regulatory regime for German DSOs will be modified. The parameters for the benchmarking model of the second regulatory period are not yet know. Cost adjusters mandatorily required by the German ordinance on incentive regulation for energy distribution (Bundesregierung (2012) § 13) include network length, volume of peak load, number of connection points, geographical characteristics of the service area. The formula finally used in the first regulatory period (starting in 2008) contained network length by means of five categories, high

[2] For instance, the government may impose a number of connection points that is currently excessive but facilitates accommodation of an increased network load. Ecological concerns may call for too much decentralization of power generation in the interest of minimizing the need for transmission lines.

and medium voltage, each weighted by length of lines and cables, and the sum of low-voltage line length. Furthermore, peak load volume was split into its high/medium and medium/low components. Additionally, the load contribution of renewables and the number of transformer stations were included, resulting in a model with 11 independent variables. Experience from the already ongoing second regulatory period (starting in 2012) in the gas sector shows that the BNetzA changed the formula considerably compared to the one applied in the first regulatory period, building again on the mandatory cost parameters. As a substitute to the German data, I use Norwegian ones. In Norway, incentive regulation with a price cap was introduced in 1997 for electricity DSOs. It is revised every five years. Individual efficiency scores were derived from a DEA benchmarking, with dimensions of output being network length, number of transformers, energy supplied, and number of customers. Clearly, with its four cost drivers the Norwegian benchmarking formula is much simpler that its German counterpart of the first regulatory period with 11 cost drivers. Therefore, chance are great to discover heterogeneity in the data. The German formula, by way of contrast, controls for many sources of heterogeneity, leaving little scope for discovering hidden heterogeneity. The associated problems thus are likely overstated for the case of Germany. Still, the analysis of the Norwegian data is worthwhile because the relative importance of the different types of heterogeneity (see Section 2 again) is likely to be comparable between the two countries. Further, the German formula is likely to be changed in the second regulatory period, requiring a specification based on the four mandatory parameters. The sample contains data on 133 Norwegian DSOs, averaged over the available years from 2004 to 2006 in order reduce the sensitivity to extreme values in one year. This data has been used in the actual benchmarking process by the Norwegian regulatory authority. Outlier detection using Cook's distance on an OLS regression (see Cook (1979)) and separate treatment of two DSOs with special characteristics led to the exclusion of five outliers. This choice seems to be the same as the one by the Norwegian regulator for benchmarking electric DSOs. The dependent variable is total cost in TNOK, which depends on energy delivered in MWh, the number of customers, the length of lines in km and the number of transformers. As the energy delivered highly correlated with the number of customers causing multicollinearity, only the number of customers is used.

Table 1 offers descriptive statistics for both the dependent and explanatory variables. The data suggest that the Norwegian power distributors are characterized by a great degree of heterogeneity. Network cost (total expenditures, totex, i.e. operating and capital costs) differ by a factor of 2,330 in Norway between the minimum and maximum value; the number of customers by a factor of 1,860.[3]

Table 1: Descriptive Statistics, Norwegian power distributors[1)]

Variable	Definition	Mean	Std. dev.	Min	Max
c	Network cost (Totex), in thousands of US$[4)]	12,231	25,188	88	203,140
nl	Network length in km	747	1,277	31	8,242
tr	Transformers	937	1,800	29	13,245
fc	Final customers[2)3)]	18,612	52,790	275	510,038

[1)] 128 observations after exclusion of outliers, average values of 2004 to 2006
[2)] Excluding secondary houses
[3)] Used as standardisation variable
[4)] 1 NOK (Norwegian Crown) = 0.15 US$ at 2005 exchange rates; 1 Euro = 1.3 US$ at 2005 exchange rates
Source: Norwegian Water Resources and Energy Directorate

The cost function to be estimated reads,

$$c_i \quad = \quad \alpha * fc_i + \beta_1 * nl_i + \beta_2 * tr_i, \tag{1}$$

with subscript $i = 1, ..., N$ representing DSOs. All terms are divided by the number of final customers, fc_i, making c_i the average cost per final customer (fc_i). Although average cost might well depend on fc_i, this variable is dropped from the equation in order to reflect the standardization applied by the German regulatory authority.[4] Note that equation (1) does not qualify as a minimum cost function because it does not contain relative factor prices. However, capital user cost do not enter the income statements of either Norwegian and German DSOs, while the wage rate is practically the same within a given country.

[3] It can be assumed that the German electricity distribution sector with its roughly 800 DSO is characterised by a similar degree of heterogeneity.

[4] This approach is not beyond controversy, as the alternative of a double-log function is more established for accounting differences in the size of organizations.

4 Stochastic Frontier Analysis, the Latent Class Extension, and Quantile Regressions

4.1 Stochastic Frontier Analysis (SFA)

The parametric benchmarking model used by the Germany regulatory authority is stochastic frontier analysis.[5] The cost frontier is calculated using a maximum likelihood estimation (MLE) and can be written (dropping the time subscript, since only a cross section is used) as

$$c_i \;=\; \alpha + \beta' X_i + \overbrace{u_i + v_i}^{\varepsilon_i}, \quad \text{with} \;\; u_i = N^{+}(0, \sigma_u^2) \quad \text{and} \;\; v_i = N(0, \sigma_v^2). \tag{2}$$

For Bayesian inference (the preferred approach of this paper), the crucial property of c_i is that it is normally distributed,

$$c_i \;\sim\; N(\alpha + \beta' X_i + u_i, \sigma^2), \tag{3}$$

where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2. The error term ε_i in (2) is divided into two additive components, random noise v_i and cost inefficiency u_i. Random noise captures separable heterogeneity; It is normally distributed $v_i \overset{iid}{\sim} N[0, \sigma_v^2]$ with mean zero and variance σ_v^2. DSO-specific inefficiency u_i is assumed to follow a half-normal distribution with support $[0, \infty]$, with larger values indicating higher cost inefficiency.[6] The compound error term ε_i has to have positive skewness to indicate the existence of inefficiency.[7]

In theory, panel data models[8] can capture some unobserved heterogeneity by tracking inefficiency over time. However, they require assumptions regarding the change of inefficiency. For instance, if the industry under consideration is a regulated monopoly, part of the inefficiency may be time invariant but another part arguably is time variant since incentives to minimize costs are weakened (see e. g. Kopsakangas-Savolainen and Svento (2011)). Moreover, panel data estimation

[5] Kumbhakar and Lovell (2000) offer an introduction to stochastic frontier models.

[6] Other possible distributions for the inefficiency term are truncated-normal, exponential or gamma. Exponential distributions were assumed in the model of the second regulatory period of gas DSOs in Germany, whereas half-normal distributions were used in the first regulatory periods of both electricity and gas.

[7] If the residuals are not skewed in the positive direction, the SFA model is reduced to an OLS regression, implying zero inefficiency (see Kumbhakar and Lovell (2000)).

[8] Panel data models are for example applied by Farsi and Filippini (2004) to analyse heterogeneity in Swiss DSOs.

is complex and is based upon additional assumptions that often cannot be tested.[9] Therefore, an alternative suitable for dealing with cross-sectional data is pursued here.

4.2 Latent Class SFA

The latent class SFA extends (3) by

$$c_i \sim N(\alpha + \beta_j' X_i + u_i|_j, \sigma^2), \tag{4}$$

where now the uniform vector of parameters β is replaced by a set of vectors β_j and u_i by a vector $u_i|_j$ of inefficiency terms conditional on latent classes. A particular DSO may be assigned to a class with probability one (the classical case) or a probability between zero and one (the generalized case). The latter case implies that no single technology is assumed to be appropriate for all DSOs. Rather, it admits of a range of possible coefficients and thus efficiency scores for a particular DSO.

Class membership probabilities $p_{(i|j)}$ can be parametrized by a multinomial logit model

$$p_{(i|j)} = \frac{exp(\delta_j' q_i)}{\sum_{j=1}^{J} exp(\delta_j' q_i)}, \tag{5}$$

where q_i is a vector of firm-specific variables and δ_j its corresponding coefficients determining class membership. A variant is not to specify any regressors q_i, letting class membership be determined by maximum likelihood only. This amounts to assuming a set of values for δ_j, calculating class probabilities as in (5) with $q_i = 1 \forall i$, and estimating (4). This procedure results in a set of efficiency scores for each DSO. There are two ways to choose a score from this set. One is to retain the one with the highest posterior membership probability. In the absence of credible prior information, the uniform or Bernoulli distribution can be used as the prior,

$$\hat{p}\_(i|j) \sim dbern(\frac{1}{j}), \tag{6}$$

with $\hat{p}\_(i|j)$ denoting prior probabilities. The other alternative is to estimate an expected value of $p_{(i|j)}$ along with α and β_j from (5).

[9] See Shuttleworth (2005) for a general critique and Greene (2005) for applications of panel data models.

This framework requires the number of classes to be determined a priori. In contrast to traditional latent class modeling, which is a panel data method, the availability of cross-sectional data only calls for a more restrictive approach here. Heterogeneity will be admitted only in those parameters where differences in the technology are suggested by a continuous quantile regression and where differences in technology are acknowledged by the (German) regulator. To my knowledge, this approach has not been applied before.

4.3 Quantile Regression

Koenker (1978) defines the quantile regression problem as follows

$$\min_{\beta\in\Re} \left[\sum_{i\epsilon\{i:c_i\geqslant\beta' X_i\}} \theta|c_i - \beta' X_i| \quad + \sum_{i\epsilon\{i:c_i<\beta' X_i\}} (1-\theta)|c_i - \beta' X_i| \right], \tag{7}$$

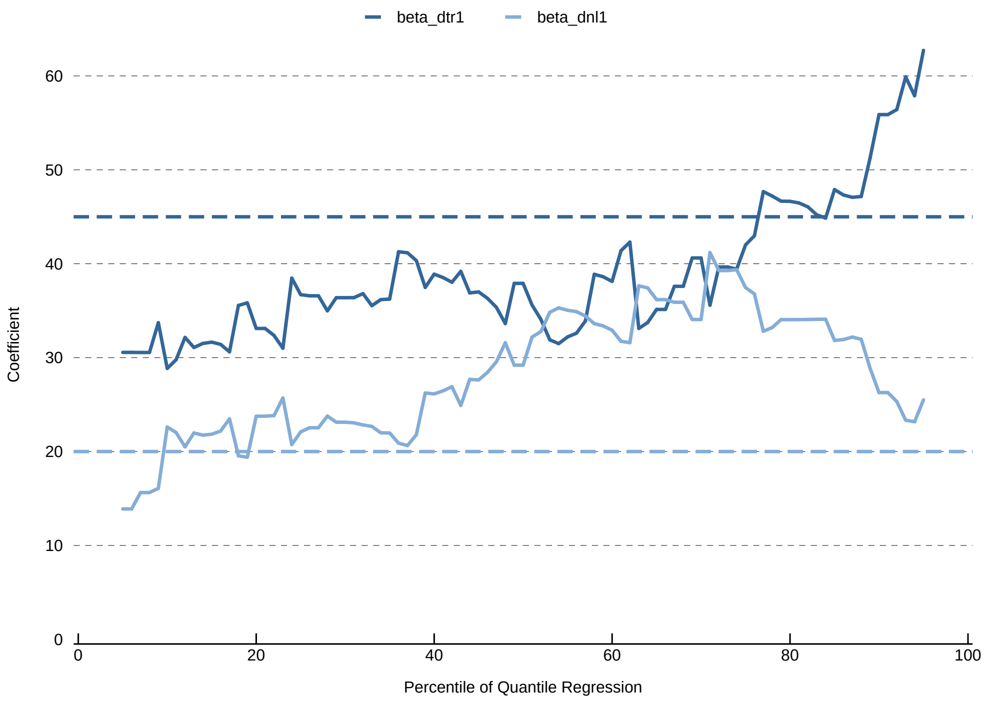
where θ symbolises the quantile to be estimated, with $\theta\epsilon(0,1)$. In particular, $\theta = 0.5$ fits a median regression. Residuals are calculated as $u_i = c_i - \beta' X_i$. The method minimizes the weighted sum of absolute residuals rather than the sum of the squared residuals as OLS. It is therefore less sensitive to outliers. Estimation of (7) for a set of θ values allows the calculation of coefficients at different points of the distribution of c_i, facilitating testing for parameter homogeneity.
Similar to Christensen (2004), I calculate standard errors for β using bootstrapping[10] to test whether coefficients differ by quantile, which are defined by values for θ between .05 and .95 and a percentile grid. In this way, the cost distribution provides an indication of technological heterogeneity.

5 Results

Figure 1 shows the values of two β coefficients as a function of θ as defined below eq. (7). The dark graph corresponds to the cost impact of the number of transformers per final customer (tr), the light one to the cost impact of network length per final customer (nl).
The standard SFA method results in coefficients for standardised transformers and network length and of 46 and 21, respectively, indicated by the dark and light dashed horizontal lines of figure 1. Visual inspection already suggests that these uniform estimates fail to do justice to

[10] More details on this can be found in Rogers (1993).

Figure 1: Coefficients of Transformers and Network Length (Standardized) at Different Cost Quantiles



parameter heterogeneity likely reflecting technological heterogeneity of DSOs. To test whether the differences are significant, I compare the respective coefficients at the values $\theta = .10$ (a percentile commonly used for cost frontiers in quantile regressions) and $\theta = .50$ (the median regression). A Wald test rejects the hypothesis of pairwise equality with a two-sided p-value of 0.036. Therefore, in order to obtain unbiased coefficients and hence efficiency scores, it is crucial to account for parameter heterogeneity, at least in the case of Norwegian DSOs.

To illustrate the importance of technological heterogeneity, I allow two latent classes for standardised transformers while allowing only one for standardised network length. For simplicity, class membership is either 0 or 1 [that is, computed as in (6)]. To obtain posterior estimates, Monte Carlo Markov Chain (MCMC) algorithms were run for 50,000 iterations, with the first 20,000 samples discarded as burn-in. Table 2 compares the coefficients pertaining to the standard SFA and those pertaining to a SFA with two latent classes for the case of the standardised number of transformers.

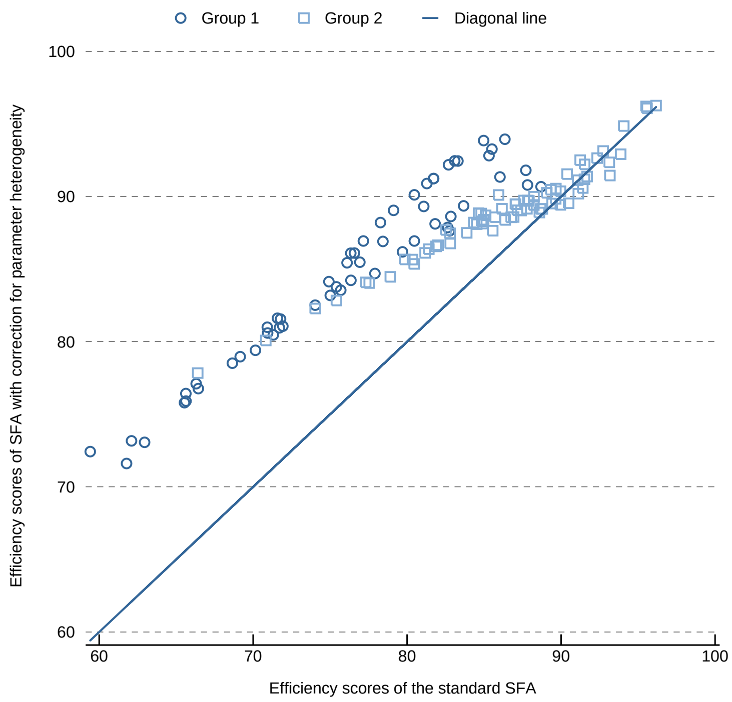
Table 2: Comparison of SFA and Latent Class SFA

Variables	**SFA**				**SFA latent class**			
	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
Constant	0.28	0.35	-0.43	0.97	0.81	0.30	0.19	1.41
network length (standardised)	20.73	6.22	8.54	33.59	23.50	5.23	13.09	33.88
transformers (standardised)	46.38	7.43	31.77	61.10	-	-	-	-
transformers latent class no. 1[1)]	-	-	-	-	32.92	6.53	20.05	45.53
transformers latent class no. 2[2)]	-	-	-	-	52.18	6.56	39.03	64.12
Efficiency scores	81.96	8.22			87.15	5.13		
DIC	393.9				343.8			
Obs.	128				128			

[1)] Latent class of 57 DSOs [2)] Latent class of 71 DSOs

In the case of two latent classes, 57 DSOs are assigned to class no. 1 (low cost) and 71 DSO to class no. 2 (high cost). The results are striking. While the constant lacks statistical significance in the conventional SFA, it is clearly significant in the latent class alternative. This implies that conventional SFA yields too low cost efficiency scores to begin with. Next, standardized network length has too low cost impact of 20.73 rather than 23.50 if estimated using conventional SFA; the difference is again statistically significant. This means that DSOs with customers far away from generating sites are punished by low efficiency scores if this technological peculiarity is not taken into account in the cost function. Of course, the major difference is observed with regard to the standardized number of transformers. According to the left-hand side of Table 2, their cost impact amounts to NOK 46 per 1,000 units. However, admitting two latent classes results in an estimate of only NOK 33 for the low-cost category but NOK 52 for the high-cost one. Since the 95% confidence intervals of the two estimates overlap somewhat, this difference fails to reach statistical significance. Still, there are several indications to the effect that latent class SFA estimation is preferable to the conventional one. First, standard deviations of all parameters are smaller. This is to be expected because the model is permitted to fit the observed data more closely. However, this gain is indeed significant, as indicated by the deviance information criterion (DIC), the Bayesian equivalent to AIC. A finding of great importance is that the mean efficiency score is 82 percent under conventional SFA but increases by a full 5 percentage points to 87 under SFA with just two latent classes for the number of transformers. As shown in figure 2, it is the DSOs with efficiency score below 90 if estimated by conventional SFA that benefit most, by up to 10 percentage points. Up to about 75 percent efficiency, both the low-cost class no. 1 and the high-cost class no. 2 experience this upward shift; beyond that value, all but the very high-ranked DSOs of class 1 continue to exhibit a somewhat reduced shift while those of class 2 approach equality between the two types of scores. In the case of Germany, this would correspond to an increase of allowed network costs in the range of millions of euros for bigger German DSOs.

Figure 2: Efficiency scores of the SFA and latent class SFA with 2 classes



6 Conclusion

Following the first regulatory period in Germany various regulated DSOs took court action to appeal against their individual efficiency score. Many of them argued that their specific structural situation is not adequately accounted for within the benchmarking process. One possible solution to account for heterogeneity in technologies is to apply panel data methods. However, due to data limitation issues and political requirements towards model specifications, these methods often are not feasible in a regulatory framework. I developed in this paper a flexible latent class framework for cross-sectional data, which allows for a range of coefficients in parameters where heterogeneity has been detected or structural change is desired. Results from Norwegian data show the importance of accounting for heterogeneity. The model presented in the paper could therefore offer an approach towards increased credibility of the results. As the data used by the German regulator is not publicly available, it is not clear if - but likely that - the problem of heterogeneity also is present in Germany. An appropriate analysis should therefore be conducted.

References

Bundesregierung (2012), Verordnung über die Anreizregulierung der Energieversorgungsnetze (Anreizregulierungsverordnung - ARegV).

Christensen, E. W. (2004), Scale and scope economies in nursing homes: a quantile regression approach., Health economics, 13(4):363–77.

Cook, D. R. (1979), Influential in Linear Observations Regression, Journal of the American Statistical Association, 74(365):169– 174.

Cullmann, A., M. Filippini and M. Farsi (2009), Unobserved Heterogeneity and International Benchmarking in Public Transport, CEPE Working paper series, 18(65).

Eurostat (2013), Electricity prices for household consumers.

Farsi, M. and M. Filippini (2004), Regulation and measuring cost-efficiency with panel data models: application to electricity distribution utilities, Review of Industrial Organization, 25(1):1–19.

Greene, W. (2005), Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, Journal of Econometrics, 126(2):269–303, doi:10.1016/j.jeconom.2004.05.003.

Koenker, R. (1978), Regression quantiles, Econometrica: Journal of the Econometric Society, 46(1):33–50.

Kopsakangas-Savolainen, M. and R. Svento (2011), Observed and unobserved heterogeneity in stochastic frontier models: An application to the electricity distribution industry, Energy Economics, 33(2):304–310.

Kumbhakar, S. C. and C. A. K. Lovell (2000), Stochastic Frontier Analysis, Cambridge University Press, Cambridge.

Manski, C. F. (2007), Identification for Prediction and Decision, vol. 45, Harvard University Press.

Rogers, W. (1993), Quantile regression standard errors, Stata Technical Bulletin, 2(9):133–137.

Shuttleworth, G. (2005), Benchmarking of electricity networks: Practical problems with its use for regulation, Utilities Policy, 13:310–317, doi:10.1016/j.jup.2005.01.002.