Discussion Papers

Astrid Cullmann

Benchmarking and Firm Heterogeneity in Electricity Distribution: A Latent Class Analysis of Germany

Berlin, April 2009
Opinions expressed in this paper are those of the author and do not necessarily reflect views of the institute.
Benchmarking and Firm Heterogeneity in Electricity Distribution:
A Latent Class Analysis of Germany*

Astrid Cullmann

DIW Berlin
Department of Industry, Innovation and Service
Mohrenstrasse 58
10117 Berlin

acullmann@diw.de
 tel.: +49-30-89789-679
fax.: +49-30-89789-103

*I am indebted to Christian von Hirschhausen, Matthias Walter, Borge Hess and the DIW IO seminar for their helpful comments and support.
Abstract

In January 2009 Germany introduced incentive regulation for the electricity distribution sector based on results obtained from econometric and nonparametric benchmarking analysis. One main problem for the regulator in assigning the relative efficiency scores are unobserved firm-specific factors such as network and technological differences. Comparing the efficiency of different firms usually assumes that they operate under the same production technology, thus unobserved factors might be inappropriately understood as inefficiency. To avoid this type of misspecification in regulatory practice estimation is carried out in two stages: in a first stage observations are classified into two categories according to the size of the network operators. Then separate analyses are conducted for each sub-group. This paper shows how to disentangle the heterogeneity from inefficiency in one step, using a latent class model for stochastic frontiers. As the classification is not based on a priori sample separation criteria it delivers more robust, statistical significant and testable results. Against this background we analyze the level of technical efficiency of a sample of 200 regional and local German electricity distribution companies for a balanced panel data set (2001-2005). Testing the hypothesis if larger distributors operate under a different technology than smaller ones we assess if a single step latent class model provides new insights to the use of benchmarking approaches within the incentive regulation schemes.

Keywords: stochastic frontiers, latent class model, electricity distribution, incentive regulation

JEL classification: C24, C81, D24, L94
1 Introduction

Electricity distribution and transmission are traditional network industries and are therefore characterized by the typical properties of natural monopolies. In contrast to the generating and supply segments in the electricity market where competition has been introduced, the distribution sector is highly regulated. The last decades of utility regulation were characterized by traditional cost of service regulation schemes by which companies recovered their costs with a risk-free fixed rate of return (Farsi et al., 2007; Joskow, 2006). Hence, distribution companies had little or even no incentive for cost minimization. Therefore, incentive regulation schemes have become increasingly important in Europe with the goals of reducing costs and increasing efficiency. Across European countries price or revenue cap regulation is extensively used in electricity distribution. Within this framework price/revenue caps are set based on the formula $RPI - X$ (see Beesley and Littlechild, 1989). The maximum rate of price increase equals the inflation rate of the retail price index ($RPI$) less the expected efficiency savings ($X$). Thus, regulated distribution companies have incentives to increase their profits by improving their productivity at a higher rate than the assigned X-factor (Farsi et al., 2007). In 2009 Germany has begun to implement the framework in the electricity and natural gas distribution sector.

The determination of the X-factors for setting price/revenue caps is usually based on empirical results obtained from benchmarking analysis. The efficiency performance of the companies is therefore evaluated against a reference, best practice, performance. This framework is favored by European regulators concerned about the robustness and reliability of the empirical outcome of individual efficiency estimation.\(^1\) The development and advancement of benchmarking models for a consistent practical application has therefore been an important research aspect in efficiency analysis. The empirical literature (see Jamasb and Pollitt, 2001; Farsi et al., 2007) can be divided into two major groupings: nonparametric and parametric methods. While the nonparametric methods construct the reference technology, the efficiency frontier, by means of linear programming methods, parametric approaches assume a functional form for the underlying production process.\(^2\)

Empirical evidence shows that the individual efficiency estimates are sensitive to the adopted benchmarking models and approaches (Farsi and Filippini, 2004). Thus, the choice of the empirical approach has a strong impact on the price/revenue cap setting.

---

\(^1\)Shuttleworth (2005) gives a critical summary of the implementation of benchmarking analysis in price cap regulation.

\(^2\)Our empirical analysis is solely based on the parametric approach as it is able to better account for unobserved heterogeneity within the panel data framework.
(and therefore the economic and financial conditions of the companies). Further, implementation is always affected by the challenge of data availability. Some firm-specific factors are either unobservable, too complex to be accounted for appropriately, or no data is available for the empirical analysis. However, these unobserved characteristics between firms may have an important impact on the underlying production process and therefore the reference or efficiency frontier.\footnote{Alternative econometric approaches have been proposed in the literature to improve benchmarking methodology. These new strategies are mainly based on panel data when companies have been observed over time, attempting to isolate the unobserved firm-specific factors that are not under managerial control from real inefficiencies, see Farsi et al. (2006).}

In the traditional framework it is assumed that firms operate under the same production technology. Therefore these unobserved factors might be inappropriately understood as inefficiency. To avoid these types of misspecification estimation is often carried out in two stages in regulatory practice. First, defining different categories based on a priori sampling criteria and then separate analyses for each sub-group. In Germany, under the assumption that the size of the network operators implies different technological and network characteristics, observations are classified ex-ante into two categories: The German Incentive Regulation (AREgV (§24)) determines for the German electricity distributors that network operators with fewer than 30,000 customers connected directly or indirectly to their distribution system can choose to take part in a simplified procedure. Thus, an ex-ante sampling splitting between large and small distributors is present in the German regulation. This shows the implicit assumption that omitted unobserved technological differences between larger and smaller operators might be present and if not taken into account, inappropriately labeled as inefficiency.\footnote{Other general empirical studies apply the two-stage approaches to decrease the probability of misspecification, see e.g. Grifell and Lovell (1997) for efficiency measurement of banks.}

Following Orea and Kumbhakar (2004) and Greene (2005b) we suggest to use a single-stage approach to account for heterogeneity within the stochastic frontier analysis. The single-stage approach is based on a latent class framework for stochastic frontiers which has not been implemented so far for performance measurement of the German electricity distribution sector. Here, in contrast to traditional models, different technologies and efficiency frontiers for different classes or groups of firms can be identified without the need for a priori sampling separation information (Greene, 2005a). This is motivated by the fact that a priori selection might be arbitrary, lacks statistical foundation and is therefore not testable. We apply the latent class model for stochastic frontiers using a multi-input multi-output parametric input distance function for a balanced panel data set (2001-2005) for 200 regional and local German distributors. We first assess if our model confirms that large distributors operate under a different technology than the smaller network operators. This implies that unobserved factors
are present and have to be taken into account within the nationwide benchmarking. Secondly, we analyze the level of technical efficiency of our sample of network operators. We derive that a single step latent class model is able to provide new insights to the future use of benchmarking approaches within the incentive regulation schemes and could serve as additional control instrument for the regulator.

The paper is structured as follows: Section 2 presents the econometric specification and Section 3 the data. Section 4 summarizes the main empirical results of the distance functions estimation under different econometric assumption. Section 5 concludes.

## 2 Model specification

### 2.1 Distance function approach

Within a technical production setting, the majority of applied parametric efficiency analyses uses the production function to describe the underlying technology of different firms. Single output Cobb-Douglas or translog functional forms are most widely assumed but become critical, when firms produce more than one output (Coelli, 2000). Applied work has managed the issue by either aggregating the different outputs into a single index, or capturing multi-output production via estimating a multi-output cost frontier function.\(^5\)

Another approach to model multi-output production is the concept of parametric distance functions (Coelli, 2000). This approach has been proposed by Shephard (1970) who derives a distance function representation of a multi-output technology as a primal alternative that requires no aggregation, price and cost information and behavioral assumption; see Coelli (2000) for a detailed description on the econometric estimation of the distance function representation.\(^6\) We apply a parametric frontier input distance function to model the customers' supply and the physical amount of electricity delivered to final customers. The input distance function is defined on the input set as

\[
d_i(x, y) = \max \{ \rho : (x/\rho) \in L(y) \}
\]

(1)

and considers how much the input vector \(x\) may be proportionally contracted by the scalar distance \(\rho\) with the output vector held fixed (Coelli, 2000).\(^7\) \(d_i(x, y)\) will assume


\(^6\)For a discussion on advantages and disadvantages of the use of distance functions see Coelli (2000), and Coelli and Perelman (2000).

\(^7\)It is assumed that the technology satisfies the standard axioms: \(d_i(x, y)\) is non-decreasing, pos-
a value greater than or equal to one if the input vector \( x \) is an element of the feasible input set \( L(y) \). In addition, \( d_i(x, y) = 1 \) if it is located on the inner boundary of the input set.\(^8\)

The translog functional form is widely used for the distance function approximation in empirical application due to its flexibility for econometric estimation. However, we employ the stricter Cobb Douglas functional form excluding the squared and cross terms of the exogenous regressors. A constant elasticity of substitution and constant scale properties are therefore assumed because we want to capture the parameter heterogeneity to define different technologies. The econometric models for SFA including parameter heterogeneity are already characterized by a sophisticated stochastic component. Moreover, parameter heterogeneity concerning the squared and cross terms of the regressors does not have a real economic interpretation. For the case of \( M \) outputs and \( K \) inputs the Cobb Douglas input distance function is specified for the \( i \)-th firm as

\[
\ln d_i = \alpha_0 + \sum_{m=1}^{M} \gamma_m \ln y_m + \sum_{k=1}^{K} \beta_k \ln x_k. \tag{2}
\]

To obtain the frontier surface (the transformation function) one would set \( d_i = 1 \), so the left side equals zero (Coelli and Perelman, 2000). The restriction for homogeneity of degree +1 in inputs is

\[
\sum_{k=1}^{K} \beta_k = 1. \tag{3}
\]

A convenient approach of imposing homogeneity constraints follows Coelli and Perelman (2000) considering that homogeneity implies that for any \( w > 0 \)

\[
d_i(wx, y) = wd_i(x, y). \tag{4}
\]

Therefore, one of the inputs may be arbitrarily chosen, such as the \( K \)-th input and set \( w = 1/x_K \). This yields

\[
d_i(x/x_K, y) = d_i(x, y)/x_K \tag{5}
\]

\(^8\)Applications of this concept to estimate parametric distance functions using econometric methods can be found in Coelli and Perelman (2000) for railways, Färe et al. (1993) for electricity generation, Saal et al. (2007) for water and sewerage industry, and Growitsch et al. (forthcoming) for European electricity distribution.
and the Cobb Douglas input distance function becomes

$$\ln x_K = \alpha_0 + \sum_{m=1}^{M} \gamma_m \ln y_m + \sum_{k=1}^{K} \beta_k \ln \left(\frac{x_k}{x_K} \right) - \ln d_i$$

(6)

by dividing equation (2) by an optional input and some rearranging; $\ln d_i$ is a non-negative variable which can be associated with technical inefficiency $u_i$. Given the stochastic error $v_i$ this model can be formulated in the common SFA form with the combined error term $v_i - u_i$ (see Section 2.2). Technical efficiency is the ratio of observed output to frontier output. A radial input-oriented measure of technical efficiency is then obtained by

$$TE = \frac{1}{d_i} = exp(-u_i).$$

(7)

The distance function provides a promising new solution to the single output restriction of the standard production functions. One concern in the econometric estimation is potential regressor endogeneity which may introduce possible simultaneous equation bias. Some authors have proposed instrumental variables estimation (see Atkinson and Primont, 2002). However, Coelli (2000) found that under an assumption of cost minimization behavior, distance functions do not face such bias and that OLS provides consistent estimates of the parameters of an input distance function. A second issue is that estimated input distance functions often fail to satisfy the concavity and quasi-concavity properties implied by economic theory. This sometimes leads to surprising conclusions regarding the effects of input and output changes on productivity growth and relative efficiency levels. Therefore, the starting point before any interpretation of inefficiencies is to check and to test for the properties.

For the interpretation of the empirical estimates of a distance function it is important to keep in mind the duality between the cost and the input distance functions (Färe and Primont, 1995). For instance, the derivative of an input distance function with respect to a particular input is equal to the cost share of that input. This implies that the expected sign of the coefficients of the inputs should be positive. Moreover, the elasticity of an input distance function with respect to any output is equal to the negative value of the cost elasticity of that output. This implies that the sign of the coefficients of the outputs should be negative. Given that all the variables are in logarithmic form, these coefficients can be directly interpreted as elasticities.

9Ratios on inputs appear on the right side of the estimating equation which may involve simultaneous feedback problems because these input variables are assumed to be endogenous.

10Regularity conditions could also be imposed by estimating the model in a Bayesian framework (O’Donnell and Coelli, 2005).
2.2 Stochastic frontiers and latent classes

2.2.1 Panel data approaches for stochastic frontiers

SFA belongs to the parametric benchmarking methods. Contrary to the nonparametric approaches, a functional relationship for the underlying production technology is assumed. Within the benchmarking process we compare some measure of actual performance against a reference technology (the stochastic frontier). The distance to the production frontier can be interpreted as a common measure of technical inefficiency. In the SFA framework the error term is divided into two uncorrelated components: a one-sided non-negative disturbance \( u_i \), half-normally distributed, representing the inefficiency; and a symmetric disturbance \( v_i \), assumed to be normally distributed, and capturing random noise in the sample (Greene, 2007). The most general cross-sectional formulation is

\[
\begin{align*}
y_i &= \beta' x_i + v_i - u_i \\
u_i &\sim N^+[0, \sigma^2_u] \\
v_i &\sim N[0, \sigma^2_v]
\end{align*}
\]

(8)

where \( x_i \) represents the set of explanatory variables and \( y_i \) the observed production of a firm. This model can be estimated using the maximum likelihood approaches.

A number of different stochastic frontier models for panel data have been proposed. The first models define the random or fixed effect as the inefficiency component meaning that the models deduce the efficiency estimates from the individual firm-specific effects.\(^{11}\) These traditional models assume a common technology/frontier encompassing every sample observation. This may be inappropriate in the sense that the estimated technology is not likely to represent the “true” technology for all observations (Farsi et al., 2006). Thus, the estimate of the underlying technology may be biased. In addition, as unobserved heterogeneity was not accounted for in the econometric models, parameter estimates also may have been biased. Moreover, since all time-invariant heterogeneity was covered by the inefficiency part, these models have a tendency to underestimate firms’ performance (Farsi et al., 2007).

European regulators, implementing the efficiency estimates into regulatory practice, are concerned about the robustness and the reliability of the empirical outcome of the individual efficiency estimation. Robust and consistent specification and models are indispensable for a trustable and effective regulation as the choice of the empirical approach has a strong impact on the financial situation of the network operators. Thus

\(^{11}\)See Pitt and Lee (1981) for the random effects SFA model and Schmidt and Sickles (1984) for the fixed effects SFA model.
over the last decades we observed a research necessity concerning mainly two aspects for parametric stochastic efficiency analysis: first, how unobserved firm-specific factors may influence the underlying production process, and second how to model different technologies for different groups of firms.

With regard to the first aspect a wide range of newer models attempts to separate unobserved heterogeneity from inefficiency. One can model heterogeneity in the stochastic part, in the mean or the variance of the inefficiency distribution $u_i$. However, it became more important to model both heterogeneity in the stochastic part and firm-specific heterogeneity in the production or cost function of the underlying production process. Kumbhakar (1991) and Greene (2005b) have suggested extending the original stochastic frontier model by adding an individual time-invariant random or fixed effect. The basic underlying assumption is the existence of firm-specific and time-invariant factors that cannot be captured by environmental variables due to the variation of the latter over time and/or omitted variables. With the additional inclusion of heterogeneity terms by means of the random firm-specific effect $\alpha_i$ the model is expected to provide a finer distinction between inefficiency and other unexplained factors (Greene, 2005b).

With regard to the second aspect other formulations of the stochastic frontier model were proposed in the literature that allow not only the constant but the entire function to vary more generally across firms: the random parameter as well as latent class models for stochastic frontiers. Instead of the continuous parameter variation, the latent class formulation can be interpreted as an approximation where the parameter variation is treated as generated by a discrete distribution (Greene, 2007). With these newer models unobserved differences in technologies may be accounted for which previously were inappropriately labeled as inefficiency. The unobserved firm-specific heterogeneity could therefore be applied to marginal products and costs represented by the coefficients of the production cost or distance function.

2.2.2 Latent class specification

The latent class framework for SFA accounts for specific technological characteristics of the observations in the sample; firms are classified into a set of different technologies and efficiency distributions. However the specific classification is à priori unknown...
(Greene, 2007; Orea and Kumbhakar, 2004), unlike the two-step approach that classifies the sample observations à priori into categories using exogenous sample separation information (Orea and Kumbhakar, 2004). In this model all parameters vary by class standing for the different technologies of the different classes. Others have modeled technology heterogeneity within SFA via the latent class formulation (Orea and Kumbhakar, 2004; Greene, 2005b; Caudill, 2003; Corral and Alvarez, 2008). Figure 1 shows a sample of firms operating under different technologies (Technology A and Technology B). Assuming in this framework a common technology, and therefore frontier, for the companies would result in biased estimates of the distance function and efficiency. The latent class model, in general, is a stochastic frontier model of the following form

\[
\ln(y_{it}|j) = f(x_{it}, \beta_j) + v_{it}|j - u_{it}|j
\]  

where \(j\) indicates the class or regime and \(J\) the total number of classes or regimes. Class membership is unknown. One assumes that there is a latent sorting of the observations in the data resulting in \(J\) classes (Greene, 2007). For one specific observation from class \(j\) the model is characterized by the conditional density \(g(.)\) determined by the class specific parameter vector \(\beta_j\).

\[
g(y_{it}|x_{it}, \text{class}_j) = f(\beta_j, y_{it}, x_{it})
\]

The contribution of the company $i$ to the conditional likelihood (conditional on class $j$) is

$$P(i|j) = \prod_{t=1}^{T} P(i, t|j)$$

(11)

The unconditional likelihood for individual $i$ is an average over the $J$ classes. It can be shown that the likelihood function can be expressed by (see Greene, 2005)

$$\log LF = \sum_{i=1}^{N} \log(\sum_{j=1}^{J} P_{ij} \prod_{t=1}^{T} LF_{ijt})$$

(12)

The class probabilities can be parameterized by a multinomial logit model:

$$P_{ij} = \frac{\exp(\delta_j' q_i)}{\sum_{j=1}^{J} \exp(\delta_j' q_i)}$$

(13)

where $q_i$ is a vector of firm-specific but time-invariant variables. These variables, called separating or switching variables, are included to identify any regularity in classifying the sample by means of the estimated coefficients of latent class probability functions $\delta_j$ (Greene, 2007). A positive sign of the coefficient suggests that the larger the variable the higher the probability that a firm belongs to this class. Similarly, the significantly negative value of a coefficient indicates that the probability of membership in this class decreases when the variable increases.

Under the maintained assumptions, maximum likelihood techniques will give asymptotically efficient estimates of all parameters. Greene (2002) points out that the technology as well as the probability to belong to a certain class are estimated simultaneously. All observations in the sample are used to estimate the underlying technology for each class. This can be viewed in opposition to the standard two-step approaches, where observations that are allocated to a specific class equal one, and zero for the others, are therefore excluded to estimate other class frontiers (Orea and Kumbhakar, 2004). The estimated parameters can be used to compute the conditional posterior class probabilities. In addition, Greene (2007) suggests that the class probabilities apply unchanged to all years of the observation period.

In standard SFA, the individual efficiency is estimated to the common frontier, since all firms are assumed to operate under the same technology. The latent class specification estimates as many frontiers as the number of classes (see Figure 1 with two different classes). There is no unique technology against which inefficiency is computed. There are different methods to measure the efficiency level of an individual firm (see Orea and Kumbhakar, 2004, for a summary): first, the highest posterior probability for
class membership can be taken, and firms’ inefficiency is computed using the frontier assigned for that class as its reference technology (most likely frontier is used); second, technologies from every class are taken into account, weighted with the respective probabilities.\textsuperscript{16}

An ongoing discussion in the literature concerns the determination of the number of classes (Greene, 2007). For estimation we assume that the number of classes is known, but as Greene (2007) has shown, there is no reason to expect this. Using the likelihood ratio test from a \( J \) class model to a \( J - 1 \) class model would lead to an ambiguous number of degrees of freedom. When we test up from \( J - 1 \) to a \( J \) class model and the correct model has \( J \) classes, the \( J - 1 \) class model is inconsistent. The empirical solution in the literature is to apply information criteria such as the Akaike Information criteria (Greene, 2007). Our empirical model defines two latent classes (see Section 4.1).

3 Data

Two sets of variables are required to estimate the latent class model. First, the variables in the production frontier model have to be defined including appropriate inputs and outputs for the process of electricity distribution. Second, the variables in the class probabilities have to be chosen to determine if observable information helps to classify the companies into different classes with different underlying technologies. The variables to describe the underlying production process are defined in the same manner for the different groups of companies. In the empirical benchmarking literature, a variety of specifications is used depending on what is being investigated. The choice of variables for input and output to describe the underlying production process and technology must account for the international experience with electricity distribution benchmarking (Cullmann and Hirschhausen, 2008a,b). Further, it is constraint by data availability. In this respect, Germany has to be ranked among the least developed for data collection. We therefore depend upon a limited data set of physical and technical data. We must define a simple production process describing the basic input transformation of the German companies. As a result we limit ourselves to the estimation of technical inefficiency which does not require any data on costs or prices; however as shown in Section 2, conclusions regarding cost efficiency can be drawn due to the distance function specification.

\textsuperscript{16}This is the strategy suggested by Greene (2002) to find firm-specific estimates of the parameters of the stochastic frontier model. The magnitude of the difference depends on the relative importance of the posterior probability of the most likely cost frontier; the higher the posterior probability, the smaller the differences.
The sample includes 200 companies and covers a five-year observation period from 2001 to 2005. We estimate a base model using the traditional input variables (labor and grid size) and the outputs are units sold and the number of customers:

* Labor input ($x_L$) is estimated by the number of workers. We are aware of the criticism of this choice of variable due to the potentially distorting effect of outsourcing: a utility can improve its efficiency simply by switching from in-house production to outsourcing. Some of the utilities have their own generating plants and we only dispose of employment data covering all workers in the electricity utility. To get an approximation of workers employed in electricity distribution, we subtract one employee for each 20 GWh electricity produced (following Auer, 2002).

* Capital input ($x_{NL}$) is approximated by the length of the existing electricity cables and lines. We differentiate between voltage levels (high, medium and low voltage) by introducing a cost factor for each type of line following standard practice used by the German network association (VDN): factor 5 for high voltage, 1.6 for medium voltage, and 1 for low voltage cables.

* Delivery ($y_D$) is defined as the annual amount of electricity sold to all final customers (household, industrial and others) in MWh.

* Customers ($y_C$) is defined as the sum of industrial, households and other customers.

* Year dummies ($d_1 – d_4$) 2001 to 2004.

We also include variables as determinants of the latent class probabilities in order to analyze whether they deliver useful information in classifying the sample:

* Delivery to other distribution companies (dummy)

* Electricity generating activities (dummy)

* Operating high voltage cables (dummy)

* Operating high voltage aerial lines (dummy)

* Annual investment in 1000 Euro

* Annual revenue out of domestic sales in 1000 Euro

* Delivery to households/total delivery in MWh
* Investment per km network in 1000 Euro/km

* Revenue per unit delivery in 1000 Euro/MWh

* Share of cables in total network in km

* Losses in MWh

* Density of inhabitants per km operation area.

For the \( q \) variables we use firm average values over the five-year observation period.

To summarize, we designed a model to describe the production process, including two input variables, two output variables and year dummies to capture the time dimension. We define labor \((x_L)\) as the numeraire input. By dividing the remaining input over the labor input and rearranging we define the following Cobb-Douglas input frontier distance function (see Section 2)

\[
-\ln x_{it,L} = \alpha_0 + \alpha_{NW} \ln \left( \frac{x_{it,NW}}{x_{it,L}} \right) + \gamma_{C} \ln(y_{it,C}) + \gamma_{D} \ln(y_{it,D}) + d_1 + d_2 + d_3 + d_4 - \ln d_{it}
\]  

with

\[
\ln d_{it} = \varepsilon_{it} = v_{it} - u_{it}
\]

4 Empirical results

4.1 Estimation results for latent classes

To test the hypothesis if larger and smaller network operators in Germany operate under different technologoes we estimate a latent class model with two different classes to model parameter heterogeneity. Within this framework we allow firms to have different underlying production technologies, caused by unobserved technological network differences or variations in the customer structure. Estimating a latent class model requires the a priori determination of the number of classes (see Section 2.2). Following Orea and Kumbhakar (2004) and Greene (2007) we applied the Akaike information criteria which favors the definition of two classes in contrast to one single class. The empirical results with more than two classes (up to four) did not converge to a meaningful solution within the maximum likelihood estimation, therefore we stayed in the framework of two different classes. This is in our view appropriate to test our hypotheses of different technologies for small and large distributors.
Table 1 shows the regression results of the distance function estimation for the two different classes. All variables are median-corrected to avoid outliers in the sample having a large impact on the estimation outcome. The estimated coefficients of the first-order terms have the expected signs and are statistically significant (see Section 2.1). Thus the estimated distance function appears to reasonably fit the observed data. The prior class probabilities show a quite equal latent sorting of the observations into both classes with a slightly higher amount of companies belonging to the first class: 57% in Class 1 and 43% in Class 2. The characteristics of both classes are shown in Table 2.

We start by characterizing both classes and calculate summary statistics of important physical and technical data of the companies differentiated between the two classes (see Table 2). We can derive one clear trend: Class 1 incorporates larger distribution companies with a higher number of employees, higher amount of delivery sold, more final customers and a larger network; Class 2 seems to include the smaller German distributors. With regard to the number of final customers we note that the separation reflects approximately the classification in the ordinance with 30.000 customers connected (AregV §24). This separation already gives a first insight that the size of the network operators matters in identifying the technology.

Next, we want to determine whether the production structures of both classes (larger firms vs. smaller) differ and may be characterized by parameter heterogeneity. Table 1 shows that the year dummies in both classes are insignificant which suggests no important technology shifts within the observation period. In both classes the input and output coefficients have the expected signs and are all significant. The coefficients of first-order output variables represent the cost elasticities with respect to the corresponding outputs. The coefficients of first-order input variables show the cost shares of the respective inputs. However, we note that the coefficients differ significantly for the two classes: Class 1 is characterized by higher capital intensity with a coefficient of 0.81 vs. Class 2 of 0.759. Larger firms operating larger networks and in particular more cost intensive high voltage networks (up to 110 kV) and are therefore characterized by a more capital intensive distribution.17

For an input distance function representation the elasticity of scale (RTS) is measured by the negative of the inverse of the sum of the output elasticities (Färe and Primont, 1995). As the output weights do not sum to unity this can be interpreted as reflecting the effect of non-constant returns to scale (Saal et al., 2007). The sum of the coefficients of the two output variables varies (1.033 for Class 1 and 0.51 for

---

17 The homogeneity of degree one assumption involves that the input coefficients sum up to one. Thus we obtain a labor share of 0.19 and 0.24 respectively. The electricity distribution sector is obviously characterized by a high capital cost share.
Class 2). This result suggests the presence of important increasing returns to scale for Class 2 while Class 1, the larger companies, operates closely under constant returns to scale. We notice that the output elasticities with respect to customers differ significantly between the two classes (-0.458 in Class 1 vs. -0.132 in Class 2) indicating that connecting customers is more cost intensive for Class 1. This can be explained by the fact that larger network operators have a higher share of industrial customers for which connection to the network is more costly than for household customers.

Clearly, the empirical evidence shows that the two groups operate under different technologies. The latent class specification leads to different technological production frontiers as references for the different companies. Estimating a common frontier without modeling the parameter heterogeneity would produce biased estimates and therefore inconsistent individual efficiency measures. The hypothesis that larger and smaller network operators are characterized by different technologies can be confirmed. The latent class model is able to estimate the technologies and the class probabilities simultaneously. In contrast to the two-stage approach, all observations of the sample are used to determine the underlying technologies for each class. This overcomes the implicit restriction of the two-stage approach which precludes using observations that were allocated to one subgroup to determine the efficiency frontiers of the other groups (Orea and Kumbhakar, 2004). Thus it represents a promising alternative to the two-step procedure.

4.2 Classification of the sample

Table 1 also shows the estimated coefficients of latent class probability functions \( \hat{\delta} \) together with their respective \( p \)-value. Thus, the class probabilities are not fixed but dependent on time-invariant observable characteristics of the German distribution companies. Including these variables reveals whether they deliver useful information in classifying the sample, more precisely if they provide information on the probability of a distribution company belonging to a certain class.

The empirical results suggest that the supply to other electricity distribution companies, the generation activity, the operation of high voltage cables, the percentage share of household customers in the total sum and the location in East or West Germany do not have significant impacts on the probability of belonging to a certain class. The \( p \)-value of these estimated coefficients is lower than the critical value of 0.05. However, we note that the remaining observable characteristics have an impact on classifying the sample into larger vs. smaller operators. The density has a positive impact on the probability belonging to Class 1. In other words, Class 1 consists of larger companies operating in more densely settled urban areas. Average losses of electricity (a type
<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model parameters for latent class 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.331</td>
<td>0.038</td>
<td>8.683</td>
<td>0.000</td>
</tr>
<tr>
<td>Lnx1 network</td>
<td>0.831</td>
<td>0.015</td>
<td>54.360</td>
<td>0.000</td>
</tr>
<tr>
<td>Lny1 delivery</td>
<td>-0.575</td>
<td>0.027</td>
<td>-21.434</td>
<td>0.000</td>
</tr>
<tr>
<td>Lny2 customers</td>
<td>-0.458</td>
<td>0.029</td>
<td>-15.899</td>
<td>0.000</td>
</tr>
<tr>
<td>Year-Dummy 2001</td>
<td>0.010</td>
<td>0.031</td>
<td>0.323</td>
<td>0.747</td>
</tr>
<tr>
<td>Year-Dummy 2002</td>
<td>-0.012</td>
<td>0.031</td>
<td>-0.393</td>
<td>0.695</td>
</tr>
<tr>
<td>Year-Dummy 2003</td>
<td>-0.017</td>
<td>0.031</td>
<td>-0.559</td>
<td>0.576</td>
</tr>
<tr>
<td>Year-Dummy 2004</td>
<td>-0.009</td>
<td>0.030</td>
<td>-0.285</td>
<td>0.776</td>
</tr>
<tr>
<td>Sigma</td>
<td>0.307</td>
<td>0.022</td>
<td>13.946</td>
<td>0.000</td>
</tr>
<tr>
<td>Lambda</td>
<td>1.473</td>
<td>0.331</td>
<td>4.448</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Model parameters for latent class 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.014</td>
<td>0.084</td>
<td>0.160</td>
<td>0.873</td>
</tr>
<tr>
<td>Lnx1 network</td>
<td>0.759</td>
<td>0.027</td>
<td>28.443</td>
<td>0.000</td>
</tr>
<tr>
<td>Lny1 delivery</td>
<td>-0.606</td>
<td>0.020</td>
<td>-30.119</td>
<td>0.000</td>
</tr>
<tr>
<td>Lny2 customers</td>
<td>-0.132</td>
<td>0.022</td>
<td>-6.072</td>
<td>0.000</td>
</tr>
<tr>
<td>Year-Dummy 2001</td>
<td>0.000</td>
<td>0.042</td>
<td>-0.009</td>
<td>0.993</td>
</tr>
<tr>
<td>Year-Dummy 2002</td>
<td>-0.037</td>
<td>0.042</td>
<td>-0.863</td>
<td>0.388</td>
</tr>
<tr>
<td>Year-Dummy 2003</td>
<td>-0.017</td>
<td>0.042</td>
<td>-0.414</td>
<td>0.679</td>
</tr>
<tr>
<td>Year-Dummy 2004</td>
<td>-0.006</td>
<td>0.042</td>
<td>-0.135</td>
<td>0.893</td>
</tr>
<tr>
<td>Sigma</td>
<td>0.321</td>
<td>0.040</td>
<td>8.072</td>
<td>0.000</td>
</tr>
<tr>
<td>Lambda</td>
<td>0.833</td>
<td>0.484</td>
<td>1.719</td>
<td>0.086</td>
</tr>
<tr>
<td><strong>Estimated prior probabilities for class membership</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.980</td>
<td>0.656</td>
<td>1.493</td>
<td>0.136</td>
</tr>
<tr>
<td>Dummy generation</td>
<td>-0.434</td>
<td>0.460</td>
<td>-0.944</td>
<td>0.345</td>
</tr>
<tr>
<td>Dummy EDC</td>
<td>-0.186</td>
<td>0.440</td>
<td>-0.422</td>
<td>0.673</td>
</tr>
<tr>
<td>Dummy high voltage cable</td>
<td>2.331</td>
<td>1.346</td>
<td>1.732</td>
<td>0.083</td>
</tr>
<tr>
<td>Dummy high voltage lines</td>
<td>-2.534</td>
<td>1.048</td>
<td>-2.417</td>
<td>0.016</td>
</tr>
<tr>
<td>Dummy West/East</td>
<td>-0.958</td>
<td>0.549</td>
<td>-1.743</td>
<td>0.081</td>
</tr>
<tr>
<td>Average Investment</td>
<td>-3.465</td>
<td>0.652</td>
<td>-5.315</td>
<td>0.000</td>
</tr>
<tr>
<td>Average Revenue</td>
<td>3.923</td>
<td>0.639</td>
<td>6.137</td>
<td>0.000</td>
</tr>
<tr>
<td>Ratio delivery household/total delivery</td>
<td>1.003</td>
<td>0.646</td>
<td>1.555</td>
<td>0.129</td>
</tr>
<tr>
<td>Investment per network</td>
<td>3.289</td>
<td>0.653</td>
<td>5.039</td>
<td>0.000</td>
</tr>
<tr>
<td>Revenue per delivery</td>
<td>-4.221</td>
<td>0.007</td>
<td>-583.532</td>
<td>0.000</td>
</tr>
<tr>
<td>Cable per network</td>
<td>-0.989</td>
<td>0.010</td>
<td>-99.792</td>
<td>0.000</td>
</tr>
<tr>
<td>Average losses</td>
<td>-0.572</td>
<td>0.007</td>
<td>-85.981</td>
<td>0.000</td>
</tr>
<tr>
<td>Density</td>
<td>0.184</td>
<td>0.008</td>
<td>23.590</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Prior class probabilities at data means for LCM variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>0.571</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td>0.429</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stochastic frontier model variance parameters</strong></td>
<td>Lambda</td>
<td>Sigma</td>
<td>Sigma(u)</td>
<td>Sigma(v)</td>
</tr>
<tr>
<td>Class 1</td>
<td>1.473</td>
<td>0.307</td>
<td>0.254</td>
<td>0.172</td>
</tr>
<tr>
<td>Class 2</td>
<td>0.833</td>
<td>0.321</td>
<td>0.206</td>
<td>0.247</td>
</tr>
</tbody>
</table>

Table 1: Estimation results of Model 2 (latent class specification)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Confidence Interval 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of employees</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>126</td>
<td>14</td>
<td>97 - 154</td>
</tr>
<tr>
<td>Class 2</td>
<td>74</td>
<td>8</td>
<td>59 - 90</td>
</tr>
<tr>
<td>Delivery in MWh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>626977</td>
<td>85713</td>
<td>458760 - 795195</td>
</tr>
<tr>
<td>Class 2</td>
<td>312335</td>
<td>50479</td>
<td>213268 - 411402</td>
</tr>
<tr>
<td>Final customers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>54641</td>
<td>6243</td>
<td>42388 - 66893</td>
</tr>
<tr>
<td>Class 2</td>
<td>31751</td>
<td>3142</td>
<td>25584 - 37917</td>
</tr>
<tr>
<td>Weighted km of lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>822</td>
<td>168</td>
<td>493 - 1151</td>
</tr>
<tr>
<td>Class 2</td>
<td>97</td>
<td>8</td>
<td>81 - 113</td>
</tr>
<tr>
<td>Weighted km of cables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>2188</td>
<td>408</td>
<td>1387 - 2990</td>
</tr>
<tr>
<td>Class 2</td>
<td>752</td>
<td>58</td>
<td>637 - 866</td>
</tr>
<tr>
<td>Weighted km of network</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>3010</td>
<td>504</td>
<td>2022 - 3998</td>
</tr>
<tr>
<td>Class 2</td>
<td>848</td>
<td>61</td>
<td>729 - 968</td>
</tr>
<tr>
<td>Density</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>3368</td>
<td>63</td>
<td>3244 - 3491</td>
</tr>
<tr>
<td>Class 2</td>
<td>3142</td>
<td>73</td>
<td>2999 - 3285</td>
</tr>
<tr>
<td>Unweighted km of lines</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>486</td>
<td>89</td>
<td>312 - 659</td>
</tr>
<tr>
<td>Class 2</td>
<td>75</td>
<td>6</td>
<td>63 - 87</td>
</tr>
<tr>
<td>Unweighted km of cables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>1860</td>
<td>354</td>
<td>1166 - 2554</td>
</tr>
<tr>
<td>Class 2</td>
<td>618</td>
<td>45</td>
<td>530 - 705</td>
</tr>
<tr>
<td>Unweighted km of network</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>2345</td>
<td>403</td>
<td>1553 - 3137</td>
</tr>
<tr>
<td>Class 2</td>
<td>693</td>
<td>46</td>
<td>603 - 783</td>
</tr>
<tr>
<td>Delivery inland in MWh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>622306</td>
<td>85534</td>
<td>454440 - 790171</td>
</tr>
<tr>
<td>Class 2</td>
<td>308665</td>
<td>49903</td>
<td>210728 - 406602</td>
</tr>
<tr>
<td>Delivery to industry in MWh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>198541</td>
<td>27888</td>
<td>143610 - 253073</td>
</tr>
<tr>
<td>Class 2</td>
<td>84766</td>
<td>13213</td>
<td>58836 - 110697</td>
</tr>
<tr>
<td>Delivery to households in MWh</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 1</td>
<td>165312</td>
<td>20786</td>
<td>124518 - 206105</td>
</tr>
<tr>
<td>Class 2</td>
<td>72113</td>
<td>6595</td>
<td>59170 - 85055</td>
</tr>
</tbody>
</table>

Table 2: Sample statistics of the two latent classes
of quality index) lowers the probability of being in Class 1, which indicates that the larger companies are characterized by higher quality standards, i.e. electricity losses.\textsuperscript{18} A higher average investment per network length increases the probability of being in the class of larger distributors. Thus, we conclude that larger companies on average invest more per km network.

Considering the average investment separately, without relating it to the network length, we obtain a negative sign indicating that the probability of being in Class 1 decreases. This would appear to contradict the previous result. However, the relation to the capital input defined in our analysis by the length of the network indicates more reliable results than considering it separately. The same argument applies to revenue, included first as a separate variable and then in relation to the units delivered. We argue that the revenue per unit delivered shows a more reliable picture; here we obtain a negative coefficient.\textsuperscript{19} This empirical result is unexpected as it suggests that higher revenue per unit electricity delivered decreases the probability of being in Class 1. Smaller distribution operators therefore are characterized by higher revenues per units. However we can argue that small local distributors (Stadtwerke) might be characterized by a higher cross-subsidization.\textsuperscript{20}

The latent class estimation provides empirical evidence that on the one hand we have to consider different technologies for different classes. We can also explain the classification of companies by observable characteristics that provide more sophisticated information about the groups. These are important for correctly estimating the true technology frontier for efficiency analysis.

### 4.3 Efficiency analysis

The sample statistics for the estimated efficiencies for the whole sample and for each estimated class are shown in Table 3. The values lie between 0 and 1, with no company showing full efficiency. The values of the efficiency vary from 0.647 to 0.978. The values of the mean technical efficiency indices are relatively high: 0.91. The high average efficiency is conform with the mean efficiency calculated by the German regulator for the German network operators. We observe a difference of the performance levels in the latent classes (0.90 vs. 0.92). In our sample it appears that the smaller distribution

\textsuperscript{18}This may also be explained with the higher voltage levels that seem to prevail in Class 1: electricity losses are inversely related to voltage levels.

\textsuperscript{19}The correlation of the four variables is very low (0.161 for average investment and investment per network; -0.277 for average revenue and revenue per unit delivered); therefore we can explain the different coefficients. This also ensures that we do not have any multi-collinearity problems including all variables as explanatory factors.

\textsuperscript{20}When we consider the average revenue separately it indicates that higher revenues are related to larger operators.
### Table 3: Descriptive statistics of efficiency estimates ordered by classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of observations</th>
<th>Mean</th>
<th>Std.dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent class model</td>
<td>1000</td>
<td>0.910</td>
<td>0.041</td>
<td>0.647</td>
<td>0.978</td>
</tr>
<tr>
<td>Latent class model Class 1</td>
<td>535</td>
<td>0.9</td>
<td>0.002</td>
<td>0.647</td>
<td>0.978</td>
</tr>
<tr>
<td>Latent class model Class 2</td>
<td>465</td>
<td>0.921</td>
<td>0.001</td>
<td>0.791</td>
<td>0.967</td>
</tr>
</tbody>
</table>

### Table 4: Kruskal-Wallis test

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Statistically more efficient class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 vs Class 2</td>
<td>can be confirmed</td>
</tr>
<tr>
<td>535 observations</td>
<td>465 observations</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
</tr>
</tbody>
</table>

companies are operating under increasing returns to scale, but from a pure technical efficiency perspective show a higher performance compared to larger distributors. This empirical evidence shows that accounting for the true frontier in each group is important for the benchmarking process.\(^{21}\)

The difference between the classes can be confirmed statistically by means of the Kruskal-Wallis Test, testing the hypothesis that several samples are from the same population. The results of the Kruskal-Wallis test are shown in Table 4. The \(p\)-value is 0.0001. These results indicate that we can reject the hypotheses of equal distribution. This again leads us to conclude that when assuming different technologies by capturing parameter heterogeneity in the econometric model we obtain more robust results for the individual efficiency estimates. This is due to the fact that we can adapt better the technology and therefore the production frontier to different classes of firms with different characteristics.

\(^{21}\)For comparison reasons we also estimated the true random effects model (see section 2.2). Within this framework all companies are benchmarked against the same technology (apart from the technology neutral shift captured by the individual specific randomly distributed constant). For this model specification we do not note any significant difference in the average efficiency in both classes. From a descriptive perspective we observe that within the latent class model specification some of the technological heterogeneity captured is labeled as inefficiency in the true random effects model. Accounting for unobserved factors not only in the intercept (technological neutral shift) but in different technologies and therefore frontiers produces other conclusions about firms’ performance.
Conclusions

In this paper we analyzed the technical efficiency level for a sample of 200 German electricity distribution companies subject to incentive regulation since January 2009. The new regulatory instruments are based on benchmarking procedures to determine the revenue caps of the individual companies. In the empirical application of benchmarking, regulators and researchers are always faced with the problem of a high degree of heterogeneity for environmental or network characteristics. Only some of this heterogeneity is observed and can therefore be accounted for in the econometric model. Another part will be unobserved. These unobserved characteristics between firms, that are not measured in the sample might influence the underlying production process. Therefore, the problem becomes one of modeling unobserved heterogeneity.

Comparing the efficiency of different firms usually assumes that they operate under the same production technology and therefore these unobserved factors might be understood as inefficiency. To avoid such types of misspecifications we observe that in regulatory practice estimation is often carried out in two stages. First, observations are classified into several groups assuming a priori that they operate under different technologies. Then in a second step, separate analyses are conducted for each sub-group of the sample. This paper shows how to disentangle the heterogeneity from inefficiency in one step, using a latent class model for stochastic frontiers. Within this framework the classification is not based on a priori sample separation criteria and therefore delivers more robust and statistical significant and testable results. In the latent class model the unobserved firm-specific heterogeneity is accounted for by parameter heterogeneity, identifying different technologies for companies. We show that this model will partially solve the unobserved heterogeneity problem in measuring the technical efficiency. The empirical results suggest that our proposed model is able to account for the fact that larger distributors operate under a different technology than smaller companies and that different frontiers are necessary to obtain more robust and reliable efficiency estimates. It represents a promising alternative to the traditional two step procedures.

We find that the estimated Cobb-Douglas distance function is a reasonable fit to the observed data and that the estimated input and output elasticities have the correct sign and magnitude for both classes. Determining the returns to scale we observe that the latent class specification can differentiate between large and smaller distribution companies: in the latent class model the estimated coefficients indicate that the larger distributors operate under constant returns to scale and the smaller firms under increasing returns to scale.

In addition, we test the differences in inefficiency scores between the two classes
via a Kruskal-Wallis test. The results underline the importance of modeling and estimating two classes. The latent class model can be helpful in distinguishing unobserved heterogeneity in technologies from inefficiency estimates. The results can be used as an additional instrument to reduce the information asymmetry between the regulator and regulated companies.

6 Acknowledgments

I am indebted to Christian von Hirschhausen, Matthias Walter, Borge Hess and the DIW IO seminar for their helpful comments and support.

References


