11/16

RAPPORT
REPORT

Evaluation of the StoNED Method for Benchmarking and Regulation of Norwegian Electricity Distribution Companies

Endre Bjørndal
Mette Bjørndal

SNF

Samfunns- og næringslivsforskning AS
Centre for Applied Research at NHH

SNF
SAMFUNNS- OG NÆRINGSLIVSFORSKNING AS

- er et selskap i NHH-miljøet med oppgave å initiere, organisere og utføre eksternfinansiert forskning. Norges Handelshøyskole og Stiftelsen SNF er aksjonærer. Virksomheten drives med basis i egen stab og fagmiljøene ved NHH.

SNF er ett av Norges ledende forskningsmiljø innen anvendt økonomisk-administrativ forskning, og har gode samarbeidsrelasjoner til andre forskningsmiljøer i Norge og utlandet. SNF utfører forskning og forskningsbaserte utredninger for sentrale beslutningstakere i privat og offentlig sektor. Forskningen organiseres i programmer og prosjekter av langsiktig og mer kortsiktig karakter. Alle publikasjoner er offentlig tilgjengelig.

SNF
CENTRE FOR APPLIED RESEARCH AT NHH

- is a company within the NHH group. Its objective is to initiate, organize and conduct externally financed research. The company shareholders are the Norwegian School of Economics (NHH) and the SNF Foundation. Research is carried out by SNF´s own staff as well as faculty members at NHH.

SNF is one of Norway´s leading research environment within applied economic administrative research. It has excellent working relations with other research environments in Norway as well as abroad. SNF conducts research and prepares research-based reports for major decision-makers both in the private and the public sector. Research is organized in programmes and projects on a long-term as well as a short-term basis. All our publications are publicly available.

SNF REPORT NO. 11/16

Evaluation of the StoNED Method for Benchmarking and Regulation of Norwegian Electricity Distribution Companies

Endre Bjørndal
Mette Bjørndal

SNF Project No. 5349: "Elbench 2016"

The project is financed by Skagerak Nett, Hafslund Nett, Eidsiva Nett, Lyse Elnett, BKK Nett, Energi Norge AS, NVE and Helgeland Kraft

CENTRE FOR APPLIED RESEARCH AT NHH
BERGEN, DECEMBER 2016

ISBN 978-82-491-0941-8 (Printed version)
ISBN 978-92-491-0942-5 (Online version)
ISSN 0803-4036

Abstract

We evaluate the StoNED methodology for benchmarking and regulation of network companies, and we compare StoNED to the two-stage DEA method currently used by the Norwegian regulator. We find that the estimated values for the skewness parameter in the second stage of the StoNED procedure can be inconsistent with the assumed positive skewness for the inefficiency term. Setting the skewness parameter to an arbitrary value can have significant consequences for the efficiency levels. This effect is partly neutralized by the revenue calibration performed by the NVE, depending on how the calibration is implemented. Our comparison of results from StoNED and two-stage DEA show that the efficiency scores from the two methods are highly correlated, but that the levels can differ significantly. We also interpret the StoNED coefficient estimates and compare them to the corresponding (dual) DEA estimates. Finally, we illustrate the robustness of the efficiency estimates to noise in data, as exemplified by noisy pension costs.

Contents

1 Introduction

The Norwegian electricity distribution companies have been subject to an incentive regulation scheme since 1997 (Agrell et al., 2005; Bjørndal et al., 2010; Amundsveen and Kvile, 2015). The annual revenue caps are determined based on the comparison of actual cost with a cost norm. The cost norm for each company is estimated based on the relative efficiency score obtained from Data Envelopment Analysis (DEA).

A problem with DEA-based models is that they are deterministic and do not account for noise in the data sets. In this report, commissioned by NVE, we will look at the StoNED methodology, introduced by Kuosmanen and Kortelainen (2012), as an alternative to the two-stage method. Like DEA, StoNED is non-parametric in nature, but it models noise explicitly. We will implement the StoNED model on data sets for the Norwegian electricity distribution companies, and we will make comparisons with variants of the two-stage model that is presently used by NVE.

Section 2 gives some more details on the regulation model applied by NVE, and in particular, the calibration mechanism that is used. Sections 3 and 4 describes the benchmarking models and the data that we will be using. In Section 5 we discuss some issues that need to be adressed when implementing StoNED, e.g., problems with negative skewness for the inefficiency term. Section 6 compares the StoNED results to results from the two-stage models. In Section 6.1 we compare efficiency scores, and in Section 6.2 we interpret and compare coefficient estimates, for output and geography factors. Finally, in Section 7, we discuss the robustness of the methods to noise in data, using pension costs as an illustrative case.

2 NVEs regulation scheme

In the present yardstick regulation model (Shleifer, 1985; Bogetoft, 1997), the revenue cap for firm i is set as

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

where $\alpha = 0.6$. The efficient costs C_i^* are calculated by $\theta_i C_i$, where θ_i is an estimate of company i's efficiency. This estimate is obtained via the two-stage DEA method described in Amundsveen et al. (2014).[1]

NVE calibrates the revenue caps, by adding the amount Δ_i to the efficient cost of each firm, 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. 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.

[1]The yardstick formula given by (1) 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. 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 $t-2$. In this report we assume that the average of the data for a five-year period 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.

Several calibration alternatives have beeen used since the introduction of the yardstick regulation in 2007. One alternative is allocate the industry revenue "shortfall", given by the difference $\sum C - \sum C^*$, relative to efficient costs, i.e.,

$$\Delta_i = \frac{\sum C - \sum C^*}{\sum C^*} C_i^*. \tag{2}$$

We will refer to this as calibration alternative A. In the present regulation model, the calibration takes the form

$$\Delta_i = \frac{\sum C - \sum C^*}{\sum BV} BV_i, \tag{3}$$

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. We will refer to this as calibration alternative B. A third method (C) that has been used by the regulator, is to allocate the calibrated amount relative to reported cost values, i.e.,

$$\Delta_i = \frac{\sum C - \sum C^*}{\sum C} C_i. \tag{4}$$

See Bjørndal et al. (2010) for a further discussion of the calibration methods A-C.

3 Benchmarking models

3.1 StoNED

Several variants of StoNED models have been explored. They can be models of production functions, as in Kuosmanen and Kortelainen (2012) and Johnson and Kuosmanen (2011), or cost functions, as in Kuosmanen (2012) and Kuosmanen et al. (2013). They can be additive with respect to the effect of inefficiency and noise, as in Kuosmanen and Kortelainen (2012), or multiplicative, as in Johnson and Kuosmanen (2011), Kuosmanen (2012), and Kuosmanen et al. (2013). Finally, they may include the effect of geography factors, as in Johnson and Kuosmanen (2011), Kuosmanen (2012), and Kuosmanen et al. (2013).

In the multiplicative StoNED cost model the assumed relationship between observed cost x_i and output $\boldsymbol{y}_i$ (a vector) for company i is

$$x_i = C(\boldsymbol{y}_i) e^{\boldsymbol{\delta z}_i + u_i + v_i}, \tag{5}$$

where $C(\cdot)$ is a convex function, and where u_i and v_i represent inefficiency and noise, respectively. The vector $\boldsymbol{\delta}$ represents the effect of geography factors, and $\boldsymbol{z}_i$ is the vector of geography variables. We assume, as in Kuosmanen and Kortelainen (2012) and Johnson and Kuosmanen (2011), that x and $\boldsymbol{y}$ are non-negative variables, whereas $\boldsymbol{z}$ can be positive or negative. Also, the noise terms v_i are assumed to follow a symmetric distribution with zero mean and a constant, finite variance σ_v^2, whereas the inefficiency terms u_i have an asymmetric distribution with a positive expected value μ and a finite variance σ_u^2. The multiplicative model can be rewritten as

$$\ln x_i = \ln C(\boldsymbol{y}_i) + \boldsymbol{\delta z}_i + u_i + v_i. \tag{6}$$

Alternatively, we could assume that the impact of inefficiency, noise, and the geography factors can be expressed in an additive manner, i.e., that

$$x_i = C(\boldsymbol{y}_i) + \boldsymbol{\delta z}_i + u_i + v_i. \tag{7}$$

According to the procedure presented in Kuosmanen and Kortelainen (2012) and Johnson and Kuosmanen (2011), the cost function and the parameters $\boldsymbol{\delta}$, μ, σ_u, and σ_v, are estimated in two stages:

1. Estimate the shape of the cost function, as well as the parameter vector $\boldsymbol{\delta}$, by using *Convex Nonparametric Least Squares (CNLS).*
2. Impose additional distributional assumptions about u_i and v_i and estimate values for the parameters μ, σ_u, and σ_v, using either the method of moments (Aigner et al. (1977)) or pseudolikelihood techniques (Fan et al. (1996)).

For the multiplicative model, the CNLS procedure in Stage 1 is done by solving the following optimization problem:

$$\min_{\gamma,\alpha,\beta,\delta,\epsilon} \sum_i \epsilon_i^2 \tag{8}$$

s.t.

$$\ln x_i = \ln \gamma_i + \sum_s \delta_s z_{si} + \epsilon_i \qquad \forall i \tag{9}$$

$$\gamma_i = \alpha_i + \sum_r \beta_{ri} y_{ri} \qquad \forall i \tag{10}$$

$$\gamma_i \geq \alpha_j + \sum_r \beta_{rj} y_{ri} \qquad \forall j, i \tag{11}$$

$$\beta_{ri} \geq 0 \qquad \forall r, i \tag{12}$$

The objective function (8) minimizes the sum of squared errors, where $\epsilon = u_i + v_i$. Equation (9) corresponds to the assumed cost relationship (6) in the multiplicative model, whereas (10)-(12) make sure that the cost function (10) is non-decreasing in the outputs (12), and that it is non-concave (11).

No restrictions on the sign of α means that variable returns to scale are assumed. Constant returns to scale can be imposed by adding (13). Non-decreasing or non-increasing returns to scale correspond to (14) or (15), respectively.

$$\alpha_i = 0 \qquad \forall i \tag{13}$$

$$\alpha_i \geq 0 \qquad \forall i \tag{14}$$

$$\alpha_i \leq 0 \qquad \forall i \tag{15}$$

The results from the StoNED model will be compared to NVEs current model (see Section 3.2), where constant returns to scale is assumed, and we will therefore make the same assumption in our StoNED models. I.e., we will set $\alpha_i = 0 \ \forall i$ in our analyses.

For Stage 2 of the StoNED procedure, we will only describe and use the method of moments, since it is simpler than the pseudolikelihood techniques. As in Kuosmanen and Kortelainen (2012) we assume that inefficiency and noise follow half-normal and normal distributions, respectively. Then, the parameters of the two distributions can be easily calculated as

$$\sigma_u = \sqrt[3]{\frac{M_3}{\left(\frac{4}{\pi} - 1\right)\sqrt{\frac{2}{\pi}}}}, \tag{16}$$

$$\mu = \sigma_u \sqrt{2/\pi}, \tag{17}$$

$$\sigma_v = \sqrt{M_2 - \left(\frac{\pi - 2}{\pi}\right)\sigma_u^2}, \tag{18}$$

where $M_2 = \sum_i (\epsilon_i - \bar{\epsilon})^2/n$ and $M_3 = \sum_i (\epsilon_i - \bar{\epsilon})^3/n$ are the estimated second and third moments for the distribution of the composite error terms.

The estimated cost norm for a given company i can now be calculated as

$$\hat{C}(\boldsymbol{y}_i, \boldsymbol{z}_i) = \left(\alpha_i + \sum_r \beta_{ri} y_{ri}\right) e^{-\mu + \sum_s \delta_s z_{si}}. \tag{19}$$

The term μ in (19) represents average (industry) inefficiency, and its role in the cost function formula is to shift from an average-practice to a best-practice cost function, analogous to the output-oriented model developed in Kuosmanen and Kortelainen (2012). The term $\sum_s \delta_s z_{si}$ adjusts the cost frontier for the effect of geography cost drivers, as suggested in Johnson and Kuosmanen (2011). An estimate of the cost efficiency for company i can now be obtained as

$$\hat{\theta}_i = \frac{\hat{C}(\boldsymbol{y}_i, \boldsymbol{z}_i)}{x_i}. \tag{20}$$

Alternatively, we can obtain cost efficiency estimates from the conditional mean formula developed by Jondrow et al. (1982),

$$\hat{u}_i = E(u_i|\epsilon_i) = \mu_i^* + \sigma^* \left[\frac{\phi(-\mu_i^*/\sigma^*)}{1 - \Phi(-\mu_i^*/\sigma^*)}\right], \tag{21}$$

where $\mu_i^* = \epsilon_i \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ and $\sigma^* = \sqrt{\sigma_u^2 \sigma_v^2/(\sigma_u^2 + \sigma_v^2)}$. The functions ϕ and Φ represent the standard normal density function and the standard normal cumulative distribution function, respectively. An alternative estimate of the cost efficiency for company i is

$$\hat{\theta}_i = e^{\hat{u}_i}. \tag{22}$$

3.2 Two-stage DEA

We will compare the results from the StoNED models to NVEs current benchmarking model. In Stage 1, the following DEA/CRS model is solved for each company i_0:

$$\min_{\lambda,\theta} \theta_i \tag{23}$$

s.t.

$$\theta_i x_i \geq \sum_{j=1}^{n} \lambda_j x_j \tag{24}$$

$$\sum_{j=1}^{n} \lambda_j y_{jr} \geq y_{ir} \qquad \forall r \tag{25}$$

$$\lambda_j \geq 0 \qquad \forall j \tag{26}$$

In Stage 2, the DEA efficiency scores are regressed against adjusted "relative values" of the geography variables, i.e., the following regression model is estimated:

$$\theta_i = \rho + \sum_s \varphi_s z_{si}^* + \varepsilon_i \qquad \forall i \tag{27}$$

The relative values for the geography factors equal the differences between the observed level for the company in question and the computed level for the corresponding reference company. Let λ_{ij} denote the weight of company j in the reference set of company i. Then $\psi_{ij} = \lambda_{ij}x_j / \sum_k \lambda_{ik}x_k$ is company j's share of the cost norm of company i, and we can compute the relative value of geography variable s for company i as $z^*_{si} = z_{si} - \sum_j \psi_{ij} z_{sj}$. The estimated coeffients from (27) are used to correct the efficiency scores:

$$\hat{\theta}_i = \theta_i - \sum_s \varphi_s z^*_{si} = \rho + \varepsilon_i \tag{28}$$

We will also evaluate some alternatives to the current methodology employed by NVE. One possible alternative might be to regress the DEA efficiency scores on the *absolute* level of the geography factors, i.e., replace z^*_{si} by z_{si} in (27). Another alternative is to regress the *logarithmic values* of the efficiency scores on the geography variables, i.e., use the following regression model:

$$\ln \theta_i = \rho + \sum_s \varphi_s z^*_{si} + \varepsilon_i \qquad \forall i. \tag{29}$$

Note that (29) is equivalent to

$$\theta_i = e^{\rho + \sum_s \varphi_s z^*_{si} + \varepsilon_i} \qquad \forall i, \tag{30}$$

and the adjusted efficiency scores can be calculated (analogous to (28)) from the following equation:

$$\hat{\theta}_i = \theta_i e^{-\sum_s \varphi_s z^*_{si}} = e^{\rho + \varepsilon_i}. \tag{31}$$

The geography adjustment in (31) is done in a multiplicative manner, similar to the assumption underlying the multiplicative StoNED cost function in (19).

3.3 Interpretation and comparison

Kuosmanen and Johnson (2010) have discussed the relationship between the StoNED and DEA models. They show that the DEA model is in fact a special case of the additive StoNED model. The output coefficients in StoNED are related to the shadow prices of the output constraints in the DEA model. In order to compare the StoNED and DEA coefficients in Section 6.2.1, we will briefly discuss some of the similarities between the StoNED model and the dual formulation of the DEA model.

Since we have only one input variable, the DEA model given by (23)-(26) can be simplified by combining (23) and (24) into

$$\min_{\lambda} \frac{\sum_j \lambda_j x_j}{x_i},$$

and by dropping the constant in the denominator we obtain the following LP-problem:

$$\min_{\lambda} \sum_j \lambda_j x_j \tag{32}$$

s.t.

$$\sum_j \lambda_j y_{jr} \geq y_{ir} \qquad \forall r \tag{33}$$

$$\lambda_j \geq 0 \qquad \forall j \tag{34}$$

Problem (32)-(34) has the following dual formulation:

$$\max_{p} \sum_{r} p_r y_{ir} \tag{35}$$

s.t.

$$\sum_{r} p_r y_{jr} \leq x_j \qquad \forall j \tag{36}$$

$$p_r \geq 0 \qquad \forall r \tag{37}$$

In the dual formulation a price p_r for each output is computed. The objective function seeks to maximize the total "revenue" for the evaluated company, while constraint (36) says that no company j should be rewarded more than the level of its cost. See Bjørndal et al. (2008) for a more thorough discussion of the dual formulation and its implications in the regulation context. We will now discuss the similarities between the StoNED model and the dual DEA model.

First, the concavity constraints (11) are related to the objective function in the dual DEA model. Under the CRS assumption, i.e., with $\alpha_i = 0$ for all companies i, the expression $\sum_r \beta_{rj} y_{ri}$ can be interpreted as the value of company i's output quantities evaluated with company j's prices. The inequalities in (11) express that the value of company i's output quantities, given by the variable γ_i, should be set as high as possible. This is similar to the dual DEA formulation, where the objective function (32) seeks to maximize the value of the output quantities.

Also, the regression equation (9) can be related to (36) in the dual DEA formulation for the special case where the combined effect of geography and the composite error term is non-negative for all companies, i.e.,

$$\sum_{s} \delta_s z_{si} + \epsilon_i \geq 0 \ \forall i.$$

This is true when, e.g., the model does not contain any geography variables and all the variation in the error term can be attributed to inefficiency (no noise). Then (9) implies

$$x_i = \gamma_i e^{\sum_s \delta_s z_{si} + \epsilon_i} \geq \gamma_i \ \forall i.$$

When we combine this inequality with the concavity constraints (10) and (11), we get

$$x_i \geq \sum_{r} \beta_{ri} y_{ri} \geq \sum_{r} \beta_{rj} y_{ri} \ \forall i, j,$$

which says the value of company i's output bundle cannot exceed its cost, and this must hold for any coefficient vector $\boldsymbol{\beta}_j$ that can be used to value the outputs. This is equivalent to the budget constraint (36) in the dual DEA formulation.

4 Data

Our data set was made available by NVE and is described in NVE (2012b) and NVE (2013), as well as NVE (2012a). Table 1 defines the variables, and tables 3 and 4 provide summary statistics.

On the input side, five cost elements are combined into a single cost measure. Most of the companies also owns and operates part of the regional distribution network, and NVE reallocates

[2] This variable is divided by the company's cost norm in order to ensure that the resulting variable is size-independent. The cost norm is based on five-year averages of input and output.

Table 1: Definition of variables.

Variable	Type	Sub-variable	Unit
Total cost	x	O&M costs	1000 NOK
		Value of lost load (VOLL)	1000 NOK
		Thermal power losses	1000 NOK
		Capital depreciation	1000 NOK
		Return on capital	1000 NOK
Customers	y	-	No. of customers
High voltage lines	y	-	Kilometers
Network stations	y	-	No. of stations
Avg. dist. to road	z	-	Meters
HV lines undergr.	z	-	Share of HV network (0-1)
Forest (coniferous)	z	-	Share of HV lines affected (0-1)
Geo1	z	Small scale hydro	Inst. cap. (MW) / cost norm[2]
		Average slope	Degrees (0-90)
		Deciduous forest	Share of HV lines affected (0-1)
Geo2	z	Wind / dist. to coast	$(m/s)^2/m$
		Islands	No. of islands / cost norm[4]
		HV sea cables	Share of HV network (0-1)

part of this cost to the (local) distribution activity. We have not included the reallocated cost in our analyses, hence our results may differ somewhat from the efficiency measurements published by NVE. The data for all years have been adjusted to the price level of a base year[3]. Table 2 shows the price data that we have used.

Table 2: Price data.

Base year	2011	2012
NVE rate of return (%)	5.31	4.20
Average system price (NOK / MWh)	393.46	259.90

The output variables measure the number of customers, as well as the length of the high voltage network and the number of network stations (transformers).

Five "geography" variables account for the heterogeneous conditions that the companies operate under. Data for the geography factors is not updated each year, and the data shown in Table 4 correspond to the factors included in NVE's benchmarking model for 2011 and 2012[4]. Two of the variables are combinations of several sub-variables, and the sub-variables are explained in Table 1. According to NVE (2013), the following expressions have been obtained by using principal component analysis (PCA):

$$Geo1 = -2.596 + 0.1687 \cdot Slope + 6.7132 \cdot DeciduousForest + 1090.9 \cdot SmallHydro$$

$$Geo2 = -0.645 + 0.876 \cdot WindCoast + 3197.6 \cdot Islands + 12.374 \cdot SeaCables$$

[3]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 at Nord Pool 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]I.e., the models used to calculate the revenue caps for 2013 and 2014, respectively.

Since these expressions include negative constant terms, the values of these two variables can be negative for some companies, as can be seen from Table 4.

Table 3: Summary of cost (1000 NOK, adjusted to 2012 price level) and output data.

	2007	2008	2009	2010	2011	2012	All
Cost mean	88091	94087	92654	93509	103283	91695	93887
Cost std	182765	196307	189098	186473	203995	174261	188469
Cost min	7424	7895	7155	7599	9659	8457	7155
Cost median	31188	32430	36390	32880	37296	35249	34235
Cost max	1532022	1625875	1651094	1532285	1573819	1512822	1651094
HV mean	783	785	792	798	800	806	794
HV std	1307	1303	1322	1337	1335	1344	1320
HV min	51	51	52	52	57	58	51
HV median	316	319	321	324	325	331	322
HV max	8313	8158	8395	8528	8648	8744	8744
NS mean	988	995	1000	1006	1013	1019	1004
NS std	1865	1876	1887	1891	1903	1910	1882
NS min	57	59	59	59	59	59	57
NS median	356	364	372	368	369	369	367
NS max	13401	13394	13515	13493	13525	13530	13530
Cust. mean	21973	22230	22406	22620	22885	23207	22554
Cust. std	57491	58071	58482	58992	59757	60809	58745
Cust. min	980	999	1012	1018	1016	1026	980
Cust. median	6350	6331	6372	6370	6438	6514	6400
Cust. max	533029	537534	541163	544925	552342	562501	562501

Table 4: Summary of data on geography factors.

	Mean	Std.dev.	Min	Median	Max
Distance to road	227.785370	207.817565	70.369850	142.874847	1056.444092
HV underground	0.338646	0.174896	0.057143	0.306000	0.864111
Forest	0.119059	0.099062	0.000000	0.119086	0.391629
Geo1	0.005215	1.485882	-2.063521	-0.438525	4.724301
Geo2	0.000347	1.512018	-0.642692	-0.455045	11.856144

Table 5 show correlation estimates between the variables. The cost and output variables are highly correlated with one another, due to size effects. More interestingly, we observe significant correlation between the geography variables and the input and output variables. For roads and the two composite geography variables the correlations with the cost and output variables are negative. One explanation for this could be that these effects are also related to size, since the latter group of variables are strongly correlated with size. It is not surprising that smaller companies have more difficult access to the road network, more small scale hydro and overhead lines through deciduous forest, and are more exposed to adverse coastal conditions, hence we would expect the corresponding variables to be negatively correlated with cost and output. For the underground high-voltage line and the forest variable the observed correlation is positive, since larger companies tend to be associated with cities, where the proportion of underground lines is higher. The observed positive correlation between the forest variable and cost/output variables, although only significant for three variables, is not so easy to explain.

Table 5: Correlations for 2008-2012. Correlation numbers in gray are not significant at 5 % level (one-sided tests).

	Cost	Cust.	HV	NS	Roads	HV u.	Forest	Geo1	Geo2
Cost (x)	1.00								
Cust. (y)	0.98	1.00							
HV (y)	0.94	0.88	1.00						
NS (y)	0.96	0.92	0.99	1.00					
Roads (z)	-0.11	-0.12	-0.08	-0.13	1.00				
HV undergr. (z)	0.28	0.33	0.15	0.22	-0.36	1.00			
Forest (z)	0.10	0.11	0.06	0.12	-0.48	0.18	1.00		
Geo1 (z)	-0.16	-0.16	-0.18	-0.20	0.02	-0.06	0.01	1.00	
Geo2 (z)	-0.07	-0.07	-0.08	-0.10	0.27	-0.08	-0.15	-0.13	1.00

5 Negative skewness and other implementation issues

We have estimated the StoNED model described in Section 3.1 on two data sets, and the estimation results are described by the parameters in Table 6. Both data sets are obtained by taking averages over a five year period, as in the current regulation framework, and we have implemented the multiplicative StoNED model formulations from Section 3.1. We include all the output and environmental variables described in Section 4, and we assume constant returns to scale ($\alpha_i = 0 \; \forall i$), as in the current regulation model.

Table 6: StoNED statistical parameters for StoNED models.

Period:	2007-2011	2008-2012
M_2	1.64E-02	1.92E-02
M_3 (est.)	-1.16E-04	8.87E-05
M_3 (used)	1.00E-05	8.87E-05
σ_u	0.04	0.07
σ_v	0.13	0.13
$\lambda = \sigma_u/\sigma_v$	0.28	0.56
μ	0.03	0.06
White (p)	0.50	0.66
No. of obs.	123	123
Time (sec.)	19	8

The distribution of the composite errors (ϵ_i) for the estimated models are shown in Figure 1, and the estimated values for M_2 and M_3 are shown in Table 6. We see that the estimated values for M_3 are negative for the 2007-2011 data set. Since we assume that the inefficiency terms u_i follow a positively skewed half-normal distribution, and that the distribution of the noise terms v_i is normally distributed, the distribution of the combined error terms $\epsilon_i = u_i + v_i$ should be positively skewed. A negative value of M_3 implies that the estimate of σ_u will also negative, by (16). Note that if the "true" value of σ_u is small, then it is quite probable that a negative value of M_3 would be observed in a small sample. We will handle this problem as suggested by Olson et al. (1980), i.e., by replacing the estimated value of M_3 by a small number. In the GAMS code that is made available with Kuosmanen (2012), this value is set to 0.00001. We use the same value in all our calculations, unless stated otherwise, and this will completely determine the values of σ_u and μ.

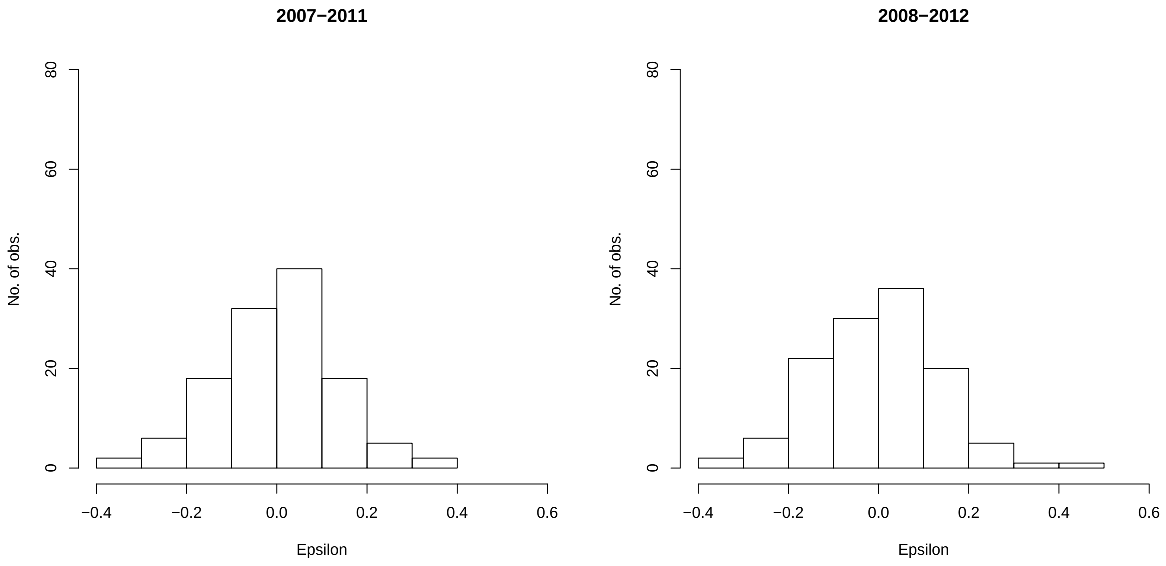

Figure 1: Distribution of estimated errors.

Figure 2 shows the StoNED efficiency scores for the 2007-2011 data set. The efficiency scores are calculated using the cost frontier approach in (20) and the conditional mean (JLMS) formula in (21). The computed value of the skewness statistic M_3 is negative, and the figure illustrates the effect, on the final efficiency scores, of choosing different values for this parameter. We see that the results are indeed influenced by this arbitrary choice. The six data series are comparable in the sense that the companies are listed in the same order. We see that all of the series show increasing efficiency scores, hence the *ranking* of the companies are not influenced by whether we use the cost frontier or the JLMS approach, or which value we use for M_3.

While the ranking of the companies is the same, we see from Figure 2 that the efficiency *levels* are indeed affected by the value of M_3. Table 7 shows average efficiency scores for the different model variants, and we see that the averages differ considerably, especially with the cost frontier approach. Moreover, Figure 2 shows that there is much less *variation* between companies with the JLMS approach than with the cost frontier approach. The effect on level and variation is important in a regulation context, since the computed efficiency scores will be used by the regulator to set revenue caps, and therefore the relative profitability of companies will be affected.

From formula (19) and (16) we see that the value of M_3 will affect the value of the cost function through the parameter μ, i.e., the average inefficiency. A change in this parameter, under the multiplicative model, will affect the computed cost and efficiency numbers in a proportional manner via the term $e^{-\mu}$. If the calibration procedure employed by the regulator is also proportional, the calibrated efficiency scores would be unaffected by the chosen value of M_3. Calibration alternative A, given by (2), corresponds to a proportional scaling of the estimated efficiency scores. To see this, insert (2) and $C_i^* = C_i\theta_i$ into (1), which results in the following formula for the revenue cap of company i:

$$R_i = \alpha C_i^* \left(1 + \frac{\sum C - \sum C^*}{\sum C^*}\right) + (1-\alpha)C_i = \alpha C_i \theta_i \frac{\sum C}{\sum C^*} + (1-\alpha)C_i \tag{38}$$

We see from (38) that the calibration is equivalent to multiplying all the efficiency scores by the same factor $\frac{\sum C}{\sum C^*}$, i.e., the inverse of the cost weighted average efficiency for the industry. The combined effect of the second StoNED stage and the regulator's calibration is thus to multiply the first-stage efficiency scores by the factor $e^{-\mu}\frac{\sum C}{\sum C^*}$. The combined effect will be such that industry revenue equals industry cost, i.e., $\sum R = \sum C$, independently of the chosen value for μ. In such a setting we might as well skip the second StoNED stage and set the revenue caps based on the average-practice frontier, i.e., with $\mu = 0$. The calibration procedure currently used by NVE is alternative B, which is is not equivalent to a proportional scaling of the efficiency scores, hence the chosen value for M_3 and μ will indeed influence the relative profitability of the companies.

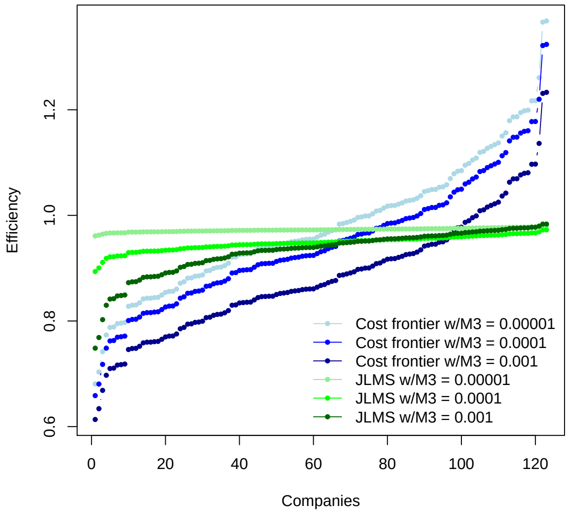

Figure 2: StoNED efficiency scores for 2007-2011 under various assumptions with respect to the skewness parameter M_3.

Table 7: Average efficiency for 2007-2011 data set under different assumptions.

	Simple avg.	Cost weighted avg.
Cost frontier w/$M_3 = 0.00001$	98.0	101.7
Cost frontier w/$M_3 = 0.0001$	94.8	98.4
Cost frontier w/$M_3 = 0.001$	88.3	91.7
JLMS w/$M_3 = 0.00001$	97.3	97.4
JLMS w/$M_3 = 0.0001$	94.8	95.2
JLMS w/$M_3 = 0.001$	93.2	94.7

Table 5 showed that the geography factors are correlated with the cost and output variables. This leads to biased efficiency estimates if two-stage methods are used, see e.g. Banker and Natarajan (2008) and Barnum and Gleason (2008). Similar biases will arise with the basic StoNED model, see Johnson and Kuosmanen (2012). Barnum and Gleason (2008) propose a *reversed* two-stage procedure in order to handle the bias problem. Similarly, Johnson and Kuosmanen (2012) propose a modified StoNED procedure that is shown (using a simulation study) to perform better than the basic StoNED model in the presence of such correlation.

Table 6 reports p-values for the White test for heteroscedasticity. This test determines whether the size of the (squared) errors can be explained by the output variables, and a low p-value indicates that the null hypothesis of homoscedasticity can be rejected. We see that the null hypothesis cannot be rejected for any of the data sets.

Computing time and numerical stability is an issue and may limit the applicability of the StoNED method. The non-linear nature of (9), combined with the large number of constraints in (11), makes (8)-(12) difficult to solve. We did our computations in GAMS, and we found that the ability of the program to solve the optimization problem at all, as well as the computing speed, is quite sensitive to the choice of solver and starting point. After some testing we chose to use the IPOPT solver, combined with the starting point $\beta_{ik} = x_i/(y_{ik}K)$, where K is the number of outputs, and $\alpha_i = \delta_s = 0$. Table 6 shows that it took 30 and 18 seconds[5] to compute the two models[6].

6 Comparison of results

6.1 Efficiency scores

We compare the two StoNED models with the two-stage procedure used by the regulator. NVE evaluates observed data for cost and output in a given year against a frontier formed by averaging data over the last five years, and our results are based on the same methodology. The non-concavity constraints (11) lead to the following alternative formula for the cost function:

$$\hat{C}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{y}_i, \boldsymbol{z}_i) = \max_j \left(\alpha_j + \sum_r \beta_{rj} y_{ri} \right) e^{-\mu + \sum_s \delta_s z_{si}} \tag{39}$$

The coefficients matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\delta}$ in (39) do not have to be based on the same data set as the one containing the data point $(\boldsymbol{y}_i, \boldsymbol{z}_i)$[7]. Hence, by storing the coefficients we can use (39) to evaluate the performance of a company in a different data set than the one used to calculate the coefficients. Specifically, we can evaluate actual performance in a given year against a frontier based on average data for a number of years, as practiced by the Norwegian regulator.

We compare the StoNED models to three different two-stage procedures. The first is the one actually used by NVE, where the efficiency scores from the the DEA model are regressed on differences between observed values for each geography variable and a weighted average that represents the level of the corresponding variable for the reference companies, see formulas (27) and (28). We have also computed an alternative where the efficiency scores are regressed on absolute levels of the geography variables. In the third variant we take logarithms of the efficiency scores and regress them on the absolute levels of the geography variables, see formulas (29)-(31).

Figure 3 illustrates the results for 2011 and 2012. In general, the StoNED results are highly correlated with results from the two-stage procedures, and this is confirmed by Table 8. Also,

[5]We used a laptop with an Intel(R) Core(TM) i7-4600U processor with a clock speed of 2.10GHz, 8GB internal memory, and 64-bit GAMS.

[6]For the pooled data sets in Section 7, where the computing times were much higher, we also tested the constraint-generation procedure in Lee et al. (2012), but we were not able to reduce the computing times in our GAMS implementation. This is due to the fact that GAMS regenerates the entire optimization problem in each iteration, even though only one or a few new constraints have been added.

[7]See the simulation study in Kuosmanen et al. (2013).

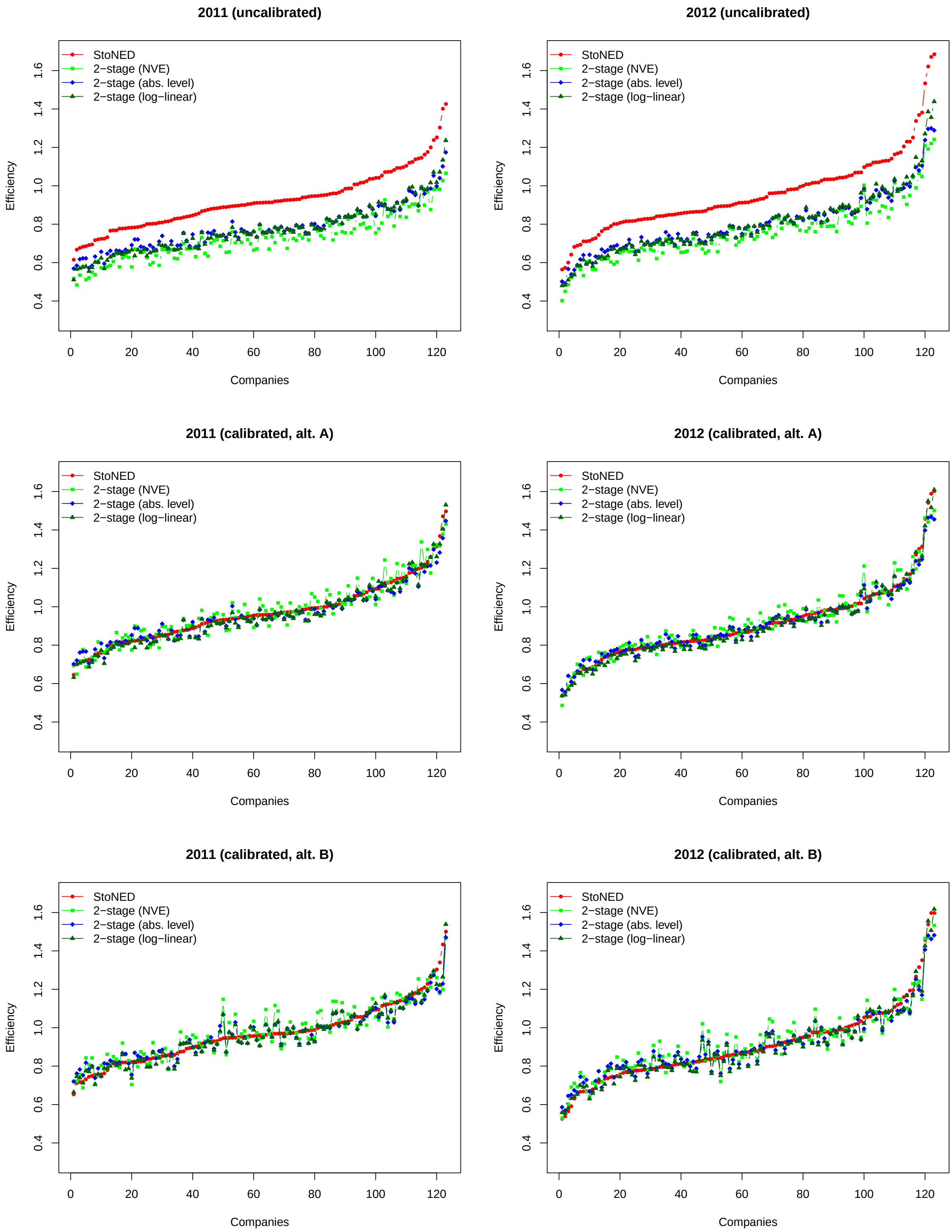

Figure 3: Efficiency scores for 2011 and 2012. Observed data for current year evaluated against 5 year average.

the correlation estimates depend on how the two stage procedure is implemented, and the results from the log-linear procedure seems to be slightly more correlated with the StoNED results than the results from NVE's procedure.

Table 8: Correlations between results from StoNED and other models.

	2011 vs 2007-2011 avg.	2012 vs 2008-2012 avg.
Ranking		
DEA	0.85	0.85
2-stage (NVE)	0.95	0.96
2-stage (abs. level)	0.97	0.98
2-stage (log-linear)	0.98	0.99
Efficiency scores		
DEA	0.88	0.89
2-stage (NVE)	0.96	0.97
2-stage (abs. level)	0.98	0.99
2-stage (log-linear)	0.99	0.99

Table 9: Average efficiency scores.

	2011 vs 2007-2011 avg.	2012 vs 2008-2012 avg.
Simple average		
StoNED	92.3	95.7
DEA	68.7	71.8
2-stage (NVE)	72.7	76.1
2-stage (abs. level)	77.8	80.3
2-stage (log-linear)	77.0	80.2
Cost weighted average		
StoNED	95.3	105.3
DEA	73.1	81.5
2-stage (NVE)	74.5	82.7
2-stage (abs. level)	81.1	88.6
2-stage (log-linear)	80.8	89.4

Figure 3 shows that the levels of the estimated efficiency scores differ between the methods. This is important, since it may affect the revenue caps via the calibration procedure that NVE applies. According to this procedure, the individual cost norms for the companies are adjusted such that the total for the industry is equal to the observed cost. The initial difference between the norm and observed cost is distributed among the companies in proportion to their capital (book) values. If the initial cost weighted average efficiency score for the industry equal is below (above) 100 %, the calibration procedure will favor companies with relative high (low) capital values. Hence, the efficiency score levels will influence the distribution of the revenue caps among the companies in the industry. Table 9 shows that the averages differ considerably, both between methods and between years. The StoNED models yield considerably higher efficiency scores than the two stage methods, and the two-stage procedures based on absolute values of geography give higher efficiency averages than the method currently used by NVE.

6.2 Coefficient estimates

6.2.1 Output variables

The output coefficients in StoNED are related to the corresponding shadow prices for outputs in DEA[8]. Table 10 shows some summary statistics for the coefficient estimates for the 2008-2012 average data set. As an example, the coefficient value for the high voltage output according to DEA ranges from 0 to 54' NOK, with a mean value of 29' NOK. The table show StoNED output coefficients for models without and with geography variables, respectively. We see that the huge variation in coefficient values is something that characterizes the StoNED methods as well, and it is not possible to conclude that one method gives less (more) variation in the coefficient values than the others. The histograms in Figures 4-6 show the entire distribution of coefficients over the companies in the data set, and they confirm that the distribution of coefficient values varies considerably between the methods.

Table 10: Output coefficients in 1000 NOK.

	Mean	Std	Min	Median	Max
HV lines					
DEA	28.996	14.088	0.000	35.723	53.811
StoNED w/o geo.	54.893	16.285	0.000	59.106	79.744
StoNED with geo.	49.030	16.587	0.000	55.794	70.935
Network stations					
DEA	16.905	15.079	0.000	18.119	54.924
StoNED w/o geo.	7.452	12.107	0.000	2.923	66.289
StoNED with geo.	11.578	14.608	0.000	3.271	60.865
Customers					
DEA	1.195	0.631	0.000	1.327	2.460
StoNED w/o geo.	2.133	0.700	0.000	2.365	3.334
StoNED with geo.	1.318	0.594	0.000	1.644	2.571

Figure 7 illustrates the virtual weights of the three outputs according to the three methods. The virtual weight of an output is found by multiplying the physical quantity of that output with its price[9]. In our regulation context we may interpret this weight as the revenue earned from each output. In the figure we show the weights as percentages of the total cost norm. Each company corresponds to a vertical bar in the diagrams, and the companies have been sorted according to the size of their cost norms. Customers receive the largest aggregate weight in all model variants, and almost all companies attribute some weight to this variable. The largest company in the data set attributes all it's weight to customers under all three model. Both of the StoNED variants attributes less of the cost norm to network stations compared to DEA, and more to high voltage lines.

Note that the size of the cost norms shown in Figure 7 does not reflect the actual industry revenue. The second-stage adjustment for geography is added to the DEA cost norm, and the StoNED (with geo.) cost norm would be adjusted via the multiplicative term $e^{\mu+\sum_s \delta_s z_{si}}$. The calibration performed by NVE would also affect the size of the final cost norm.

[8]The objective function (23) in the DEA model measures the efficiency score, and the DEA shadow prices will therefore measure the marginal effect of output changes on the efficiency scores. The StoNED coefficients, on the other hand, will measure the marginal effect of output changes on the cost norm. In order make the estimates comparable, we multiply the DEA shadow prices by the companies' reported cost numbers.

[9]See the discussion of DEA virtual weight restrictions in Sarrico and Dyson (2004), Bjørndal et al. (2008) and Bjørndal et al. (2009).

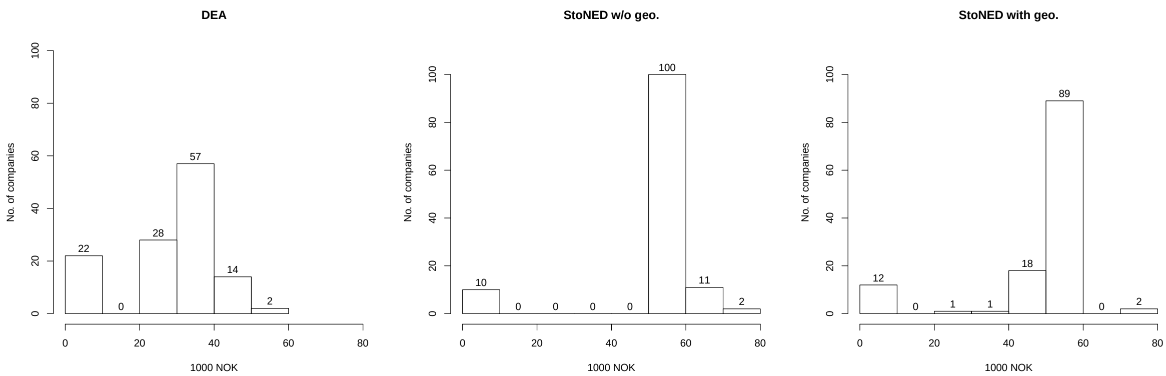

Figure 4: HV line coefficients.

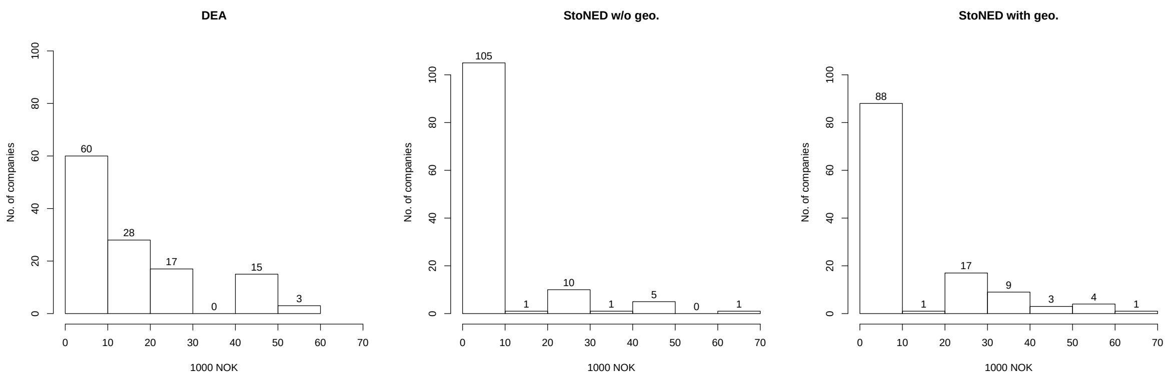

Figure 5: Network station coefficients.

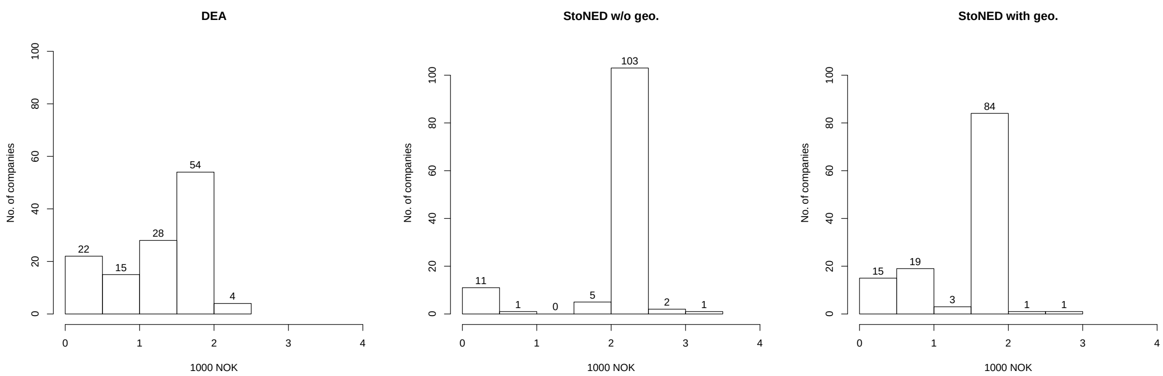

Figure 6: Customer coefficients.

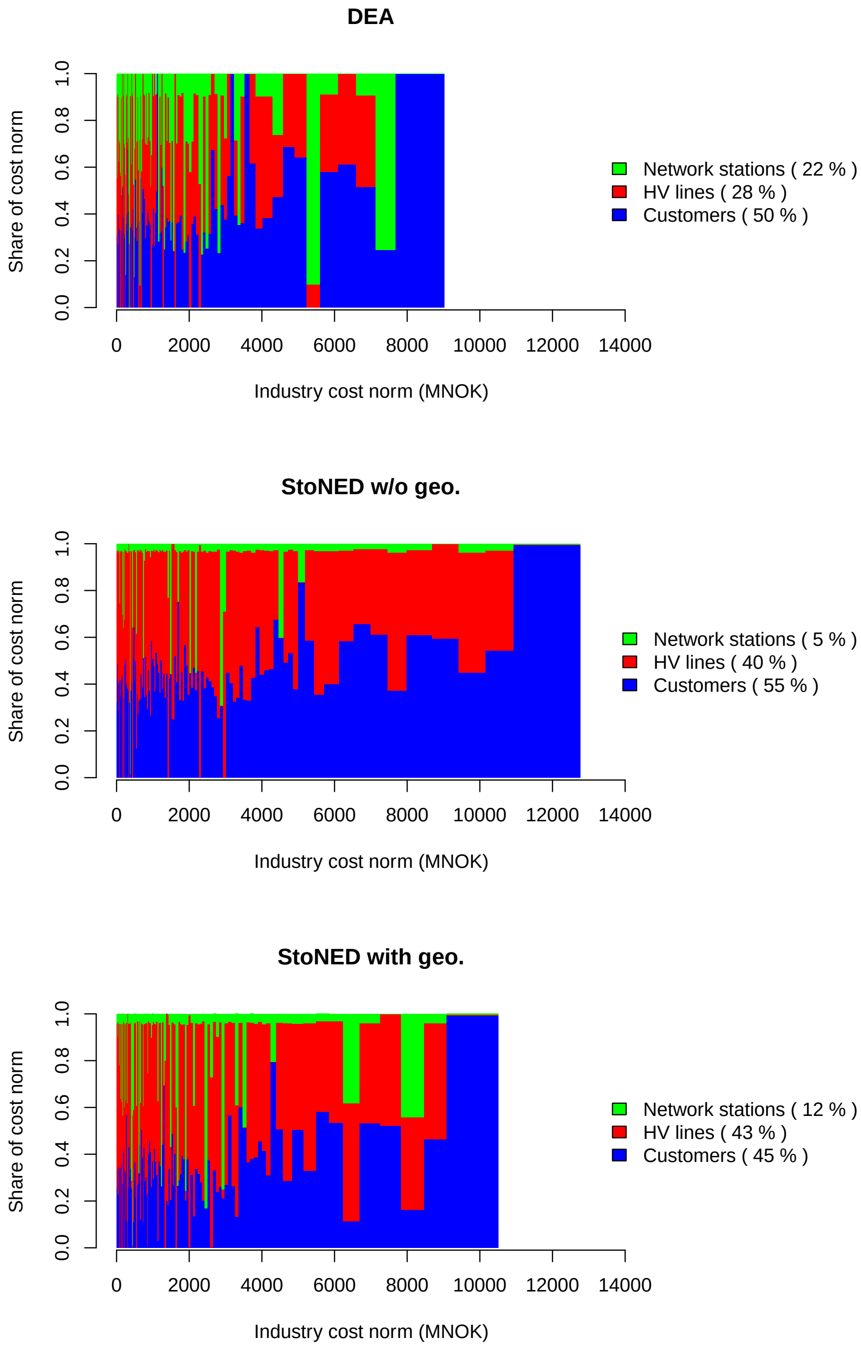

Figure 7: Virtual output weights for cost norms.

6.2.2 Geography variable coefficients

Johnson and Kuosmanen (2011) provide a formula for the standard error of the contextual variables (geography) coefficient estimates in StoNED, and we report these in the second column of Table 11[10]. We see that all the geography variables, except for the distance variable, have a significant effecct[11]. We can compare the StoNED significance results to similar results for the regression in the two-stage methods. The regression results agree with the StoNED results with respect to the distance variable, but in addition the effect of forest is insignificant.

Table 11: Coefficients for geography variables.

	Estimate	Std. Error	t value	$Pr(>\|t\|)$
StoNED				
Distance to road	0.000077	0.000051	1.50	0.14
HV underground	0.308907	0.055570	5.56	0.00
Forest	0.358375	0.122806	2.92	0.00
Geo1	0.044162	0.008695	5.08	0.00
Geo2	0.048376	0.008824	5.48	0.00
2-stage (NVE)				
Intercept	0.757795	0.013526	56.03	0.00
Distance to road	-0.000070	0.000071	-0.98	0.33
HV underground	-0.176156	0.086006	-2.05	0.04
Forest	-0.174246	0.113993	-1.53	0.13
Geo1	-0.028455	0.006972	-4.08	0.00
Geo2	-0.029196	0.006819	-4.28	0.00
2-stage (abs. level)				
Intercept	0.799424	0.034012	23.50	0.00
Distance to road	-0.000090	0.000058	-1.57	0.12
HV underground	-0.132873	0.059481	-2.23	0.03
Forest	-0.160473	0.111790	-1.44	0.15
Geo1	-0.036507	0.006596	-5.53	0.00
Geo2	-0.031258	0.006715	-4.65	0.00
2-stage (log-linear)				
Intercept	-0.236658	0.046784	-5.06	0.00
Distance to road	-0.000122	0.000079	-1.55	0.12
HV underground	-0.170597	0.081817	-2.09	0.04
Forest	-0.245131	0.153770	-1.59	0.11
Geo1	-0.050192	0.009074	-5.53	0.00
Geo2	-0.049012	0.009237	-5.31	0.00

The interpretation of the variable coefficients depends on how the respective variables are specified, as well as the type of model used. For the multiplicative StoNED model, the interpretation follows from formula (19), and the geography indices are used directly, i.e., without scaling them. The HV underground variable measures the share of underground cables in the company's high voltage network, i.e., it has values between 0 and 1. A 1 % increase in a company's share of underground cables should increase its cost norm by 0.31 %.

[10] Instead of using the formula in Johnson and Kuosmanen (2011), we can obtain the standard errors by regressing $\hat{\boldsymbol{\epsilon}} + \boldsymbol{z}\boldsymbol{\delta}$ on the geography variables $\boldsymbol{z}$, where $\hat{\boldsymbol{\epsilon}}$ is the estimate of the composite error terms from the first stage of the StoNED method.

[11] Given a 5 % significance level.

For the two-stage models the estimated coefficients measure the relationship between the DEA efficiency score from the first stage and the value of the geography variables. In the two-stage procedure, the coefficients are used to adjust the efficiency scores for the effect of geography. Since an efficiency score is the ratio between a cost norm and observed cost, the adjustment with respect to the cost norm is made relative to the observed cost. In the model based on geography differences, the coefficient value for the underground variable is -0.176156, meaning that if the share of underground cables for a company increases by 1 %-point,then the cost norm should increase by an amount equalt to 0.18 % of its actual cost, assuming that the underground share for its reference company does not change. If the second stage regression is formulated with respect to absolute levels of geography, the underground coefficient is -0.132873, indicating that a 1 %-point increase in the underground share should lead to an increase in the cost norm equal to 0.13 % of the actual cost. For the logarithmic model, the coefficient indicates an effect that is relative to the efficiency score / cost norm. The estimated coefficient indicates an increase in the cost norm of 0.17 % if the share of underground cables increases by 1 %-point.

7 Effect of noise

Under a deterministic DEA model variation in the data can easily affect the location of the efficient frontier. More noise in the data set will move the frontier away from the observed points and lower the efficiency scores. Since the StoNED model explicitly models noise, the effect of more noise is not so straightforward. If the existing noise is modelled correctly, the location of the frontier should not change. In order to isolate the effect of noise on the location of the frontier, we have performed annual efficiency evaluations for each company in the period 2008-2012 against three different data sets:

1. Average data over a five-year period, i.e., as practised in the current regulation model.
2. Annual data sets for each of the five years.
3. Pooled data set for the entire period.[12].

We want to isolate the frontier effects, and we have therefore kept the data for the evaluated companies constant in all the three alternatives. This means that in Alternative 1, the data for the frontier companies will be different from the data for the evaluated companies, and the cost estimate underlying the cost efficiency will have to be calculated using formula (39). The cost data are expressed in 2012 nominal values.

Table 12: Average efficiency scores for 2008-2012. Actual observations against three different frontiers.

Average:	Simple			Cost weighted		
Frontier:	Average	Annual	Pooled	Average	Annual	Pooled
StoNED (best-practice)	95.3	94.7	97.8	99.9	99.5	102.6
StoNED (avg.-practice)	101.1	100.6	100.7	106.0	105.6	105.5
DEA	72.0	69.8	64.6	78.5	76.2	69.9
2-stage (NVE)	76.1	74.1	68.0	79.6	77.6	71.0
2-stage (abs. level)	80.1	77.5	74.7	84.9	82.4	79.3
2-stage (log-linear)	79.9	77.7	74.9	85.1	83.0	80.3

[12]We do not consider dynamic effects explicitly. See the discussion of a StoNED panel data model in Johnson and Kuosmanen (2012)

Table 13: Summary statistics for pension provisions relative to total costs for 2008-2012.

	2008	2009	2010	2011	2012	All
No. of companies	10	11	10	10	8	18
Min.	0.043	-0.223	-0.002	-0.110	-0.796	-0.796
Median	0.102	-0.010	0.069	0.170	-0.482	0.064
Mean	0.127	0.005	0.080	0.142	-0.400	0.007
Max.	0.259	0.201	0.228	0.251	-0.005	0.259

Table 14: Parameters from StoNED calculations.

Data set:	2008	2009	2010	2011	2012	2008-2012 average	2008-2012 pooled
With reported pension provisions							
M_2	2.02E-02	2.25E-02	2.06E-02	2.42E-02	3.72E-02	1.92E-02	2.66E-02
M_3 (est.)	-5.78E-05	6.73E-04	4.76E-04	-6.03E-04	-1.58E-03	8.87E-05	-2.09E-04
M_3 (used)	1.00E-05	6.73E-04	4.76E-04	1.00E-05	1.00E-05	8.87E-05	1.00E-05
σ_u	0.04	0.15	0.13	0.04	0.04	0.07	0.04
σ_v	0.14	0.12	0.12	0.15	0.19	0.13	0.16
$\lambda = \sigma_u/\sigma_v$	0.25	1.20	1.08	0.23	0.19	0.56	0.22
μ	0.03	0.12	0.10	0.03	0.03	0.06	0.03
White (p)	0.77	0.56	0.71	0.54	0.23	0.66	0.67
Time (sec.)	27	12	9	15	11	9	759
With smoothed pension provisions							
M_2	1.97E-02	2.13E-02	2.16E-02	2.23E-02	2.64E-02	1.92E-02	2.39E-02
M_3 (est.)	-5.95E-05	6.99E-04	4.76E-04	-5.21E-04	1.39E-03	8.87E-05	5.44E-04
M_3 (used)	1.00E-05	6.99E-04	4.76E-04	1.00E-05	1.39E-03	8.87E-05	5.44E-04
σ_u	0.04	0.15	0.13	0.04	0.19	0.07	0.14
σ_v	0.14	0.12	0.12	0.15	0.12	0.13	0.13
$\lambda = \sigma_u/\sigma_v$	0.26	1.27	1.04	0.24	1.57	0.56	1.03
μ	0.03	0.12	0.10	0.03	0.15	0.06	0.11
White (p)	0.77	0.61	0.93	0.57	0.62	0.66	0.03
Time (sec.)	17	8	15	17	18	9	778
No. of companies	123	123	123	123	123	123	615

Table 14 describes the statistical parameters from the StoNED calculations. We focus on the multiplicative model variant, and with all geography factors included. We see that the single-year data sets for 2008, 2011, and 2012, as well as the pooled data set for 2008-2012, exhibit the wrong skewness, i.e. the estimated values of M_3 are negative. The White test yields p-values above 0.05 for all the data sets, hence heteroscedasticity does not seem to be a problem. We also note that the computation time for the pooled data set is almost 22 minutes!

In order to illustrate the frontier effect of smoothing / pooling, we show the average efficiency scores for the various model in Table 12. Because the value of the skewness parameter M_3 can have significant effects on the level of the efficiency scores, we also report results for the "average-practice" StoNED frontier (see Johnson and Kuosmanen (2012)), i.e., with $M_3 = \sigma_u = 0$. Under the multiplicative StoNED model, changing the assumed value of σ_u affects the efficiency scores via the average inefficiency $\mu = \hat{\sigma}_u\sqrt{2/\pi}$ and the term $e^{-\mu}$ in (19). Hence, using the average-practice frontier is equivalent to calibration of the efficiency score via a proportional scaling of the efficiency scores. We see that the average efficiency scores under average-practice StoNED are quite stable over the three frontier alternatives, and this reflects that the setting of the statistical parameters σ_u and σ_v in the second stage of the StoNED procedure determines the location of the frontiers to a large degree.

The two-stage methods are quite sensitive to how the frontier is defined, and the effects here are more predictable than with the best-practice StoNED frontier. The alternative with average data has less noise, and we would therefore expect the average efficiency scores to be higher for this alternative. This is indeed confirmed by Table 12. We also see that the pooled data set have lower average efficiency scores than the current year data sets. This is also as expected, since adding more DMUs to a data set will always result in lower (or unchanged) efficiency scores.

In order to further study the effects of noise under the different methods, we focus on the reported pension provisions, where we know that there is a considerable amount of inter-temporal noise. Table 13 shows the reported pension provisions, as share of total cost. We see that there are in total 18 companies that reported pension provisions in the period 2008-2012. We see that this cost element can be positive or negative, and that the average was close to zero.

Since there are only a few companies that report pension provisions, we can perform an experiment to see how the noise in their cost data affects the efficiency results of the other companies in the data set. We smooth the reported pension provisions by taking an average over the period 2008-2012, and then we replace the reported cost numbers by the averages. Hence, we are not changing the total cost of the industry for the this period, the only change is that we reallocate some of the cost for the 18 companies that have reported pension provisions. The StoNED calculations for the smoothed data set is described in Table 14. The skewness parameter for 2012 now is positive, and this is also the case for the pooled data set. The value of σ_u for the pooled data set changes from 0.03 to 0.14, and the assumed inefficiency μ changes from 0.03 to 0.11. The p-value for the White test is below 0.05 for the pooled data set, indicating that heteroscedasticity may be a problem.

Table 15 show the cost weighted average efficiency scores for the 105 companies that did not report any pension provisions. We see that the level of the average-practice StoNED frontier is left almost unchanged, but that the best-practice changes considerable. This is especially true for the pooled frontier, due to the increase in the value of σ_u and μ. As expected, the noise smoothing increases the efficiency scores for the remaining companies under the two-stage methods, and we see that the average effects can be considerable.

Table 15: Average (cost weighted) efficiency scores for 2008-2012. Only companies with no reported pension provisions.

Data set:	Annual			Pooled		
Pension provisions:	Reported	Smoothed	Difference	Reported	Smoothed	Difference
StoNED (best-practice)	99.5	97.3	-2.2	102.6	94.9	-7.6
StoNED (avg.-practice)	105.6	105.9	0.2	105.5	105.8	0.2
DEA	76.2	77.6	1.5	69.9	73.1	3.2
2-stage (NVE)	77.6	79.0	1.4	71.0	74.5	3.5
2-stage (abs. level)	82.4	83.1	0.7	79.3	80.6	1.3
2-stage (log-linear)	83.0	83.3	0.3	80.3	80.8	0.6

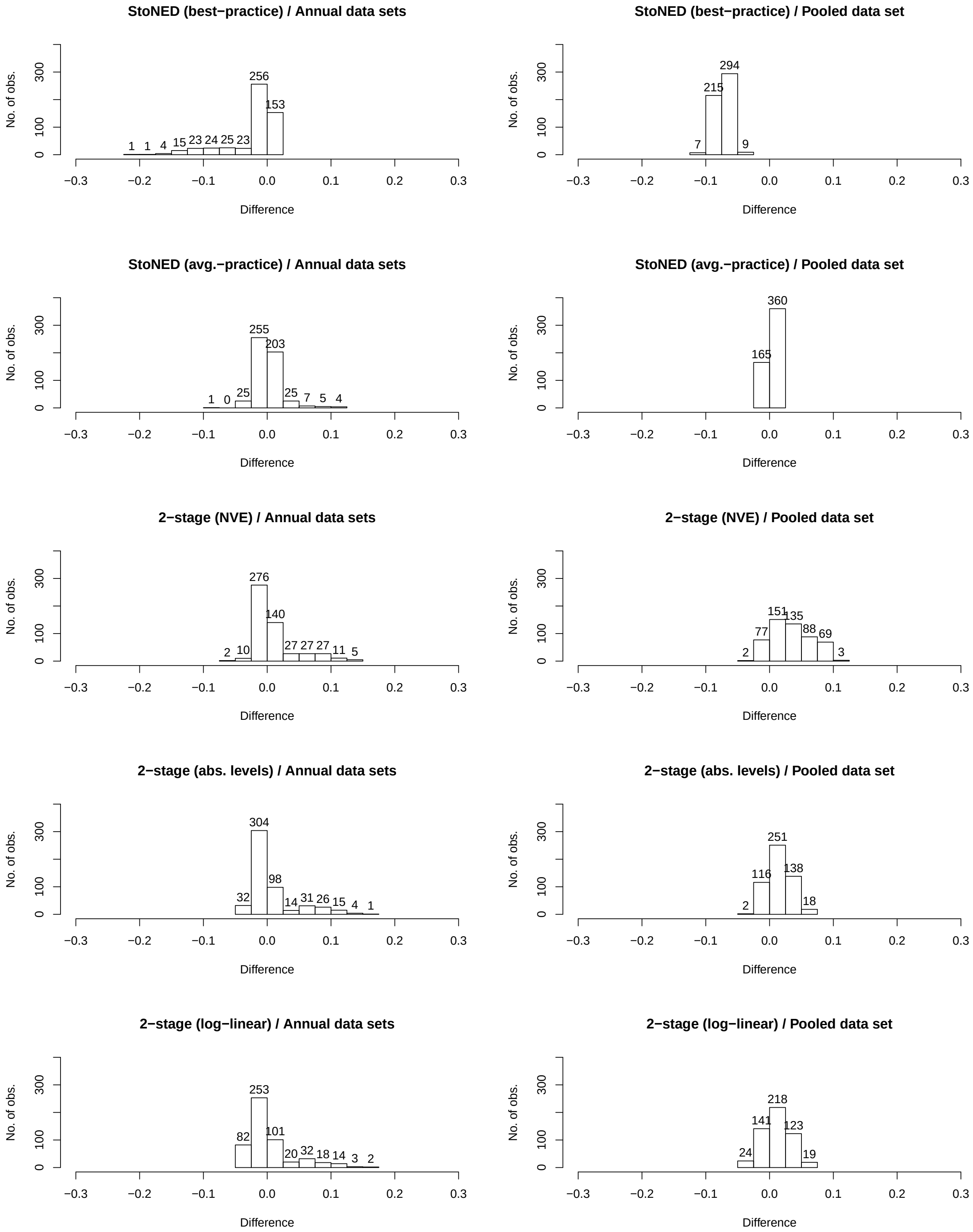

Figure 8: Effect of pension smoothing on efficiency scores for 2008-2012. Companies with no reported pension provisions.

Figure 8 illustrates the effect for individual companies. The figure confirms the mostly negative smoothing effects for the best-practice StoNED frontier, and that the average-practice frontier effects can be both positive and negative. For the pooled data set, the individual effects with the average-practice frontier are very small. It is interesting to note that while pooling reinforces the effect of noise under the DEA-based method, it seems to make the average-practice StoNED model more robust to noise.

8 Conclusions

We have implemented the StoNED models on data for Norwegian electricity distribution networks. There are many similarities between the StoNED models and the two-stage model presently used by NVE.

We saw in Section 5 that the estimated values for the skewness parameter M_3 in the StoNED second stage can be inconsistent with the assumed positive skewness for the inefficiency term. Setting the skewness parameter to an arbitrary value can have significant consequences for the efficiency levels.

The comparison in Section 6.1 showed, firstly, that the StoNED efficiency results are highly correlated with results from two-stage models. Interestingly, the correlation is even stronger if the two-stage DEA model is modified so that the regression in the second stage is done in a manner that is more in line with the assumptions of the multiplicative StoNED model. Secondly, the average efficiency scores with the StoNED method are higher than with the two-stage methods. This has implications for the calibration that is done by NVE.

The interpretation of the StoNED model is related to the dual DEA model in Stage 1 of NVEs procedure, as discussed in 3.3. Section 6.2 showed that the coefficient estimates under the various model variants are similar in the aggregate, but that there are large differences in the coefficient estimates when we look at individual companies.

In Section 7 we studied the effect of noise on the results from the various models. Noise affects the DEA-based results in a predictable way. The average effect of noise reduction (smoothing) is positive, but with significant individual variation. The noise effects under (the multiplicative) StoNED method are more unpredictable, but this is due to the somewhat arbitrary parameter choices in Stage 2 of the procedure. The average-practice frontier from Stage 1 seems to be more robust to noise effects.

References

Agrell, P. J., Bogetoft, P., and Tind, J. (2005). Dea and dynamic yardstick competition in scandinavian electricity distribution. *Journal of Productivity Analysis*, 23(2):173–201.

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

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.

Banker, R. D. and Natarajan, R. (2008). Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations Research*, 56(1):48–58.

Barnum, D. T. and Gleason, J. M. (2008). Bias and precision in the DEA two-stage method. *Applied Economics*, 40(18):2305–2311.

Bjørndal, E., Bjørndal, M., and Camanho, A. (2008). Weight restrictions on geography variables in the dea benchmarking model for norwegian electricity distribution companies. Centre for Applied Research at NHH (SNF), report 38/08.

Bjørndal, E., Bjørndal, M., and Camanho, A. (2009). Weight restrictions in the dea benchmarking modul for norwegian electricity distribution companies: size and structural variables. Centre for Applied Research at NHH (SNF), report 22/09.

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.

Fan, Y., Li, Q., and Weersink, A. (1996). Semiparametric estimation of stochastic production frontier models. *Journal of Business & Economic Statistics*, 14(4):pp. 460–468.

Johnson, A. L. and Kuosmanen, T. (2011). One-stage estimation of the effects of operational conditions and practices on productive performance: asymptotically normal and efficient, root-n consistent StoNEZD method. *Journal of Productivity Analysis*, 36(2):219–230.

Johnson, A. L. and Kuosmanen, T. (2012). One-stage and two-stage DEA estimation of the effects of contextual variables. *European Journal of Operational Research*, 220(2):559 – 570.

Jondrow, J., Lovell, C. K., Materov, I. S., and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19(23):233 – 238.

Kuosmanen, T. (2012). Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model. *Energy Economics*, 34(6):2189 – 2199.

Kuosmanen, T. and Johnson, A. L. (2010). Data envelopment analysis as nonparametric least-squares regression. *Operations Research*, 58(1):149–160.

Kuosmanen, T. and Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints. *Journal of Productivity Analysis*, 38(1):11–28.

Kuosmanen, T., Saastamoinen, A., and Sipilinen, T. (2013). What is the best practice for benchmark regulation of electricity distribution? comparison of DEA, SFA and StoNED methods. *Energy Policy*, 61(0):740 – 750.

Lee, C.-Y., Johnson, A. L., Moreno-Centeno, E., and Kuosmanen, T. (2012). A more efficient algorithm for convex nonparametric least squares. *European Journal of Operational Research*.

NVE (2012a). Forslag til endring av modeller for å fastsette kostnadsnormer - inntektsregulering av nettselskaper fra 2013. høringsdokument nr 2/2012.

NVE (2012b). Rundskriv 3/2012 om beregning av inntektsrammer og kostnadsnorm for 2013.

NVE (2013). Infoskriv 4/2013: Om beregning av inntektsrammer og kostnadsnorm for 2014.

Olson, J. A., Schmidt, P., and Waldman, D. M. (1980). A Monte Carlo study of estimators of stochastic frontier production functions. *Journal of Econometrics*, 13(1):67 – 82.

Sarrico, C. and Dyson, R. (2004). Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research*, 159(1):17 – 34.

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

We evaluate the StoNED methodology for benchmarking and regulation of network companies, and we compare StoNED to the two-stage DEA method currently used by the Norwegian regulator. We find that the estimated values for the skewness parameter in the second stage of the StoNED procedure can be inconsistent with the assumed positive skewness for the inefficiency term. Setting the skewness parameter to an arbitrary value can have significant consequences for the efficiency levels. This effect is partly neutralized by the revenue calibration performed by the NVE, depending on how the calibration is implemented. Our comparison of results from StoNED and two-stage DEA show that the efficiency scores from the two methods are highly correlated, but that the levels can differ significantly. We also interpret the StoNED coefficient estimates and compare them to the corresponding (dual) DEA estimates. Finally, we illustrate the robustness of the efficiency estimates to noise in data, as exemplified by noisy pension costs.

SNF

Samfunns- og næringslivsforskning AS
Centre for Applied Research at NHH

Helleveien 30
NO-5045 Bergen
Norway

P +47 55 95 95 00
E snf@snf.no
W snf.no

Trykk: Allkopi Bergen