Essays on efficiency and economies of scope and scale in electricity networks

Ørjan Mydland

PhD THESIS

PREFACE

This thesis is submitted in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* (*PhD*) at Norwegian School of Economics (NHH), Department of Business and Management Science. This work has been partly financed by the project "Elbench". I am grateful for the financial support and the input from the participants in the project. I would also thank my employer at INN University who granted me a research scholarship, enabling me to focus on my research.

The fulfillment of this thesis would not have been possible without the help of a number of people. First, I would like to express my gratitude to my supervisors: Mette Bjørndal, Endre Bjørndal, and Gudbrand Lien. I have learned much from the experience of Mette and Endre in research within the electricity markets. I hope we will continue working together in the future. I owe a huge thanks to Gudbrand for giving me the opportunity and encouragement to go into research. Gudbrand's experience as a researcher and his personality have had a great impact on my work.

I am thankful to Subal C. Kumbhakar at State University of New York (SUNY), Binghamton, for all his help and contributions. I enjoyed my stay (together with Gudbrand Lien) in Binghamton. The work on our papers and our meetings in seminars and conferences are always interesting and educative.

I would also like to thank Hilde Marit Kvile and Roar Amundsveen at the Norwegian Water Resources and Energy Directorate for their input, help with data, and not least for all the fun (e.g. "The Penguin").

The support from my colleagues at INN University and the people at NHH has been crucial. Erik Haugom had a great impact at the start of my PhD research. I have been impressed by the way he sets his goals in life and how he carries out the tasks needed to achieve them. I would also like to thank Steinar Veka for his interest in solving problems, even if the problems are not his own. I am grateful to my new colleague, Andrew Musau, for his contribution to my research. Joachim Thøgersen at OsloMet and Per Kristian Alnes at Østlandsforskning contributed by teaching my classes in microeconomics at INN University, and I am grateful to them. This enabled me to complete my thesis, knowing that the students were in good hands. I have also had the pleasure of working with Jonas Andersson at NHH and I am looking forward to finishing our work together.

My father, Håkon, and his wife, Torbjørg, my sister, Linda, my brother, Lars Endre, and my nephew, Stian, have all been supportive of my work, and for that I am grateful. A special thanks to my stepfather, Willy Kristiansen, for all the help he provided and for being so important in my life.

To other family and friends: I have appreciated your interest in what I have been doing and for the good times we had together when I was not working on my thesis. This has stimulated my motivation to complete my work.

The two most important persons in my life are my significant other, Tove Karin, and our daughter, Ine. We have achieved a number of goals in life together and I am very grateful for the support you two have given me in my work. To Ine, who is now a young woman of 11-years old: I am really proud of you. Your personality and the way you tackle obstacles in your life by keeping on working and doing what is expected of you, and more, are no less than inspiring to me. When I was a kid, I sometimes tried to impress the girls I liked the most, sometimes by doing odd things. My technique might have improved, but essentially, I am still the same. I love you both very much!

Finally, my sincerest apologies to anyone whom I have failed to mention but deserves some gratitude.

Lillehammer, October 2018

Ørjan Mydland

CONTENTS

ABSTRACT

The electricity market in Norway has undergone substantial changes in recent decades, which prompts the need for research on how the industry should be organized. In an indivisible electricity industry that, on the one side, consists of market-oriented competitive entities in production and power trading, and, on the other side, natural monopolies within transmission and distribution, it is of interest to perform cost analysis within the productivity and efficiency framework. The electricity industry is complex, owing to the fact that production and consumption must happen simultaneously. After the Energy Act of Norway came into force on January 1, 1991, only transmission and distribution remained regulated. The regulation of transmission and distribution serves to avoid the typical disadvantages arising from natural monopolies. Many countries have gone through the same or similar changes in their respective industries. In a developing and increasingly global industry regarding power trading, more regulation is probably needed, not less. The possible future changes in the power grid—owing to private firms and the ability of households to take advantage of the technological developments in solar and wind generation—will probably also affect the regulating task in the future. The main objective of this thesis is to improve the understanding of efficiency measures and methods, and to increase the knowledge of the market structure in the Norwegian electricity distribution industry.

In **Essay 1** "Economies of scale in Norwegian electricity distribution: A quantile regression approach", we investigate scale economies to see if the structure of the industry affects the costs. If a restructure of the industry would reduce costs for the firms, and, hence, in the industry, it would mean increased productivity and efficiency. In **Essay 2,** "Economies of Scope and Scale in the Norwegian Electricity Industry", we study scale and scope economies. Economies of scope measures the synergies of producing more than one output. Some electricity companies in Norway both generate and distribute electricity. If there exist some positive synergies from producing more than one output, it means that the cost would be higher if two separate firms produced the same amounts of output of each product as the one firm producing both products. In **Essay 3,** "Lost economies of scope and potential merger gains in the Norwegian electricity industry", I investigate what are the potential gains from merging the distribution companies in Norway. Both Essay 1 and Essay 2 state that there are economies of scale in the industry, meaning that the industry would benefit from increasing the size of the companies in terms of increased output. Because the output is given, this means that companies must merge. In Essay 2, we report that there exist economies of scope. Due to the change in the Energy Act of Norway in 2016, we find that the separation of the integrated firms producing both generation and distribution services, would increase costs to the industry and, hence, incur losses by not utilizing economies of scope. If disentangling generation and distribution of electricity would lead to more mergers of the distribution companies, it is of interest to seek the potential gains in terms of cost reductions to the industry from such mergers. In **Essay 4** "Disentangling Costs of Persistent and Transient Technical Inefficiency and Input Misallocation: The Case of Norwegian Electricity Distribution Firms", we focus on the fact that many efficiency studies neglect allocative efficiency, and only concentrate on technical efficiency. In addition, we also disentangle costs of persistent and transient inefficiency, and include determinants for both persistent and transient inefficiency.

INTRODUCTION

1 BACKGROUND AND OBJECTIVES

What is efficiency and productivity analysis? Basically, it is the relation between inputs and outputs in any kind of production of equivalent products or services. If two firms, A and B, have the same output, but firm A has lower inputs than firm B, we know that firm A is more productive than firm B. However, is firm A more efficient than firm B? Yes, in this case, where the two firms produce the same quantities of outputs, we can say that firm A is more efficient than firm B. But what if the two firms do not have the same output? If firm B has higher output than firm A, then firm A is not necessarily more efficient or more productive than firm B. Whereas productivity is simply a performance measure, given by the ratio of output per input for each firm, efficiency measures the relative performance of one firm against other firms. If firm C has the same input as firm A, but lower output, by finding the ratio of the two firms' productivity, we can measure what the output of firm C should be if it was equally productive as firm A. This gives the efficiency measure of firm C. If we can find the maximum output possible for every level of input, we have defined a production possibility frontier. Because this frontier represents the maximum output possible for each input level, it also represents maximum efficiency for each input and output level. If firms A and B are located on the frontier, even with different levels of input and output, both of them will be fully efficient, but one of them is more productive than the other. Every firm that is located below the frontier is less efficient. Figure 1 illustrate this situation, for a one-output and one-input case.

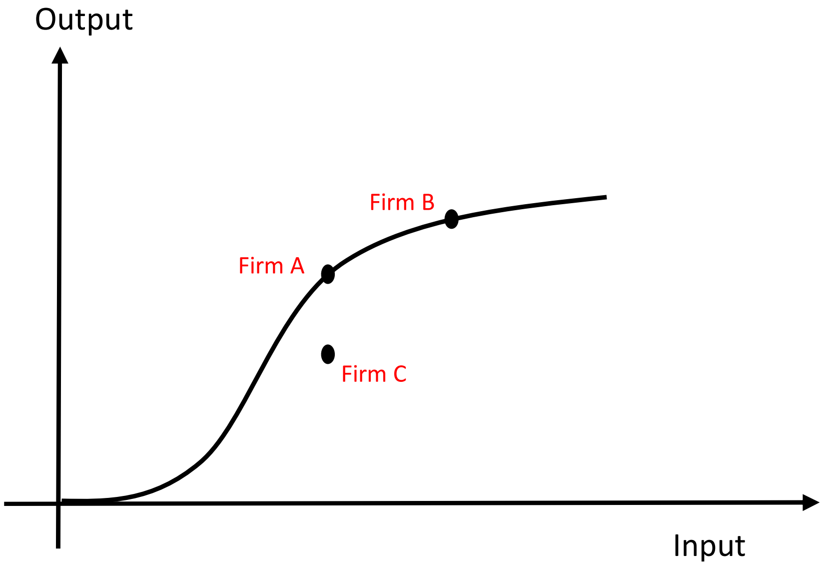


Figure 1. Production frontier, productivity and efficiency.

Firms A and B are located on the frontier, meaning that both are 100% efficient, whereas firm C has the same input as firm A, but less output, and is thereby located below the frontier, which means that the firm is less than 100% efficient. If we find the ratio of output and input for these three firms, we will find that firm A has the highest ratio and is thus the most productive firm.[1]

In practice, it is never really possible to find the "true" frontier, by which we mean that we know that, for each input level, it is not possible to have higher output, as shown in Figure 1. If we have input and output data for several firms in an industry, we can use this information to find a frontier. The firms with the highest ratio of output per input define the frontier. We do not know whether this frontier represents the "true" frontier, but we know that the firms deciding the frontier are the best firms with regard to the output – input ratios in this industry. This kind of productivity and efficiency analysis performed on an industry is often called benchmarking (Bogetoft and Otto, 2010), and refers to the firms that decide the frontier, by setting a benchmark against which every other firm can be measured.

In this thesis we focus on Norwegian electricity distribution companies that are regulated by the Norwegian Water Resources and Energy Directorate (NVE). The regulation model that NVE uses is basically the same as the simple example we outline above. However, in a

[1] Reviews of models and recent applications in general are given in, e.g., Kumbhakar and Lovell (2000), Coelli et al. (2005), Fried et al. (2008) and Kumbhakar et al. (2015a).

benchmarking model used to regulate firms in an industry, there is often more than one input and more than one output. It then becomes more difficult to calculate the productivity and efficiency measures, and more advanced estimation methods must be applied, e.g. Data Envelopment Analysis (DEA) or stochastic frontier analysis (SFA). NVE uses the non-parametric DEA model.[2] Further, the Norwegian electricity distribution companies have more than one output. The model NVE uses has three outputs, *number of network stations*, *kilometers of network,* and *number of customers*. Costs are combined into one single input, *total costs.* For the distribution companies, we assume that the output is exogenously given, meaning that the output is decided by the demand of the customers. This means that a distribution company that is not efficient can only increase their efficiency by reducing cost, e.g. see firm C in Figure 1. If this firm is a Norwegian electricity distribution company, the only way for this company to improve efficiency is to reduce their costs while producing the same output, giving a horizontal right–to–left movement toward the frontier. So, what is the focus in terms of productivity and efficiency analysis when applied to Norwegian electricity distribution companies? As should be clear, it basically concerns the cost of production. The output for each firm is fixed given the exogenously given demand and output. The only variable, in this case, is the costs and, for this reason, studies on the costs of the companies are of interest.

This introduction proceeds as follows. Section 2 starts with a brief description of the Norwegian regulation model and the industry. Further, for each of the essays, I provide some extended explanations on the methodology and the empirical results. In Section 3, implications of data are discussed, and Section 4 provides comments on future research. Section 5 contains concluding remarks.

[2] Amundsveen and Kvile (2015) give a review of the regulation model and application of DEA used by NVE.

2 METHODOLOGY AND RESULTS

Brief description of the Norwegian regulation model and the industry

After the Energy Act of Norway came into force on January 1, 1991, transmission and distribution of electricity, which are considered as natural monopolies, remained regulated whereas production and power trading became more market oriented and competition was introduced. The model used to regulate the Norwegian electricity distribution companies is a form of incentive model with a revenue cap, Coelli et al. (2003).[3, 4]

$$Revenue\ cap = (1-\rho)C + \rho C^* \quad (1)$$

where C represents the actual cost, C^* is the cost norm that is calculated in two steps; and the value of ρ determines the strength of the incentives in the regulation model. From 2009, ρ was set equal to 0.6, meaning that 40% of the revenue cap is decided by the firm's actual cost, and 60% is decided by the cost norm, calculated in the model. Some details in the regulatory model change from year to year, but, from 2007, the model has been mainly unchanged. The actual cost, C, for the company at year t is a combination of reported and calculated costs, based on accounting values from year t-2, (see Bjørndal et al. (2010)). For 2018, the cost norm, C^*, for each company is carried out in two steps, as follows.[5]

Step 1: The efficiency score for each company is calculated using data for each company for year t-2 (2016). These results are measured against the average (industry) efficiency, which is on a frontier obtained from DEA using yearly averages for the period 2012-2016.[6]

Step 2: To account for firm heterogeneity originating from firms operating in different environmental conditions, the DEA results from Step 1 are adjusted by the parameter estimates of environmental variables on costs.

[3] We can also refer to this model as a yardstick model. The regulator uses a yardstick model, using benchmarking methods to assess relative efficiency.

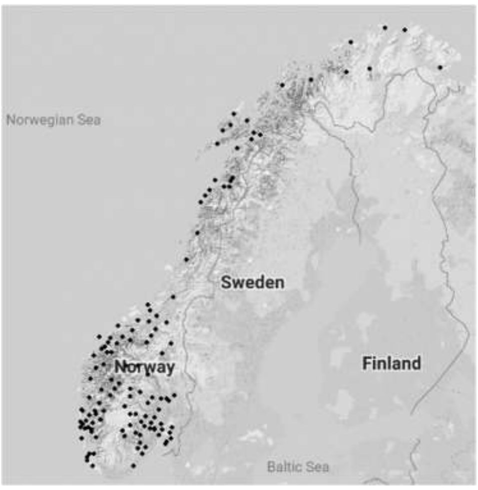
[4] For an overview on the development in the regulation model see Bjørndal et al. (2010) and Amundsveen and Kvile (2015).

[5] This information is retrieved from NVE, see Langseth (2017).

[6] This is done so that the companies can obtain super efficiency, meaning that fully efficient companies can earn more than the normal rate of return, in order to secure increasing efficiency over time, see Bjørndal et al. (2010).

NVE performs a calibration of the results from Step 1 and Step 2 to ensure that the actual cost in the industry is accounted for and that the industry in total receives a rate of return equal to the NVE rate of return (r_{NVE}).[7] Further, the calibration adjusts for the age effect from differences in capital cost in new and old networks. Finally, the revenue cap for year t-2 is based on CPI-adjusted actual costs for year t-4 as an estimate of expected costs at year t-2. The revenue cap in year t (2018) is calibrated by the difference between estimated and actual cost for year t-2 (2016) and the difference is adjusted by r_{NVE}.

Figure 2 gives an overview of the Norwegian electricity distribution industry,[8] from which we notice that *length of network*, *number of customers*, *number of network stations*, *MWh electricity*, and *operational costs* show that the size of the companies vary. The environmental variables describe the differences in the environments that the companies operate in, which can affect the cost of production. These variables also vary, implying that it is important to include them in the analysis to control for firm heterogeneity.



Data	Mean	Min	Max
Length of network, in km	854	9	11,866
Number of customers	25,600	61	696,540
Number of network stations	1,075	4	17,929
MWh electricity	683,508	4,268	18,446,274
Operational costs, 1,000 NOK	29,858	468	695,246
Environmental variables			
Proportion of underground cables	0.38	0.00	1.00
Proportion of sea cables	0.02	0.00	0.36
Proportion of air cables	0.12	0.00	0.40
Average slope of terrain	10.21	2.97	22.22
Average distance to road	64.18	1.00	126.00
Number of islands	2.62	0.00	30.00
Proportion of deciduous forest	0.08	0.00	0.31
Coastal climate	0.23	0.00	4.74

Figure 2. Overview of 120 Norwegian electricity distribution companies in 2016. Location and descriptive statistics.

Similar data are not readily available for other countries. According to the Nordic Energy Regulators, the number of electricity distribution companies in Sweden is 175, with 87 in Finland, and 84 in Denmark. The mean values for market share, based on the number of customers, are low in all three countries: 0.6% in Sweden, 0.8% in Norway, and 1.2% in

[7] NVE rate of return is a regulated rate of return and is determined annually.

[8] The data is collected by NVE. The map is created in R, using the package “ggmap” (Kahle and Wickham, 2013).

Finland. The market share of the three biggest companies in Finland, Norway, and Sweden are 41%, 33%, and 51%, respectively. [9]

Essay 1: Economies of scale in Norwegian electricity distribution: a quantile regression approach

Because the Norwegian electricity distribution industry consists of companies that are natural monopolies within their concession area, we expect to find economies-of-scale properties within the companies in the industry. It is interesting to investigate how the structure of the industry affects the cost in the industry. Large economies of scale in the industry imply that the companies should increase their outputs. However, the companies cannot increase their outputs, because their outputs are exogenously given, decided by demand by customers. To increase the outputs of companies, the companies need to merge. If the policymakers decide to change the structure of the Norwegian distribution industry, it is important that there is knowledge on how the existing structure affects the costs in the industry. By implementing a quantile regression model, we find economies of scale for different firm sizes. To our knowledge, this is the first attempt to study scale economies using quantile regression. By implementing panel data for the period 2000–2013, we retrieve updated information on the economies of scale in the Norwegian industry. Kumbhakar et al. (2015b), the last scale study for Norway, used panel data for the period 1998–2010.[10] Further, different from earlier studies of this industry, we retrieve information on how the economies-of-scale results changed over time for each quantile. Our results state that returns to scale (RTS) increase over time for all quantiles. One interpretation of these results is that the Norwegian electricity distribution companies are too small, and this is becoming increasingly more so over time. The technical explanation is that the firms in the industry are further away from the optimal scale where RTS equals unity. Figure 3 illustrates this situation.

---

[9] Information is retrieved from Nordic Energy regulators, *Economic regulation of electricity grids in Nordic country*, report 7/2011. http://www.nordicenergyregulators.org.

[10] Other economies of scale studies on the Norwegian electricity distribution industry; (e.g., Salvanes and Tjøtta, 1994; Førsund and Kittelsen, 1998; Førsund and Hjalmarsson, 2004; Growitsch et al., 2009; Miguéis et al., 2011)

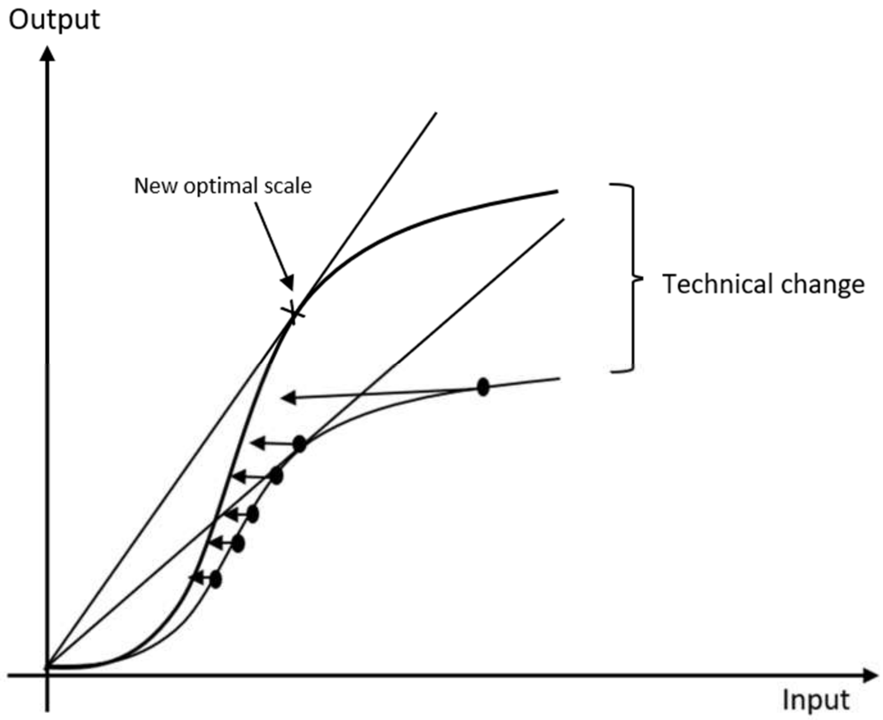


Figure 3. Technical change and movement toward the new frontier.

The black dots in Figure 3 depict firms that initially are located on the frontier. We assume that, due to technical change, the frontier shifts upward. The firms that are now not on the new frontier will try to reduce input/costs to become more efficient. Because the output is exogenously given, the movement toward the new frontier will be a horizontal movement from right to left. Figure 3 shows that the firms are now further away from the new optimal scale on the new frontier. If this scenario repeats itself over time, it explains the results in Essay 1 with increasing RTS over time.

Essay 2: Economies of scope and scale in the Norwegian electricity industry

In the economies of scope literature within studies on electricity companies, the quadratic cost function has been widely adopted (see, e.g. Kaserman and Mayo, 1991; Kwoka, 2002; Jara-Dıaz et al., 2004; Fetz and Filippini, 2010; Arocena et al., 2012; Meyer, 2012). In a dataset for studying economies of scope, there will be some firms producing more than one output, and other firms producing only one output (here output refers to a product or a service). Table 1 gives an example of what a dataset in an economies of scope study looks like.

Table 1. Example dataset in economies of scope study, quadratic cost function

	Raw data			Blanks replaced with zero		
Firm	**Cost x_1**	**Output y_1**	**Output y_2**	**Cost x_1**	**Output y_1**	**Output y_2**
1	156	100	182	156	100	182
2	105	58	52	105	58	52
3	191	84	67	191	84	67
4	170	35	73	170	35	73
5	144	67	45	144	67	45
6	222	112	172	222	112	172
7	194	34	57	194	34	57
8	111	87		111	87	0
9	162	126		162	126	0
10	167	49		167	49	0
11	165	59		165	59	0
12	165	101		165	101	0
13	132	55		132	55	0
14	183		101	183	0	101
15	157		46	157	0	46
16	200		54	200	0	54
17	144		60	144	0	60
18	155		81	155	0	81
19	138		4	138	0	4
20	166		120	166	0	120
Mean	161.35	74.38	79.57	161.35	48.35	55.70

In this example dataset, the raw data consist of firms 1-7 producing both outputs y_1 and y_2. Firms 8-13 are specialized in producing output y_1 and firms 14 – 20 are specialized in producing output y_2. The goal in an economies of scope study is to see if the output per input (cost) is higher or lower for the specialized firms compared to the integrated firms producing both outputs. If the costs are higher for the specialized firms, there exists economies of scope. To estimate (with the standard approach) the costs for all three firm types jointly, the blanks need to be filled in with zero. Then, an assumption on shared technology is imposed on the model, which is not always a suitable assumption. The alternative is to estimate three separate cost functions, one for each firm type. This implies an assumption on different technologies between the firm types, but this is impossible to test. One of the advantages of the flexible technology dummy variable approach, introduced by Triebs et al. (2016), which we adopt in Essay 2, is that it is possible to test for shared technology. Another advantage by using the dummy variable approach is that we estimate a different set of parameters for the zero values. If the number of zero values represents a large proportion of the total number of sample observations, the parameter estimates may be biased (Battese, 1997). One way of explaining this, is that the "blanks" in the output values in the dataset are not really zero, meaning they are non-existing. In this simple example dataset, we see that the mean value of the output values in Table 1 changes quite much when including the zero values. This will of course affect the results. This potential problem does not arise in Triebs et al.'s (2016) approach.

The translog cost function is far less applied in economies of scope studies within electricity markets. The problem is that it is only possible to take the logarithm of positive numbers. This is sometimes solved by replacing the zero values (blanks) in Table 1, with an arbitrarily chosen small number. Table 2 gives an example of this method. By replacing blanks (or zero) values with, say 0.0001, it is possible to take the logarithm of the data. As shown in Table 2, the small numbers may have no or a very small effect on the mean value of the outputs compared to the case where blanks are replaced by zero values in Table 1. However, Fraquelli et al. (2005) show that with a translog specification, economies of scope are very sensitive to the value of the arbitrarily selected small number. Even in this simple example the regression coefficient from estimating x_1 on y_1 and y_2 varies quite considerably when one zero is added or subtracted from the small number. By applying the flexible dummy variable specification, presented in Essay 2, the blanks can be replaced by any positive number, because outputs that do not belong to the specific firm type, decided by the dummy variables, are eliminated in the estimation.

Table 2. Example dataset in economies of scope study, translog cost function

	Raw data			Blanks replaced with 0.00001		Blanks replaced with 0.0001		Blanks replaced with 0.000001	
Firm	**Cost x_1**	**Output y_1**	**Output y_2**	y_1	y_2	y_1	y_2	y_1	y_2
1	156	100	182	100	182	100	182	100	182
2	105	58	52	58	52	58	52	58	52
3	191	84	67	84	67	84	67	84	67
4	170	35	73	35	73	35	73	35	73
5	144	67	45	67	45	67	45	67	45
6	222	112	172	112	172	112	172	112	172
7	194	34	57	34	57	34	57	34	57
8	111	87		87	0.00001	87	0.0001	87	0.000001
9	162	126		126	0.00001	126	0.0001	126	0.000001
10	167	49		49	0.00001	49	0.0001	49	0.000001
11	165	59		59	0.00001	59	0.0001	59	0.000001
12	165	101		101	0.00001	101	0.0001	101	0.000001
13	132	55		55	0.00001	55	0.0001	55	0.000001
14	183		101	0.00001	101	0.0001	101	0.000001	101
15	157		46	0.00001	46	0.0001	46	0.000001	46
16	200		54	0.00001	54	0.0001	54	0.000001	54
17	144		60	0.00001	60	0.0001	60	0.000001	60
18	155		81	0.00001	81	0.0001	81	0.000001	81
19	138		4	0.00001	4	0.0001	4	0.000001	4
20	166		120	0.00001	120	0.0001	120	0.000001	120
Mean	161.35	74.38	79.57	48.3500	55.7000	48.3500	55.7000	48.3500	55.7000
Reg.coef.		-	-	0.0206	0.2268	0.0016	0.0090	0.0012	0.0064

The flexible technology dummy variable approach is applicable to any functional form, e.g., the quadratic cost function:

$$x_1 = y_1 y_{2_{dum}}\left(\alpha_o + \beta_1 y_1 + \beta_2 y_2 + \frac{1}{2}\beta_{11} y_1^2 + \frac{1}{2}\beta_{22} y_2^2 + \frac{1}{2}\beta_{12} y_1 * y_2\right)$$
$$+ \ y_{1_{dum}}\left(\alpha_1 + \delta_1 y_1 + \frac{1}{2}\delta_{11} y_1^2\right)$$
$$+ \ y_{2_{dum}}\left(\alpha_2 + \gamma_1 y_2 + \frac{1}{2}\gamma_{11} y_2^2\right)$$

or the translog cost function:

$$lnx_1 = y_1y_{2_{dum}}\left(\alpha_o + \beta_1(lny_1) + \beta_2(lny_2) + \frac{1}{2}\beta_{11}(lny_1)^2 + \frac{1}{2}\beta_{22}(lny_2)^2 + \beta_{12}(lny_1 * lny_2)\right) + y_{1_{dum}}\left(\alpha_1 + \delta_1(lny_1) + \frac{1}{2}\delta_{11}(lny_1)^2\right) + y_{2_{dum}}\left(\alpha_2 + \gamma_1(lny_2) + \frac{1}{2}\gamma_{11}(lny_2)^2\right)$$

where $y_1y_{2_{dum}}$, $y_{1_{dum}}$ and $y_{2_{dum}}$ represents the dummy variable for the integrated firm with two outputs, the specialized firm with only output y_1, and the specialized firm with only output y_2, respectively.

The results from Essay 2 show evidence of economies of scope and scale. We reject the null hypothesis when testing for shared technology for the different firm types, implying that the standard quadratic cost-function approach is not recommended. Further, in one of the models, we find a clear relationship between firm size and economies of scope, more specifically that the cost of separating the vertically integrated firms is costlier for the smallest firms in the industry.

Essay 3: Lost economies of scope and potential merger gains in the Norwegian electricity industry

In Essay 3 I investigate what are the potential gains from merging the electricity distribution companies in Norway. Most efficiency studies focus on what can be gained by a firm improving, whereas in this study, I focus on the improvement on the industry level from firms merging. Both Essay 1 and Essay 2 state that there exist economies of scale in the industry, meaning that the industry would benefit from increasing the size of the companies in terms of increased output. Because output is given, this means that companies must merge. In Essay 2, we report that there exist economies of scope. Due to the change in the Energy Act of Norway in 2016, we find that the separation of the integrated firms, producing both electricity and distribution services, increases costs to the industry, implying lost economies of scope. If disentangling generation and distribution of electricity will lead to more mergers of the distribution companies, it is of interest to seek the potential gains in terms of cost reductions to the industry from these actions. An efficiency analysis on the Norwegian electricity distribution industry, including the integrated firms that are affected by the amendments in the Norwegian Energy Act, provides answers to the changes in costs to the industry. In a recent study on potential merger gains in the Norwegian electricity industry, Saastamoinen et al. (2017) focus on potential merger gains from companies that are located geographically close. Recently, there have been mergers between Norwegian electricity distribution companies that were not located geographically close.

Figure 4. Map of South Norway. Norgesnett (four merged companies).

Figure 4 illustrates four Norwegian electricity distribution companies, Askøy Nett AS, Gauldal Nett AS, Follo Nett AS and Fredrikstad Nett AS, which from July 1, 2018 merged into Norgesnett AS. As can be seen, these companies are not located near to each other. I provide a method of investigating the optimal merger combination to the industry where the restrictions on proximity are relaxed. However, the results show quite small potential merger gains compared with the loss in not utilizing economies of scope presented in Essay 2.

Essay 4: Disentangling costs of persistent and transient technical inefficiency and input misallocation: The case of Norwegian electricity distribution firms

In section 1 of this introduction, the term efficiency was introduced. Note that this referred to technical efficiency. Cost efficiency (CE) consists of technical efficiency (TE) and allocative efficiency (AE). Formally, the relationship is $CE = TE * AE$. All firms that are located on the frontier, are fully technically efficient. To explain AE, it is useful to start with the standard microeconomic theory on cost minimization. Let us assume that two inputs, x_1 and x_2 are needed to produce output y. The input prices for inputs 1 and 2 are w_1 and w_2, respectively. The condition for cost minimization is given by

$$\frac{MP_{x_1}}{MP_{x_2}} = \frac{w_1}{w_2} \tag{2}$$

where $MP_{x_j}, j = 1,2$, is the marginal product of input j given by $\partial y/\partial x_j$. The marginal rate of technical substitution (MRTS) is given by $\frac{MP_{x_1}}{MP_{x_2}}$ and measures the substitution between the two inputs, which indicates how much input x_2 must increase to keep output y constant if input x_1 is decreased (see Mas-Colell et al. (1995)). In Figure 5 (a), the cost-minimized solution for one firm is described. At point A, the slope of the isoquant, given by the MRTS, equals the slope of the isocost line, given by the input price ratio.[11] In this situation, the firm has no misallocation of inputs and is therefore allocatively efficient. However, it is not possible to tell if this firm is technically efficient without measuring its relative performance against other firms. We do not know the "true" frontier and so we need more firms to find the CE and to identify the TE. In Figure 5 (b), there are six firms (A-F) that produce the same amounts of output. It is important to note that the frontier is not exactly the same as the isoquant in (a). Whereas the isoquant shows different combinations of inputs to produce a fixed amount of output for one firm (or identical firms), the frontier in (b) shows the firms that have the lowest input to produce a fixed amount of output. This means that firms E and F produced the same amount of output as the other firms, but they used more input to do so. Firms A-D are all located on the frontier; hence, they are technically efficient. However, only firm A is cost efficient and, therefore, both technically and allocatively efficient. Firm F is allocatively efficient but not technically efficient. Finally, firm E is neither technically nor allocatively efficient.

[11] $C = w_1x_1 + w_2x_2 \Rightarrow x_2 = \frac{C}{x_2} - \frac{w_1}{w_2}x_1 \Rightarrow \frac{\partial x_2}{\partial x_1} = -\frac{w_1}{w_2}$

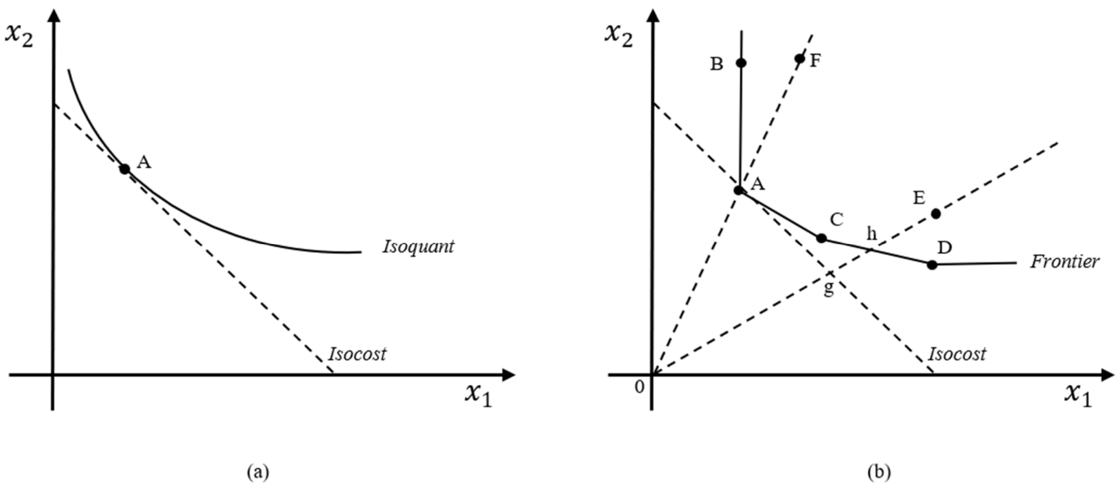


Figure 5. Cost minimization and cost efficiency.

The CE for firm E in Figure 5 (b) can be measured by the ratio $\frac{0g}{0E}$. AE is measured by the ratio $\frac{0g}{0h}$ and TE measures the distance from point E to the frontier given by the ratio $\frac{oh}{0E}$. To summarize, CE can be decomposed in the following way.

$$CE = TE * AE$$
$$\Downarrow$$
$$\frac{0g}{0E} = \frac{0h}{0E} * \frac{0g}{0h}$$

In Essay 4, we report that the costs to the industry arising from input misallocation ranges, on average, from 9.0% to 11.3%. This means that, even if all firms in the Norwegian electricity distribution industry where technically efficient, the cost in the industry would still be 9.0%-11.3% too high.

Essay 4 makes some important contributions on the modeling aspect. Filippini et al. (2018) argue that regulators may fail to set optimal efficiency targets if they are unable to identify systematic shortfalls in managerial capabilities that generate persistent inefficiency and to distinguish these from non-systematic management problems in the short run. Our findings emphasize that future efficiency studies should disentangle persistent and transient technical inefficiency. This is supported by Kumbhakar and Lien (2017). Further, in our study we also include determinants for inefficiency, both in the persistent and transient components of technical inefficiency.

3 IMPLICATIONS ON DATA

As highlighted by Coelli et al. (2003), when doing empirical economics research, the first rule is "garbage inn, garbage out". This applies no matter the method used, whether DEA or SFA. In the four essays presented in this thesis, the data are a crucial factor. It is a great advantage that these data, are also used in the regulation model, meaning that the regulator and the companies in the industry are also interested in the data being correct.[12] From the point of view of the regulator and the companies, it is an advantage that the data are used in research because the researcher probably will control the data carefully. The Norwegian data used in the regulation models are open access, and the model, data, and results are published on the regulator's web site. This is not the case for this industry in all other countries. In the selection of methodology, the availability of data and likely noise in the data play a key role.

A regulator can determine the amount of data, together with the details that the regulated companies are obligated to report. There is a trade-off between the desire to obtain enough data and the details needed for use in the regulation model or in empirical research, and the time and effort that current companies must put in to provide the data required.

Capital can be a challenging input to define in empirical research, especially because the characteristics of capital are not the same as for most other inputs. Most inputs like labor, fuel, and power are purchased and transformed into output within one period of production. However, capital is often transformed into output for many years into the future. This makes it challenging to decide how the input should be measured in each accounting period. In electricity transmission and distribution, investments in the power grid are expected to last for 20-30 years. A common method to allocate the cost for all the years in the life span of capital investments is to use the depreciation cost for each year. However, Coelli et al. (2003, p. 110) states that this might be problematic.

- "Price inflation will make the quantities (that is, the depreciation cost) of new capital items appear larger than identical capital items purchased in previous years.

- Different firms could assume different asset lives or use different depreciation patterns, such as declining balance, or use accelerated depreciation to minimize tax payments."

[12] Note that the data on generation of electricity in Essay 2 "Economies of scope and scale in the Norwegian electricity industry" is not retrieved from data used in the model used to regulate the Norwegian electricity distribution companies.

These are valid points to keep in mind when performing analysis, and evaluating results based on capital measures.

In the regulation model in Norway, NVE uses the following outputs: *number of network stations*, *number of customers* and *kilometers of network*. These variables are the main cost drivers in the industry. Because the outputs are exogenously given, it is reasonable to model cost as input(s). In the current model, all costs are summarized into one input. Table 3 shows the input factors in the current regulation model.

Table 3. Input factors in the DEA model used by the Norwegian regulator[13]

Input	**Input price**
Labor (number of man-years)	Company specific average wage
Capital, book values (NOK)[14]	Depreciation factor + r_{NVE}
Goods and services (NOK)	1
Power losses (MWh)	Base on Nord Pool Spot's system price of power
Value of lost load (VOLL)	1

In applying parametric methods in our estimations, we were concerned about strong multicollinearity between the three outputs. In Essay 1, we dropped *number of network stations* from the analysis to avoid strong multicollinearity. Further, the danger of summing all inputs into one measure of total costs is that we can hide low or negative correlation between some of the cost elements.

It would be interesting for regulators and researchers to have datasets from different countries and to compare the various methods of analysis on these different electricity data. Further, it would be possible to measure the relative performance of the companies in one country against companies in other countries. This would strengthen the estimated efficiency measures, and it could provide useful knowledge about how well the regulators of the companies in the different countries were performing.

Finally, in this section, I would like to comment on the development within business analytics. It is interesting to consider how developments within areas such as "business intelligence,"

[13] Source: Bjørndal et al. (2010).
[14] NOK = Norwegian Kroner. 1 GBP = 10.74 NOK, 1 EUR = 9.59 NOK, 1 USD = 8.19 NOK, on September 15, 2018.

“artificial intelligence,” and “big data” might open up new possibilities within the collection and control of data. Of course, “garbage in, garbage out” still applies. However, we expect that these developments can have significant effects on the efficiency of collection, control, analysis, and reporting of data by researchers, regulators, and companies.

4 FUTURE RESEARCH

An interesting question to address for future research is to investigate which firms in the industry have experienced technical change in recent years, and if there are differences associated with characteristics such as the size of firms.

It would be valuable to conduct an economies-of-scale study on the Norwegian electricity distribution companies within a spatial econometric framework to test the effect of the neighboring companies on scale properties. Orea et al. (2018) present a method in which they combine a spatial econometric approach with SFA to control for unobserved environmental conditions when measuring efficiency of electricity distribution utilities.

Further, it is valuable for future merger analysis to develop a method that facilitates the testing of all possible merger combinations. Currently, there are some problems in performing efficiency analysis on an industry level applying parametric methods such as SFA. It would be worthwhile to expand the range of empirical methodologies other than nonparametric methods such as DEA to further test the existing merger results.

In Essay 4, we include determinants of persistent and transient technical inefficiency. An interesting expansion on this modeling framework would be to include determinants for allocative inefficiency. This would provide important knowledge to the industry on how to identify the misallocation in the existing inputs, and how to find the optimal input mixture to become allocatively efficient.

Finally, it would be interesting to check the ideas for future research that have emerged from this thesis, as well as the answers provided through our analysis, against data from other countries. To check for robustness in the methods and to compare results from the Norwegian electricity industry against other countries, a cross-country study would be of value.

5 CONCLUDING REMARKS

Because the electricity market in Norway and throughout the world has changed rapidly in recent decades, and is likely to change further, there is a need for more knowledge on the market, the industry, regulation models, and on methods on productivity and efficiency analysis. The power network, as we know it today, may well change due to increasing needs for power. Moreover, developments in the technology of solar and wind generation of electricity–which have made remarkable progress, thereby resulting in lower costs–have created greater opportunities for investments by private customers. An increase in the supply of electricity by private investments is likely to affect the organization of the network and the regulation of the industry in the future. It is unlikely that the transmission or distribution network will become redundant, but if new solutions to the network begin to play a larger role in the future, the need to regulate the distribution services is likely to increase to make the system function. The work presented in this thesis expands our knowledge on the electricity industry in Norway and contributes to broadening existing methods of analysis of efficiency and economies-of-scale and -scope studies.

6 REFERENCES

Amundsveen, R. and Kvile, H. M. (2015). The development and application of an incentive regulation model – a balancing act. *The ICER Chronicle* 3: 17–23.

Arocena, P., Saal, D. S. and Coelli, T. (2012). Vertical and horizontal scope economies in the regulated US electric power industry. *Journal of Industrial Economics* 60(3): 434–467.

Bjørndal, E., Bjørndal, M. and Fange, K. A. (2010). Benchmarking in Regulation of Electricity Networks in Norway – An Overview [Internet]. I: Energy, Natural Resources and Environmental Economics. Heidelberg: Springer, 317–342. DOI: 10.1007/978-3-642-12067-1_18

Bogetoft, P. and Otto, L. (2010). Benchmarking with DEA, SFA, and R (Vol. 157). Springer Science & Business Media. Springerhttp://dx.doi. org/10.1007/978-1-4419–7961-2_10.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J. and Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*, 2nd Ed. Springer Science & Business Media. DOI: 10.1007/b136381

Coelli, T., Estache, A., Perelman, S. and Trujillo, L. (2003). A primer on efficiency measurement for utilities and transport regulators. Tech. Rep. 129, World Bank Publications

Fetz, A. and Filippini, M. (2010). Economies of vertical integration in the Swiss electricity sector. *Energy Economics* 32(4): 1325–1330.

Filippini, M., Greene, W. and Masiero, G. (2018). Persistent and transient productive inefficiency in a regulated industry: electricity distribution. *Energy Economics* 69: 325–334.

Fraquelli, G., Piacenza, M. and Vannoni, D. (2005). Cost savings from generation and distribution with an application to Italian electric utilities. *Journal of Regulatory Economics* 28(3): 289–308.

Fried, H. O., Lovell, C. A. K. and Schmidt, S. S. (Eds.). (2008). *The measurement of productive efficiency and productivity growth*. Oxford University Press, New York.

Førsund, F. R. and Hjalmarsson, L. (2004). Are all scales optimal in DEA? Theory and empirical evidence. *Journal of Productivity Analysis* 21: 25–48.

Førsund, F. R. and Kittelsen, S. A. C. (1998). Productivity development of Norwegian electricity distribution utilities. *Resource and Energy Economics* 20: 207–224.

Growitsch, C., Jamasb, T. and Pollit, G. P. (2009). Quality of service, efficiency and scale in network industries: An analysis of European electricity distribution. *Applied Economics* 41: 2555–2570.

Jara-Dıaz, S., Ramos-Real, F.J. and Martınez-Budrıa, E. (2004). Economies of integration in the Spanish electricity industry using a multistage cost function. *Energy Economics* 26(6): 995–1013.

Kahle, D. and Wickham, H. (2013). ggmap: *Spatial Visualization with ggplot2*. R Journal, 5(1).

Kaserman, D. L. and Mayo, J. W. (1991). The measurement of vertical economies and the efficient structure of the electric utility industry. *Journal of Industrial Economics* 39(5): 483–502.

Kumbhakar, S. C. and Lien, G. (2017). Yardstick regulation of electricity distribution – Disentangling short-run and long-run inefficiency. *Energy Journal* 38(5): 17–37.

Kumbhakar, S. C., Wang, H. and Horncastle, A. P. (2015a). *A practitioner's guide to stochastic frontier analysis using Stata*. Cambridge University Press.

Kumbhakar, S. C., Amundsveen, R., Kvile, H. M. and Lien, G. (2015b). Scale economies, technical change and efficiency in Norwegian electricity distribution, 1998–2010. *Journal of Productivity Analysis* 43(3): 295–305.

Kumbhakar, S. C. and Lovell, C. K. (2000). *Stochastic production frontier*. Cambridge University Press.

Kwoka, J. E. (2002). Vertical economies in electric power: Evidence on integration and its alternatives. *International Journal of Industrial Organization* 20(5): 653–671.

Langset, T. (2017). Infoskriv ETØ-4/2017: Om beregning av inntektsrammer og kostnadsnorm for 2018. Oslo.

Mas-Colell, A., Whinston, M. D. and Green, J. R. (1995). *Microeconomic theory* (Vol. 1). New York: Oxford University Press.

Meyer, R. (2012). Economies of scope in electricity supply and the costs of vertical separation for different unbundling scenarios. *Journal of Regulatory Economics* 42(1): 95–114.

Miguéis, V. L., Camanho, A. S., Bjørndal, E. and Bjørndal, M. (2011). Productivity change and innovation in Norwegian electricity distribution companies. *Journal of the Operational Research Society* 63(7): 982–990.

Orea, L., Álvarez, I. C. and Jamasb, T. (2018). A spatial stochastic frontier model with omitted variables: electricity distribution in Norway. *Energy Journal* 39(3): 93–116.

Salvanes, K. G. and Tjøtta, S. (1994). Productivity differences in multiple output industries: an empirical application to electricity distribution. *Journal of Productivity Analysis* 5: 23–43.

Saastamoinen, A., Bjørndal, E. and Bjørndal, M. (2017). Specification of merger gains in the Norwegian electricity distribution industry. *Energy Policy* 102: 96–107.

Triebs, T. P., Saal, D. S., Arocena, P. and Kumbhakar, S. C. (2016). Estimating economies of scale and scope with flexible technology. *Journal of Productivity Analysis* 45(2): 173–186.

ESSAYS OF THE THESIS

ESSAY 1

Ørjan Mydland, Erik Haugom, Gudbrand Lien

Economies of scale in Norwegian electricity distribution: a quantile regression approach

Applied Economics, forthcoming

APPLIED ECONOMICS, 2018
https://doi.org/10.1080/00036846.2018.1450481

Economies of scale in Norwegian electricity distribution: a quantile regression approach

Ørjan Mydland[a,b], Erik Haugom[a] and Gudbrand Lien[a]

[a]Faculty of Economics and Social Sciences, Inland Norway University of Applied Sciences, Lillehammer, Norway; [b]Department of Business and Management Science, Norwegian School of Economics, Bergen, Norway

ABSTRACT
In this article, we investigate scale economies in Norwegian electricity distribution companies using a quantile regression approach. To the best of our knowledge, this is the first attempt to apply this estimation technique when analysing scale economies. We estimate the cost elasticities of the two output components: *network length* and *number of customers*, to calculate returns to scale. Our results show large potential of scale economies, particularly for the smallest companies. We also find that returns to scale is increasing over time. These findings have important implications for policymakers when they are deciding the structure of the industry in the future.

KEYWORDS
Norwegian electricity distribution companies; policymakers; economies of scale; returns to scale; quantile regression

JEL CLASSIFICATION
C23; D00; D39; D42; D29

I. Introduction

In this article, we investigate whether there are any unexploited economies of scale among Norwegian electricity distribution companies across company size. To address this issue, we apply a quantile regression model to a unique data set with detailed information on the characteristics of the distribution companies over the period from 2000 to 2013. To the best of our knowledge, this is the first attempt in the literature to apply this method when analysing economies of scale. The quantile regression approach is particularly useful in our context as it allows us to examine the economies of scale across the whole distribution of companies. That is, while ordinary least square estimates only provide the mean effect of the output components, and a corresponding mean estimate of any returns to scale, a quantile regression model provides the corresponding effects across all quantiles of interest (see, e.g. Davino, Furno, and Vistocco 2014). It is, thus, a more flexible modelling approach, which also will provide more relevant results for policymakers as the effects can be examined directly by company size.

Our results show that there is a large potential for economies of scale in the Norwegian electricity distribution sector, and particularly so for small companies. That is, the returns to scale estimates based on our quantile regression are, in general, higher for the lower quantiles (the smallest companies measured in total costs), compared to the higher quantiles (the biggest companies). This means that the efficiency gains from increasing the output components for small companies are higher than for big companies. An implication of this finding is that it is beneficial from a productivity point of view that small companies should increase in size before big companies do so. Put differently, small companies should either expand or merge.

Our results also show that the returns to scale has increased over time. This means that the companies, in general, have become less productive and/or efficient under the given regulations. This finding holds across all quantiles, but particularly so for big companies. For policymakers, this means that there may be a need to revise the current regulations to change this negative trend.

A range of studies already exist concerning economies of scale in the electricity distribution industry (e.g. Giles and Wyatt 1993; Filippini 1996; Burns and Weyman-Jones 1996; Kumbhakar and Hjalmarsson 1998; Yatchew 2000; Kwoka 2005; Farsi et al. 2010; Tovar and Ramos-Real 2011; Alaeifar, Farsi, and Filippini 2014) from Norway (e.g. Salvanes and

CONTACT Ørjan Mydland orjan.mydland@inn.no Faculty of Economics and Social Sciences, Inland Norway University of Applied Sciences, Pb 952, NO-2604 Lillehammer, Norway

Tjøtta 1994; Førsund and Kittelsen 1998; Førsund and Hjalmarsson 2004; Growitsch, Jamasb, and Pollit 2009; Miguéis et al. 2011; Kumbhakar et al. 2015). The findings of these studies are somewhat mixed, both around the world and in Norway, but most provide at least some evidence of unexploited scale economies, including the recent Norwegian study of the period 1998–2010 by Kumbhakar et al. (2015).

In this study, we model the electricity distribution sector as a single-input multi-output production process, characterizing the production process with a cost function. To control for firm heterogeneity, we introduce several environmental variables in our models. Moreover, as the effects from the multi-output production process may vary in size and nature across the distribution of the total costs (TOTEX), we use a quantile regression approach. This method allows us to estimate a range of conditional quantile functions and, hence, provides a more complete picture of the conditional density of the covariate effects.'[1] [2]

The remainder of the article is organized as follows. Section II gives an overview of the Norwegian electricity distribution industry. Section III presents model specification. Section IV describes the Norwegian electricity distribution data followed by discussion of the econometric models and cost elasticities in Section V. Empirical results are presented in Section VI and Section VII provides our concluding comments.

II. The Norwegian electricity distribution industry

The Norwegian electricity sector has undergone significant reorganization and restructuring. Until 1991, Norwegian national, county and municipal governments largely owned the sector, with electricity generation, wholesale and retail supply, and transmission and distribution activities more or less woven together. The basic premise of the 1991 restructuring under the *Energy Act* was to unbundle services in the value chain of delivering electricity to consumers and to expose some parts of the industry, including electricity generation, wholesale and retail supply, to greater competition. As the distribution of electricity displays the characteristics of a natural monopoly, this part of the industry continued to consist of monopoly producers, with each distribution company operating its own concession area (see Salvanes and Tjøtta 1998). As part of the 1991 *Energy Act*, Norway's regulatory agency, the Norwegian Water Resources and Energy Directorate (in Norwegian: *Norges vassdrags- og energidirektorat* or NVE), regulates the distribution companies. The purpose of the regulation is to introduce competition through the regulation model given that the companies do not actually compete with each other directly. The regulation of the distribution companies works through a revenue cap system. NVE uses Data Envelopment Analysis to calculate each company's efficiency level. The most efficient companies constitute a production possibility frontier and all of the other companies get a measure of efficiency in relation to this frontier. The efficiency scores determine 60% of the revenue cap which the regulator calculate for each company. As discussed by Kumbhakar et al. (2015), until 2013, the revenue cap regulation system was specified in such a way that it discouraged mergers among the distribution companies. The current structure of the electricity distribution network is partly the result of the preexisting structure of the locally owned, vertically integrated and regulated power sector prior to the 1991 reforms. Consequently, the discussion above and the ongoing policy debate about the structure of the electricity distribution network in Norway (e.g. Reiten, Sørgard, and Bjella 2014) make it interesting to question whether there are too many distribution companies in Norway and how we can best describe the economies of scale across distribution company size.

III. Model specification

Based on NVE's current regulatory model, firm outputs in the Norwegian electricity industry consist of the numbers of costumers, the length of wires to transport electricity and the amount of electricity

[1]For example, the effects from the various outputs (such as length of network or number of customers) could be different in the lower and upper tail of the TOTEX distribution.

[2]The Skewness and kurtosis indicate that our dependent variable, total costs (TOTEX) and ln(TOTEX) are not normally distributed. This is supported by the Shapiro–Wilk test which rejects the null-hypothesis of normal distribution in both TOTEX and ln(TOTEX). See Table A1 in the appendix for test statistics.

delivered. We assume that the outputs of distribution companies are exogenous, whereas cost is endogenous. The cost function is then given by

$$C = f(Y_i) \quad (1)$$

where C is the total costs and Y_i is the output variables. By taking the natural logarithm of Equation (1), we obtain

$$\ln C = ln\ f(Y_i) + v \quad (2)$$

The above cost function is made stochastic by including the standard error term v. We define returns to scale (RTS) as follows:

$$RTS = \frac{1}{\sum_i \partial lnC/\partial lnY_i} \quad (3)$$

where $\sum_i \partial lnC/\partial lnY_i$ is the sum of the cost elasticity's for input 1 and 2.

By adding a time component (t) to the model, we can find technical change (TC), defined as

$$TC = -\frac{\partial lnC}{\partial t} \quad (4)$$

If TC is larger, equal or smaller than zero, it means there exists positive, zero or negative technical change, respectively.

IV. Data

The data comprise economic and technical information on Norwegian electricity distributors from 2000 to 2013, as collected by the NVE and historically used to implement income regulation in the industry. In total, there are 1750 firm-year observations, constituting an unbalanced panel of 133 Norwegian distribution companies.

We specify a model with a cost function with two outputs based on NVE's current regulatory model. The outputs are the *length of network* (N) and the *number of customers* (Q), representing the main cost drivers in the industry.[3] The output variable *length of network* is the length of the (high-voltage) distribution network in kilometres. The *number of customers* is the total number of entities (both households and firms) that pay the net rent tariff. The endogen variable in our model is *total costs* (*TOTEX*), and include capital expenditure (*CAPEX*), controllable operational expenditure (*OPEX*) and external costs of interruptions for customer.[4] For the entire industry in 2013, total costs were about 12.2 billion Norwegian kroner (NOK),[5] while average total cost per company was approximately 100 million NOK.[6] Furthermore, in 2013, total cost for the largest company (Hafslund) was 1.4 billion NOK and only about 3.1 million NOK for the smallest company (Modalen Kraftlag).

Table 1 presents descriptive statistics for the total cost and output variables, all of which exhibit significant dispersion (large standard deviations). We include various environmental variables (represented by the Z variables in Table 1), which are likely to have an impact on each firm's total costs. However, the output variables N and Q will also, to some extent, capture the heterogeneity of firms. For instance, it is obvious that the *length of network* relative to the *number of customers* will be higher in rural areas than in urban ones, because rural areas generally have a lower population and a more dispersed pattern of settlement. By including the environmental variables, *proportion of underground cables* ($Z1$), *proportion of sea cables* ($Z2$), *proportion of air cables* ($Z3$), *average slope in terrain* ($Z4$), *average distance to road* ($Z5$), *number of islands* ($Z6$), *proportion of deciduous forest* ($Z7$) and *coastal climate* ($Z8$), we control better for firm-specific costs and thus firm heterogeneity.

V. Econometric models and cost elasticities

We use panel data, but to simplify the notation, we omit the subscript i $(i = 1, 2, \ldots, N)$ (indicating the distribution company) and the subscript t $(t = 1, 2, \ldots T)$ (indicating time). However, we

[3]In the regulatory model of NVE, the number of network stations (NT) is also included. However, as one referee pointed out, there is a strong multicollinearity between the output variables. The variance inflation factors (VIF) between the output variables reduce from (NT = 42.76), (Q = 18.04), (N = 16.92) to (Q = 6.60), (N = 6.61) when we drop *number of network stations* from the model to avoid strong multicollinearity.

[4]CAPEX includes annual depreciations and return on book values including 1% working capital. OPEX includes operational and maintenance costs. External costs of interruptions for customers consist of a cost of interrupted effect in the powerlines and losses per MWh (300 NOK). All distribution companies in Norway are by law obligated to report numbers on production and costs; see the website of the Norwegian Ministry of Petroleum and Energy: https://www.regjeringen.no/no/dokumenter/forskrift-om-okonomisk-og-teknisk-rappor/id507169/.

[5]1 USD = 7.72 NOK, 1 EUR = 9.56 NOK per 1 February 2018.

[6]All costs are in 2010 NOK. The evolution of total costs from 2000 to 2013 is presented in Table A2 in the appendix.

Table 1. Descriptive statistics.

Variables	Label	Mean	SD	Min	Median	Max
Length of network, in km	*N*	714	1188	10	306	8744
Number of customers	*Q*	19,601	50,837	348	6037	570,179
Total costs, 1000 NOK (2010)	*TOTEX*	88,439	177,682	2400	32,841	1,748,090
Year	*t*			2000		2013
Environmental variables						
Proportion of underground cables	Z1	0.30	0.19	0.00	0.25	1.00
Proportion of sea cables	Z2	0.02	0.04	0.00	0.00	0.37
Proportion of air cables	Z3	0.12	0.10	0.00	0.12	0.40
Average slope of terrain	Z4	10.21	3.65	2.97	9.95	22.22
Average distance to road	Z5	64.18	36.20	1.00	65.00	126.00
Number of islands	Z6	2.62	5.49	0.00	0.00	30.00
Proportion of deciduous forest	Z7	0.08	0.09	0.00	0.03	0.31
Coastal climate	Z8	0.23	0.64	0.00	0.02	4.74

include the time variable (t) in the model specification.

In our models to be estimated, we use $C = TOTEX$. Outputs Y_1 *and* Y_2 are now represented by N and Q, respectively.

We express the econometric specification of the translog function (TL) as

$$\begin{aligned}\ln C = \alpha + \beta_1 \ln N + \beta_2 \ln Q + 0.5\beta_3(\ln N)^2 \\ + 0.5\beta_4(lnQ)^2 + \beta_5(\ln N \ln Q) + \beta_6 t \\ + 0.5\beta_7(t)^2 + \beta_8(\ln Nt) + \beta_9(\ln Qt) \\ + \beta_{10}Z1 + \beta_{11}Z2 + \beta_{12}Z3 + \beta_{13}Z4 \\ + \beta_{14}Z5 + \beta_{15}Z6 + \beta_{16}Z7 + \beta_{17}Z8 \\ + v\end{aligned} \tag{5}$$

where $\ln N$, and $\ln Q$ are vectors of *length of network* and *number of customers*, for all distribution companies over the sample period, respectively. α and β_1–β_{17} are parameters to be estimated. $Z1$–$Z8$ are control variables (the environmental variables).

Using this specification, we calculate the elasticities for each cost driver as follows:

$$\frac{\partial \ln C}{\partial lnN} = \varepsilon_N = \beta_1 + \beta_3 \ln N + \beta_5 \ln Q + \beta_8 t \tag{6}$$

$$\frac{\partial \ln C}{\partial lnQ} = \varepsilon_Q = \beta_2 + \beta_4 \ln Q + \beta_5 \ln N + \beta_9 t \tag{7}$$

and the corresponding RTS is given by

$$RTS_{TL} = \frac{1}{(\varepsilon_N + \varepsilon_Q)} \tag{8}$$

Firm heterogeneity, unobserved factors and fixed effects

The data used in this study include several firm-specific environmental variables. By introducing all the eight environmental variables, we control for firm heterogeneity. One could argue that to fully control for possible unobserved factors such as economic, organizational, regulatory and additional environmental factors, one should estimate a fixed effect model. However, our main focus in this article is to show that the quantile regression approach is a better method to analyse economies of scale for various outputs components by company size. To introduce fixed effects to the quantile estimation is problematic due to the number of observations. Hence, we estimate our models by including the environmental variables that control for firm heterogeneity. To check the robustness of our results, we estimate three different models: fixed effect, true fixed effect stochastic frontier analysis (SFA) and finite mixture model (latent class model). The estimated cost elasticities and RTS from each of the models are presented in Table A7 in the Appendix.

Quantile regression and returns to scale

The estimate of RTS in Equation (8) draws on the conditional mean (OLS) estimates in Equation (5). However, in our case, it is particularly interesting to consider the effects in the tails of the distribution of the dependent variable, which will reveal the range of differences in the individual covariates and the range of the values of RTS more generally.

In the general case, the simple linear quantile regression model is as follows:

$$Y_{i,q} = \alpha_q + \beta_q X_i + \varepsilon_{i,q} \tag{9}$$

where the distribution $\varepsilon_{i,q}$ is left unspecified.[7] The expression for the conditional q (quantile), $0 < q < 1$, is defined as any solution to the minimization problem (see Koenker and Bassett 1978):

$$\min_{\alpha,\beta} \sum_{n}^{i=1} \left(q - 1_{Y_i \le \alpha_q + \beta_q X_i}\right)\left(Y_i - \left(\alpha_q + \beta_q X_i\right)\right) \tag{10}$$

where

$$1_{Y_i \le \alpha_q + \beta_q X_i} = \begin{cases} 1 \; if \; Y_i \le \alpha_q + \beta_q X_i \\ 0 \text{ otherwise} \end{cases} \tag{11}$$

The least absolute error (the conditional median) is a special case, but the quantile regression method explicitly allows us to model all relevant quantiles of the distribution of the dependent variable.

In our case, we define the quantile regression version of the TL cost function using the following notation:

[7]Not to be confused with the standard error term from OLS, as the distributional properties are not intended to meet the same criteria as standard regression models.

$$\ln C_q = \alpha_q + \beta_{1,q}\ln N + \beta_{2,q}\ln Q + 0.5\beta_{3,q}(\ln N)^2 + 0.5\beta_{4,q}(lnQ)^2 + \beta_{5,q}(\ln N \ln Q) + \beta_{6,q}t + 0.5\beta_{7,q}(t)^2 + \beta_{8,q}(\ln Nt) + \beta_{9,q}(\ln Qt) + \beta_{10,q}Z1 + \beta_{11,q}Z2 + \beta_{12,q}Z3 + \beta_{13,q}Z4 + \beta_{14,q}Z5 + \beta_{15,q}Z6 + \beta_{16,q}Z7 + \beta_{17,q}Z8 \quad (11)$$

and the corresponding RTS for each quantile as

$$RTS_{TL,q} = \frac{1}{(\varepsilon_{N,q} + \varepsilon_{Q,q})} \quad (12)$$

where q is a given quantile between 0 and 1. The two elasticities ($\varepsilon_{N,q},\ \varepsilon_{Q,q}$) are calculated in the same way as Equations (6) and (7), but for each quantile.

VI. Results and discussion

Results for OLS TL function

We initially tested the translog cost function model against the restricted and more parsimonious Cobb–Douglas cost function model. Statistical testing, using Wald likelihood-ratio tests (Wald 1943), rejected the Cobb–Douglas model at the 1 per cent level of significance. To account for heterogeneity across distribution companies, we included firm-specific environmental (Z) variables. This is because it is important to account for firm-specific cost factors, including the impact on costs of demographic, geographic and climatic factors (Growitsch, Jamasb, and Wetzel 2012).

As shown in Table 2, the output elasticity of *length for network* is 0.374, meaning that if you increase the output value *length for network* by 1 per cent, the total costs will increase by 0.374 per cent. For *the number of customers*, the cost elasticity is 0.544.[8]

Evaluated at the means of the variables, RTS exceed unity (1.089).[9] This suggests the presence of scale economies. In this context, our results support existing findings for the electricity distribution sector in Norway (Førsund and Hjalmarsson 2004; Growitsch, Jamasb, and Wetzel 2012; Kumbhakar et al. 2015) and in Sweden (Kumbhakar and Hjalmarsson 1998). However, as depicted by the histogram in Figure 1, there is a large variation in the RTS estimates across the companies. We investigate this further by applying quantile regression on our model.

Table 2. Cost elasticities and returns to scale.

Components	Mean	Std. error
Cost elasticity with respect to output		
Length of network	0.374	0.009
Number of customers	0.544	0.013
Returns to scale (RTS)	1.089	0.001
Technical change (TC)	0.001	0.001

Heteroscedasticity consistent standard errors (Davidson and Mackinnon, 1995).

In Table 2, we also present the value of technical change. The value is small, but positive, meaning that the production possibility frontier 'shifts up' indicating that at least some of the distribution companies gets more productive over time.

Results for quantile TL function

We estimate the model in Equation (12) for 19 quantiles ranging from 0.05 to 0.95. From all of the observations in each quantile (1750 × 19), we calculate the elasticities and take the median elasticity for each quantile.[10,11]

The upper panels in Figure 2 depict the parameter estimates for each of the two cost drivers. That is, for each of the two coefficients, we plot 19 distinct quantile estimates. For each quantile, we interpret these point estimates as the impact of a one-unit change in the covariate on *total costs, ceteris paribus*. We also plot 95% confidence intervals for the quantile regression estimates.[12] In the traditional OLS approach, we would obtain the average percentage

[8]Data and model specifications are available from the authors upon request. Test statistics are available in Table A3 in the Appendix.
[9]RTS > 1 is to be interpreted as follows: if you double your inputs, you will more than double your output. At optimal scale RTS = 1, meaning that if you double your inputs, your outputs will also double.
[10]For the interested reader, tables with mean, median, standard errors, p-values and confidence intervals for each quantile for each cost elasticity and RTS are to be found in the Appendix, see Tables A4–A6.
[11]Note that since we apply panel data from 133 firms, ranging from year 2000 to year 2013, some of the firms that have changes in input and output values can appear in different quantiles. However, since our goal is to estimate cost elasticities and RTS on the industry, this will not violate the results obtained in our estimation. If someone would like to use this method to derive firm-specific measures, this issue should be handled with caution. By reducing the number of quantiles, the number of 'shifts' will be reduced.
[12]The 95% confidence intervals are computed by the Delta-method. The Delta-method is a method for deriving the variance of a function of asymptotically normal random variable with known variance using a Taylor series expansion; see https://cran.r-project.org/web/packages/modmarg/vignettes/delta-method.html#fn7 for more details.

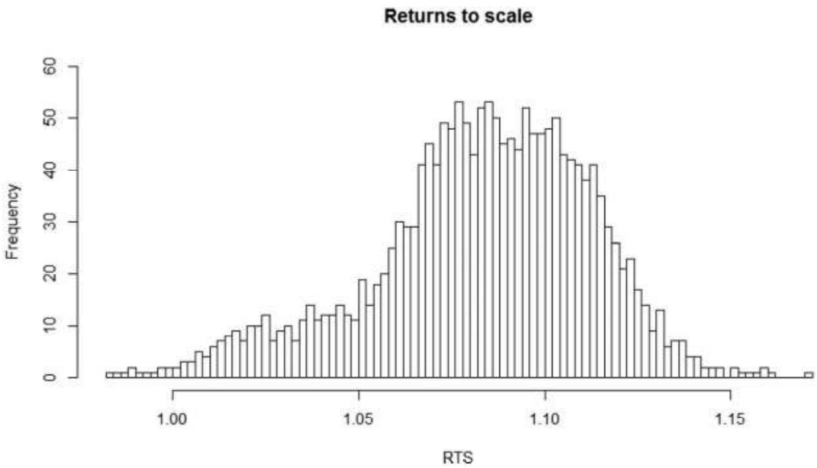


Figure 1. Histogram of returns to scale for all companies, estimated using the OLS translog function.

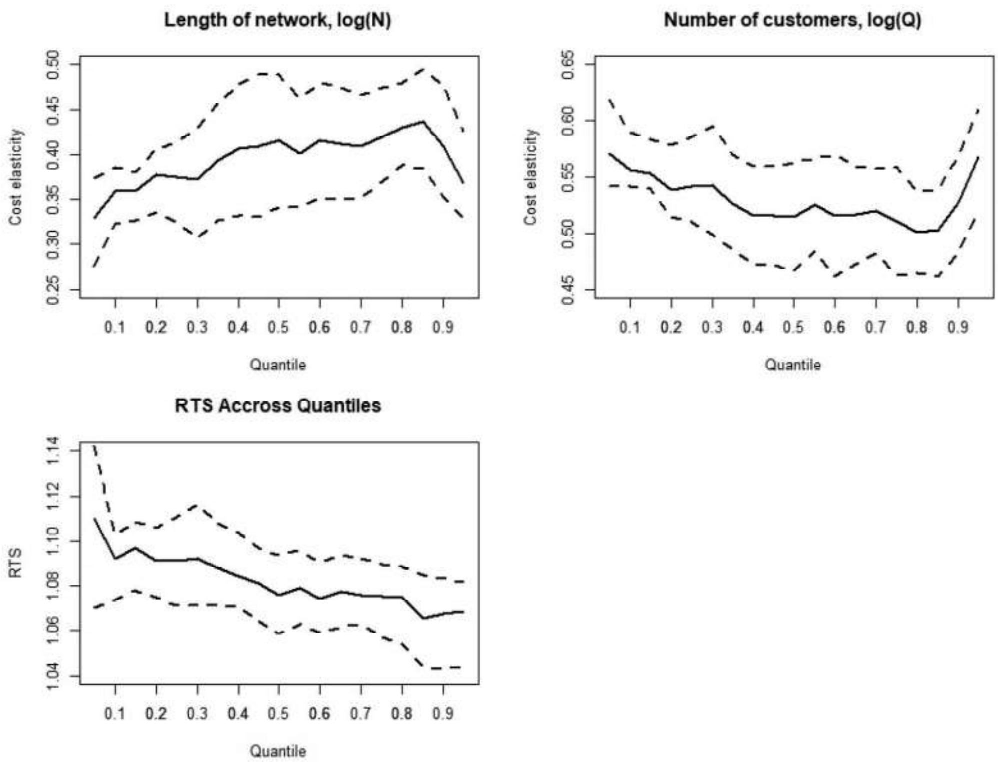


Figure 2. Median cost elasticity of outputs (multiplied by −1) estimates and RTS across quantiles, based on the TL function. The scattered lines are 95% confidence intervals.

cost increase for a 1 per cent increase in the cost driver. For example, an OLS parameter of 0.2 (in absolute terms) for the variable measuring *length of network* (log(N)) could be interpreted as meaning that for a 1 per cent increase in the *length of network*, total costs increase on average by 0.2 per cent, *ceteris paribus*. However, with our quantile regression approach, we can now see how much total costs will increase for a 1 per cent increase in any of the covariates within each quantile.

As shown in the upper left panel of Figure 2, the parameter estimates of the *length of network* range from about 0.33 for the lower quantiles to about 0.44 for the upper quantiles. For the highest quantile, the estimated cost elasticity for the *length of* network is 0.37. There is a small upward trend, meaning that the increase in costs associated with a 1 per cent increase in the *length of network* is somewhat lower for small distribution companies. The *length of network* describes the size of the network for each firm and an intuitive economic interpretation of this can be that small distribution companies have spare capacity, while larger distribution companies are producing nearer their capacity limit.

In the upper right panel, we present the quantile regression estimates for the *number of customers*. We see, from the figure, that there are only minor differences in the estimated cost elasticities across the quantiles. For the smallest quantile and the highest quantile, the range is 0.571 and 0.568, respectively. The minimum cost elasticity of 0.501, we find in the 0.8 quantile, gives the curve a somewhat u-shape. It is likely that the smaller companies generally have their concession area in rural areas, where the population is lower and settlement is more scattered. Increasing the number of customers can then be more costly if any new customers are located far from existing customers. In urban areas where the larger distribution companies are typically located, it is more common for households to reside in closely located apartment buildings and houses, so the dispersion of residences is generally over a much smaller area. This can explain the downward trend of the cost elasticities. However, we see that for quantiles above 0.8, the biggest companies, the cost elasticities are increasing with firm size.

The lower right panel reports the RTS across quantiles. As shown, the results indicate that there is a potential for scale economies, particularly for small companies. The RTS exceeds one for all quantiles but is larger in magnitude for the smallest companies. In this situation, firms should increase production. If we consider the Norwegian electricity distribution industry as an autarky with fixed exogenous given demand, the only way for distribution companies to increase their production is to merge with other distribution companies. Based on these plots, we can conclude that there is large potential for scale economies among Norwegian electricity distribution companies, and particularly for small companies.[13]

We have checked the robustness of our results by estimating three different models: *fixed effect*, *true fixed effect SFA* and *finite mixture model* (latent class model). We find that the mean values of cost elasticities and RTS are robust across the three model specifications. RTS exceeds unity in all models.[14]

RTS development across quantiles over time

In Figure 3, we present the development in RTS across the quantiles over time. As shown, the difference in RTS for small distribution companies (lower quantiles) compared with large distribution companies (higher quantiles) appears to decline over time. For example, the range in RTS for lower/higher quantiles is larger in 2000 than in 2013. The RTS for each quantile exceeds one in all time-periods. This supports our earlier results, suggesting that smaller distribution companies should merge.

Figure 3 also show that RTS is increasing over time, which is an interesting result. If we regard firms with RTS greater than one as being too small, then an increasing RTS over time suggests that the firms are getting smaller and smaller relative to the optimal scale over time.[15] One interpretation of this could be that only some of the companies cause the technical change (reported in Table 2). All the other companies will then be less efficient as the frontier shifts up. Due to the regulation, they have incentives to increase their efficiency. To get more efficient, the companies have to reduce their inputs and keep the outputs constant. We regard the demand for distribution of electricity services as given, so to increase their outputs is not an option. As the companies move towards the new frontier as they are getting more efficient by reducing their inputs, they end up further away from the point of optimal scale. This will lead to an increase in RTS.[16]

[13]All Data and R-code used in this analysis are available on request.
[14]The results are presented in Table A7 in the Appendix.
[15]We would like to thank the referee who correctly pointed out that the increase in RTS over time could be affected by firms merging over time. Since we use unbalanced panel data in our analysis, this could be the case. However, to check this we have also run the model with a balanced panel data, and the effect on RTS increasing over time is higher; see Figure A1 in the appendix.
[16]There are several reasons that firms experience different RTS (see, e.g. Coelli et al. 2005).

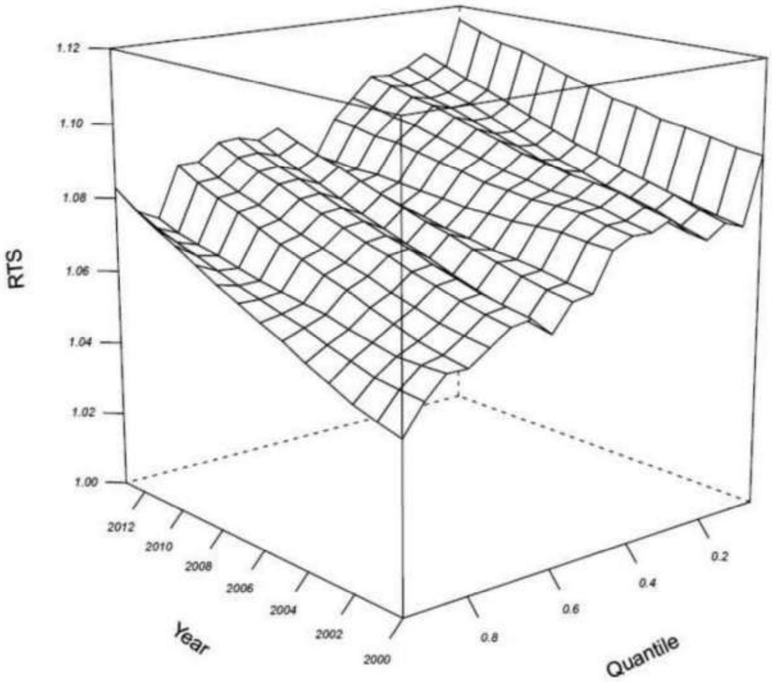


Figure 3. Returns to scale (RTS) across quantiles and over time.

VII. Concluding comments

In this study, we investigate scale economies in Norwegian electricity distribution companies using a quantile regression approach. We find potential for scale economies and especially for the smaller companies. By applying quantile regression, we find how the costs will be effected by an increase in outputs for each quantile, meaning that we can tell how the results vary with firm size. Our results show highest RTS for the smallest companies, which suggests that the smallest firms in the industry should increase their outputs. Since the demand for distribution services is fixed, the distribution companies cannot readily increase their production of distribution services. This implies that there are too many small distribution companies in the Norwegian electricity industry and it would be expedient if the smallest companies would merge. Our results also suggest that RTS is increasing over time, implying that it is getting more and more expedient, from a cost minimization point of view, to increase outputs for the smallest distribution companies. There might be several reasons for RTS to increase over time. One explanation supported by our results is that some of the companies in the industry experience positive technical change, while others do not. An interesting question to address for future research would be to find out which firms in the industry experience technical change, and if there are differences between firm size.

The electricity industry plays an important part in the economy, not only in Norway but also in most countries. This naturally leads to political debate. We believe that this article brings useful knowledge when political strategy and visions within the electricity industry are being compiled in the future.

Acknowledgements

The work in this study is a part of the project *Benchmarking for Regulation of Norwegian Electricity Networks* (ElBench). The project is funded by Norwegian Water Resources and Energy Directorate, Energy Norway and six electricity companies in Norway (Agder Energi Nett, Skagerak Nett, Eidsiva Nett, Hafslund Nett, BKK Nett and Istad Nett). Support from the Research Council of Norway through project 228811 is gratefully acknowledged.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The project is funded by Norwegian Water Resources and Energy Directorate, Energy Norway and six electricity

companies in Norway (BKK Nett, Eidsiva Nett, Hafslund Nett, Helgeland Kraft, Lyse Elnett and Skagerak Nett).

References

Alaeifar, M., M. Farsi, and M. Filippini. 2014. "Scale Economies and Optimal Size in the Swiss Gas Distribution Sector." *Energy Policy* 65: 86–93.

Burns, P., and T. G. Weyman-Jones. 1996. "Cost Functions and Cost Efficiency in Electricity Distribution: A Stochastic Frontier Approach." *Bulletin of Economic Research* 48: 41–64.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell, and G. E. Battese. 2005. *An Introduction to Efficiency and Productivity Analysis*. New York: Springer.

Davidson, R., and J. G. MacKinnon. 1995. "Estimation And Inference In Econometrics." *Econometric Theory* 11 (3): 631–635.

Davino, C., M. Furno, and D. Vistocco. 2014. *Quantile Regression – Theory and Applications*. Wiley & Sons.

Farsi, M., M. Filippini, M. A. Plagnet, and R. Saplacan. The Economies of Scale in the French Power Distribution Utilities. *CEPE Working Paper* No. 73; 2010.

Filippini, M. 1996. "Economies of Scale and Utilization in the Swiss Electric Power Distribution Industry." *Applied Economics* 28: 543–550.

Førsund, F. R., and L. Hjalmarsson. 2004. "Are All Scales Optimal in DEA? Theory and Empirical Evidence." *Journal of Productivity Analysis* 21: 25–48.

Førsund, F. R., and S. A. C. Kittelsen. 1998. "Productivity Development of Norwegian Electricity Distribution Utilities." *Resource and Energy Economics* 20: 207–224.

Giles, D., and N. Wyatt. 1993. "Economies of Scale in the New Zealand Electricity Distribution Industry." In *Models, Methods and Applications of Econometrics*, editor P. C. B. Phillips, 370–382. Oxford: Blackwell.

Greene, W. 2005. "Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier." *ModellJournal of Econometrics* 126 (2): 269-303.

Growitsch, C., T. Jamasb, and G. P. Pollit. 2009. "Quality of Service, Efficiency and Scale in Network Industries: An Analysis of European Electricity Distribution." *Applied Economics* 41: 2555–2570.

Growitsch, C., T. Jamasb, and H. Wetzel. 2012. "Efficiency Effects of Observed and Unobserved Heterogeneity: Evidence from Norwegian Electricity Distribution Networks." *Energy Economics* 34 (2): 542–548.

Koenker, R., and G. Bassett. 1978. "Regression Quantiles." *Econometrica* 46: 33–50.

Kumbhakar, S. C., and L. Hjalmarsson. 1998. "Relative Performance of Public and Private Ownership under Yardstick Competition: Electricity Retail Distribution." *European Economic Review* 42: 97–122.

Kumbhakar, S. C., R. Amundsveen, H. M. Kvile, and G. Lien. 2015. "Scale Economies, Technical Change and Efficiency in Norwegian Electricity Distribution 1998–2010." *Journal of Productivity Analysis* 43 (3): 295–305.

Kwoka, J. 2005. "Electric Power Distribution: Economies of Scale, Mergers and Restructuring." *Applied Economics* 37: 2373–2386.

McLachlan, G., and D. Peel. 2000. *Finite Mixture Models*. New York: Wiley & Sons.

Miguéis, V. L., A. S. Camanho, E. Bjørndal, and M. Bjørndal. 2012. "Productivity Change and Innovation in Norwegian Electricity Distribution Companies." *Journal of the Operational Research Society 63(7):* 982–990.

Reiten, E., L. Sørgard, and K. Bjella. 2014. "Et Bedre Organisert Strømnett." In *Report Ordered by the Ministry of Energy and Oil*, The Ministry of Energy and Oil, Oslo.

Salvanes, K. G., and S. Tjøtta. 1994. "Productivity Differences in Multiple Output Industries: An Empirical Application to Electricity Distribution." *Journal of Productivity Analysis* 5: 23–43.

Salvanes, K. G., and S. Tjøtta. 1998. "A Test for Natural Monopoly with Application to Norwegian Electricity Distribution." Kluwer Academic Publisher. In *Review of Industrial Organization*, 669–685.

Tovar, B., and F. J. Ramos-Real. 2011. "Fagundes De Almeida E. Firm Size and Productivity, Evidence from the Electricity Distribution Industry in Brazil." *Energy Policy* 39: 826–833.

Wald, A. 1943. "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large." *Transactions of the American Mathematical Society* 54: 426–482.

Yatchew, A. 2000. "Scale Economies in Electricity Distribution: A Semiparametric Analysis." *Journal of Applied Econometrics* 15: 187–210.

Appendix

Table A1. Descriptive statistics and Shapiro–Wilk test for normal distribution.

	Descriptive statistics		Shapiro–Wilk test, H_0: normal distribution	
Variable	Skewness	Kurtosis	*z*	*p*-Value
TOTEX	5.158	37.278	16.195	0.000
ln(TOTEX)	0.841	3.835	9.958	0.000

Skewness and kurtosis indicate that our dependent variable (TOTEX) and ln(TOTEX) are not normally distributed. This is supported by the Shapiro–Wilk test which rejects the null-hypothesis of normal distribution in both TOTEX and ln(TOTEX).

Table A2. Evolution of total costs (all prices are in 2010 Norwegian Kroner).

	Year	Mean	SD	Min	Median	Max	CPI
Total costs; 1000 NOK (2010)	2000	61,381	94,795	5682	28,335	665,174	1.617
	2001	75,068	137,870	5831	27,469	840,839	1.523
	2002	75,494	134,700	3070	28,483	859,966	1.445
	2003	73,841	129,683	2559	29,670	857,095	1.373
	2004	88,041	189,225	2491	30,980	1,656,051	1.314
	2005	85,317	176,397	2695	30,215	1,436,703	1.261
	2006	90,766	187,983	2399	31,316	1,523,005	1.204
	2007	90,028	189,119	2691	32,995	1,620,911	1.148
	2008	96,627	202,989	2807	33,707	1,716,352	1.088
	2009	95,726	197,202	3066	37,414	1,748,090	1.042
	2010	97,069	195,365	3059	35,365	1,637,310	1.000
	2011	108,389	214,386	3306	40,438	1,660,168	0.964
	2012	98,326	187,757	3162	37,881	1,614,091	0.934
	2013	107,435	207,284	3447	41,636	1,634,736	0.903

The consumer price index (CPI) is retrieved from Statistics Norway, Table 03363-*Other services with wages as dominating price factor*. http://www.ssb.no/en.

Table A3. Wald likelihood-ratio test.

Wald likelihood-ratio test, Cobb–Douglas versus Translog function.		
F	DF	*p*-Value
24.75	7	0.000

Table A4. Results of quantile estimation, cost elasticities for *network length* (N).

Median	Mean	Std. error (Delta-method)	*p*-Value (Delta-method)		95% Confidence interval (Delta-method)	
0.05	0.329	0.324	0.025	0.000	0.275	0.373
0.10	0.359	0.354	0.016	0.000	0.323	0.385
0.15	0.359	0.353	0.014	0.000	0.327	0.380
0.20	0.378	0.371	0.018	0.000	0.336	0.407
0.25	0.375	0.369	0.022	0.000	0.326	0.413
0.30	0.373	0.369	0.031	0.000	0.307	0.430
0.35	0.393	0.391	0.033	0.000	0.326	0.456
0.40	0.406	0.405	0.037	0.000	0.331	0.448
0.45	0.410	0.410	0.040	0.000	0.331	0.489
0.50	0.415	0.415	0.037	0.000	0.342	0.488
0.55	0.402	0.402	0.031	0.000	0.342	0.426
0.60	0.415	0.416	0.033	0.000	0.351	0.480
0.65	0.411	0.412	0.032	0.000	0.350	0.474
0.70	0.409	0.410	0.029	0.000	0.353	0.467
0.75	0.419	0.422	0.027	0.000	0.370	0.474
0.80	0.430	0.434	0.024	0.000	0.388	0.480
0.85	0.436	0.440	0.029	0.000	0.384	0.496
0.90	0.410	0.415	0.032	0.000	0.353	0.477
0.95	0.368	0.378	0.025	0.000	0.329	0.426

The Delta-method is a method for deriving the variance of a function of asymptotically normal random variable with known variance using a Taylor series expansion; see https://cran.r-project.org/web/packages/modmarg/vignettes/delta-method.html#fn7 for more details.

Table A5. Results of quantile estimation, cost elasticities for *number of customers* (Q).

Quantiles	Median	Mean	Std. error (Delta-method)	*p*-Value (Delta-method)	95% Confidence interval (Delta-method)	
0.05	0.571	0.582	0.061	0.000	0.725	0.965
0.10	0.556	0.566	0.060	0.000	0.738	0.972
0.15	0.553	0.562	0.063	0.000	0.736	0.984
0.20	0.539	0.547	0.074	0.000	0.710	1.000
0.25	0.542	0.548	0.077	0.000	0.654	0.957
0.30	0.543	0.547	0.070	0.000	0.653	0.927
0.35	0.526	0.527	0.080	0.000	0.564	0.880
0.40	0.516	0.516	0.074	0.000	0.547	0.837
0.45	0.516	0.516	0.071	0.000	0.461	0.738
0.50	0.514	0.515	0.081	0.000	0.409	0.726
0.55	0.525	0.525	0.068	0.000	0.392	0.657
0.60	0.515	0.515	0.076	0.000	0.328	0.626
0.65	0.517	0.516	0.078	0.000	0.248	0.555
0.70	0.520	0.519	0.069	0.000	0.230	0.501
0.75	0.511	0.510	0.072	0.000	0.171	0.455
0.80	0.501	0.500	0.079	0.001	0.108	0.417
0.85	0.503	0.500	0.071	0.001	0.097	0.374
0.90	0.527	0.526	0.074	0.003	0.077	0.368
0.95	0.568	0.564	0.098	0.016	0.045	0.428

Table A6. Returns to scale (RTS) across quantiles.

Quantiles	Median	Mean	Std. error (Delta-method)	*p*-Value (Delta-method)	95% Confidence interval (Delta-method)	
0.05	1.110	1.104	0.018	0.000	1.070	1.142
0.10	1.092	1.087	0.007	0.000	1.074	1.102
0.15	1.097	1.092	0.008	0.000	1.078	1.108
0.20	1.091	1.089	0.008	0.000	1.075	1.106
0.25	1.091	1.090	0.010	0.000	1.071	1.111
0.30	1.092	1.093	0.011	0.000	1.071	1.116
0.35	1.088	1.089	0.009	0.000	1.071	1.108
0.40	1.084	1.086	0.008	0.000	1.071	1.104
0.45	1.081	1.080	0.008	0.000	1.065	1.097
0.50	1.076	1.076	0.009	0.000	1.059	1.094
0.55	1.079	1.079	0.008	0.000	1.063	1.096
0.60	1.075	1.074	0.008	0.000	1.059	1.090
0.65	1.077	1.077	0.008	0.000	1.062	1.094
0.70	1.076	1.077	0.007	0.000	1.063	1.092
0.75	1.076	1.072	0.008	0.000	1.057	1.090
0.80	1.075	1.070	0.009	0.000	1.054	1.088
0.85	1.065	1.063	0.010	0.000	1.044	1.085
0.90	1.067	1.062	0.010	0.000	1.043	1.083
0.95	1.068	1.062	0.009	0.000	1.045	1.082

Table A7. Cost elasticities, returns to scale and technical change for alternative models.

					Finite mixture model[c]			
	Fixed effect[a]		True FE SFA[b]		1 Class		2 Class	
Model Class	Mean	Std. error	Mean	Std. error	Mean	Std. error.	Mean	Std. error.
Cost elasticity								
Length of network	0.33	0.05	0.38	0.16	0.32	0.02	0.31	0.02
Number of customers	0.46	0.05	0.54	0.16	0.45	0.02	0.57	0.02
Returns to scale (RTS)	1.28	0.10	1.09	0.22	1.34	0.03	1.15	0.01
Technical change (TC)	0.003	0.002	−0.001	0.005	0.006	0.003	−0.005	0.003

[a]Fixed effect panel data estimator.
[b]True fixed effect stochastic frontier panel data estimator by Greene (2005).
[c]Finite mixture model, or latent class model (McLachlan and Peel 2000). Total costs (TOTEX) is used as independent variable for class probabilities. The average total cost for sample '1. class' is NOK 21,080,000, while for '2. class' is NOK 165,163,000.

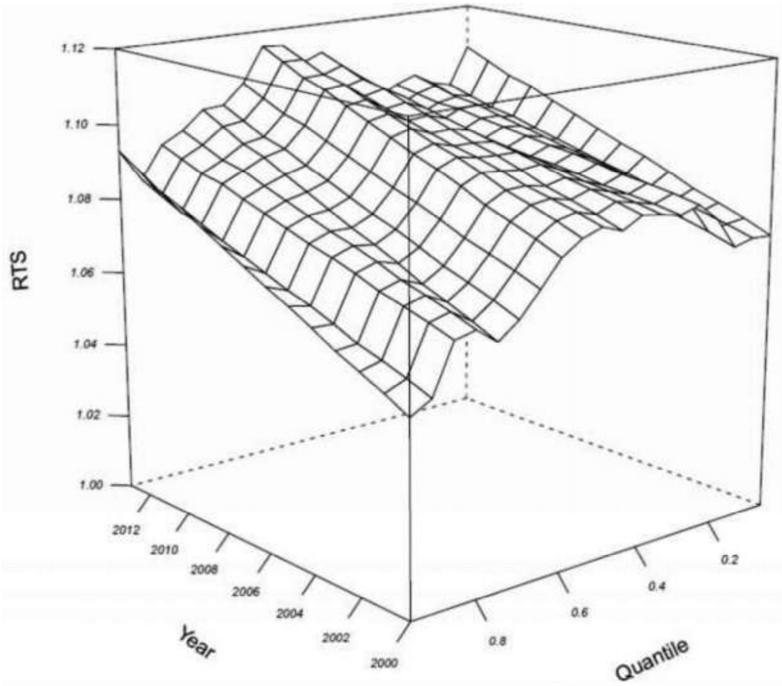


Figure A1. Returns to scale (RTS) across quantiles and over time. Balanced panel data. Number of observations reduces from 1750 to 1414.

ESSAY 2

Ørjan Mydland, Subal C. Kumbhakar, Gudbrand Lien, Roar Amundsveen, Hilde Marit Kvile

Economies of scope and scale in the Norwegian electricity industry

Submitted to Journal of Regulatory Economics

Economies of Scope and Scale in the Norwegian Electricity Industry

Ørjan Mydland*
Faculty of Economics and Organization Science, Inland Norway University of Applied Sciences, Lillehammer, Norway
and Department of Business and Management Science, Norwegian School of Economics, Bergen, Norway.
Postal address: Inland Norway University of Applied Sciences, Faculty of Economics and Organization Science, Pb 952, NO-2604 Lillehammer, Norway, Email: orjan.mydland@inn.no / orjan.mydland@nhh.no, Phone: +47 909 35 042

Subal C. Kumbhakar
Department of Economics, State University of New York – Binghamton, Binghamton, NY 13902, United States, kkar@binghamton.edu

Gudbrand Lien
Faculty of Economics and Organization Science, Inland Norway University of Applied Sciences, Lillehammer, Norway, gudbrand.lien@inn.no

Roar Amundsveen
The Norwegian Water Resources and Energy Directorate, Oslo, Norway, roam@nve.no

Hilde Marit Kvile
The Norwegian Water Resources and Energy Directorate, Oslo, Norway, hkv@nve.no

ABSTRACT

In this paper, we use panel data for the period 2004–2014 to investigate economies of scope and scale in the Norwegian electricity industry, with a focus on the distribution and generation of electricity. We examine economies of scope and scale in unbundled and vertically integrated firms using both quadratic and translog cost functions. We implement a new method of estimating economies of scope and scale in which the technologies for unbundled (specialized) and integrated firms are different. Our results show evidence of economies of scope and scale which have important policy implications.

Keywords: Cost function, Economies of scope, Economies of scale, Flexible technology

1. INTRODUCTION

Traditionally, the organizational structure of electricity industries worldwide has been one of vertical integration. Electric utilities have typically performed all services from generation to distribution in the electric supply chain. Vertical integration creates the potential for economies of scope, which may reduce the total cost of providing services. However, electricity industries worldwide have undergone profound changes involving strict separation of these services. In Norway, the main motivation for the unbundling of services is to increase competition in the electricity industry, avoid cross-subsidization, and ensure the distribution system operators (DSOs) focus only on network operations.

The deregulation and market reforms in Norway during the 1990s have led to a more market-oriented environment. The basic idea behind the deregulation of this sector was to introduce competition where possible, namely in generation and supply (retail). Distribution and transmission services were regulated because they are natural monopolies. To promote efficient markets, regulatory rules for unbundling were developed. Accounting unbundling was introduced in the 1990s. The European Union's Third Energy Package, which consists of directives aimed at unbundling energy generation and supply interests from the distribution and transmission of electricity, was adopted in 2009. The directives impose rules on legal unbundling for firms with more than 100,000 customers. In 2016, the Norwegian parliament amended the Energy Act, with the changes taking effect from 2021. The amended legislation will introduce legal and functional unbundling for all firms involved in electricity distribution regardless of the number of customers. Ownership unbundling was also considered, but found to be infeasible because ownership is mostly public. Functional unbundling was introduced as an alternative solution, which implies that the distribution company shall have its own board separate from the other companies in the group and reduces the possibility of interference with decisions made by DSO managers. In addition, goods and services are required to be purchased in the market, not within the group. The rationale behind these rules is to minimize the risk of cross-subsidization (which will lead to less-efficient electricity markets) and strengthen the neutrality of the DSO in its operations (by limiting the possibility of favoring other services in the group).

The new legislation will obviously reduce the potential for cost savings from economies of scope. However, society is supposed to benefit from economies of scale and increased competition in the power market. These gains might be greater than the loss from not utilizing economies of scope. We believe that this view is held by policy makers because the new rules were implemented without referring to any economies of scope studies for Norway or any other country that generates electricity mainly from hydropower.

In our analysis, we examine how these rules affect the Norwegian electricity industry. Baumol et al. (1982) pointed out that economies of scope can exist because of synergies in the joint utilization of labor and capital. The type of labor required in the distribution and generation of electricity might be quite similar. Furthermore, combining all elements of electricity supply into one value chain from electricity generation to distribution may minimize production costs. Examples of positive synergies are advertising and billing costs, and what Waldman and Jensen (2001) called "massed reserves," which means that multioutput firms can exploit the same reserve capacity during emergency repairs and maintenance.

In this study, we estimate three random effects cost models using data from 212 Norwegian electricity firms observed over a period of 11 years. We estimate a quadratic and a translog cost function using Triebs et al.'s (2016) flexible technology approach. The flexible technology approach is useful because it gives a more realistic estimate for the effect of separation, and is more realistic because it actually allows the technologies of integrated and separated firms to be different. To compare results, we also estimate a common cost function model. For the scope and scale measures, we do not follow the standard practice of presenting the results at the mean or median values of output. In our analysis, we seek to identify the costs and benefits of separating an integrated firm into two specialized firms to see how the new legislation for Norwegian electricity firms will affect the economic situation in the industry. Economies of scope estimates are often sensitive to the actual output values used in the estimation. Therefore, we present scope and scale estimates for all combinations of output values from all 42 integrated firms in our dataset.

The remainder of the paper is organized as follows. Section 2 presents a brief survey of the literature. Section 3 describes the model specifications and methods. Section 4 describes the data and Section 5 presents the results. In Section 6, we present a summary of our main results and conclusions.

2. LITERATURE REVIEW

It is somewhat surprising that considering its policy importance, there is little research on economies of scope in the electricity industry in Norway. We are aware of only one recent report from the Norwegian Water Resources and Energy Directorate (NVE) that briefly addresses the topic of economies of scope in the Norwegian electricity industry. Nevertheless, NVE (2015) finds that the operational costs of vertically integrated companies are 15% lower than those of other companies. One reason for the lack of studies in this area may be that it is difficult to obtain data suitable for analyzing economies of scope. By examining both economies of scale and scope, this study provides new insights for both policy makers and the electricity industry.

Although many scope studies have been conducted within energy markets in other countries, including the markets for electricity, gas, water, and coal, only a few have focused on economies of scope and scale in the electricity industry. Mayo (1984) and Chappell and Wilder (1986) found evidence of economies of scope in the US electricity and gas markets. Fraquelli et al. (2004) and Piacenza and Vannoni (2004) examined the Italian electricity, gas, and water distribution markets, while Farsi et al. (2009) examined the corresponding Swiss markets. Garcia et al. (2007) studied North American water utilities, and Carvalho and Marques (2014) studied Portuguese water utilities.

Based on our knowledge, there have been five scope studies of US electricity markets. Using cross-sectional data examining US electric utilities, Kaserman and Mayo (1991), Kwoka (2002), and Arocena et al. (2012) used data from 1981, 1989, and 2001, respectively. Meyer (2012a) and Triebs et al. (2016) examined the US electricity market with panel data covering the periods 2001–2008 and 2000–2003, respectively. These studies provide empirical evidence for the existence of significant economies of vertical integration between generation and transmission/distribution in electricity supply companies. The scope estimates ranged from 4% to 27%.[1]

Four studies of economies of scope of the European electricity industry exist in the literature; all of them used panel data. Jara-Dıaz et al. (2004) analyzed Spanish electricity generation and distribution companies for the period 1985–1996. Piacenza and Vannoni (2009) examined the Italian electricity market for the years 1994–2000, while Fetz and Filippini (2010) investigated Swiss generation and distribution companies for the period 1997–2005. Gugler et al.

[1] A summary of the most important previous empirical economies of scope and scale studies within the electricity sector is presented in Table A1 in the Appendix. Meyer (2012b) provided a review of the theoretical and empirical literature within the field of vertical economies and the costs of separating electricity supply.

(2017) studied 28 electricity generation and transmission firms from 16 European countries for the period 2000–2010. These European studies reported evidence of economies of scope, ranging from 6.5% to 60%, which is higher than the estimates for the US. As this brief review shows, no scientific published economies of scope studies of the electricity market exist for Norway or Scandinavia.

The estimation approach in the previous economies of scope studies mentioned above (except Piacenza and Vannoni (2009) and Triebs et al. (2016)) used either a quadratic or a translog cost function for each firm type (integrated and specialized firms), jointly or separately. The quadratic function's violation of the linear homogeneity (in input prices) property is discussed in Farsi et al. (2008) and Triebs et al. (2016). We do not use input prices in our cost models because there is no input price variation cross-sectionally in our data and the temporal variation can be captured in the time dummies or the time trend in the models. In Norway, union agreements regarding wages and social benefits are centralized at a national level. Thus, the assumption of constant input prices across firms is a reasonable assumption in a small country such as Norway.[2] As a result, homogeneity (in input prices) violation is not a problem in our models. In estimating a cost function that includes multiple firm types jointly, a common technology among firm types is assumed. The question is whether the technology used by the specialized utilities is identical to that used by the utilities providing more than one service. If the technologies are different, and one assumes a common technology, the results are likely to be invalid. For instance, results suggesting the presence of economies of scope may actually be a result of scale economies. One way to get around this issue is to perform separate estimations for each firm type. This allows the technology

[2] For fixed input (factor) prices, the cost function is written as a function of outputs. For example, see Varian (1992). Temporal variations in input prices are captured by the time dummies or trend included in the cost function.

to be different between the firms, which may also affect the results through the firm's ability to utilize factors of production. Triebs et al. (2016) introduced a method that allows us to test for differences in technology. Another advantage of this method is that it avoids the problem of zero values for output in a translog function.[3] Previous studies have shown that replacing zero values by some arbitrary number can influence the results (e.g., see Pulley and Humphrey, 1993). However, the flexible technology approach introduced by Triebs et al. (2016) avoids the zero-value problem by allowing the technologies of the specialized firms to be different from the integrated firms.

In addition, there might also be a problem with zero values when using a quadratic function. If the number of zero values represents a large proportion of the total number of sample observations, the parameter estimates may be biased (Battese, 1997). This potential problem does not arise in Triebs et al.'s (2016) approach.

3. MODEL SPECIFICATION AND METHOD

In this section, we describe the specifications of three models and the estimation method used in this study. We start by describing Model 1, which uses a common quadratic cost function, followed by a description of Models 2 and 3, which use a flexible technology approach (Triebs et al., 2016). Model 2 is a quadratic cost function and Model 3 is a translog cost function. Before introducing the models, we provide definitions of scope and scale economies. For economies of scope, we measure the difference between the cost of one firm producing two outputs and the costs (sum) of

[3] In scope studies, one or several outputs are zero for specialized firms. This is a problem in the translog function approach because the logarithm of zero is not defined (missing values will be created). The common way to handle this problem is to replace zero values with a small number.

two specialized firms producing the same outputs (see Baumol et al., 1982; Panzar and Willig, 1981).[4] Economies of scope are measured as:

$$Scope = \frac{(C_D(D)+C_G(G))-C_I(D,G)}{C_I(D,G)}, \tag{1}$$

where $C_D(D)$ is the estimated cost for the specialized firms in *distribution* and is usually obtained by setting the output of *generation* (G) to zero in $C(D,G)$, i.e., $C_D(D) = C(D,0)$. Likewise for the specialized firms in *generation*, the estimated cost is $C_G(G) = C(0,G)$, and for the *integrated* firms with positive outputs in both *distribution* and *generation*, the estimated costs are $C_I(D,G) = C(D,G)$. If the scope measure is positive (or negative), economies (or diseconomies) of scope exist.

Following Baumol et al. (1982), global economies of scale in a multioutput setting are defined as:

$$Scale = \frac{C(D,G)}{D\frac{\partial C(D,G)}{\partial D}+G\frac{\partial C(D,G)}{\partial G}}. \tag{2}$$

If the scale measures are greater than, equal to, or less than unity, the returns to scale (RTS) are increasing, constant, or decreasing, respectively.

3.1 Model 1: Common Quadratic Cost Function

Electricity generation in Norway is mainly based on hydropower. In our analysis, we use panel data, but to simplify the notation, we drop the subscripts *i* and *t*, where *i* denotes the firm, *i* = 1,...,n and *t* denotes time, *t* = 1,…,t. The common quadratic cost function is specified as follows:

[4] To avoid any confusion, in previous studies, the term "economies of vertical and horizontal integration" was used instead of "economies of scope and scale." We believe these terms are synonymous.

$$C = \alpha_0 + \alpha_D Ddum + \alpha_G\, Gdum \ + \beta_1 L + \beta_2 Q + \frac{1}{2}\beta_{11}L^2 + \frac{1}{2}\beta_{22}Q^2 + \beta_3 E + \beta_4 N + \frac{1}{2}\beta_{33}E^2 + \frac{1}{2}\beta_{44}N^2 + \frac{1}{2}\beta_{12}L * Q + \frac{1}{2}\beta_{13}L * E + \frac{1}{2}\beta_{14}L * N + \frac{1}{2}\beta_{23}Q * E + \frac{1}{2}\beta_{24}Q * N + \frac{1}{2}\beta_{34}E * N \quad (3)$$

where *C* is total operational cost, and *Ddum* and *Gdum* are dummy variables representing the distribution and generation companies, respectively. Kilometers of high-voltage network (L) and number of customers (Q) represent the electricity distribution outputs $D(L, Q)$. Megawatt hours of produced electricity (E) and number of generators (N) are the outputs in electricity generation $G(E, N)$. Note that we allow the cost functions to differ only in the intercepts, i.e., α_D and α_G are coefficients for the dummy variables representing the distribution and generation companies, respectively. The coefficients α_D and α_G represent the fixed costs of only the distribution and only the generation companies, respectively, over the fixed cost of both the distribution and generation (integrated) companies.

A single cost function in eq. (3) is estimated using all the data (pooled) and then the costs for the specialized and integrated firms are obtained from the estimated cost function by using their respective output values. Thus, substituting eq. (3) into eq. (1), the scope measure from Model 1 is:

$$Scope_{Model\ 1} = \frac{\alpha_0 + \alpha_D + \alpha_G - \frac{1}{2}\beta_{13}L*E + \frac{1}{2}\beta_{14}L*N + \frac{1}{2}\beta_{23}Q*E + \frac{1}{2}\beta_{24}Q*N + \frac{1}{2}\beta_{34}E*N}{C_I(D,G)}. \quad (4)$$

If the numerator of eq. (4) is positive, economies of scope exist, meaning that an integrated firm producing both outputs has lower costs compared with a situation where the production is separated across two specialized firms. If the numerator is negative, diseconomies of scope exist, meaning that costs are lower if we have two specialized firms compared with one integrated firm.

3.2 Models 2 and 3: Flexible Technology Approach

In Model 2, we use a quadratic cost function and add flexibility to it by allowing the technology to vary across specialized and integrated firms. We do so by introducing dummies for specialized and integrated firms so that the technologies are different. Thus, the specification of Model 2 is:

$$\begin{aligned}
C = Idum\Big(&\alpha_0 + \beta_1 L + \beta_2 Q + \frac{1}{2}\beta_{11}L^2 + \frac{1}{2}\beta_{22}Q^2 + \beta_3 E + \beta_4 N + \frac{1}{2}\beta_{33}E^2 \\
&+ \frac{1}{2}\beta_{44}N^2 + \frac{1}{2}\beta_{12}L * Q + \frac{1}{2}\beta_{13}L * E + \frac{1}{2}\beta_{14}L * N + \frac{1}{2}\beta_{23}Q * E \\
&+ \frac{1}{2}\beta_{24}Q * N + \frac{1}{2}\beta_{34}E * N\Big) \qquad (5) \\
&+ Ddum\left(\alpha_D + \delta_1 L + \delta_2 Q + \frac{1}{2}\delta_{11}L^2 + \frac{1}{2}\delta_{22}Q^2 + \frac{1}{2}\delta_{12}L * Q\right) \\
&+ Gdum\left(\alpha_G + \gamma_1 E + \gamma_2 N + \frac{1}{2}\gamma_{11}E^2 + \frac{1}{2}\gamma_{22}N^2 + \frac{1}{2}\gamma_{12}E * N\right)
\end{aligned}$$

By introducing the dummy variables *Idum*, *Ddum*, and *Gdum* for the integrated firms and the two specialized firms in distribution and generation, respectively, Model 2 in eq. (5) combines three separate cost functions, one for each firm type. The dummy variable approach makes it possible to estimate the three cost functions jointly.[5]

Model 3 is also specified using the flexible technology dummy variable approach as in Model 2, but with a translog cost function. The specification of Model 3 is:

[5] This is equivalent to stacking the three cost functions and then estimating the stacked cost function as a single cost function.

$$
\begin{aligned}
lnC = {} & Idum\Big(\alpha_0 + \beta_1(lnL) + \beta_2(lnQ) + \frac{1}{2}\beta_{11}(lnL)^2 + \frac{1}{2}\beta_{22}(lnQ)^2 \\
& + \beta_3(\ln\mathrm{E}) + \beta_4(\ln\mathrm{N}) + \frac{1}{2}\beta_{33}(lnE)^2 + \frac{1}{2}\beta_{44}(lnN)^2 + \beta_{12}(lnL * lnQ) \\
& + \beta_{13}(lnL * lnE) + \beta_{14}(lnL * lnN) + \beta_{23}(lnQ * lnE) + \beta_{24}(lnQ * lnN) \\
& + \beta_{34}(lnE * lnN)\Big) \\
& + Ddum\Big(\alpha_D + \delta_1(lnL) + \delta_2(lnQ) + \frac{1}{2}\delta_{11}(lnL)^2 + \frac{1}{2}\delta_{22}(lnQ)^2 \\
& + \delta_{12}(lnL * lnQ)\Big) \\
& + Gdum\Big(\alpha_G + \gamma_1(lnE) + \gamma_2(lnN) + \frac{1}{2}\gamma_{11}(lnE)^2 + \frac{1}{2}\gamma_{22}(lnN)^2 \\
& + \gamma_{12}(lnE * lnN)\Big).
\end{aligned} \tag{6}
$$

The three models presented in eqs (3), (5), and (6) are made stochastic by introducing the error term u_i. To control for firm heterogeneity, we include random effects in all three models. Thus, in our estimation, we replace the constant terms $\alpha_0, \alpha_D,$ and α_G in Model 1 in eq. (3), Model 2 in eq. (5), and Model 3 in eq. (6) with $\epsilon_\tau = (\alpha_\tau + w_i)$, where the subscript τ = 0, D, or G. w_i is a time-invariant, firm-specific random term that controls for firm heterogeneity.[6,7] Note that we omitted the time subscript from u_i and both the firm and time subscripts from the variables in the cost functions.

[6] The Breusch and Pagan Lagrange multiplier test for random effects (against a standard OLS regression) rejects the null hypothesis at the 0.000 level of significance in all three models.

[7] We have low intertemporal variation in some variables in our data. Thus, we are more interested in between than within variation. In addition, we use an unbalanced panel where a portion of the sample has four or fewer observations per firm (i.e., panel data with a short time-series component). In cases such as this, based on Clark and Linzer (2015), a fixed effect model exacerbates measurement error bias and the random effect model is preferable. The fixed effect model will therefore not be appropriate in our analysis.

How can one estimate the model in eq. (6) given that there are no output data for the generation (distribution) utilities? If one estimates the technologies separately, then no output data on electricity generation (E = N = 0, meaning that the firm is a distribution utility) will be used to estimate the technology for the distribution utilities. To mimic this, in estimating eq. (6), we replace the output data for electricity generation (E = N = 0) for the distribution utilities by any positive numbers so that the log of these numbers can be defined. These numbers then disappear when multiplied by *Ddum*, and the model reduces to one, for which only the output data on the distribution utilities are used; the same applies for the generation utilities. Furthermore, this approach applies for any kind of functional form specified, including the quadratic model specification in eq. (5). For further details, see Triebs et al. (2016).

The scope measure expression in both Models 2 and 3 is different from the scope measure in the common quadratic cost function in Model 1 presented in eq. (3). From the general definition of scope in eq. (1), we obtain: $\frac{(C_{Ddum}(D)+C_{Gdum}(G))-C_{Idum}(D,G)}{C_{Idum}(D,G)}$. See Fuss and Waverman (1981) for more on this.

The dummy variable specifications of the translog and quadratic cost functions make it possible to test whether the common technology assumption in Model 1 is appropriate. We can do this by performing a joint likelihood ratio test with the following restrictions on both Models 2 and 3:

$$\beta_1 = \delta_1,\ \beta_2 = \gamma_2,\ \beta_3 = \gamma_1,\ \beta_4 = \gamma_2,\ \beta_{11} = \delta_{11},\ \beta_{22} = \delta_{22},\ \beta_{33} = \gamma_{11},\ \beta_{44} = \gamma_{22}\ . \tag{7}$$

Note that the technology, e.g., for the generation companies is obtained by imposing the above restriction together with *Gdum* = 1, which also implies *Ddum* = 0 and *Idum = 0*. Failure to reject the null hypothesis with the above restrictions indicates the presence of a shared technology for all firm types. The specification in Model 1 does not permit us to impose these cross-equation restrictions and thus test for the presence of a shared technology for all firm types.

4. DATA

The data comprise economic and technical information on Norwegian electricity companies from 2004 to 2014 and are collected by the NVE.[8] In total, there are 1,494 firm-year observations constituting an unbalanced panel of 212 Norwegian electricity companies. Table 1 presents the descriptive statistics of the variables used in our analysis.

Total operational cost for each firm consists of the sum of material costs, salaries and other personnel costs (including pension costs), other operating expenses, losses on receivables, losses on disposal of fixed assets, internally priced services, and allocated overhead costs. All costs are adjusted for inflation by an industry commodity price index, where wages are the main cost.[9] The output variables for electricity distribution are kilometers of high-voltage network (km network) and number of customers. We also considered including the number of network stations as a proxy for electricity delivered. However, this will cause multicollinearity between the output variables. The ratio between the number of customers and number of network stations, and that between the number of customers and km network, captures the same effect. In urban areas, the number of customers is high compared with the number of network stations and km network, while in rural areas, the situation is the exact opposite. The output variables for electricity generation are electricity production in megawatt-hours (electricity MWh) and number of generators. In a scope study, the minimum value of the output variables will naturally equal zero because by definition,

[8] The data used in this study are confidential. Readers who want to gain access to the data must apply to the NVE for permission; see www.nve.no for details.

[9] The price index was retrieved from Statistics Norway, Table 03363. http://www.ssb.no/en.

specialized firms do not produce some outputs. The minimum values for distribution and generation outputs in parentheses are the minimum output values, given that the outputs are not zero.

To control for time effects, we also include a time component in our estimation. Note that because the cost function does not include input prices, which only change over time, the time components capture input price effects as well as other effects (technical change) that shift the cost function. In other words, technical change cannot be separated from any temporal changes in input prices.

Our data consist of three types of firms: integrated firms with positive outputs for both distribution and generation, and specialized firms that have positive output only for distribution or generation. There are large variations in firm size in our data. For example, the lowest total operational cost is 255,000 Norwegian kroner (NOK), while the highest total operational cost is about 275 million NOK. Large variation also exists in the outputs.

Table 1: Descriptive Statistics

Variable	Mean	St.Dev.	Min	Median	Max
Total operating costs (1,000 NOK):					
Distribution firms	30,773	37,364	4,126	16,117	274,822
Generation firms	19,065	20,054	255	12,272	146,887
Integrated firms (distribution and generation)	30,754	21,213	2,485	22,952	91,701
Outputs distribution firms:					
Km network	479	520	0 (37)	269	2,909
Number of customers	13,943	22,315	0 (178)	6646	134,854
Outputs generation firms:					
Electricity MWh	276,804	296,476	0 (2,391)	134,428	1,081,649
Number of generators	8.07	9.73	2	5	68
Outputs integrated firms:					
Km network	486	337	31	368	1,185
Number of customers	7,939	6,236	391	6,335	25,748
Electricity MWh	91,430	108,907	3,861	14,640	535,554
Number of generators	4.35	2.73	2	3	15
Time			2004		2014
Firm type observations:					
Integrated firms (distribution and generation)	316 observations, 42 firms				
Specialized firms (distribution)	671 observations, 77 firms				
Specialized firms (generation)	507 observations, 93 firms				
Total firms	1,494 observations, 212 firms				

Note: Numbers in parentheses are the minimum positive outputs for distribution and generation.

5. RESULTS

Using the estimated parameters from the three models, we calculated the marginal costs at the mean values of our data for each of the four outputs (Table 2). These are the derivatives of the estimated cost function with respect to each output.[10] In Model 1, we assume joint technology and do not distinguish between integrated or specialized firms. All of the marginal costs have positive signs as expected, but only *km of network* is significant at the 1% level. This result shows that increasing *km of network* by one, will increase cost by 51.86 (1000 NOK), ceteris paribus. In Model 2, we use the flexible technology dummy variable approach, and even though the cost elasticities for each firm type are presented in different columns, the three cost functions are estimated simultaneously. Compared with Model 1, the estimated marginal costs of *km of network* in Model 2 are lower for both integrated and specialized firms in distribution. In Model 2, all but *Number of generators* for the integrated firms are significant at the 5% level. In Model 3, we estimate a translog cost function, meaning that the data are in log form. To simplify comparison across models, we convert the cost elasticities to marginal costs for the results in Model 3. The marginal cost for *number of generators* for the integrated firms in Model 3 has a negative sign, which is counterintuitive. However, this is not statistically significant. In Models 2 and 3, the estimated cost elasticities for the specialized firms in generation (*electricity MWh* and *number of generators*) are significant at the 1% level.

[10] For the interested reader, the parameter estimates are available in the Appendix (see Table A2).

Table 2: Marginal Costs

	Model 1	Model 2			Model 3		
	All firms	Integrated	Distribution	Generation	Integrated	Distribution	Generation
Km network (L)	51.86** (8.63)	30.00** (10.06)	27.71** (5.11)		24.60* (10.74)	30.82** (6.77)	
Number of customers (Q)	0.31 (0.50)	1.24* (0.57)	1.17** (0.15)		1.50** (0.57)	0.76** (0.18)	
Electricity MWh (E)	0.04 (0.03)	0.06* (0.03)		0.04** (0.01)	0.02 (0.02)		0.03** (0.002)
Number of generators (N)	905.96 (665.31)	400.90 (897.95)		999.57** (139.17)	-321.01 (922.37)		1116.97** (122.32)
Observations	1494	316	671	507	316	671	507

Notes: Standard error in parentheses. ** and * indicates significance at 0.01 and 0.05 levels respectively.
Model 1: Quadratic function, combined, random effects. Model 2: Quadratic function, separate technology, random effects.
Model 3: Translog function, separate technology, random effects.
In Model 3, the data are in log-form but we have converted the results into marginal costs.

To test the restrictions presented in eq. (7) for Models 2 and 3, we perform a joint likelihood ratio test. The test results are presented in Table 3. At the 5% significance level, we reject the null hypothesis of shared technology in both Models 2 and 3. Thus, the results based on a common technology are likely to be incorrect. In the reminder of the paper, we focus on the economies of scope and scale results from Models 2 and 3.

Table 3: Test for Common Technology: Likelihood Ratio Test

	Model 2			Model 3		
$\boldsymbol{H_0}$:	χ^2	DF	P	χ^2	DF	P
$\beta_1 = \delta_1$ $\beta_2 = \delta_2$ $\beta_3 = \gamma_1$ $\beta_4 = \gamma_2$ $\beta_{11} = \delta_{11}$ $\beta_{22} = \delta_{22}$ $\beta_{33} = \gamma_{11}$ $\beta_{44} = \gamma_{22}$	22.83	8	0.005	47.64	8	0.001

The natural policy question to answer from this exercise is "What are the costs or benefits from separating one integrated firm into two specialized firms?" Alternatively, we could ask what would happen to costs if two specialized firms became one integrated firm. However, this is not relevant for the Norwegian situation. Policy makers have decided to change the Norwegian Energy Act so that all integrated firms within the industry will be strictly separated into specialized firms by 2021. In previous economies of scope studies, it is normal to either use median or mean values of the data to calculate the scope measures. Alternatively, a two-by-two table that gives different scope estimates for different combinations of output levels can be constructed. However, these approaches are not always recommended because they are likely to use combinations of output values that do not actually exist in the applied data set, and which might not even be realistic in the real world.[11] To overcome this problem we calculated economies of scope by using the parameter estimates combined with all actual output combinations for the integrated firms in our data set. We retrieve 316 economies of scope estimates for each of the two models from an 11-year period for the 42 integrated firms. The distribution of these 316 estimated economies of scope measures for each model is presented in percentiles in Table 4. The 1% percentile estimate represents the smallest 1% of economies of scope estimates from the distribution. For Model 2, we see that the 1% percentile gives diseconomies of scope (−0.10), meaning that for these integrated firms, costs will be reduced by 10% if they are separated into two specialized firms. The scope measure differs between the models. In Model 2, we find economies of scope at the 25% percentile and above while in Model 3, we find economies of scope at the 75% percentile and

[11] For the integrated firms in our data set, median L = 383.5 and median E = 53,016. However, this company (with this output combination) does not exist in our data set. This output combination for distribution and generation might not ever exist or might not even be feasible in the Norwegian electricity industry.

above. At the median, the estimated economies of scope estimates are 10%, and −4% for Models 2, and 3, respectively.[12]

Table 4: Economies of Scope Results for Integrated Firms

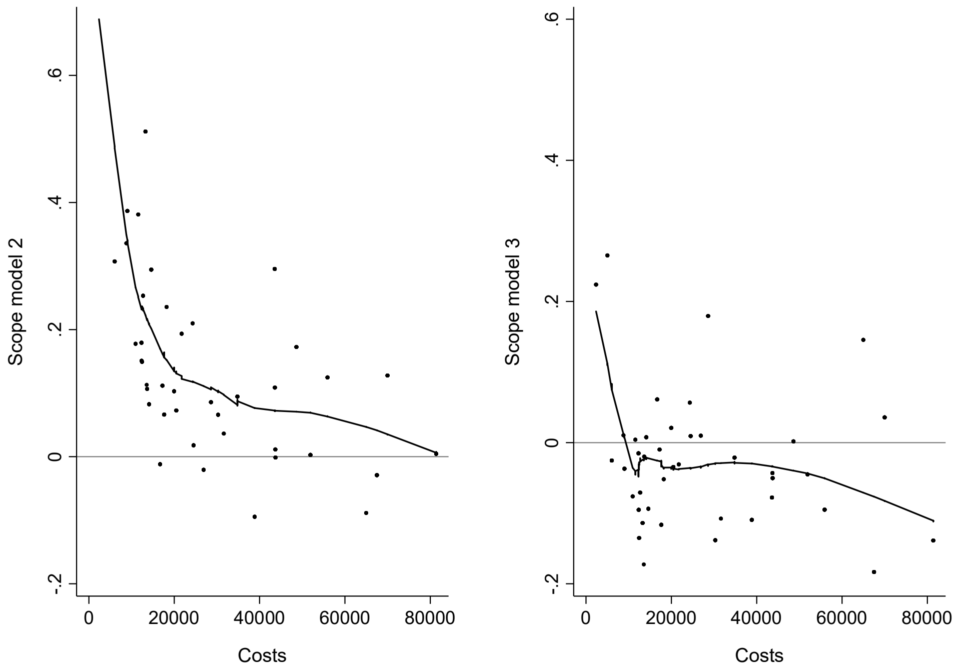
	Economies of scope	
Percentiles	Model 2	Model 3
1%	-0.10	-0.22
5%	-0.08	-0.18
10%	-0.05	-0.15
25%	0.02	-0.10
50% (median)	0.10	-0.04
75%	0.21	0.02
90%	0.38	0.12
95%	0.53	0.24
99%	0.92	0.47
Mean	0.18	-0.02
Standard deviation	0.37	0.13

Table 4 shows the distribution of scope estimates using all the real output combinations for the integrated firms from Models 2 and 3. This gives us an overview of the present economies of scope in the Norwegian electricity industry. However, it is interesting to determine what characterizes the firms with diseconomies and economies of scope. In Figure 1, a plot of the economies of scope estimates by firm size shows that there is a clear relationship between firm size and economies of scope. Figure 1 uses the scope measures in Table 4, but we plot the mean scope measures for each firm from each model against total operational costs (we use total operational cost for each observation as a proxy for firm size for each time period in our data, 42

[12] We also estimated the models without random effects, and the scope estimates in the models at median values change from: Model 2: 0.10 (random effect) to 0.18 (pooled OLS), Model 3: –0.04 (random effect) to 0.07 (pooled OLS). The economies of scope results increase in all three models if we do not control for firm heterogeneity.

firms and 316 observations over 11 years). The scope estimates from Models 2 and 3 are presented in the left and right panels, respectively. The solid black line in each panel is a fitted line between economies of scope and total operational costs to show the trend in the results. In both models, there is a negative relationship between firm size and economies of scope. For Model 2, there is an unambiguous relationship between firm size and economies of scope. The largest economies of scope estimates are for the smallest firms, meaning that the costs of separating an integrated firm into two specialized firms are highest for the smallest firms. The bigger the firms, the lower the economies of scope; in both models, we see that there are even diseconomies of scope for some of the biggest firms, meaning that the operational costs will be lower if production is separated across two specialized firms. In Model 3, there is no clear relationship between firm size and economies of scope. Although there is a negative trend, we also note that even for the smallest firms, there are both economies and diseconomies of scope.

Figure 1: Economies of Scope and Firm Size Using Total Operational Costs as Proxy for Firm Size



Notes: The solid black line is computed by locally weighted scatterplot smoothing (LOWESS).

To further investigate the firm's characteristics, we examine the economies of scope results for different output levels in our sample of Norwegian integrated firms. In Table 5, we present the mean values of the scope estimates for the firms within each output combination. To measure the output value for distribution of electricity, we use km of network, and for generation of electricity, we use electricity in MWh. In the upper and lower parts of Table 5 we see the results from Models 2 and 3, respectively. For Model 2, we can see the same trend in the results as that in Figure 1, where we used total costs as a proxy for firm size. We find the highest economies of scope values for the lowest output combination values, suggesting that the smallest firms in terms of output, have the highest economies of scope. For Model 3, the results are more dispersed, as in Figure 1, and we do not find a clear trend. The overall economies of scope estimates are lower those that in Model 2, and the highest diseconomies of scope are found at the lowest and the highest output level combinations. The highest economies of scope estimates for Model 3 are close to the median of the output distribution.

Table 5: Economies of scope and firm output

Model 2. Generation (*MWh*)	Model 2. Distribution (*Km network*) 200	400	600	800	1000	1200
5000	51% (2)	18% (1)	10% (1)	*	*	*
10000	51% (6)	17% (13)	18% (4)	*	*	*
20000	36% (13)	14% (22)	10% (9)	*	1% (1)	*
50000	28% (30)	10% (35)	8% (3)	*	-2% (8)	*
100000	29% (2)	3% (20)	5% (12)	-0.8% (12)	-8% (10)	-3% (10)
300000	*	0.7% (19)	9% (6)	4% (11)	11% (7)	7% (27)

Model 3. Generation (*MWh*)	Model 3. Distribution (*Km network*) 200	400	600	800	1000	1200
5000	-11% (2)	-2% (1)	2% (1)	*	*	*
10000	2% (6)	-5% (13)	5% (4)	*	*	*
20000	6% (13)	-4% (22)	2% (9)	*	-4% (1)	*
50000	-5% (30)	-6% (35)	7% (3)	*	-6% (8)	*
100000	-9% (2)	-5% (20)	2% (12)	-3% (12)	-10% (10)	-18% (10)
300000	*	2% (19)	-2% (6)	-6% (11)	-8% (7)	-8% (27)

Notes: * means no observations.

In Table 6, we present the economies of scale results for the integrated and specialized firms in distribution and generation from Models 2 and 3. The economies of scale estimates in Model 2 follow the definition in eq. (2). For Model 3 with the translog function, we have used the inverse of the sum of the cost elasticities to compute the economies of scale measures.[13] Except for the smallest percentiles, both models show increasing RTS across all three firm types, which is consistent with previous scale studies using data from Norwegian electricity distribution firms

[13] $RTS = \frac{1}{(el_{IL}+el_{IQ}+el_{IE}+el_{IN})}$, where $el_{IL}, el_{IQ}, el_{IE}, and\ el_{IN}$ are the cost elasticities for the integrated firms considering the outputs L, Q, E, and N, respectively.

(e.g., see Mydland et al., 2016; Kumbhakar et al., 2015). An interesting research topic would be to examine how costs in the industry change if separating the integrated firms leads to more mergers among the distribution companies.[14] By doing a merger analysis, one could see whether the cost savings from optimal merges can balance the loss of not utilizing economies of scope. This topic is left for future research.

[14] The regulated distribution companies can represent "safe income" for an integrated firm because 40% of the revenue cap is decided by the costs in the distribution company. By unbundling distribution and generation into two totally separated firms, one can avoid any cross subsidizations or anticompetitive behavior. This might lead to an increased interest in merging more distribution companies.

Table 6: Economies of Scale Results for Integrated and Specialized Firms

	Integrated firms	
Percentiles	Model 2	Model 3
1%	0.82	0.82
5%	0.92	0.84
10%	0.94	0.92
25%	0.99	1.01
Median	1.07	1.29
75%	1.21	1.69
90%	1.79	2.12
95%	2.67	2.94
99%	2.99	2,65
Mean	1.95	1.70
St. dev.	2.82	1.64

	Specialized firms, distribution	
Percentiles	Model 2	Model 3
1%	1.03	0.89
5%	1.04	0.98
10%	1.06	1.02
25%	1.09	1.12
Median	1.15	1.23
75%	1.26	1.36
90%	1.44	1.48
95%	1.58	1.53
99%	1.92	1.65
Mean	1.21	1.24
St. dev.	0.19	0.18

	Specialized firms, generation	
Percentiles	Model 2	Model 3
1%	0.02	0.72
5%	1.10	0.76
10%	1.22	0.78
25%	1.33	0.85
Median	1.50	1.01
75%	1.81	1.31
90%	2.34	1.73
95%	2.83	2.01
99%	4.35	2.37
Mean	1.80	1.15
St. dev.	4.38	0.40

6. SUMMARY AND CONCLUDING REMARKS

In our study, we found evidence of scope and scale economies in the Norwegian electricity industry. Using the new flexible technology approach, we obtained consistent estimates of scope and scale economies. One important point is that this approach provides the possibility to test whether specialized and integrated firms share the same technology. By applying the flexible technology approach for both the quadratic and translog cost functions using new data from the Norwegian electricity industry, we examined how the results vary among the functional forms and model specifications. However, we found a negative relationship between firm size in terms of total costs, and economies of scope in our models, which suggests that for the smallest companies in the industry, the policy decision on strict separation between generation and distribution is costly. In Model 2 the results show that firms characterized by low output values in both the distribution and generation of electricity have the highest economies of scope. In Model 3, we also find evidence of economies of scope, but the firm's characteristics in terms of output values are not as clear as those in Model 2.

From a political perspective, it might be desirable to separate generation and distribution because natural monopolies in electricity distribution are being regulated, whereas electricity generation occurs in competitive markets. If unbundling distribution and generation leads to increased competition, less cross-subsidization, and more productive DSOs, this might be beneficial for the electricity industry as a whole. However, as this study shows, the policy decision on strict separation also comes with a cost. Moreover, this cost seems to differ across firms. The overall conclusion of this paper is that for the larger firms, there are no incentives to keep the firms integrated, but for the smaller firms, distribution and generation should not be unbundled.

In policy making, all pros and cons should be considered before a political action is taken. One of the cons of introducing strict separation between distribution and generation of electricity is the cost of not utilizing economies of scope. In this paper, we provide new insights into this issue, and our results are useful for the formulation of future political strategy and visions within the energy sector.

For future research, it would be interesting to combine the analysis on both economies of scope and merger gains from utilizing economies of scale to determine how we can expect the cost structure in the Norwegian electricity market to evolve. If the policy of separating integrated firms into specialized firms leads to more mergers, the net cost changes might be positive in the long run, provided that the companies utilize unexploited economies of scale.

ACKNOWLEDGMENTS

The work in this study is a part of the project "Benchmarking for Regulation of Norwegian Electricity Networks" (ElBench), which is funded by Norwegian Water Resources and Energy Directorate, Energy Norway and six electricity companies in Norway (Agder Energi Nett, Skagerak Nett, Eidsiva Nett, Hafslund Nett, BKK Nett, and Istad Nett).

REFERENCES

Arocena, P., Saal, D.S., and T. Coelli (2012). Vertical and horizontal scope economies in the regulated US electric power industry. *Journal of Industrial Economics* 60(3): 434–467.

Battese, G.E. (1997). A note on the estimation of Cobb–Douglas production functions when some explanatory variables have zero values. *Journal of Agricultural Economics* 48(1–3): 250–252.

Baumol, W.J., Panzar, J.C., and R.D. Willig (1982). *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt Brace Jovanovich.

Carvalho, P. and R.C. Marques (2014). Computing economies of vertical integration, economies of scope and economies of scale using partial frontier nonparametric methods. *European Journal of Operational Research* 234(1): 292–307.

Chappell, H.W. and R.P. Wilder (1986). Multiproduct monopoly, regulation, and firm costs: Comment. *Southern Economic Journal* 52(4): 1168–1174.

Clark, T.S. and D.A. Linzer (2015). Should I use fixed or random effects? *Political Science Research and Methods* 3(2): 399–408.

Evans, D.S. and J.J. Heckman (1984). A test for subadditivity of the cost function with an application to the Bell system. *American Economic Review* 74(4): 615–623.

Farsi, M., Fetz, A., and M. Filippini (2008). Economies of scale and scope in multi-utilities. *Energy Journal* 123–143.

Fetz, A. and M. Filippini (2010). Economies of vertical integration in the Swiss electricity sector. *Energy Economics* 32(4): 1325–1330.

Fuss, M.A. and L. Waverman (1981). "Regulation and the multi-product firm: The case of telecommunications in Canada," *Studies in Public Regulation*, Fromm, G. (ed.). Cambridge, MA: MIT Press, 277–328.

Fraquelli, G., Piacenza, M., and D. Vannoni (2004). Scope and scale economies in multi-utilities: Evidence from gas, water and electricity combinations. *Applied Economics* 36(18): 2045–2057.

Garcia, S., Moreaux, M., and A. Reynaud (2007). Measuring economies of vertical integration in network industries: an application to the water sector. *International Journal of Industrial Organization* 25(4): 791–820.

Gugler, K., Liebensteiner, M., and Schmitt, S. (2017). Vertical disintegration in the European electricity sector: Empirical evidence on lost synergies. *International Journal of Industrial Organization* 52(5): 450-478.

Jara-Dıaz, S., Ramos-Real, F.J., and E. Martınez-Budrıa (2004). Economies of integration in the Spanish electricity industry using a multistage cost function. *Energy Economics* 26(6): 995–1013.

Kaserman, D.L. and J.W. Mayo (1991). The measurement of vertical economies and the efficient structure of the electric utility industry. *Journal of Industrial Economics* 39(5): 483–502.

Kumbhakar, S.C., Amundsveen, R., Kvile, H.M., and G. Lien (2015). Scale economies, technical change and efficiency in Norwegian electricity distribution, 1998–2010. *Journal of Productivity Analysis* 43(3): 295–305.

Kwoka, J.E. (2002). Vertical economies in electric power: evidence on integration and its alternatives. *International Journal of Industrial Organization* 20(5): 653–671.

Mayo, J.W. (1984). Multiproduct monopoly, regulation, and firm costs. *Southern Economic Journal* 51(1): 208–218.

Meyer, R. (2012a). Economies of scope in electricity supply and the costs of vertical separation for different unbundling scenarios. *Journal of Regulatory Economics* 42(1): 95–114.

Meyer, R. (2012b). Vertical economies and the costs of separating electricity supply-a review of theoretical and empirical literature. *Energy Journal* 33(4): 161–185.

Mydland, Ø., Haugom, E., and G. Lien (2016), Economies of scale in Norwegian electricity distribution: a quantile regression approach. Working paper. Lillehammer University College.

Norwegian Water Resources and Energy Directorate (NVE) (2015). Utvikling i Nøkkeltal for Nettselskap Med Fokus På Nettstruktur. Report from the Norwegian Water Resources and Energy Directorate (in Norwegian: *Norges vassdrags- og energidirektorat*). http://publikasjoner.nve.no/rapport/2015/rapport2015_28.pdf.

Panzar, J.C. and R.D. Willig (1981). Economies of scope. *American Economic Review* 71(2): 268–272.

Piacenza, M. and D. Vannoni (2004). Choosing among alternative cost function specifications: An application to Italian multi-utilities. *Economics Letters* 82(3): 415–422.

Piacenza, M. and D. Vannoni (2009). Vertical and horizontal economies in the electric utility industry: An integrated approach. *Annals of Public and Cooperative Economics* 80(3): 431–450.

Pulley, L.B. and D.B. Humphrey (1993). The role of fixed costs and cost complementarities in determining scope economies and the cost of narrow banking proposals. *Journal of Business* 66(3): 437–462.

Triebs, T.P., Saal, D.S., Arocena, P., and S.C. Kumbhakar (2016). Estimating economies of scale and scope with flexible technology. *Journal of Productivity Analysis* 45(2): 173–186.

Varian, H.R. (1992). *Microeconomic Analysis*. (3rd ed). New York: W.W. Norton, 67.

Waldman, D.E. and E.J. Jensen (2001). *Organization: Theory and Practice*. New York: Addison Wesley Longman.

APPENDIX

Table A1. Summary of Previous Empirical Scope and Scale Studies of the Combined Generation and Transmission/Distribution Electricity Companies

Author(s)	Data	Functional form	Established method	Economies of scope and scale*
Kaserman and Mayo (1991)	Cross-section (1981, US)	Quadratic cost function	OLS	Economies of scope (EOS) = 0.12 (at mean)
Kwoka (2002)	Cross-section (1989, US)	Quadratic cost function	OLS	EOS = 0.27 (at median). Reports substantial costs of vertical integration and highest for the smallest utilities
Jara-Dıaz et al. (2004)	Panel-data (1985–1996, Spain)	Quadratic cost function together with cost share equations	Seemingly unrelated regressions (SUR)	EOS = 0.065–0.28. Economies of scale returns to scale (RTS) = 1.07.
Piacenza and Vannoni (2009)	Panel-data (1994–2000, Italy)	Multiproduct & multistage Box–Cox transformed cost function	Nonlinear SUR	EOS = 0.24. RTS = 0.96. Reports findings of both vertical integration gains and horizontal economies of scope
Fetz and Filippini (2010)	Panel-data (1997–2005, Switzerland)	Quadratic cost function	Random effects GLS and Random Coefficient model	EOS = 0.50–0.60 (at median). RTS = 1.40–1.70 (at median). Presence of considerable economies of vertical integration and economies of scale for most companies
Arocena et al. (2012)	Cross-section 2001, US)	Quadratic cost function together with cost share equations	SUR	EOS = 0.04–0.10. RTS = 1.01–1.03. Reports positive sample mean estimates of both vertical and horizontal economies
Meyer (2012a)	Panel-data (2001–2008, US)	Quadratic cost function	OLS	EOS = 0.19–0.26, when separating generation from distribution and retail. Reports that if generation and transmission remain integrated but are separated from distribution and retail, EOS = 0.08–0.10.
Triebs et al. (2016)	Panel-data (2000–2003, US)	Flexible technology translog cost functions with different specifications	SUR	EOS = 0.04 (0.40 when zeros are replaced by small numbers in the common cost function model). RTS = 1.10–1.13. Reports evidence of economies of scale and vertical economies of scope.
Gugler et al. (2017)	Panel-data (2000-2010, 16 European countries)	Multistage quadratic cost function	Nonlinear SUR	EOS = 0.14-0.20. Reports that at the median integrated utilities have EOS = 0.14 while large scale utilities have EOS = 0.20.

*Estimates of economies of scale (measured by RTS) are for integrated firms.

Table A2: Parametric Estimate

	Model 1	Model 2			Model 3		
	All firms	Integrated	Distribution	Generation	Integrated	Distribution	Generation
L	25.67	29.28	22.29		0.23	0.96	
	(7.68)	(24.72)	(8.77)		(1.98)	(0.98)	
L*L	-0.02	-0.04	-0.01		0.87	0.22	
	(0.01)	(0.07)	(0.01)		(0.85)	(0.37)	
Q	1.50	1.16	1.52		-0.09	-0.99	
	(0.19)	(1.07)	(0.20)		(1.74)	(0.72)	
Q*Q	-0.000	0.000	-0.000		0.51	0.29	
	(0.00)	(0.00)	(0.00)		(0.54)	(0.24)	
E	0.04	0.11		0.04	-0.04		-0.70
	(0.01)	(0.05)		(0.01)	(1.13)		(0.31)
E*E	-0.000	0.000		-0.000	0.01		0.09
	(0.00)	(0.00)		(0.00)	(0.09)		(0.03)
N	758.60	367.60		782.46	-1.48		-0.49
	(219.60	(1727.86)		(222.34)	(1.84)		(0.39)
N*N	-24.79	-266.76		-26.63	-0.04		-0.03
	(6.46)	(245.35		(6.44)	(0.35)		(0.09)
L*Q	0.001	-0.001	0.001		-0.64	-0.20	
	(0.00)	(0.01)	(0.00)		(0.65)	(0.28)	
L*E	-0.000	-0.001			0.09		
	(0.00)	(0.00)			(0.21)		
L*N	4.40	19.03			-0.25		
	(3.36)	(9.45)			(0.43)		
Q*E	0.000	0.000			-0.06		
	(0.00)	(0.00)			(0.18)		
Q*N	-0.17	-0.55			0.42		
	(0.00)	(0.45)			(0.34)		
E*N	0.004	-0.03		0.004	-0.07		0.08
	(0.001)	(0.02)		(0.00)	(0.15)		(0.04)
α_D	-368.50						
	(1,719.07)						
α_G	530.92						
	(2,046.40)						
t	-744.05	-495.12	-1049.50	-549.24	0.002	0.08	0.09
	(317.06)	(676.33)	(470.86)	(557.39)	(0.09)	(0.03)	(0.04)

t^2	121.86	– 65.04	225.46	34.34	0.003	0.004	-0.01
	(54.45)	(109.90)	(81.90)	(94.84)	(0.00)	(0.00)	(0.00)
t*L	3.83	-0.72	2.24		-0.001	0.008	
	(0.63)	(2.08)	(0.86)		(0.02)	(0.01)	
t*Q	-0,16	0.12	-0.13		-0.007	-0.02	
	(0.02)	(0.11)	(0.02)		(0.02)	(0.01)	
t*E	0.004	0.01		0.01	0.006		-0.01
	(0.00)	(0.11)		(0.00)	(0.01)		(0.00)
t*N	-38,84	35.07		-25.05	-0.001		0.02
	(29.41)	(203.19)		(32.90)	(0.02)		(0.01)
Const.	2,930.17	4,088.680	3575.15	4317.90	7.85	8.28	10.40
	(1995.43)	(3,714.10)	(1974.99)	(2095.94)	(6.64)	(1.70)	(1.75)
Observations	1494	316	671	507	316	671	507
Log-likelihood	Model 1: -15670.5	Model 2: –15654.7			Model 3: –304.37		
R^2	Model 1: 0.89	Model 2: 0.89			Model 3: 0.82		

Standard errors in parentheses. Model 1: Quadratic function, combined, random effects. Model 2: Quadratic function, separate technology, random effects. Model 3: Translog function, separate technology, random effects.
R^2 is calculated as squared correlation between estimated and observed costs, which is the equivalent to R^2 from OLS.

In Table A2, we present the parameter estimates for the three random effects models. Model 1 in the first column is a quadratic cost function and assumes shared technology between the specialized and integrated firms. Models 2 and 3 are the quadratic cost function and the translog cost function, respectively, both using the flexible technology approach. Note that even though the parameter estimates for each firm type are presented in different columns, the cost function for integrated, distribution, and generation firms is estimated jointly in both Models 2 and 3. In Model 1, the constant term refers to α_o in eq. (1), and in Models 2 and 3, it refers to *Idum*, *Ddum*, and *Gdum* from eqs (4) and (5). In both Models 1 and 2, we use a quadratic cost function, which means that the estimation results are in levels. In the translog cost function in Model 3, we have taken the log of the variables, and we cannot compare the results in Models 1 and 2 directly.

ESSAY 3

Ørjan Mydland

Lost economies of scope and potential merger gains in the Norwegian electricity industry

Submitted to Empirical Economics

Lost economies of scope and potential merger gains in the Norwegian electricity industry

Ørjan Mydland*
Faculty of Economics and Organization Science, Inland Norway University of Applied Sciences, Lillehammer, Norway
and Department of Business and Management Science, Norwegian School of Economics, Bergen, Norway. Postal address: Inland Norway University of Applied Sciences, Faculty of Economics and Organization Science, Pb 952, NO-2604 Lillehammer, Norway, Email: orjan.mydland@inn.no / orjan.mydland@nhh.no, Phone: +47 909 35 042

ABSTRACT

In 2016, the Norwegian Parliament amended the Energy Act, with changes taking effect from 2021. The amended legislation will introduce strict separation of all *generation* and *distribution* companies within the electricity industry in Norway. Economies of scope studies from Norway show evidence of economies of scope. Further, the companies in the industry could utilize the economies of scale potential if they merged. In this paper, we perform merger analysis to investigate best- and worst-scenario outcomes regarding the cost effects in the industry from the amendments in the Norwegian Energy Act. By providing a method of testing for optimal mergers, we can present the best merger combination to the Norwegian electricity industry.

Keywords: Potential merger gains, Energy, Policy, Energy Act, Lost economies of scope, Optimized mergers.

1. Introduction

The organizational structure of electricity industries has been one of vertical integration. Up until the Energy Act of Norway came into force on January 1, 1991, electricity utilities in Norway performed all of the services in the power delivery value chain, from generation of electricity to supply. In a vertically integrated utility, the management benefits in terms of both short- and long-term planning by controlling the entire value chain. The main motivation for the unbundling of services is to increase competition in the electricity industry, to avoid cross-subsidization, and to make the *distribution* system operators focus only on network operations.

The deregulation in Norway during the 1990s have led to a more market-oriented environment in the electricity industry. The basic idea behind the deregulation of this sector was to separate the services in the value chain, regulate the monopolistic services of transporting electricity to consumers, and expose others (generation, wholesale, and supply) to competition. To secure efficient markets, regulatory rules for unbundling were developed. The European Union's Third Energy Package from 2009 consists of directives meant to effectively unbundle energy production and supply interests from the *distribution* of electricity. The directives impose rules on ownership unbundling for the largest companies in *generation* and *distribution* of electricity. In 2016, the Norwegian Parliament amended the Energy Act, with changes to take effect from 2021. The amended legislation will introduce strict separation of all *generation* and *distribution* companies within the electricity industry in Norway. The main motivations for these changes in the Energy Act are as follows. (1) To increase competition in the market. Generation of electricity is exposed to competition whereas distribution of electricity is not, because distributors have the character of a natural monopoly. However, if the integrated firms (electricity firms with both generation and distribution of electricity) prioritize the distribution of electricity from their own generation, this could lead to market failure because integrated firms have an advantage compared with firms that only have generation of electricity (specialized firms). (2) To avoid cross-subsidization. Because it is not possible to impose competition directly into the market for the distribution of electricity, the regulator, the Norwegian water and energy directorate (NVE), make the distribution companies compete in a benchmarking model. Sixty percent of the revenue cap (the total amount of income each distribution company can charge their customers in net rent each year) is decided by the efficiency score for the company. Forty percent of the distribution company's income is decided by the annual actual costs for each company. Because a portion of the company's income is decided by its actual costs, this might be regarded as "safe income" that, no matter what, the company will retrieve from its customers. This can be problematic for the following two reasons. First, if there is some cross-subsidization within the integrated firms, between distribution of electricity and generation of electricity, this will decrease competition because the integrated firms have some advantages over the specialized firms. Second, previous studies of the Norwegian electricity distribution industry have shown that the distribution companies are too small, and several of them should merge to utilize unexploited economies of scale in the industry (Kumbhakar et al., 2015; Mydland et al., 2018a). If the possibility to cross-subsidize and to prioritize their own generation of electricity leads the companies to be less willing to sell or merge their distribution services with other companies, this is problematic because the Norwegian distribution electricity industry does not utilize economies of scale. (3) To make the distribution companies focus on network operations only. To concentrate on both generation and distribution might be challenging. The policy makers believe that, by unbundling generation and distribution, the leaders of the distribution companies will be more able to focus on distribution services, and to increase efficiency.

Although all of the arguments above seem valid, we also know that the policy of unbundling generation and distribution of electricity comes with a cost. Mydland et al. (2018b) report that, on average, the costs increase by 8% if the generation and distribution of electricity is strictly separated in the Norwegian electricity industry. This cost increase is caused by not utilizing economies of scope.[1] The aim for this study is to seek the potential merger gains in the Norwegian electricity industry for the 55 distribution companies that will be affected by the changes in the Energy Act of Norway. To perform this study, we follow the framework introduced by Bogetoft and Wang (2005), based on a standard *data envelopment analysis* (DEA), imposing different model specifications and assumptions on returns to scale (RTS).

The remainder of the paper is organized as follows. Section 2 gives an overview of previous literature. In Section 3, the framework and model specifications in this analysis are presented. The data are presented in Section 4 and, in Section 5, the results are presented. Section 6 provides conclusions.

2. Literature overview

There is only one recent published study on potential merger gains in the Norwegian electricity distribution industry, Saastamoinen et al. (2017), which investigates potential merger gains between all distribution companies in Norway. They use both DEA and StoNED models to estimate potential merger gains. The study reports some merger gains and illustrates that potential gains from the mergers may vary on the assumption of the regulatory model about the production technology. Saastamoinen et al. (2017) seek merger gains between all distribution companies in Norway, whereas our study seeks to find potential merger gains based on data for the companies that will be affected by the amended Energy Act of Norway. Further, Saastamoinen et al. (2017) concentrate on potential merger gains between firms located geographically close.

In the Norwegian electricity industry, there exist examples of firms merging without any

[1] Economies of scope, sometimes called economies of vertical integration, means that one firm producing two products can do this at a lower cost than if two companies produce one of the product each.

geographical proximity. Hence, it is interesting to find potential merger gains without considering geographical proximity. One could argue that, if mergers were beneficial for the distribution companies in Norway, they would already be merged. However, there are only a few examples of recent mergers between Norwegian electricity distribution companies. Even if there are potential merger gains, they might not be feasible because the companies in the industry do not consider them as beneficial. The ownership structure in the Norwegian electricity industry consists of many municipalities. These might find importance in owning their own distribution companies to have security of supply, secure jobs, and be able to control dividends from the companies. However, some of the motivation for the policy makers to introduce strict separation between generation and distribution of electricity is that the distribution companies might be more willing to merge after the reform.

Whereas Saastamoinen et al. (2017) use data for the period 2004–2012, we add two more years, 2004–2014.

Agrell and Teusch (2016) study the effect from realized mergers within the Norwegian electricity industry in the period 1994–2004. They report only small gains from these realized mergers. There are few merger studies within the electricity sectors also for other countries. Bagdadioglu et al. (2007) look at Turkish electricity companies for the period 1999–2003. Kwoka and Pollitt (2010) report no merger gains from US panel data for the period 1994–2003. Çelen (2013) also uses panel data from the US for the period 2002–2009 and reports no evidence of merger gains. Almost all these previous studies apply DEA in their analysis. Only Çelen (2013) conducts the merger analysis with stochastic frontier analysis (SFA). As pointed out in Bogetoft and Otto (2010), there are some problems with the SFA approach. The main problem is that one cannot rule out a "learning effect" (explained later in this paper), and it is difficult to identify the pure merger gains from the results from the analysis. Saastamoinen et al. (2017) apply the StoNED model to their analysis in addition to the more common DEA model. Their analysis shows that the results are quite sensitive to model specifications and the method of estimation. In Table 1, we present a brief overview of the previous studies.

Table 1. Summary of previous merger studies on the electricity distribution industry.

Author(s)	Data	Viewpoints	Est. method/ model	Merger gains*
Bagdadioglu et al. (2007)	Panel data (1999–2003, Turkey)	Effects on productive efficiency	DEA	Yes
Kwoka and Pollitt (2010)	Panel data (1994–2003, US)	Effects on productive efficiency	DEA	No
Çelen (2013)	Panel data (2002–2009, US)	Effects on productive efficiency	SFA	No
Agrell and Teusch (2016) Working paper	Panel data (1994–2004, Norway)	Effects on productivity efficiency from realized mergers	DEA	Reports rather small gains from realized mergers in the period 1994–2004. Largest effect from internal efficiency increases within the firms
Saastamoinen et al. (2017)	Yearly average (2004–2012, Norway)	Effects on productivity efficiency with different assumptions on the regulatory model and the production technology	DEA, StoNED	Reports some merger gains and illustrates that potential merger gains from the mergers may vary on the assumption of the regulatory model about the production technology

3. Model specification and conceptual framework

This section introduces the Bogetoft and Wang framework for measuring merger gains. In our model, we use total costs as an input measure because, in distribution of electricity in Norway, the demand and, hence, the outputs are exogenous.

3.1. Mergers

The merger gains can be explained by the reduction in costs relative to the outputs, after one or more companies have merged. This is what is called unadjusted gains, and it can be decomposed into several measures. Following Bogetoft and Wang (2005), the unadjusted gains are defined by

$$E^H = c\left(\sum_{k\in H} y^k\right) \,/\, \sum_{k\in H} x^k \tag{1}$$

where c(.) is the estimated cost function; y^k is the output vector for the i-th company; x^k is the actual costs for the i-th company; and H denotes the set of merging companies (|H| = 2 means that the measure shows the effect of two companies merging). If E^H is less, greater or equal to unity, it indicates potential gains, losses, or no effect from the mergers, respectively.

The learning effect is each company's individual learning effect, meaning the potential to be more efficient. Hence, this is not linked to the merger. Therefore, in our study, we rule out this effect. The learning effect is given by

$$LE^H = \sum_{k\in H} c(y^k) \,/ \sum_{k\in H} x^k \tag{2}$$

By ruling out the learning effect, we get the adjusted gains, which is given by

$$E^* = c\left(\sum_{k\in H} y^k\right) \,/\, \sum_{k\in H} c(y^k) \tag{3}$$

The adjusted gains can further be decomposed into the harmony effect and the size effect. The harmony effect is given by

$$HA^H = c\left(\frac{1}{|H|}\sum_{k\in H} y^k\right) \,/\, \left(\frac{1}{|H|}\sum_{k\in H} c(y^k)\right) \tag{4}$$

where we find the possible reduction in the average input in the production of the average output. The size effect is given by

$$SI^H = c\left(\sum_{k\in H} y^k\right) / \left(|H|\, c\left(\frac{1}{|H|}\sum_{k\in H} y^k\right)\right) \tag{5}$$

The learning effect LE^H is a measure of the cost reduction when we assume that every firm learns best practices (become efficient) but do not merge with any other firm. HA^H is the harmony effect and is a measure of the minimal cost of the average output vector compared with the average of the costs adjusted for LE^H. The size effect SI^H is a measure of the cost of operating at full (merged) scale compared with operating at average scale of the original firms. The harmony effect, also called the scope effect, is linked to the ability of the companies to better utilize inputs after the merger, whereas the size effect, also called the scale effect, is the movement toward optimal scale from a merger. By using equations (4) and (5) above, we can readily see that $E^* = HA^H * SI^H$.

It is easy to illustrate the conceptual framework of potential merger gains. In Fig. 1, we see two initial firms (A and B) and their production possibility set. There are two outputs (y_1 and y_2) relative to costs (x). If firms A and B merge into firm C, the unadjusted merger gains in equation (1) correspond to the distance between C and C''. However, we want to rule out the learning effect, see equation (2), meaning firm A's and B's movement to the frontier to point A' and B', respectively. The adjusted (true) merger gains in equation (3) correspond to the distance between C' and C''. This measure can further be decomposed into the harmony effect (equation (4)) and the size effect (equation (5)). Note that under constant returns to scale (CRS), the size effect is zero, hence, $E^* = HA^H$.

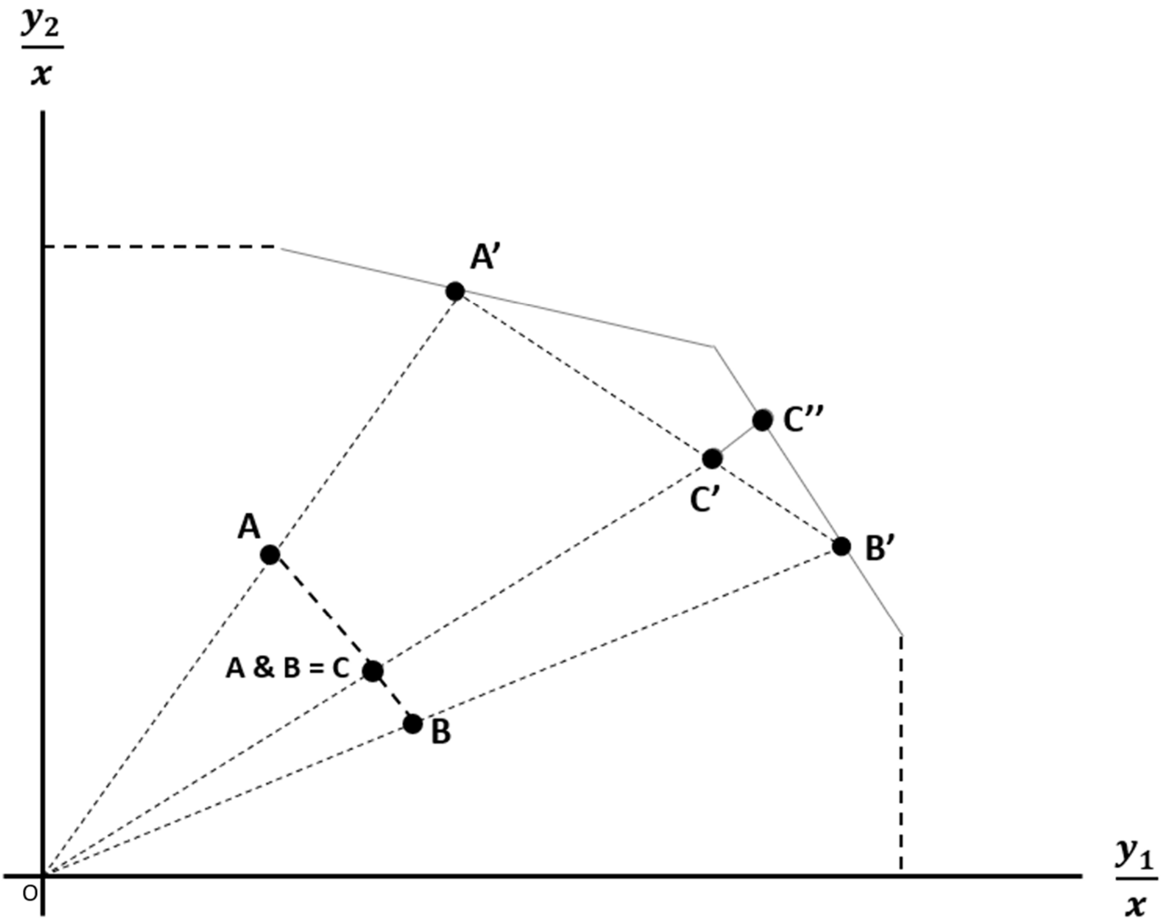


Fig. 1. Conceptual framework merger gains.

3.2. DEA estimator

To find the estimated costs and the efficiency scores for each firm, the following minimization problem must be solved:

$$
\begin{aligned}
&\min_{e,\lambda^1,\dots,\lambda^K} e \\
&s.t. \quad ex_i^o \geq \sum_{k=1}^{K} \lambda^k x_i^k, \quad \forall i = 1,\dots,m \\
&y^{oj} \leq \sum_{k=1}^{K} \lambda^k y_j^k, \qquad \forall j = 1,\dots,n \\
&\lambda^k \geq 0, \qquad \forall k = 1,\dots,K.
\end{aligned}
\tag{6}
$$

This is an input-oriented DEA model, similar to the DEA model used in the regulation of the Norwegian electricity distribution companies by NVE. The input orientation is because we assume the outputs to be exogenous (demand is fixed), meaning that the companies can only adjust their inputs (costs) to be more efficient. In (6), CRS is assumed, as in the regulator model. From the regulators standpoint, it might be desirable to treat all the firms the same in regard to size, hence, CRS is appropriate. This gives the firms incentives to merge if they are too small or split up if they are too big. However, variable returns to scale (VRS) might be more realistic to the actual technology in the industry. To implement VRS in the model, the convexity constraint $\lambda^k \geq 0$ changes to $\sum \lambda^k = 1$. Introducing decreasing returns to scale (DRS) or increasing returns to scale (IRS) would favor the small or the large firms, respectively, and is therefore not interesting in this analysis.[2,3]

By solving the minimization problem in equation (6), we find the piecewise linear

---

[2] DRS (IRS) are sometimes referred to as non-increasing (non-decreasing) returns to scale.

[3] The free disposable hull (FDH) and the free requirement hull (FRH) are two more possible assumptions on scale (not always called DEA-models). However, to implement these assumptions, the convexity constraint must be avoided; see Bogetoft and Otto (2010).

production frontier illustrated in Fig. 2. The frontier is determined by the efficient companies (VRS) or company (CRS). All deviations from the frontier are regarded as due to inefficiency. Therefore, all companies represented by the black dots, along the frontier in Fig. 2, are 100% efficient, and all companies below the frontier are less efficient.

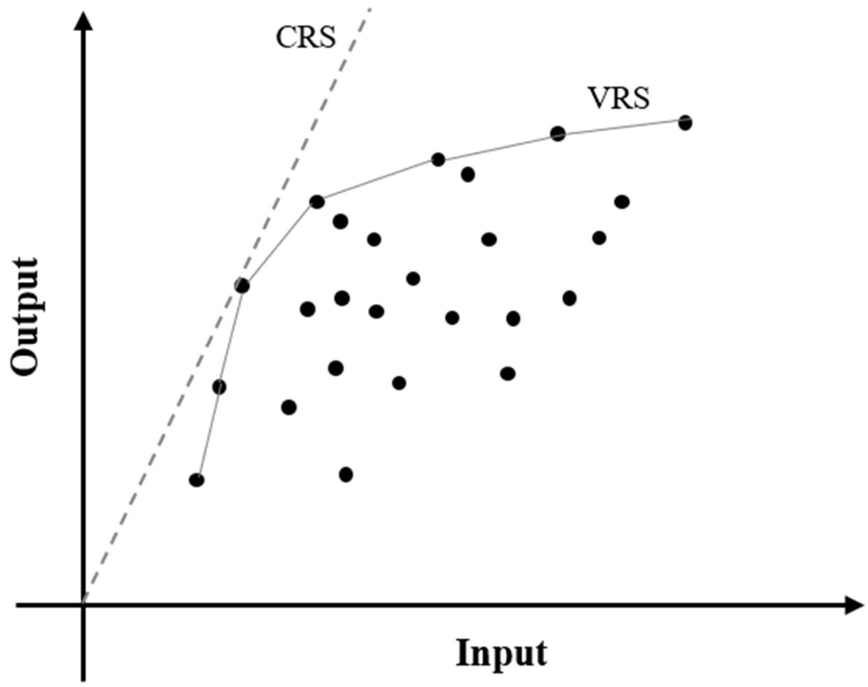


Fig. 2. Illustration of DEA.

3.3. Adjusting for firm heterogeneity

The Norwegian electricity distribution companies are operating in different environments. The weather conditions, the amount of rainfall/snowfall or the steepness of the terrain are examples of variables that can affect the costs of the firm. To take the environmental variables into account, but still following the Bogetoft and Wang (2005) framework, the costs are adjusted. To do this, we follow the procedure by Barnum and Gleason (2008). By finding the parameter estimates for the environmental variables by using a cost function model, the observed costs in our data can be adjusted:

$$lnx_i = \alpha + \hat{\boldsymbol{y}}_i{}'\boldsymbol{\theta} + \mathbf{z}_i{}'\boldsymbol{\delta} + u_i \tag{7}$$

where $\hat{\boldsymbol{y}}_i$ is a vector of logged output values for the i-th firm; $\boldsymbol{\theta}$ contains the parameter estimates; $\mathbf{z}$ is the vector of the environmental variables; and $\boldsymbol{\delta}$ is the vector of parameter estimates for these variables, which gives the effect on costs from the environmental variables. The observed costs are adjusted, as indicated in the following equation:

$$\tilde{x}_i = \exp(\ln(x_i) - \boldsymbol{z}_i{}'\boldsymbol{\delta}) = x_i \exp(-\boldsymbol{z}_i{}'\boldsymbol{\delta}) \tag{8}$$

4. Data

The data include yearly averages for 55 integrated firms during the period 2004–2014. There are two main reasons why we use average data instead of yearly data. First, by averaging, we reduce the effect of noise in the data; see Kuosmanen et al. (2013). Second, by using yearly data, we would obtain potential merger gains at different points in time. This would be difficult to analyze and to give meaningful conclusions. Further, the regulator in Norway also applies average data to secure a more stable cost frontier; see Amundsveen and Kvile (2016).

The integrated firms have positive outputs in both generation and distribution of electricity over the whole period. However, in this analysis, the aim is to investigate potential merger gains between the distribution companies. In Table 2, descriptive statistics for the costs and outputs for these firm's distribution services are presented. The outputs representing the main cost drivers in the industry are the *number of network stations* (NT), the *length of network* (N), and the *number of customers* (Q). *Length of network* is the length of the (high-voltage) distribution network in kilometers. The *number of customers* is the total number of entities (firms and households) that pay the network tariffs. The single input is *total costs* (TOTEX). For the entire industry in 2013, *total costs* were about 13.5 billion Norwegian kroner (NOK).[4] whereas *total cost* on average was approximately 110 million NOK. In 2013, the largest

[4] 1 USD = 8.53 NOK, 1 EUR = 10.01 NOK on July 3, 2018.

company (Hafslund) had a total cost of 1.6 billion NOK while the smallest company (Modalen Kraftlag) had a total cost of only about 3.4 million NOK. Considering the outputs and costs of the firms in the industry, it is clear that there are large variations in firm size. In Table 2, the environmental variables are also included, and there is large variation in the different variables describing the environment in which each company operates. For example, the *length of network* (km) range from about 32 km to over 6,000 km, and the environmental variable that gives the proportion of sea cables shows a range from 2% to 30%. For underground cables, the range is from 24% to 51%, and so on. This shows the existence of firm heterogeneity for which we should control.

Table 2. Descriptive statistics. Yearly average 2004–2014.

Variables	Label	Mean	SD	Min	Median	Max
Number of network stations	NT	507	511	30	318	2,974
Length of network, in km	N	484	469	34	329	2,938
Number of customers	Q	8,334	10,077	385	4,961	56,758
Total costs, 1,000 NOK (2010)	TOTEX	48,738	51,631	2,912	31,610	285,327
Firm-type observations		55 distribution companies that are integrated firms (positive outputs in both generation and distribution of electricity)				
Environmental variables						
Underground cables	Z1	0.24	0.11	0.07	0.21	0.51
Sea cables	Z2	0.02	0.04	0.00	0.01	0.30
Overhead cables	Z3	0.09	0.09	0.00	0.06	0.39
Average slope of terrain	Z4	11.38	3.54	4.69	10.98	22.22
Average distance in km to road	Z5	281.20	232.50	84.69	178.86	924.09
Number of islands	Z6	2.80	5.76	0.00	0.00	25.00
Deciduous forest	Z7	0.09	0.09	0.00	0.05	0.31
Coastal climate	Z8	0.17	0.42	0.00	0.02	2.30

4.1. Pairwise mergers based on minimal distance

In this study, pairwise mergers of the 55 electricity distribution companies are examined. First, the mergers are decided by minimization of the geographical distance between all the 55 distribution companies. This is done by listing all possible combinations (55*54 / 2 = 1485 combinations) and then finding the 27 pairs that gives the total minimum distance in kilometers when each firm is only merged once.[5]

In Fig. 3, the map of Norway with all the pairwise mergers is presented. The map shows the 27 mergers decided by minimization of geographical distance. Because we have 55 companies, and due to the large distances between distribution companies in the north of Norway, we find that around Alta, three companies are merged. We want to find out how the cost in the industry is affected if all these mergers where implemented.

[5] The distance is measured by using the haversine formula, see Robusto (1957).

Fig. 3. Geographical presentation of pairwise merges by minimization of geographical distance. (South of Norway left. North of Norway right.)

4.2. Pairwise mergers based on maximum cost reduction

Because we have examples of mergers between distribution companies in Norway that are not located geographically close, it is also interesting to find the potential merger gains if we optimize the merger combinations while relaxing the geographical restrictions, meaning that one merger can consist of two companies located in different parts of the country. By doing this, we can find the total maximum potential merger gain in the electricity distribution industry in Norway. Hence, the goal is to maximize the sum of gains from all the pairwise mergers.

While keeping the initial frontier constant, we run all the 1 485 possible combinations of pairwise mergers. By sorting the result on potential merger gains, we can find the maximum cost reducing combination of potential merger gains when each firm is only merged once. This gives us approximately the optimal potential pairwise merger gains from the companies included in this analysis.[6]

5. Results

5.1. Potential gains from pairwise mergers based on minimal distance

In this section, the potential merger gains from the DEA analysis where geographical distance is minimized are presented. The analysis is conducted with the assumptions of CRS and VRS in the model. In addition, in both assumptions on technology, the models are estimated with initial observed costs and with observed costs adjusted by the environmental variables, following the procedure described in equations (7) and (8). This gives us four different DEA models.

The results from the four model specifications in this analysis are presented in Table 3. In the first column, the results from the CRS model with observed costs are presented. In the second column, the results from the CRS_Z model with adjusted costs are presented. When assuming CRS, the size effect will, of course, be zero. It does not matter how big the companies get after the merger because returns to scale are constant. Therefore, in the CRS models, we see that E* = HA. In a convex function, HA will, by construction, always be less than or equal to one; see equation (4). Here this means that HA will never be less than zero (the cost reduction is non-negative). Hence, because E* = HA, the size effect when assuming CRS will never be different from zero. For the two CRS models, the E* is the interesting result, and we have presented the minimum, maximum, mean, and median values of the potential merger gains on

[6] Let v(i, k) be the merger gains from merging firm i and k. In theory we can have a situation where v(i , k) + v(j , l) > v(i , j) + v(k , l) even if v(i , j) gives the highest pairwise merger gains. In our analysis v(i , j) will be chosen first, making v(i, k) and v(j , l) not possible.

cost reduction. The standard deviation is also presented to show the variation in the results. At the bottom of each column, we have calculated the total cost effect in 1000 NOK to the industry if all mergers presented in section 4.1 were realized. The mean value for the CRS model shows that the potential merger gains gives a reduction in cost of 1.18%, the maximum cost reduction is 3.79%, whereas the minimum value is zero. We see that the total potential merger gains from these firms merging (see Fig. 3) represents a cost reduction of about 27 million NOK. In the CRS_Z model, where the environmental variables have been used to adjust costs, we see that overall the effect on reduced costs is lower. The total potential cost reduction to the industry is 10.5 million NOK.

In the VRS and VRS_Z models, we see that the mean results of E* in both models gives a cost increase to the industry. However, we can also see that the median represents a cost reduction in the industry. We can observe that the minimum value is very high, meaning high cost increases. In the two models where we assume VRS, the harmony effect can be different from the adjusted gains effect because the size effect now also applies. We see from the harmony effect that both the mean and median in the VRS- and the VRS_Z-models are positive, meaning cost reductions. The size effect is negative at the mean but positive at the median. The minimum value in the VRS model gives a cost increase of almost 43% and 21% in the VRS_Z model. The total cost effect on the industry is negative in both VRS models owing to the gains for several mergers being negative, meaning it will lead to a cost increase for these companies and the industry, (cost increase of 116 million NOK in the VRS model, and a cost increase of 40 million NOK in the VRS_Z model).

Table 3. Potential gains from pairwise mergers based on minimal distance

		CRS	CRS_Z	VRS	VRS_Z
E*					
	Min	0.00%	0.00%	−35.40%	−17.85%
	Max	3.79%	2.39%	8.24%	6.19%
	Mean	1.18%	0.67%	−3.04%	−0.10%
	Median	1.11%	0.23%	2.23%	1.92%
	SD	0.0011	0.0072	0.1073	0.0571
HA					
	Min	0.00%	0.00%	0.00%	0.00%
	Max	3.79%	2.39%	6.20%	7.67%
	Mean	1.18%	0.67%	1.97%	1.60%
	Median	1.11%	0.23%	1.64%	1.11%
	SD	0.0011	0.0072	0.0168	0.0185
SI					
	Min	0.00%	0.00%	−42.81%	−20.53%
	Max	0.00%	0.00%	4.62%	4.12%
	Mean	0.00%	0.00%	−5.04%	−2.41%
	Median	0.00%	0.00%	0.92%	1.48%
	SD	0.0000	0.0000	0.1183	0.0649
1000 NOK		27,631	10,467	−116,564	−40,072

Notes: E* = Adjusted merger gains; HA = Harmony effect; SD = standard deviation; SI = size effect.
Negative numbers mean cost increases whereas positive numbers mean cost reductions to the industry.

The results from the VRS models tell us that several of the suggested merger combinations will lead to high cost increases if the mergers are being completed. Further, this indicates that some of the firms are too big initially to have any positive effect on a merge with another firm, which is surprising because these firms are natural monopolies. This also contradicts the findings in other studies on scale economies on the Norwegian electricity distribution industry; see Kumbhakar et al. (2015) and Mydland et al. (2018a). However, the findings in this analysis correspond to the results in Saastamoinen et al. (2017). They report very high size-related diseconomies (cost increases) in both the DEA and the StoNED models when assuming VRS. This might suggest that the assumption of VRS is not an appropriate assumption with respect to the technology of the firms in this merger analysis. The Norwegian regulator assumes CRS in the regulation model because this gives the firms incentives to merge if they are too small.

5.2. Potential gains from pairwise mergers based on maximum cost reduction

In this section, we focus on the results from finding the optimized merger combinations in the CRS and the CRS_Z models. In Table 4, the results from the optimized mergers are presented. Because we assume CRS technology, the SI effect will be zero, and, thereby, the HA effect will be equal to E*. We, therefore, drop HA and SI from the presentation in Table 4. Note that the merging firms are not necessarily the same in the CRS and the CRS_Z models.[7] It is an interesting result that even when we optimize the merger combinations, we still find that some mergers will give zero effect on costs. However, compared with the situation where the mergers were decided by minimal geographical distance, we see that these results give higher cost reductions to the industry. In the CRS model, we find that if all these mergers were implemented, the total cost reduction in the Norwegian electricity distribution industry would be 68 million NOK and 28 million NOK in the CRS_Z model. When we adjust the costs by the environmental variables in the CRS_Z model, we see that the potential merger gains are lower. This is because when we adjust the cost for the environmental variables, on average, the cost for each firm reduces by about 22%. This means that we reduce the cost base and the potential gains with respect to cost reductions before we run the DEA merger analysis.

---

[7] Figs. A1 and A2 in the Appendix show which companies are merging in each model.

Table 4. Potential cost reduction from optimized merger combinations. CRS and CRS_Z.

		CRS	CRS_Z
E*			
	Min	0.00%	0.00%
	Max	11.28%	6.50%
	Mean	4.08%	2.13%
	Median	4.51%	2.11%
	SD	0.0342	0.0175
1000 NOK		68,369	28,388

Notes: E* = Adjusted merger gains.

6. Conclusions

This study provides insight into the potential merger gains for the electricity distribution companies in Norway that are affected by the changes in the Norwegian Energy Act. The results show evidence of potential merger gains for some of the companies in this study. However, the potential gains are not very high. The sum of the costs for all firms included in this study is about 2.5 billion NOK. When we optimize the merger combinations, the highest total potential merger gain on costs to the industry is about 68 million NOK. This represents a potential cost reduction of less than 3% from the total costs in the industry.

Mydland et al. (2018b) reports that, on average, the costs will increase by 8% if generation and distribution of electricity are strictly separated in the Norwegian electricity industry. This means that, by comparing the results from these two studies, we see that the losses from not utilizing economies of scope, will lead to a cost increase in the industry, even if all the suggested optimized merger combinations were implemented. The potential merger gains are only one of the reasons for the action of the policy makers to amend the Energy Act of Norway. However, this seems to be an important reason in justifying the new reform taking effect from 2021 for all electricity companies in Norway.

To find the optimal mergers in this study, we have relaxed the restriction on the geographical proximity by checking all possible pairwise merger combination across Norway. This gives us the optimal mergers within the Norwegian electricity distribution industry. In future work, it would be interesting to see the potential merger gains if also the restriction on pairwise mergers where relaxed. By finding potential merger gains from companies merging with more than one company, we might find different potential merger gains.

In this study we hold the initial frontier constant while performing the merger analysis on all companies in the Norwegian industry. In reality the frontier might shift if one or several merges would happen between some of the firms in the industry. If so, then the results from a merger analysis on the other firms in the industry, that has not yet merged, might change. It would be interesting to do a merger analysis where we take into account that mergers between firms might affect the potential merger gains for other firms.

Acknowledgements

The work in this study is a part of the project *Benchmarking for Regulation of Norwegian Electricity Networks* (ElBench). Participants in the project is Norwegian Water Resources and Energy Directorate, Energy Norway and six electricity companies in Norway (BKK Nett, Eidsiva Nett, Hafslund Nett, Helgeland Kraft, Lyse Elnett and Skagerak Nett).

Thanks to Eirik Hoel Høiseth for help with a mathematical problem in a previous version of this paper.

Compliance with Ethical Standards

Conflict of Interest: The authors declare that there exists no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

Agrell, P., Teusch, J., 2016. Merger incentives under yardstick competition: Theory and evidence from Norway. (Unpublished manuscript). http://www.cresse.info/uploadfiles/2016_pa3_pa4.pdf (accessed 29 September 2016).

Amundsveen, R., Kvile, H.M., 2016. Balancing incentives: The development and application of a regulatory benchmarking model, in: Greene, W.H., Khalaf, L., Sickles, R.C., Veall, M., Voia, M.-C. (Eds.), Productivity and Efficiency Analysis. Springer Proceedings in Business and Economics, Springer, New York.

Bagdadioglu, N., Price, C.W., Weyman-Jones, T.G., 2007. Measuring potential gains from mergers among electricity distribution companies in Turkey using a non-parametric model. Energy J. 28, 83–110.

Barnum, D.T., Gleason, J.M., 2008. Bias and precision in the DEA two-stage method. Appl. Econ. 40, 2305–2311.

Bogetoft, P., Wang, D., 2005. Estimating the potential gains from mergers. J. Product. Anal. 23, 145–171.

Bogetoft, P., & Otto, L., 2010. Merger Analysis, in: Benchmarking with DEA, SFA, and R. Springer Science & Business Media, New York, p. 282.

Çelen, A., 2013. The effect of merger and consolidation activities on the efficiency of electricity distribution regions in Turkey. Energy Policy 59, 674–682. https://doi.org/10.1016/j.enpol.2013.04.024

Kumbhakar, S.C., Amundsveen, R., Kvile, H.M., Lien, G., 2015. Scale economies, technical change and efficiency in Norwegian electricity distribution, 1998–2010. J. Product. Anal. 43, 295–305.

Kuosmanen, T., Saastamoinen, A., Sipiläinen, T., 2013. What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. Energy Policy 61, 740–750. https://doi.org/10.1016/j.enpol.2013.05.091

Kwoka, J., Pollitt, M., 2010. Do mergers improve efficiency? Evidence from restructuring the US electric power sector. Int. J. Ind. Organ. 28, 645–656. https://doi.org/10.1016/j.ijindorg.2010.03.001

Mayo, J.W., 1984. Multiproduct monopoly, regulation, and firm costs. South. Econ. J. 51, 208–218.

Mydland, Ø., Haugom, E., Lien, G., 2018a. Economies of scale in Norwegian electricity distribution: A quantile regression approach. Appl. Econ. https://doi.org/10.1080/00036846.2018.1450481

Mydland, Ø., Haugom, E., Kumbhakar, S.C., Lien, G., Amundsveen, R., Kvile, H.M., 2018b. Economies of scope and scale in the Norwegian electricity industry. Working paper. Inland Norway University of Applied Sciences and Norwegian School of Economics. Lillehammer, Norway.

Robusto, C. C., 1957. The cosine-haversine formula. The American Mathematical Monthly, 64(1), 38-40.

Saastamoinen, A., Bjørndal, E., and Bjørndal, M. 2017. Specification of merger gains in the Norwegian electricity distribution industry. Energy Policy, 102, 96–107. https://doi.org/10.1016/j.enpol.2016.12.014

Appendix

Fig. A1. CRS. Optimized mergers, with assumption on constant returns to scale.

Fig. A2. CRS_*Z*. Optimized mergers, with assumption on constant returns to scale where costs are adjusted for environmental variables.

ESSAY 4

Subal C. Kumbhakar, Ørjan Mydland, Andrew Musau, Gudbrand Lien

Disentangling Costs of Persistent and Transient Technical Inefficiency and Input Misallocation: The Case of Norwegian Electricity Distribution Firms

Submitted to The Energy Journal

Disentangling Costs of Persistent and Transient Technical Inefficiency and Input Misallocation: The Case of Norwegian Electricity Distribution Firms

Subal C. Kumbhakar[a], Ørjan Mydland[b], Andrew Musau[b], Gudbrand Lien[b*]

[a] Department of Economics, State University of New York, Binghamton, NY13902-6000, USA
[b] Inland School of Business and Social Science, Inland Norway University of Applied Sciences, Pb 952, NO-2604 Lillehammer, Norway
* Corresponding author. Tel.: +47 9248 8335, E-mail address: Gudbrand.Lien@inn.no

Abstract

Numerous studies have focused on estimating technical inefficiency in electricity distribution firms. Most of these studies did not distinguish between persistent and transient technical inefficiency. Furthermore, almost none of the studies estimated the cost of input misallocation arising from non-optimal use of inputs. The cost function used to model inefficiency assumes that all firms are fully efficient allocatively, which is a very strong assumption.

In this study, we estimate both the persistent and transient components of technical inefficiency and input misallocation of Norwegian electricity distribution firms, using panel data from 2000 to 2016. Our modelling and estimation strategy is to use a system approach, consisting of the production function and the first-order conditions of cost minimization. Input misallocation for each pair of inputs is modelled via the first-order conditions of cost minimization. We also estimate the costs of each component of technical inefficiency and input misallocation by deriving the cost function for a multiple-output separable production technology. Our modelling and estimation strategy handles endogeneity of inputs. Finally, we allow for the determinants of persistent and transient technical inefficiency. Our results show that the costs of input misallocation of Norwegian electricity distribution firms are non-negligible.

Keywords: Cost and production functions, Allocative and technical inefficiency, Determinants of inefficiency, Norwegian electricity distribution firms

JEL Classification: C31; D21

1 Introduction

Application of stochastic frontier analysis (SFA) to study inefficiency of electricity generation and/or distribution firms has overwhelmingly focused on technical inefficiency. It is often assumed that all firms are fully allocatively efficient, that is, there is no input misallocation (under or overutilization) that results from failure to minimize costs exactly because of some institutional, structural, and managerial problems. Furthermore, few studies go further to distinguish between persistent (long-term/time-invariant) and transient (short-term/time-variant) technical inefficiency. There is also a lack of empirical research investigating the extent to which these omissions matter. The purpose of this paper is to address these questions using panel data for Norwegian electricity distribution firms.

Electricity distribution firms in Norway have the characteristics of natural monopolies within their service territories. As a part of greater market orientation introduced to the industry during the 1990s, the firms are regulated. The Norwegian regulator, Norwegian Water Resources and Energy Directorate (NVE), uses a benchmarking model to estimate each firm's technical efficiency score, while not identifying input misallocation. The efficiency scores, as the regulator calculate for each firm, determine sixty per cent of the firms' revenue cap. From a regulator's point of view, the main focus is to motivate the firms to increase productivity and efficiency without going into micro management. However, in reality, the firms could reduce costs in production by changing its input allocation. This could be important for the society, the consumers of electricity and the owners of the electricity distribution firms. Because the NVE's regulation model is a one-year model, it does not distinguish between transient and persistent technical efficiency. Again, this might not be a direct problem within the task given to the regulator, but if the goal is to minimize overall (economic) costs, one should distinguish between different sources of inefficiency because it is likely that the solution to the sources of inefficiency might differ, and disentangling these sources might influence the overall technical inefficiency score. Also ignoring persistent inefficiency, for example, might affect the estimates of the technology. In this study, we apply panel data that make it possible to disentangle persistent (time-invariant) and transient (time-variant) inefficiency. Ignoring input misallocation can be more serious (see Kumbhakar and Wang, 2006a) because the estimated technology parameters may be inconsistent.

Previous studies that estimate both technical inefficiency and input misallocation have widely adopted the dual approach that utilizes the duality between the cost and production functions. However, estimation of a cost function with both technical inefficiency and input misallocation is quite complex (Kumbhakar and Wang, 2006b). Estimation of the production function alone cannot accommodate input misallocation. The alternative is to use a primal system that uses the production function and the first-order condition for cost minimization. Schmidt and Lovell (1979) used this approach, which is flexible enough to incorporate both technical inefficiency and input misallocation, and costs therefrom. However, they used a Cobb–Douglas (CD) production function. Subsequently, Kumbhakar and Wang (2006b) extended their modelling approach and used a translog production function. Although they used panel data, their model is essentially cross-sectional. An advantage of this approach is that in addition to estimating technical inefficiency and input misallocation, the model takes into account endogeneity of inputs. Some studies have used the production function and the first-order conditions for profit maximization or a profit function and the implied demand system using Hotelling's lemma to estimate the profit loss from each component (see Kumbhakar et al., 2015 and references therein). This approach takes into account endogeneity of both inputs and outputs. Another approach is to study the observed demands to determine whether these are the cost-minimizing demands consistent with observed prices. Alternatively, one can seek to find the set of prices that would make observed demand cost minimizing. In all of the described approaches above, a number of variations have been presented over the years (for more details, see Kumbhakar, 2015, ch. 8, and Greene, 1993).

To the best of our knowledge, there exists only one recent SFA study on technical inefficiency and input misallocation within the electricity distribution industry. Nemeto and Goto (2006) used a CES cost frontier to study technical inefficiency and input misallocation in the Japanese electricity transmission and distribution industry. Using panel data of Japanese utilities for the period 1981–1998, they reported that technical inefficiency raises costs by 1–28%, while input misallocation raises costs by 8–30%.

Our approach is an extension of Kumbhakar (1988), and it involves estimating a production function frontier together with the first-order conditions of cost minimization. Kumbhakar (1988) introduced a flexible functional form of production technology, which permits elasticity of output to vary across firms, and introduced input misallocation separate from random errors in optimization. In our paper, we consider some extensions of the Kumbhakar (1988) model to study inefficiency in the Norwegian electricity distribution industry where we disentangle persistent and transient technical inefficiency and at the same time estimate input misallocation for each pair of variable inputs. We also estimate the costs of each of these inefficiency components. Furthermore, our model allows for multiple inputs and outputs, handles endogeneity of inputs and includes the determinants of persistent and transient technical inefficiency.

Our study has two main contributions. First, we extend the sparse SFA literature on modelling both technical inefficiency and input misallocation in electricity distribution industries. Second, we extend the modelling and estimation approach by including additional inefficiency components, decomposing them into persistent and time-varying components and also include inefficiency determinants for both of these.

The remainder of the paper is organized as follows. The model specifications and estimation method are described in Section 2. Section 3 describes the data, and Section 4 presents the results. In Section 5, we present a summary of our main results and concluding remarks.

2 Model specification and estimation method

2.1 The model

We consider the production function used in Kumbhakar (1988) extended to accommodate a generalized error specification

$$F(Y_{it}) = \alpha_0 \left(\prod_{j=1}^{J} X_{jit}^{\alpha_j}\right) e^{-u_{i0}-u_{it}+v_{it}} \tag{1}$$

where Y_{it} is output for firm *i* and time *t* $(i = 1, \cdots, N;\ t = 1, \cdots, T)$, X_j are inputs $(j = 1, \cdots, J)$, and α_0 and α_j are the parameters to be estimated.[1] v_{it} is the noise term that captures exogenous shocks unknown to the producer, u_{i0} is persistent inefficiency and u_{it} is transient inefficiency. We extend the model to incorporate a multiple-output separable production technology. Assuming a translog functional form of $F(Y_{it})$; that is

$$ln\, F(Y_{it}) = lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\,\beta_3 lnY_{1it} \cdot lnY_{2it} \tag{2}$$

and substituting (2) into the log form of (1), we can rewrite (1) as

$$lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\,\beta_3 lnY_{1it} \cdot lnY_{2it} = \ ln\alpha_0 + \sum_{j=1}^{J} \alpha_j lnX_{jit} - u_{i0} - u_{it} + v_{it} \tag{3}$$

If inputs are exogenous, direct estimation of the production function parameters is possible by the maximum likelihood method using distributional assumptions on the inefficiency and

[1] This is referred to as the generalized production function and was introduced by Zellner and Revankar (1969).

noise components when there is a single output. However, for regulated industries such as the electricity distribution firms that we consider, output (outputs) is (are) treated as exogenous, and inputs are endogenous.[2] In this case, estimating equation 3 will result in inconsistent parameter estimates, even if there is a single output. If outputs are exogenous (not a choice variable), maximization of profit is the same as minimization of cost. Since the distribution industry is a service industry, and outputs then are exogenously determined, it is a standard practice to estimate the technology using the dual cost function. One can then use the duality results to derive the underlying features of the production function. For the production function in equation 1, one can derive the cost function analytically. That is, from the parameters of the cost function, we can derive the parameters of the production function, and vice versa.

From microeconomic theory, the firm is said to be allocatively efficient (no input misallocation) if it equates the marginal rate of technical substitution between each pair of inputs with the ratio of input prices. We therefore model input misallocation as

$$\frac{MP_{X_{jit}}}{MP_{X_{1it}}} = k_{ji}\left(\frac{w_{jit}}{w_{1it}}\right)e^{\eta_{jit}}\; j = 2, \cdots, J \tag{4}$$

where the factors of proportionality k_{ji} are firm and input specific, η_{jit} are random errors in cost minimization, $MP_{X_{jit}}$ are the marginal products of X_{jit}, and w_{jit} are the input prices. Apart from input misallocation that arises from a non-optimal mix of inputs, the equation underlines the fact that some inefficiency may also arise from uncontrolled random exogenous shocks; e.g., uncertainty in input and output prices, quality of inputs, etc. Solving X_{jit} from equations (3) and (4) yields the input demand functions (see Appendix A3) that can be used to derive the cost function, which is

[2] For example, the output of Norwegian electricity distribution companies is decided by their customers. This means that there is no point for the companies to increase their amount of distribution services, if the customers do not demand this increase.

$$lnC_{it} = -\frac{1}{r}\left(ln\ \alpha_0 + \sum_{j=1}^{J} \alpha_j\ ln\ \alpha_j\right) + \frac{1}{\gamma}\left(lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\ \beta_3 lnY_{1it} \cdot lnY_{2it}\right) +$$

$$\frac{1}{\gamma}\sum_{j=1}^{J} \alpha_j\ ln\ w_{jit} + ln\left(\alpha_1 + \sum_{j=2}^{J} \alpha_j\ e^{-\eta_{jit}}\right) + \frac{1}{\gamma}\sum_{j=2}^{J} \alpha_j\ \eta_{jit} + \frac{1}{\gamma}(u_{i0} + u_{it}) - \frac{1}{\gamma}(v_{it}) + E_{it} \quad (5)$$

where

$$E_{it} = ln\left[\alpha_1 + \sum_{j=2}^{J} \alpha_j / (k_{ji} e^{\eta_{jit}})\right] + \frac{1}{\gamma}\sum_{j=2}^{J} \alpha_j\ ln\ k_{ji} - ln\left[\alpha_1 + \sum_{j=2}^{J} (\alpha_j\ e^{-\eta_{jit}})\right] \quad (6)$$

From the expression of the cost frontier in (5), it is evident that any kind of inefficiency leads to an increase in cost. The cost of technical inefficiency is given by $\frac{1}{\gamma}(u_{i0} + u_{it})$, whereas the cost of input misallocation is given by (6). As the derivation of the cost function uses the FOCs of cost minimization, the resulting cost function depends on input misallocation, and persistent and transient technical inefficiency. In particular, the above cost function shows the amount of cost (normally) increase because of input misallocation, and persistent and transient technical inefficiency. Once the relevant parameters are estimated, the increase in cost because of input misallocation, and persistent and transient technical inefficiency can be computed for each firm.

2.2 Estimation method

Equation (3) can be rewritten as (Appendix A1)

$$lnX_{1it} \equiv \delta_0 + \sum_{j=2}^{J} \delta_j ln\left(\frac{X_{jit}}{X_{1it}}\right) + \mu_1 lnY_{1it} + \mu_2 lnY_{2it} + \mu_3 lnY_{1it} * lnY_{2it} + \epsilon_{it} \quad (7)$$

where $\epsilon_{it} = u_{io}^* + u_{it}^* + v_{it}^*$, $u_{i0}^* = \frac{1}{\gamma}u_{i0}, u_{it}^* = \frac{1}{\gamma}u_{it}, v_{it}^* = \frac{1}{\gamma}v_{it}$ and $\gamma = \frac{1}{\mu_1}$. From this specification, we can compute $\alpha_j = -\delta_j \cdot \gamma = -\frac{\delta_j}{\mu_1}$ for $j = 1, \dots, J, \alpha_1 = \gamma - \sum_{j=2}^{J} \alpha_j =$

$\frac{1}{\mu_1}\left(1 + \sum_{j=2}^{J} \delta_j\right)$, and $\gamma = \sum_{j=1}^{J} \alpha_j$, where γ defines the returns to scale (RTS). Similarly, the FOCs for cost minimization (equation 4) can be expressed as

$$ln\left(\frac{x_{1it} \cdot w_{1it}}{x_{jit} \cdot w_{jit}}\right) \equiv \gamma_j + \theta_{ji} + \eta_{jit} \tag{8}$$

where $\gamma_j = \ln\left(\frac{\alpha_1}{\alpha_j}\right)$ for $j = 1, \ldots, J$ (see Appendix A2). The parameter θ_{ji} can be directly calculated as $\hat{\theta}_{ji} = \overline{ln\delta_{ji}} - \hat{\gamma}_j$, where $\overline{ln\delta_{ji}}$ is the average across time of $ln\delta_{jit}$; i.e., $\frac{1}{T}\sum_{t=1}^{T} ln\delta_{jit}$, where $ln\delta_{jit} = ln\left(\frac{x_{1it} \cdot w_{1it}}{x_{jit} \cdot w_{jit}}\right)$. Furthermore, $\theta_{ji} = \ln k_{ji}$, which implies $k_{ji} = \exp\{\theta_{ji}\}$. This allows us to estimate $\eta_{jit} = \ln \delta_{jit} - \hat{\gamma}_j - \hat{\theta}_{ji}$. We can therefore substitute the parameter estimates of α_j, k_{ji} and η_{jit} into (6) to calculate the cost of input misallocation. The cost of technical inefficiency as defined earlier is the sum of persistent inefficiency (u_{i0}^*) and transient inefficiency (u_{it}^*). Note that in (8), we find that $\frac{x_{jit}}{x_{1it}}$ are functions of input price ratios and input misallocation. Thus, $\tilde{x}_{jit} = \frac{x_{jit}}{x_{1it}}$ in (7) is exogenous as long as input misallocation is independent of u_{it} and v_{it}, which is a common assumption. Therefore, we can estimate (7) without any endogeneity problems because our outputs are exogenous. However, estimation of (7) will not help us to estimate the cost of input misallocation.

In our study, we define the composite error term ε_{it} in (7) across four different model specifications. The numeraire input (X_1) in all the models is capital. Model 1 is the SFA model with a normal distribution of the noise term (v_{it}) and half-normal distribution of the efficiency term (u_{it}) (Aigner et al., 1977). Model 2 is the "true fixed-effects" (TFE) model with determinants for inefficiency, whereas Model 3 is the "true random-effects" (TRE) model with determinants for inefficiency (both introduced by Greene, 2006a, 2006b). Model 4 is a generalized TRE (GTRE) model with determinants for inefficiency (Badunenko and

Kumbhakar, 2017; Lai and Kumbhakar, 2018; Lien et al., 2018). The model specifications are summarized in Table 1.

Table 1: Econometric specification of stochastic frontier models

	Model 1	Model 2	Model 3	Model 4
ε_{it}	$\varepsilon_{it} = u_{it} + v_{it}$	$\varepsilon_{it} = b_i + u_{it} + v_{it}$	$\varepsilon_{it} = b_i + u_{it} + v_{it}$	$\varepsilon_{it} = b_i + u_{io} + u_{it} + v_{it}$
Firm effect		$b_i \sim fixed\ parameter$	$b_i \sim iidN(0, \sigma_b^2)$	$b_i \sim iidN(0, \sigma_b^2)$
Noise	$v_{it} \sim iidN(0, \sigma_v^2)$	$v_{it} \sim iidN(0, \sigma_v^2)$	$v_{it} \sim iidN(0, \sigma_v^2)$	$v_{it} \sim iidN(0, \sigma_v^2)$
Transient inefficiency	$u_{it} \sim iidN^+(0, \sigma_u^2)$	$u_{it} \sim iidN^+(0, \sigma_u^2(z_{it})) = N^+(0, exp(\theta_{u0} + \theta'_{uit} z_{it}))$	$u_{it} \sim iidN^+(0, \sigma_u^2(z_{it})) = N^+(0, exp(\theta_{u0} + \theta'_{uit} z_{it}))$	$u_{it} \sim iidN^+(0, \sigma_u^2(z_{it})) = N^+(0, exp(\theta_{u0} + \theta'_{uit} z_{it}))$
Persistent inefficiency				$u_{i0} \sim iidN^+(0, \sigma_u^2(z_i)) = N^+(0, exp(\vartheta_{u0} + \vartheta'_{ui} z_i))$
Main reference(s)	Aigner et al. (1977)	Greene (2006a, b)	Greene (2006a, b)	Badunenko and Kumbhakar (2017), Lai and Kumbhakar (2018), Lien et al. (2018)

Details on the estimation issues can be found in the relevant papers (main references) cited in Table 1. Model 1 corresponds to the specification in Kumbhakar (1988), and we include this as a benchmark in our study. This model does not account for the panel structure of the data. Models 2 and 3 are panel data models that separate firm heterogeneity from transient inefficiency. The former models firm heterogeneity as a fixed effect, whereas the latter models it as a random effect. Both models specify the time-invariant component (b_i) as firm heterogeneity, whereas, for example, in Kumbhakar and Heshmati (1995) this is considered as persistent inefficiency. Model 4 is a generalized version of the TRE (Model 3). It decomposes the error term into four components, where the first captures random firm heterogeneity; the second, persistent inefficiency; the third, transient inefficiency; and the final, random shocks.

Models 2 and 3 allow for determinants of transient inefficiency, whereas Model 4 in addition allows for determinants of persistent inefficiency.

Models 1 and 2 are estimated by standard one-step maximum likelihood techniques. Model 3 is estimated using a simulation-based one-step maximum likelihood estimator (Filippini and Greene, 2016). Estimation of Model 4 can be done in several ways. One is the single-step full maximum likelihood procedure, first proposed in Colombi et al. (2014), and extended in Badunenko and Kumbhakar (2017) and Lai and Kumbhakar (2018) to allow for heteroscedasticity in the error components. In this study, Model 4 is estimated using a multi-step procedure introduced by Lien et al. (2018), which is a modified version of Kumbhakar et al. (2014), where the inefficiency components are not necessarily independently and identically distributed (iid) but their variances are functions of exogenous determinants. All four models are estimated in Stata.

3 Data

The data used in this study are an unbalanced panel of 146 Norwegian electricity distribution firms for the years 2000 to 2016 compiled by the Norwegian Water Resources and Energy Directorate (NVE). The data contain economic and technical information on the firms. We used 2,143 firm-years.

We use three inputs and two outputs. The inputs are capital, labour and operational cost, and the outputs are the total number of customers and the size of the network, defined as the length of the high-voltage power lines in kilometres. We define capital as the aggregate book value of all assets owned by the firm, labour as number of man-years, and operational cost as total cost minus capital cost, labour cost and value of lost load. [3] We use the regulatory rate of return as the input price of capital. This rate is calculated annually by NVE using the weighted average cost of capital method (Amundsveen and Kvile, 2016). The price of labour is a year-specific variable calculated by the regulatory agency as the annual average cost per man-year (yearly average for the firms in the survey). The price of labour is measured in 2015 Norwegian kroner (NOK).[4] Finally, we use the consumer price index compiled by Statistics Norway as the price index for operational cost.

We include two environmental variables in the analysis, one as a determinant of transient inefficiency and the other as a determinant of persistent inefficiency. NVE uses several environmental variables in its regulation model, and these are intended to account for heterogeneity in the firms' production environments. In this study, we used the proportion of underground cables as a determinant of persistent inefficiency, and the value of lost load per kilometre network as a determinant of transient inefficiency. The value of lost load represents the estimated amount in NOK that customers receiving electricity through contracts would be willing to pay to avoid a disruption in the service, and this variable varies from year to year for each firm.[5] As the value of lost load also varies with firm size, we divided it by the firms' kilometre network to get a standardized measure.

Firms with fewer than two consecutive years of observations are dropped from the analysis because this is a requirement for panel data models that exploit the within variation in the data. Table 2 presents descriptive statistics of the inputs, input prices, outputs and environmental variables.

[3] We subtract labour cost from operational cost because this is included as an input variable in our analysis. Similarly, the value of lost load is subtracted from operational cost because we include this as a determinant of transient inefficiency.

[4] 1 USD = 8.14 NOK, 1 EUR = 9.50 NOK on June 25, 2018.

[5] The Value of lost load is calculated as lost load times a unit price, with different unit prices for various customer groups (Bjørndal et al., 2010).

Table 2: Descriptive statistics (N=2143, Year 2000–2016)

Variable	Name	Mean	SD	Min	Max
Input					
Capital (book-value, 1000 NOK)	X_1	265,290	570,248	1,605	6,873,112
Man-years (numbers)	X_2	31.1	48.9	0.6	396.3
Operational costs (1000 NOK)	X_3	24,957	65,173	168	901,652
Price of input					
Price of capital	W_1	0.072	0.016	0.042	0.100
Price of labour (per man-year)	W_2	601	94	439	720
Price of operational costs (CPI)	W_3	1.145	0.103	0.965	1.325
Output					
Number of customers	Y_1	20,176	54,751	4	696,540
Network (Km)	Y_2	727	1,252	5	11,866
Environmental (Z) transient variable					
Value of lost load per km of network (1000 NOK)	Z_1	2.70	2.92	0.000	31.32
Environmental (Z) persistent variable					
Proportion of underground cables	Z_2	0.318	0.204	0.000	1.000

4 Results

Table 3 presents estimates of the input elasticities. Although the estimates exhibit some variation depending on the model, we observe that across all models, the elasticity of capital is the largest of the three, ranging from 0.67 to 0.85. This is followed by the elasticity of labour (man-years) and finally the elasticity of operational costs, ranging from 0.56 to 0.64 and 0.43 to 0.53, respectively.

Table 3: Input elasticities

Input elasticities	Model 1	Model 2	Model 3	Model 4
Capital	0.847	0.672	0.676	0.788
Man-years	0.563	0.638	0.642	0.636
Operational costs	0.533	0.442	0.428	0.460

Table 4 presents the estimates of transient, persistent and overall technical efficiency across all four models mentioned in Table 1. The mean transient (and overall) technical efficiency for Model 1 is 0.76. Because this model does not account for unobserved firm heterogeneity, we expected it to overestimate the level of inefficiency relative to the other models, and this is what we observe. The mean transient technical efficiency for Model 3 is 0.93, a value slightly higher than that found by Kumbhakar and Lien (2017) in their study of Norwegian electricity distribution firms using the same model but over a shorter time frame (2000–2013). However, transient efficiency in Model 4 is 0.92, which is higher compared with Kumbhakar and Lien, who also use the same model but do not include determinants of inefficiency. However, persistent inefficiency is virtually the same in both studies. The mean transient efficiency in the TFE model (Model 2) is 0.85, which is within the range of values found by Kumbhakar et al. (2015), who used Norwegian data for the period 1998 to 2010 and used the same estimator. In this study, technical efficiency scores range from 0.82 to 0.87. Filippini et al. (2018), using data for electricity distribution firms in New Zealand for the years 2000 to 2011, found mean transient efficiency varying between 0.70 to 0.88 depending on the model, whereas persistent efficiency varied from 0.88 to 0.94.

Models 2 through 4 in Table 4 show that the value of lost load per kilometre of network increases the variance of transient inefficiency (illustrated by the positive θ coefficient), which implies increased inefficiency (or decreased efficiency). Similarly, for persistent inefficiency in Model 4, we observe that higher proportions of underground cables increase inefficiency.

Figure 1 graphically exhibits the marginal effects of value of lost load per kilometre of network on transient inefficiency, as well as the marginal effects of the proportion of underground cables on persistent inefficiency, showing a positive association that is increasing in both cases. The value of lost load represents a cost to the distribution companies resulting from a disruption in the service. High rates of disruption, more likely due to lack of maintenance, represent higher costs to the companies and thus will have a negative impact on the efficiency scores. It is therefore intuitive that the value of lost load per kilometre of network has a positive and increasing effect on the marginal effect on transient inefficiency.

There is no intuitive explanation of why the proportion of underground cables has a positive and increasing marginal effect on persistent inefficiency. One could think that the firms investing in newer network solutions also had higher efficiency. One reason for these results might be that the firms that have a high proportion of underground cables operate in an environment that makes this type of network necessary. If so, it is not the underground cables that increase persistent inefficiency but rather the environment. Kuosmanen (2012) reported that for Finnish electricity distribution firms, the proportion of underground cables has a highly significant positive effect on the total cost, which is consistent with our findings.

Table 4: Estimates of technical efficiency and determinants of technical inefficiency

Statistics	Model 1	Model 2	Model 3	Model 4
Transient technical efficiency				
Mean	0.757	0.849	0.933	0.921
SD	0.121	0.083	0.083	0.028
P5	0.561	0.686	0.760	0.872
P95	0.913	0.943	0.982	0.953
Persistent technical efficiency				
Mean				0.949
SD				0.064
P5				0.868
P95				0.981
Overall technical efficiency				
Mean	0.757	0.849	0.933	0.874
SD	0.121	0.083	0.083	0.065
P5	0.561	0.686	0.760	0.764
P95	0.913	0.943	0.982	0.928
Determinants of transient technical inefficiency				
Value of lost load per km network		0.073^{***}	0.091^{***}	0.126^{***}
Determinants of persistent technical inefficiency				
Proportion of underground cables				7.177^{*}

* Significant at $p > 0.10$, ** Significant at $p > 0.05$, *** Significant at $p > 0.01$.

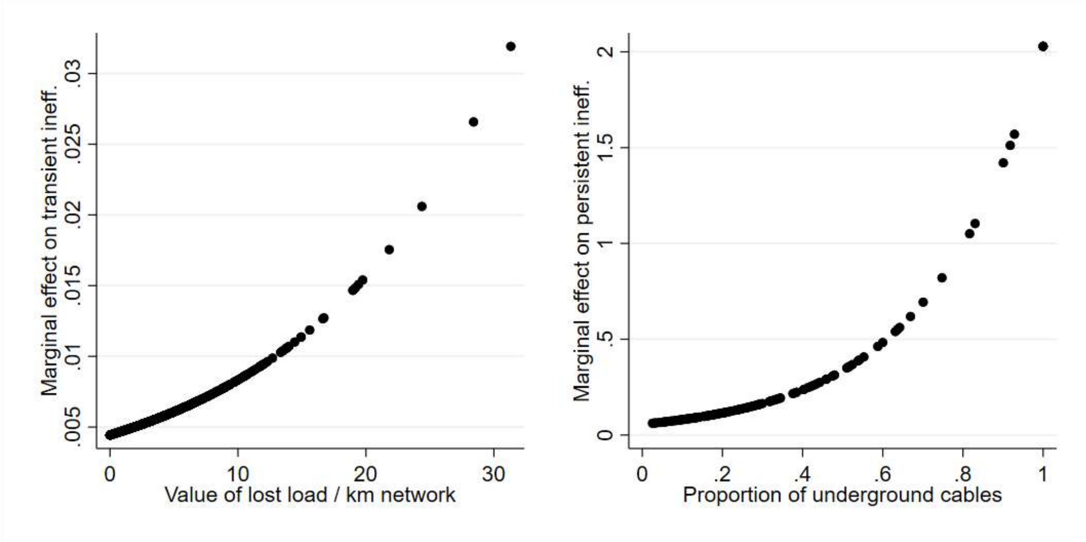


Figure 1: Scatterplot of marginal effects of lost load and proportion of underground cables on expected value of inefficiency, based on Model 4. The left plot shows marginal effects of value of lost load per km network on transient inefficiency. The right plot shows marginal effects of proportion of underground cables on persistent inefficiency.

The estimates of input misallocation, where capital is used as the numeraire, are shown in Table 5. K_2 and K_3 respectively indicate whether the capital/labour and capital/operational costs ratios deviate from unity (or the optimal proportion). Our findings show that the capital/labour ratio is on average less than one, indicating excessive use of labour relative to capital. The results show large variations both between and within the estimation models. For example, in Model 4, we observe that more than 15% of the firms have a K_2 value greater than 1, indicating over-capitalization for these firms. The capital/operational costs ratios, however, show excessive use of operational costs relative to capital. This ratio is less than unity for firms at the 95th percentile across all models, illustrating the robustness of the mean values.

Our finding that increased capital investments are required in the electricity distribution industry is not surprising. In an evaluation report commissioned by the Norwegian Ministry of Energy and Oil, Reiten et al. (2014) clearly stressed the need for such investment in the electricity distribution sector in the next few decades.

Table 5: Estimates of input misallocation of labour and operational costs

	Misallocation in labour (K_2)				Misallocation in operational costs (K_3)			
Statistics	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Mean	0.564	0.806	0.806	0.685	0.487	0.510	0.491	0.452
SD	0.408	0.584	0.584	0.496	0.210	0.219	0.211	0.195
P5	0.260	0.372	0.372	0.316	0.227	0.237	0.228	0.210
P95	1.650	2.360	2.360	2.010	0.827	0.865	0.833	0.767

As shown above, input misallocation increases cost; therefore, it is interesting to examine its effect on costs. Table 6 presents estimates of the overall cost of inefficiency, which is decomposed into technical and input misallocation costs. Apart from highlighting that there are substantial costs arising from technical inefficiency, the findings show that input misallocation also poses a significant challenge to the industry. Eliminating technical inefficiency can reduce costs by between 6.7% (Model 3) and 24.3% (Model 1), while for input misallocation, the corresponding cost reductions are between 9.0% (Model 2) and 11.3% (Model 1). From Model 4, we observe that the cost of input misallocation is larger than the cost of technical inefficiency for 28% of observations in the sample, implying that for more than a quarter of all firms, input misallocation is the main cost challenge. Nemeto and Goto (2006) find that for Japanese electricity transmission and distribution firms observed over the period 1981 to 1998, the cost of technical inefficiency increased overall cost by between 1% and 20%, whereas for input misallocation, the increase was between 8% and 30%.[6] As is evident from the lower section of Table 6, when input misallocation is taken into account, the overall cost of inefficiency is significant for the sample of firms, ranging from 16.1% (Model 3) to 35.6% (Model 1).

[6] Our findings that both technical inefficiency and input misallocation contribute to the overall cost inefficiency is consistent with a similar study (but not with the same estimation model) by Brissimis et al. (2010) of the European banking industry.

Table 6: Estimates of cost inefficiency

	Model 1	Model 2	Model 3	Model 4
Estimates of cost of technical inefficiency				
Mean	0.243	0.151	0.067	0.142
SD	0.121	0.083	0.083	0.102
P5	0.087	0.057	0.018	0.076
P95	0.439	0.314	0.240	0.276
Estimates of cost of input misallocation				
Mean	0.113	0.090	0.094	0.109
SD	0.085	0.096	0.099	0.096
P5	0.016	0.000	0.002	0.010
P95	0.276	0.292	0.303	0.307
Estimates of overall cost of inefficiency				
Mean	0.356	0.240	0.161	0.251
SD	0.144	0.124	0.149	0.150
P5	0.177	0.086	0.033	0.115
P95	0.565	0.473	0.454	0.484

5 Summary and concluding remarks

In this study, we have estimated persistent and transient technical inefficiency and input misallocation using a panel of Norwegian electricity distribution firms for the years 2000 to 2016. Our modelling and empirical strategy was to formulate and estimate a primal system consisting of the production function (generalized Cobb–Douglas) and the first-order conditions of cost minimization. We estimated the costs of technical inefficiency and input misallocation by deriving the cost function for a multiple-output separable production technology, extending the model in Kumbhakar (1988).

The results show that there exist non-negligible costs of input misallocation for Norwegian electricity distribution firms and call into question a commonly imposed modelling assumption under the SFA framework, that all firms are fully allocatively efficient. Even if we assumed that all firms are technically efficient, the costs to the industry arising from input misallocation would be too high, ranging on the average from 9.0% to 11.3% in our analysis. The robustness of these estimates across the different model specifications emphasize the importance of estimating input misallocation in general, also in studies of electricity distribution firms.

Beyond the modelling aspects, our results may also have important implications for regulators of electricity generation and distribution firms from across the world. The priority so far has been to identify the best method for estimating technical efficiency for benchmarking; see, e.g., Bogetoft and Otto (2011). The question that this study poses is as follows: given that the goal of regulation is cost minimization, is it not imperative that allocative efficiency also should be included in the benchmarking?

The results from the generalized true random effects model show evidence of persistent inefficiency. Filippini et al. (2018) argued that regulators may fail to set optimal efficiency targets if they are unable to identify systematic shortfalls in managerial capabilities that generate persistent inefficiency and to distinguish these from non-systematic management problems in the short run. For firms, however, investment decisions could be delayed and incentives for innovation weakened. Therefore, in line with Kumbhakar and Lien (2017), our findings further emphasize that future efficiency studies need to disentangle persistent and transient technical inefficiency.

Acknowledgements

The project is funded by Norwegian Water Resources and Energy Directorate, Energy Norway and six electric utilities in Norway (BKK Nett, Eidsiva Nett, Hafslund Nett, Helgeland Kraft, Lyse Elnett and Skagerak Nett).

References

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

Amundsveen, R., Kvile, H. M. (2016). Balancing incentives: The development and application of a regulatory benchmarking model. In: W. Greene, L. Khalaf, R. Sickles, M. Veall, M. C. Voia (eds.) *Productivity and Efficiency Analysis*. Springer Proceedings in Business and Economics. Springer, Cham.

Badunenko, O., Kumbhakar, S. C. (2017). Economies of scale, technical change and persistent and time-varying cost efficiency in Indian banking: Do ownership, regulation and heterogeneity matter? *European Journal of Operational Research* 260, 789–803.

Bjørndal, E., Bjørndal, M. and Fange, K. A. (2010). Benchmarking in Regulation of Electricity Networks in Norway – An Overview [Internet]. I: Energy, Natural Resources and Environmental Economics. Heidelberg: Springer, 317–342. DOI: 10.1007/978-3-642-12067-1_18.

Bogetoft, P., Otto, L. (2011). *Benchmarking with DEA, SFA, and R*, International Series in Operations Research & Management Science, 157, Springer.

Brissimis, S. N., Delis, M. D., Tsionas, E. G. (2010). Technical and allocative efficiency of European banking. *European Journal of Operational Research* 204, 153–163.

Colombi, R., Kumbhakar, S. C., Martini, G., Vittadini, G. (2014). Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis* 42, 123–136.

Filippini, M., Greene, W., Masiero, G. (2018). Persistent and transient productive inefficiency in a regulated industry: electricity distribution. *Energy Economics* 69, 325–334.

Greene, W. (1993). The Econometric Approach to Efficiency Analysis. In: Fried, H. O., Schmidt, S. S., Lovell, C. A. K (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications* (pp. 68–119), Oxford University Press.

Greene, W. (2006a). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis* 23, 7–32.

Greene, W. (2006b). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* 126, 269–303.

Kumbhakar, S. C. (1988). On the estimation of technical and input misallocation using stochastic frontier functions: The case of US class 1 railroads. *International Economic Review* 29(4), 727–743.

Kumbhakar, S. C., Heshmati, A. (1995). Efficiency measurement in Swedish dairy farms: An application of rotating panel data, 1976–88. *American Journal of Agricultural Economics* 77, 660–674.

Kumbhakar, S. C., Wang, H. (2006a). Pitfalls in the estimation of cost function ignoring allocative inefficiency: A Monte Carlo analysis. *Journal of Econometrics* 134, 317–340.

Kumbhakar, S. C., Wang, H. (2006b). Estimation of technical and allocative inefficiency: A primal system approach. *Journal of Econometrics* 134, 419–440.

Kumbhakar, S. C., Wang, H., Horncastle, A. P. (2015). *A Practitioner's Guide to Stochastic Frontier Analysis using Stata*. Cambridge University Press.

Kumbhakar, S. C., Lien, G. (2017). Yardstick regulation of electricity distribution – Disentangling short-run and long-run inefficiency. *Energy Journal* 38(5), 17–37.

Kumbhakar, S. C., Amundsveen, R., Kvile, H. M., Lien, G. (2015), Scale economies, technical change and efficiency in Norwegian electricity distribution 1998–2010. *Journal of Productivity Analysis* 43(3), 295–305.

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.

Lai, H. P., Kumbhakar, S. C. (2018). Panel data stochastic frontier model with determinants of persistent and transient inefficiency. *European Journal of Operational Research* 271(2), 746-755.

Lien, G, Kumbhakar, S. C., Alem, H. (2018). Endogeneity, heterogeneity, and determinants of inefficiency in Norwegian crop-producing farms. *International Journal of Production Economics* 201, 53–61.

Mydland, Ø., Haugom, E., Lien, G. (2018). Economies of scale in Norwegian electricity distribution: A quantile regression approach. *Applied Economics* (forthcoming).

Nemeto, J., Goto, M. (2006). Measurement of technical and allocative efficiencies using CES cost frontier: a benchmarking study of Japanese transmission-distribution electricity. *Empirical Economics* 31, 31–48.

Reiten, E., Sørgard, L., Bjella, K. (2014). Et bedre organisert strømnett. In: Report Ordered by the Ministry of Energy and Oil, The Ministry of Energy and Oil, Oslo.

Schmidt, P., Lovell, C. A. K. (1979). Estimating technical and input misallocation relative to stochastic production and cost frontiers. *Journal of Econometrics* 9(3), 343–366.

Zellner, A., Revankar, N. S. (1969). Generalized production functions. *Review of Economic Studies* 36, 241–250.

Appendix

A1. Derivation of Equation 3

By taking the log of equation 1, we have

$$\ln F(Y_{it}) = ln\alpha_0 + \sum_{j=1}^{J} \alpha_j lnX_{jit} - u_{i0} - u_{it} + v_{it} \tag{A11}$$

Inserting equation 2 into the left-hand side of equation A11 yields

$$lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\beta_3 lnY_{1it} \cdot lnY_{2it} = ln\alpha_0 + \sum_{j=1}^{J} \alpha_j lnX_{jit} - u_{i0} - u_{it} + v_{it}$$

which can be rewritten as

$$ln\alpha_0 + \alpha_1 lnX_{1it} = -\sum_{j=2}^{J} \alpha_j lnX_{jit} + lnY_{1it} + \beta_2 lnY_{2it}$$

$$+\frac{1}{2}\beta_3 lnY_{1it} \cdot lnY_{2it} + u_{i0} + u_{it} - v_{it} \tag{A12}$$

By adding and subtracting $\sum_{j=2}^{J} \alpha_j lnX_{1it}$ to the right-hand side of Equation A12, we can restate the equation as

$$ln\alpha_0 + \alpha_1 lnX_{1it} = -\sum_{j=2}^{J} \alpha_j lnX_{jit} + lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\beta_3 lnY_{1it} \cdot lnY_{2it}$$

$$+u_{i0} + u_{it} - v_{it} + \left(\sum_{j=2}^{J} \alpha_j lnX_{1it} - \sum_{j=2}^{J} \alpha_j lnX_{1it}\right).$$

Taking $-\sum_{j=2}^{J} \alpha_j lnX_{1it}$ to the left-hand side and simplifying, noting that $lnx - \ln y = \ln(x/y)$, we have

$$ln\alpha_0 + lnX_{1it}\left(\sum_{j=1}^{J} \alpha_j\right) = -\sum_{j=2}^{J} \alpha_j \left(\frac{lnX_{jit}}{lnX_{1it}}\right) + lnY_{1it}$$

$$+\beta_2 lnY_{2it} + \frac{1}{2}\beta_3 lnY_{1it} \cdot lnY_{2it} + u_{i0} + u_{it} - v_{it} \tag{A13}$$

Finally, expressing equation A13 in terms of lnX_{1it} yields equation 3.

A2. Derivation of Equation 8

The marginal product of input X_{1it} is given by

$$MP_{X_{1it}} = \frac{\partial F(Y_{it})}{\partial X_{1it}} = \frac{\partial\left[\alpha_0\left(\prod_{j=1}^{J} X_{jit}^{\alpha_j}\right)e^{-u_{i0}-u_{it+v_{it}}}\right]}{\partial X_{1it}}$$

$$= \alpha_0\alpha_1 \frac{X_{1it}^{\alpha_1}}{X_{1it}}\left(\prod_{j=2}^{J} X_{jit}^{\alpha_j}\right) e^{-u_{i0}-u_{it+v_{it}}} \tag{A22}$$

The marginal product of X_{jit} is similarly defined by changing the subscripts 1 in the expression in equation A21 to j and amending the product within the parenthesis to include all terms except the j_{th}. Inserting these marginal products into equation 4 and cancelling out the common terms yields

$$\frac{MP_{X_{jit}}}{MP_{X_{1it}}} = \frac{\alpha_j X_{1it}}{\alpha_1 X_{jit}} = k_{ji}\left(\frac{w_{jit}}{w_{1it}}\right) e^{\eta_{jit}}\ j = 2, \cdots, J$$

which can be rearranged to

$$\frac{X_{1it}w_{1it}}{X_{jit}w_{jit}} = k_{ji}\left(\frac{\alpha_1}{\alpha_j}\right) e^{\eta_{jit}} \tag{A22}$$

Taking the log of equation A22 results in the expression in equation 8.

A3. Derivation of the input demand functions

Equation A22 equates the marginal rate of technical substitution between inputs X_{jit} and X_{1it} to the slope of the isocost line, which is the ratio of the input prices; that is (w_{jit}/w_{1it}). We can solve for X_{1it} in the equation to obtain the expansion path, which represents all combinations of inputs that are cost minimizing.

$$X_{1it} = k_{ji}\left(\frac{\alpha_1 X_{jit} w_{jit}}{\alpha_j w_{1it}}\right) e^{\eta_{jit}} \tag{A31}$$

Substituting equation A31 into the production function in equation 3 (i.e., substituting the expression for X_{1it}) and solving for X_{jit} yields the following log efficient input demand functions

$$lnX_{jit} = \; a_j + \sum_{j=2}^{J}\left(\frac{\alpha_j}{\gamma} - \delta_{ij}\right) lnk_{jit}$$

$$+\frac{1}{\gamma}\left(lnY_{1it} + \beta_2 lnY_{2it} + \frac{1}{2}\,\beta_3 lnY_{1it} \cdot lnY_{2it}\right)$$

$$+\sum_{j=1}^{J}\left(\frac{\alpha_j}{\gamma}\right)\, ln\left(\frac{w_{jit}}{w_{1it}}\right)$$

$$+\sum_{j=2}^{J}(\alpha_j/\gamma - \delta_{ij})\eta_{jit} + \frac{1}{\gamma}(u_{i0} + u_{it} - v_{it}) \tag{A32}$$

where

$$a_j = ln\alpha_j - \frac{1}{\gamma}\left(ln\alpha_0 + \sum_{j=1}^{J}\alpha_j\, ln\alpha_j\right), \gamma = \sum_{j=1}^{J}\alpha_j \text{ , and}$$

$$\delta_{ij} = \begin{cases} 1 & if\ j = i \\ 0 & otherwise, \end{cases} \qquad j = 1,2,\cdots,J.$$