Forthcoming in the Review of Industrial Organization, 2004.

REGULATION AND MEASURING COST EFFICIENCY WITH PANEL DATA MODELS: APPLICATION TO ELECTRICITY DISTRIBUTION UTILITIES

Mehdi Farsi Massimo Filippini

Center for Energy Policy and Economics

Federal Institute of Technology

ETH Zentrum, WEC, 8092 Zurich, Switzerland

and

Department of Economics

University of Lugano

Via Ospedale 13, 6900 Lugano, Switzerland

April 2004

Correspondence: Massimo Filippini, CEPE, ETH Zentrum, CH-8092, Zurich, Switzerland

Phone: +41-1-632-0649, Fax: +41-1-632-1050, E-mail: filippin@cepe.mavt.ethz.ch

ABSTRACT

This paper examines the performance of panel data models in measuring cost efficiency of electricity distribution utilities. Different cost frontier models are applied to a sample of 59 utilities operating in Switzerland from 1988 to 1996. The estimated coefficients and inefficiency scores are compared across different specifications. The results indicate that while the average inefficiency is not sensitive to the econometric specification, the efficiency ranking varies significantly across models. The reasonably low out-of-sample prediction errors suggest that panel data models can be used as a prediction instrument in order to narrow the information gap between the regulator and regulated companies.

Keywords: Cost efficiency; Electricity utilities; Incentive regulation; Yardstick competition.

1. INTRODUCTION

Transmission and distribution of electricity have been considered as natural monopolies, thus less affected by the recent waves of deregulation in power industry. However, as competition is introduced into generation sector, regulatory reform and incentive regulation of distribution utilities have become more common. In traditional cost-of-service regulation systems companies recover their costs with a risk-free fixed rate of return and therefore have little incentive to minimize costs. The incentive-based schemes on the other hand, are designed to provide incentive for cost-efficiency by compensating the company with its savings. A variety of methods are proposed in the literature. Main categories of incentive-based schemes used for electricity utilities are: price or revenue cap regulation schemes, sliding-scale rate of return, partial cost adjustment, menu of contracts, and yardstick regulation.[1] Jamasb and Pollitt (2001) provide an extensive survey of different regulation practices in electricity markets around the world. Virtually all the models used in practice, are based on 'benchmarking' that is, measuring a company's efficiency against a reference performance. Inefficiency can be resulted from technological deficiencies or non-optimal allocation of resources into production. Both technical and allocative inefficiencies are included in cost-inefficiency, which is by definition, the deviation from minimum costs to produce a given level of output with given input prices.

In benchmarking applications the regulator is generally interested in obtaining a measure of firms' inefficiencies such as X-factors in price cap regulation, in order to reward (or punish) companies accordingly. The inefficiency estimates can have great financial sequences for the firms and therefore, their reliability is crucial for an effective regulation system. In particular, if the estimated inefficiency scores are sensitive to the benchmarking method, a more detailed analysis to justify the adopted model is required. However, in most

[1] See Joskow and Schmalensee (1986) for a review of regulation models.

cases it is difficult to identify the ‘right’ model among the set of legitimate ones. Bauer et al. (1998) propose a series of criteria that can be used to evaluate if the results obtained from different methods are “mutually consistent”, that is, lead to comparable inefficiency scores and ranks. It is recommended that rather than using the inefficiency estimates in a mechanical way, the benchmarking analysis should be used as a complementary instrument in incentive regulation schemes.

Using a cross section of 63 power distribution utilities in Europe, Jamasb and Pollit (2003) show that there are substantial variations in estimated efficiency scores and rank orders across different methods.[2] There is a common perception that the estimation results can be improved using panel data. In contrast with cross-sectional data, panels provide information on same companies over several periods. Moreover, panel data models can better control for unobserved heterogeneity among companies. This perception is supported by suggestive evidence. For instance, after reviewing their previous empirical literature on measuring inefficiency with panel data, Kumbhakar and Lovell (2000)[3] conclude that different approaches are likely to generate rather similar efficiency rankings, especially at the top and bottom of the distribution. However, none of the cited works is related to electricity distribution.[4]

These results raise an important question as to whether (or to what extent) the problems reported by Jamasb and Pollitt (2003) are due to the limitations associated with cross-section data models. This question is addressed in the present paper by using several alternative

[2] Other authors like Horrace and Schmidt (1996), Street (2003) and Jensen (2000) reported substantial errors and inconsistency problems in the estimation of individual efficiency scores in cross sectional data.

[3] See page 107.

[4] There is some evidence that the reliability of benchmarking analysis depends on the nature of production. For instance Gong and Sickles (1989), using Monte Carlo simulations, find that with complex production functions most models show a poor performance, while with simpler production forms the results are more reasonable.

models to a panel of power distribution utilities in Switzerland. Given that the validity of benchmarking methods in electricity industry has been put into question,[5] studying the reliability of these methods using better data (panels in this case) can bring some light into a controversial debate. Moreover, as an increasing number of regulators throughout Europe and elsewhere have access to panel data sets, they will soon face the question if the longitudinal data can help evaluate the regulated companies' inefficiency.

The main goal of this paper is to study the sensitivity of inefficiency estimates in panel data models. We focus on parametric cost frontier methods mainly because these methods are more easily adaptable to panel data.[6] It should be noted that in practice, most of the regulators use deterministic frontier methods like Data Envelopment Analysis. This is mainly because such methods require a relatively low number of observations. However, the effects of unobserved differences among companies, which are particularly important in network industries, are completely ignored in these models.

In this paper, different stochastic frontier models are applied to a sample of distribution utilities in Switzerland. The sample is an unbalanced panel of 59 companies over a period of 9 years (a total of 380 observations). The inefficiency scores estimated from four different models are compared. In particular, the resulted efficiency rankings obtained from random effect and fixed effect models are analyzed. Although there is a reasonably good correlation between the estimates obtained from certain models, individual inefficiency scores and ranks change quite significantly from one model to another. The estimated measures of inefficiency are therefore sensitive to econometric specification and should not be used as a direct instrument in benchmarking. These results suggest that the sensitivity problems in

[5] For instance see Shuttlewoth (2003) and Irastorza (2003) for criticisms of benchmarking approaches in electricity industry.

[6] Other advantages of the parametric approaches are described in section 2.

benchmarking electricity utilities are not limited to cross-sectional data and cannot be completely resolved by using panel data models. However, our analysis of prediction errors indicates that panel data models can be used for predicting utilities costs with a rather reasonable error, suggesting that benchmarking analyses can be used as a complementary instrument for monitoring utilities' performances.

The rest of the paper proceeds as follows: Section 2 provides a brief review of cost-frontier models and presents the specification used in this paper. The data are described in section 3. Section 4 presents the estimation results and discusses their implications. The main conclusions are summarized at the end.

2. METHODOLOGY

A frontier cost function defines minimum costs given output level, input prices and the existing production technology. It is unlikely that all firms will operate at the frontier. Failure to attain the cost frontier implies the existence of technical and allocative inefficiency. This section provides a description of the cost frontier models and the specification used in this paper. The adopted methodology is based on a comparison of different models with respect to the estimated cost function parameters, estimated inefficiency scores, and the out-of-sample prediction errors. The main goal is to study the limitations of different models in benchmarking and the sensitivity of inefficiency scores to model selection.

2.1. Cost frontier models

There are several cost frontier methods to estimate the cost efficiency of individual firms. Two main categories are non-parametric methods originated from operations research, and econometric approaches namely stochastic cost frontier models.[7] In non-parametric approaches like Data Envelopment Analysis, the cost frontier is considered as a deterministic function of the observed variables but no specific functional form is imposed. Moreover, non-parametric approaches are generally easier to estimate.[8] Parametric methods on the other hand, allow for a random unobserved heterogeneity among different firms but need to specify a functional form for the cost function. The main advantage of such methods over non-parametric approaches is the separation of the inefficiency effect from the statistical noise due to data errors, omitted variables etc. The non-parametric methods' assumption of a unique deterministic cost frontier for all companies is unrealistic. Another advantage of parametric methods is that these methods allow statistical inference on the significance of the variables included in the model, using standard statistical tests. In non-parametric methods on the other hand, statistical inference requires elaborate and sensitive re-sampling methods like bootstrap techniques.[9] Given the above discussion we decided to focus on the parametric approaches.

In this paper we consider the estimation of a deterministic and three versions of a stochastic frontier cost function using panel data. It should be noted that the theoretical development of stochastic frontier models in panel data has been subject of a great body of literature.[10] We contend that some of the recent developments like the models proposed by

---

[7] See Kumbhakar and Lovell (2000) for an extensive survey of parametric methods and Coelli et al. (1998), chapter 6, and Simar (1992) for an overview of non-parametric approaches.

[8] See Coelli et al. (2003) for more details.

[9] These methods are available for rather special cases and have not yet been established as standard tests. See Simar and Wilson (2000) for an overview of statistical inference methods in non-parametric models.

[10] See Kumbhakar and Lovell (2000) for a review and Greene (forthcoming) and Tsionas (2002) for some recent developments.

Greene (forthcoming) can be useful in benchmarking. However, a critical analysis of these models needs further research and is beyond the scope of this paper.

The deterministic approach adopted in this paper can be formulated as:

$$\ln C_{it} = \ln C(y_{it}, w_{it}) + u_{it} \quad u_{it} \geq 0 \qquad i = 1, 2, ..., N \;\; and \;\; t = 1,2,\ldots,T \qquad (1)$$

where C_{it} is observed total cost in year t, y_{it} is a vector of outputs, w_{it} is an input price vector and u_{it} is a positive one-sided disturbance capturing the effect of inefficiency. N represents the number of firms and T the number of years in the sample. Firms can therefore operate on or above the cost frontier but not below it. One interesting method proposed for estimating equation (1) is Greene's (1980) version of Richmond's (1974) Corrected Ordinary Least Squares model. A functional form for the cost function is assumed, and parameter estimates are obtained using ordinary least squares method. The intercept is corrected by shifting the value of the intercept such that all residuals are positive and at least one is zero.

The main shortcoming of this method is that it confounds inefficiency with statistical noise: the entire residual is classified as inefficiency. Nevertheless many studies have used this approach.[11] It should be noted that the COLS method does not consider the panel aspect of the data because it considers the repeated observations of a given firm as independent observations. These problems can be partly overcome using the stochastic cost frontier approach suggested by Pitt and Lee (1981) who extended the original model of Aigner et al. (1977) to panel data setups. This model can be written as follows:

$$\ln C_{it} = \ln C(y_{it}, w_{it}) + u_i + v_{it} \quad u_i \geq 0 \qquad i = 1, 2, ..., N \;\; and \;\; t = 1,2,\ldots,T \qquad (2)$$

In this specification the error term is composed of two uncorrelated parts: The first part u_i, is a one-sided non negative disturbance reflecting the effect of inefficiency (including both

[11] See for example Wagstaff (1989) and Filippini and Maggi (1993).

allocative and technical inefficiencies), and the second component v_{it} , is a symmetric disturbance capturing the effect of noise. Usually the statistical noise is assumed to be normally distributed, while the inefficiency term u_i is assumed to follow a half-normal distribution.[12] This model with a normal-half-normal composite error term can be estimated using Maximum Likelihood Estimation method. Consistent with Kumbhakar and Lovell (2000) we refer to this model as the MLE model.

Compared to a deterministic approach the main advantage of this stochastic cost frontier model is the separation of the inefficiency effect from the statistical noise. However, this method is subject to the potential criticism of having arbitrary assumptions about the distribution of the random terms. These assumptions can be relaxed by rewriting equation 2 as:

$$\ln C_{it} = \ln C\left(y_{it}, w_{it}\right) + \alpha + u_i^* + v_{it} \qquad \text{with} \qquad u_i = u_i^* - \min\{u_i^*\}, \tag{3}$$

and using a feasible Generalized Least Squares method as proposed by Schmidt and Sickles (1984).[13] The resulting model is referred to as the GLS model.

The remaining restrictive assumption is that the two random components be uncorrelated with each one of the explanatory variables. This implies that the firm's inefficiency is uncorrelated with its observed characteristics included in the cost function. In the real world however, many of these factors may affect the firm's inefficiency. Schmidt and Sickles (1984) propose a solution around this assumption.[14] In their model the overall residual w_{it} is composed of two terms ($w_{it} = u_i + v_{it}$): a symmetric disturbance v_{it}, like

---

[12] Other extensions of this model have also considered exponential and truncated normal distributions for the inefficiency term. See for instance Battese (1992) and Battese and Coelli (1992).

[13] See also Kumbhakar and Lovell (2000).

[14] For a presentation of this method see also Simar (1992).

previous models, and a one-sided fixed component u_i, that represents cost inefficiency. The latter component can be identified by a fixed effects specification with no assumption on the distribution of u_i.[15] Inefficiency scores are estimated as the distance to the firm with the minimum fixed effect, that is: $u_i - \min\{u_i\}$. The resulting model is a fixed-effects model and is labeled as the FE model in the rest of the paper.

The fixed effects approach controls for unobservable firm specific effects, such as inefficiency, that are not captured by control variables. There are however, two limits to this approach: First, the time invariant variables are captured by the fixed effects and cannot be included in the model. This implies that the inefficiency estimators include the variations in time-invariant firm characteristics. Moreover, inefficiency is assumed to be constant over time. Notice that this assumption can be relaxed in the random effects models discussed above.[16]

The main advantage of the fixed-effects specification is that the estimations are unbiased even if explanatory variables are correlated with firm-specific dummies. However, the inefficiency measures may be confounded with time-invariant factors, which could not be included in the model. The choice between random effects and fixed effects models also depends on whether or not firms belong to the same population.[17] The random effects model is a legitimate specification to the extent that the heterogeneity among companies is limited to a single population.

---

[15] In this approach the term stochastic is referred to the fact that the model is stochastic (presence of a symmetric component of the disturbance v_{it}) but not the inefficiency term u_i. In the approach suggested by Aigner et al. (1977) both the model and the inefficiency term are stochastic.

[16] Battese and Coelli (1992) propose a method. See also Coelli, Rao and Battese (1998) for a summary.

[17] See Baltagi (2001) and Hsiao and Sun (2000) for detailed discussions on fixed *vs* random effects.

2.2. Specification of the Cost Function

Electricity distribution utilities operate in networks with different shapes, which directly affect the costs. As discussed in Robert (1986), Salvanes and Tjϕota (1994) and Thompson (1997), the cost function should take into account differences in network characteristics, load factor and other factors that are unrelated to cost-efficiency but affect the costs. The specification used here draws basically from the model used by Filippini (1998). The output is measured by the total number of kWh delivered. Inputs to the electricity distribution process consist primarily of labor, capital and the power purchased from the generator. The costs of distribution utilities consist of two main parts: the costs of the purchased power and the network costs including labor and capital costs. There are therefore two alternatives for measuring cost efficiency in power distribution utilities: total costs approach and network costs approach. The network costs approach has a practical advantage in that the estimated average costs can be directly used in a price-cap formula.[18] However, this approach neglects the potential inefficiencies in the choice of the generator. In this paper we use the first approach based on the total costs function. The firm's total cost of distributing electricity can be represented by:

$$\mathrm{C} = \mathrm{C}(Y, P_K, P_L, P_P, LF, CU, AS, HGRID, DOT, DW, T) \qquad (4)$$

where C represents total cost; Y is the output in kWh; P_K , P_L and P_P are respectively the prices of capital, labor and input power; and T is a time variable representing the linear trend in technological progress.

In addition to the above variables that are generally included in a cost function model, the following six output (and network) characteristics are included in the model: LF is the 'load factor' defined as the ratio of utility's average load on its peak load; AS the size of the

---

[18] Notice that the price cap is generally applied to the network access.

service area served by the distribution utility; and *CU* is the number of customers. The load factor captures the impact of the intensity of use on costs.[19] This variable is expected to have a negative effect on total costs, because a relatively high value of load factor represents a higher capacity utilization, thus lower fixed costs for producing a given output. Obviously, the service area and the number of customers are expected to have a positive effect on costs.

HGRID is a binary indicator to distinguish the utilities that operate a high-voltage transmission network in addition to their distribution network. Some of the utilities in our sample are involved in auxiliary services such as installation of electric appliances. Both these types are expected to be more costly compared to other companies. The utilities whose share of auxiliary revenues is more than 25 percent of total revenues are distinguished by dummy variable *DOT*. The maintenance costs and damage risks of power lines are generally higher in forests. Binary indicator *DW* represents the cases in which more than 40 percent of the service area is covered by forests.

It should be noted that there are other output characteristics that are not considered in the cost function (4). An important unobserved dimension of the output in our sample concerns the quality of service. Different utilities may deliver different levels of quality, which affect their costs. In power utilities, output quality is usually measured by the number of interruptions that are not related to rare natural accidents. In Switzerland however, there has been practically no such outage cases. The high level of quality is obviously related to the tight regulation and high quality standards applied in Switzerland. We therefore contend that the quality differences between the utilities in our sample are not significant. Moreover, the stochastic models used in our analysis control for unobserved variations that are not correlated with the included explanatory variables.

---

[19] See Foreman-Peck and Waterson (1985) for a discussion of the role of load factor in cost models.

The regularity conditions require that the cost function in equation (4) be linearly homogeneous in input prices, non-decreasing in input prices and concave.[20] The translog model and Cobb-Douglas form are two main functional forms commonly used in the literature. Translog form does not impose any technological restriction and allows the economies of scale, size and density vary with output. These values are assumed constant in the Cobb-Douglas functional form. In this paper, the Cobb-Douglas form is used for two main reasons. First, because of the large number of parameters[21] in translog model there is a considerable risk of near-multicollinearity, especially given that different output variables *Y*, *AS* and *CU* are highly correlated.[22] Moreover, the estimation of scale economies is of secondary importance in this paper. The assumption that scale-economies do not vary with output (implicit in Cobb-Douglas form) can therefore be justified to the extent that it does not affect the inefficiency estimators.

The Cobb-Douglas specification of the cost function in (4) can be written as:

$$\ln\left(\frac{\mathrm{C}}{P_P}\right) = \beta_0 + \beta_Y \ln Y + \beta_K \ln\left(\frac{P_K}{P_P}\right) + \beta_L \ln\left(\frac{P_L}{P_P}\right) + \gamma_1 \ln LF$$
$$+ \gamma_2 \ln AS + \gamma_3 \ln CU + \delta_1 HGRID + \delta_2 DOT + \delta_3 DW + \tau_{\mathrm{T}} T \qquad (5)$$

Linear homogeneity in input prices is imposed by dividing money values by the price of the input power.

[20] See Cornes (1992) for a discussion of the properties of cost functions.

[21] In our specification the number of parameters in translog model would be 40.

[22] Our preliminary analysis suggests that this problem creates technical difficulties in our maximum likelihood estimations.

3. DATA

The data used in this paper consists of an unbalanced panel of 59 Switzerland's distribution utilities over a 9-year period from 1988 to 1996. The sample includes 380 observations with a minimum of four observations per company. The original data set has been prepared and analyzed by Filippini (1998) and completed by Wild (2001) and Filippini and Wild (2001). These data are mainly based on the information from the annual reports of the Swiss Federal Statistical Office, the Swiss Federal Energy Office, and the Swiss Cities Association. A mail survey from the utilities has been used to complete the data.[23] The sample does not include the utilities that generate more than 20 percent of their input power. There are about 900 electricity distribution companies in Switzerland. This sector is characterized by a large number of small companies along with a relatively small number of large firms. The 59 companies included in this study deliver about a third of Switzerland's electricity consumption.[24] The sample used in this study can therefore be considered as a representative sample of relatively large distribution utilities in the country. In spite of a considerable degree of variation in costs and other characteristics, the sample represents relatively similar companies compared to the entire sector.

Table 1 gives the summary statistics of the key variables used in the analysis. All money values are converted to 1996 Swiss Francs using the global consumer price index. The cost of purchased electricity is included in total costs. For those companies that produce part of their power the average price of input electricity is assumed to be equal to the price of purchased power. Labor price is defined as the average annual salary of the firm's employees.

---

[23] See Filippini and Wild (2001) for a more detailed description of data sources.

[24] The total power delivered by the companies in our sample in 1993 is 13,250 GWh, which is about a third of the total 43,000 GWh electricity consumption in Switzerland in that year.

Capital expenditure is approximated by the residual costs that is, total costs minus labor and purchased power costs. Because of the lack of inventory data the capital stock is measured by the capacity of transformers.[25]

Table 1. Descriptive statistics (380 observations)

	Mean	Standard Deviation	Minimum	Maximum
Total annual costs per kWh output (CHF)	.188	.0303	.128	.323
Annual output (*Y*) in GigaWh	263.51	390.36	17	2301.5
Number of customers (*CU*)	26975.6	36935.8	2461	220060
Load Factor (*LF*)	.5541	.0727	.3219	.9817
Service Area (*AS*) in km^2	15,467	35,376	176	198,946
Average annual labor price (P_L) per employee (CHF 1000)	101.27	32.55	43.36	253.89
Average capital price (P_K) in CHF per kVoltAmpere installed capacity	95.06	39.35	32.08	257.98
Average price of input power (P_P) in CHF/kWh	.105	.0210	.0583	.161
High-voltage network dummy (*HGRID*)	.35	.4776	0	1
Auxiliary revenues more than 25% (*DOT*)	.397	.490	0	1
Share of forest area more than 40% (*DW*)	.261	.440	0	1

- All monetary values are in 1996 Swiss Francs (CHF), adjusted for inflation by Switzerland's global consumer price index.

25 Transformer is the main device used to transfer electricity in the network. This is basically a device to convert current variations to voltage variations and vice versa.

4. ESTIMATION RESULTS

The estimated parameters of the cost frontier are listed in table 2. In the OLS model data are pooled across different years and the estimators are based on an implicit assumption that the unobserved random variations are not firm-specific. The other three models have the advantage of accounting for firms' heterogeneity. In the random-effects model (GLS) it is assumed that firms' unobserved heterogeneity is uncorrelated with their observed characteristics. The MLE model imposes an additional restriction that firms' unobserved heterogeneity has a half-normal distribution. Both these assumptions are relaxed in the fixed-effects specification. This model however is mainly based on the variations within firms and cannot estimate the effect of time-invariant factors.

The first observation on the estimation results (table 2) is that virtually all coefficients are highly significant and have the expected signs. The within R^2 values are reported for the two panel models estimated by the least squares method. These values show that the adopted specification has a relatively high explanatory power.[26] As it can be seen in the table, the fixed-effect estimators for output (*Y*) and costumers (*CU*) coefficients are quite different from other models. This contrasting difference suggests that the estimations could be sensitive to firm-specific characteristics. This result is not surprising in network industries. Any correlation between random effects and other explanatory variables may bias the estimation results. Therefore, in the absence of information regarding the unobserved heterogeneity among firms, the fixed-effect model can provide more reliable estimates for the factors that

[26] It should be noted that the R^2 value in the OLS model does not consider the panel structure of the data and thus, cannot be used to evaluate the model's explanatory power. In fact the extremely high value of R^2 in the OLS model can be explained by the relatively low within variations in costs.

vary over time. This advantage however, is hardly clear in our sample: A more careful examination of the results shows that other coefficients are rather similar among different models. Moreover, even though output coefficients *Y* and *CU* are different in the FE model, their sum is quite similar among different models. This result suggests that the value of economies of scale is robust to the econometric specification. Finally, the results of the Hausman specification test indicate that the FE and RE estimates are not significantly different at 5 percent significance level (p-value of .055). Overall, these results suggest that the cost function estimations are not sensitive to the unobserved heterogeneity among companies.

Table 2. Cost frontier parameters

	OLS		**Random-Effects (GLS)**		**Random-Effects MLE (Half-Normal)**		**Fixed-Effects**	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
lnY	0.851	0.017	0.780	0.032	0.767	0.040	0.677	0.055
lnCU	0.084	0.017	0.153	0.033	0.163	0.048	0.251	0.096
lnAS	0.044	0.004	0.051	0.009	0.047	0.013	-	-
lnLF	-0.243	0.037	-0.239	0.039	-0.230	0.023	-0.213	0.044
lnPL	0.067	0.011	0.041	0.014	0.039	0.014	0.038	0.016
lnPK	0.200	0.009	0.174	0.010	0.171	0.005	0.169	0.010
HGRID	0.063	0.012	0.075	0.027	0.098	0.039	-	-
DOT	0.033	0.010	0.050	0.022	0.040	0.028	-	-
DW	0.014	0.010	0.012	0.023	0.062	0.028	-	-
T	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.002
Constant	-2.236	0.233	-0.793	0.369	-0.653	0.430	-	-
R Square	0.995		0.711				0.714	

- The reported R^2 values for GLS and FE models are based on the within variations.

It should be noted that both random and fixed effects models used in this analysis, assume that the firm-specific stochastic component represents the efficiency differences among firms.[27] This assumption leads to an overestimation of inefficiency in the FE model for the following reasons. First, the fixed firm-specific effects capture both observed and unobserved time-invariant factors. Moreover, since the fixed effects do not follow any distribution and efficiency is estimated compared to the best observed practice (the firm with the minimum fixed effect), the estimators are sensitive to outliers. In fact, the problem of outlier firms is transferred from the cost function to efficiency estimators. To illustrate this fact, several specifications were compared. We started from a 'naïve' OLS model that ignores the time-invariant factors and refined this model in several steps until all time-invariant factors were included. The inefficiency scores obtained from these models were compared to those obtained from a fixed-effects model. The results (not given here) indicate that as the OLS model becomes more 'refined' the estimated inefficiency scores show less correlation with those obtained from the fixed-effects model.[28] These results suggest that as far as inefficiency scores are concerned the performance of a fixed-effect model is quite poor (even compared to a naïve OLS model).

In order to see the limitations of these models we also study the inefficiency estimates obtained from different models. Table 3 gives the summary statistics of the inefficiency scores resulted from different models. The inefficiency score is defined as $\exp(U_i)$, where U_i is the inefficiency term obtained from the regression model. In the COLS model where the inefficiency term is time-variant, the company's inefficiency U_i is assumed to be the average

---

[27] In fact with the exception of a few recent developments (cf. Greene, forthcoming; Tsionas, 2002), this is a general assumption in panel data frontier models.

[28] This is also the case for the mean and median inefficiency scores (and other quartiles). That is, as the OLS model get more refined the summary statistics decrease and get farther from that of the fixed-effects model.

of u_{it} over the entire sample period. The scores therefore represent the ratio of a company's actual costs to a minimum level that would have been achieved had the company operated as cost-efficient as the 'best practice' observed in the sample. The excessively large values resulted from the fixed-effect model confirms the poor performance of this model in estimating inefficiencies.

Table 3. Summary statistics of inefficiency scores

	COLS	GLS	FE	MLE
Minimum	1.07	1	1	1.07
Maximum	1.46	1.38	2.14	1.36
Average	1.23	1.16	1.35	1.15
Median	1.22	1.16	1.31	1.13
95 percentile	1.41	1.32	1.66	1.30
Number of firms	59	59	59	59

In practice, benchmarking is usually based on efficiency ranking of companies. The correlation matrix between the ranks obtained from different models is given in table 4. The ranks are obtained by comparing the firms' average inefficiency scores over the sample period. These results indicate a relatively high correlation between rankings from GLS and MLE models.

Table 4. Correlation between inefficiency ranks from different models

	COLS	**GLS**	**FE**	**MLE**
COLS	1			
RE (GLS)	0.936	1		
FE	0.447	0.514	1	
RE (MLE)	0.838	0.895	0.417	1

To see the individual differences the GLS and MLE models are compared regarding the inefficiency scores. The FE model is not considered here because as discussed above, its efficiency scores are significantly different from other models. Table 5 summarizes the results. These results indicate that the maximum difference of cost-inefficiency between the two models is about 9 percent. This difference is quite significant noting that both models are based on the same Cobb-Douglas functional form and their only difference is in the distribution of the efficiency term. A closer look at the rankings highlights these differences. Our results show that changing from one model to another results in significant changes in rankings. For instance for more than half of the companies in the sample changing the model from GLS to MLE implies a change of 4 places or more in their ranks, and for about 25 percent of companies this means a change of 9 places or more. Even the ranking quartile changes considerably. Change of the model from MLE to GLS results in a change in ranking quartile for 20 companies out of 59. This change results in a change of quartile for about a third of the companies in the first quartile (the 25 percent most efficient firms). Moreover, different models do not identify the same companies as the best and worst practices. The best practice as identified by the MLE method is ranked 17 in the GLS model, whereas the GLS model's best practice is ranked 7 by the MLE model. These differences are as more striking as the two models differ only in their assumption on the distribution of the inefficiency term.

Therefore, the mutual consistency conditions proposed by Bauer et al. (1998)[29] are not satisfied. These results show the sensitivity of the benchmarking method in our sample. In contrast with the general contention in previous studies that different approaches give rather similar inefficiency rankings, this analysis suggests that rankings may be sensitive to the adopted model. Therefore, a direct use of inefficiency estimates in benchmarking regulation

[29] According to Bauer et al., to be mutually consistent, different approaches should have the following conditions: inefficiency scores should have comparable means, standard deviations and other distributional properties; the ranking order of the firms should be approximately the same; and the 'best-practice' and 'worst-practice' firms should be mostly the same.

of network industries may be misleading. In usual cases where the choice of the appropriate model is not clear, a sensitivity analysis could be used to study the robustness of the results and the limitations of different models.

Table 5. Summary statistics of the absolute value of difference in inefficiency scores

Model:	GLS and COLS	GLS and MLE
Minimum	0.01	0.00
Maximum	0.14	0.09
Mean	0.07	0.03
Median	0.07	0.02
95 percentile	0.12	0.08
N	59	59

Cost frontier models can alternatively be used by the regulator to predict the costs of individual companies. Three specifications, OLS, fixed effects (FE) and random effects (GLS), are compared with respect to their predictive power. Predictions are considered in two directions: out-of-sample prediction which consists of predicting the costs of a given firm using the estimations obtained from other firms, and forecasting that involves the prediction of costs in a year using the estimation based on the data previous to that year. One, two and three-year-ahead forecasts are considered. In all predictions the actual values of explanatory variables are used. Prediction errors are defined as the predicted total costs minus the actual costs divided by the actual costs. In the RE model the forecasts are based on the optimal predictor given in Bailli and Baltagi (2000).[30] The prediction errors of different models are compared. The results are summarized in table 6. The two-year-ahead forecast errors are not listed.

[30] See page 256 of Bailli and Baltagi (1999) for more details.

Table 6. Prediction errors

Type of prediction	Out-of-sample			1-year-ahead			3-year-ahead		
Model	OLS	GLS	FE	OLS	GLS	FE	OLS	GLS	FE
Average error (absolute value)	7.37	7.57	11.8	5.91	3.02	3.10	6.24	4.18	4.58
Maximum error (absolute value)	25.2	27.3	39.0	17.0	10.3	10.3	16.7	14.6	16.4
95 percentile error (absolute value)	17.8	20.1	31.1	13.5	6.82	7.66	14.1	10.9	13.0
Average prediction bias (average error)	.34	.98	1.08	-.03	0.08	.24	-1.89	-1.87	-1.39
Number of predictions	380	380	380	52	52	52	52	52	52

- Errors are given in percentage of the actual costs.

As expected, the out-of-sample estimation errors are significantly higher in the fixed-effects model. Interestingly, even the forecasting performance of the GLS model is comparable to that of the FE model. This implies that in our sample the FE model does not provide a significant predictive advantage over other models. The results show that from a practical standpoint, the prediction errors are generally within an acceptable range. The GLS model shows the best performance. As it can be seen in table 6 (see the middle column), the GLS model's one-year-ahead prediction errors with an average absolute value of 3 percent and an average value of less than 0.1 percent are particularly low. While the maximum error in this case is 10.3%, for 95 percent of the companies the prediction error is limited to 6.8%.

These results suggest that the panel data models can predict individual companies' total costs with a rather reasonable precision. Therefore, the regulator can use these models to predict a confidence interval for the costs of each one of the firms. Acceptable intervals for revenue and price caps can be calculated accordingly. Using such predictions along with other monitoring instruments, the regulator can hold the companies within a reasonably well-predicted range of cost-efficiency.

5. CONCLUSION

Four different parametric cost frontier models are applied to a panel data set of electricity distribution utilities in Switzerland. A comparison of the estimation results indicates significant differences among different models. This result can be explained by the strong unobserved heterogeneity among distribution companies, which is a common characteristic of network industries. These differences are particularly important for the estimates of inefficiency scores. While the summary statistics of inefficiency estimates are not sensitive to model specification, the efficiency ranks change quite significantly from one model to another. The alternative models are not found to be 'mutually consistent' with respect to inefficiency measures. These results point to an important shortcoming of the benchmarking methods in network industries. Furthermore, the results confirm that the sensitivity problems reported in the previous literature are not limited to cross sectional data. Given that the regulators actually use some of these methods in practice, the analysis in this paper has an important implication suggesting that benchmarking analysis should be applied with caution. In particular, it is recommended that several models be used and compared. A sensitivity analysis should be performed to identify the limitations of different models.

This paper also uses different cost frontier panel data models to predict the firms' costs. In particular, the out-of-sample prediction performance of different models is studied. The prediction errors are within acceptable limits from a practical point of view. The results suggest that stochastic frontier models can be used to gain information about costs of individual firms. Moreover, certain models with apparent limitations in the estimation of cost-frontier parameters have a relatively good performance in predicting costs. Although this conclusion may be limited to the data used in this paper, the results suggest that cost frontier models can be used as a complementary control instrument in order to narrow the information gap between the regulator and regulated companies. An interesting example is provided by Antonioli and Filippini (2001) in the regulation of water supply in Italy, where a yardstick competition model in line with Schleifer (1985) has been applied. This regulation method is based on an interactive approach: The company proposes its tariff in the first stage. The regulator estimates a price cap for the firm using a benchmarking analysis and adjusting for observed differences among companies. The proposed tariff is approved if it does not exceed an acceptable range around the estimated price cap. Otherwise, the tariffs can be renegotiated with the requirement that the company justify its excessive tariff before any revision. The interaction between the regulator and companies may be helpful in the face of information asymmetry. In order to provide a disincentive to renegotiate a penalty can be imposed on the companies that do not accept the first-stage prices. In case of disagreement the regulator performs a more detailed analysis with additional information from individual companies and offers a new price. The probability of disagreement and the flexibility of the regulator depend on the prediction power of the adopted econometric models.

Acknowledgements

The authors are grateful to Jörg Wild for useful discussions and his help in preparing the data. We benefited from the suggestions of the editor and two anonymous referees, which are gratefully acknowledged. We would also like to thank the participants of the Annual Conference of the International Association for Energy Economics (Prague, June 2003) particularly Tooraj Jamasb and Einar Hope for their helpful comments. The views expressed in this paper are strictly personal and the responsibility for all remaining errors lies solely with the authors.

References

Aigner, D., C. A. K. Lovell and P. Schmidt (1977) 'Formulation and Estimation of Stochastic Frontier Production Function Models', *Journal of Econometrics*, **6**, 21-37.

Antonioli, B., and M. Filippini (2001) 'The Use of a Variable Cost Function in the Regulation of the Italian Water Industry', *Utilities Policy*, **10** (3-4): 181-187.

Baillie, Richard T., and Badi H. Baltagi (1999) 'Prediction From the Regression Model with One-way Error Component' in C. Hsiao, K. Lahiri, L-F. Lee, and M. H. Pesaran, eds., *Analysis of Panels and Limited Dependent Variable Models*, Cambridge: Cambridge University Press, pp. 255-267.

Baltagi, Badi H. (2001) *Econometric Analysis of Panel Data*, Second Edition, Chichester, England: John Wiley and Sons.

Battese, George E. (1992) 'Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics', *Agricultural Economics*, **7**, 185-208.

Battese, George E., and Tim Coelli (1992) 'Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India', *Journal of Productivity Analysis*, **3**, 153-169.

Bauer, Paul W., Allen N. Berger, Gary D. Ferrier, and David B. Humphrey (1998) 'Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods', *Journal of Economics and Business*, **50**, 85-114.

Coelli, Tim, D. S. Prasada Rao, and George E. Battese (1998) *An Introduction to Efficiency and Productivity Analysis*, Norwell, MA: Kluwer Academic Publishers.

Cornes, R. (1992) *Duality and Modern Economics*, Cambridge: Cambridge University Press.

Filippini, Massimo (1998) 'Are Municipal Electricity Distribution Utilities Natural Monopolies?', *Annals of Public and Cooperative Economics*, **2**, 157-174.

Filippini, M., and R. Maggi (1993) 'Efficiency and Regulation in the Case of the Swiss Private Railways', *Journal of regulatory Economics*, **5**, 199-216.

Filippini, Massimo, and Jörg Wild (2001) 'Regional Differences in Electricity Distribution Costs and Their Consequences for Yardstick Regulation of Access Prices', *Energy Economics*, **23**, 477-488.

Foreman-Peck, J., and M. Waterson (1985) 'The Comparative Efficiency of Public and Private Enterprise in Britain: Electricity Generation Between the World Wars', *Economic Journal*, **95**, 83-95.

Gong, B. H., and R. C. Sickles (1989) 'Finite Sample Evidence on the Performance of Stochastic Frontier Models Using Panel Data', *Journal of Productivity Analysis*, **1** (3), 229-261.

Greene, William (forthcoming) ‘Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model’, *Journal of Econometrics*.

Greene, William H. (1980) ‘Maximum Likelihood Estimation of Econometric Frontier Functions’, *Journal of Econometrics*, **13**, 27-56.

Horarce, W. C., and P. Schmidt (1996) ‘Confidence Statements For Efficiency Estimates From Stochastic Frontier Models’ *Journal of Productivity Analysis*, **7**, 257-282.

Hsiao, Cheng, and Baohong Sun (2000) ‘To Pool or Not to Pool Panel Data’, in J. Krishnakumar and E. Ronchetti, eds., *Panel Data Econometrics: Future Directions*, Amsterdam: North Holland Publishing Company, pp. 181-198.

Irastorza, Veronica (2003) ‘Benchmarking for Distribution Utilities: A Problematic Approach to Defining Efficiency’ *The Electricity Journal*, **16** (10), 30-38.

Jamasb, Tooraj, and Michael Pollitt (2003) ‘International Benchmarking and Regulation: An Application to European Electricity Distribution Utilities’ *Energy Policy*, **31** (2003): 1609-1622.

Jamasb, Tooraj, and Michael Pollitt (2001) ‘Benchmarking and Regulation: International Electricity Experience’ *Utilities Policy*, **9** (3), 107-130.

Jensen, Uwe (2000) ‘Is it Efficient to Analyse Efficiency Rankings?’ *Empirical Economics*, **25**, 189-208.

Joskow, P. J., and R. Schmalensee (1986) ‘Incentive Regulation for Electric Utilities’, *Yale Journal of Regulation*, **4** (1), 1-49.

Kumbhakar, Subal C., and C. A. Knox Lovell (2000) *Stochastic Frontier Analysis*, Cambridge: Cambridge University Press.

Pitt, M., and L. Lee (1981), ‘The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry’, *Journal of Development Economics*, **9**, 43-64.

Richmond, J. (1974) ‘Estimating the Efficiency of Production’, *International Economic Review*, **15**, 515-521.

Roberts, M. J. (1986) ‘Economies of Density and Size in the Transmission and Distribution of Electric Power’, *Land Economics*, **62**, 378-387.

Rossi, Martin A., and Christan A. Ruzzier (2000) ‘On the Regulatory Application of Efficiency Measures’ *Utilities Policy*, **9** (2), 81-92.

Salvanes, K.G. and Tjϕota, S. (1994) ‘Productivity Differences in Multiple Output Industries: An Empirical Application to Electricity Distribution’, *The Journal of Productivity Analysis*, **5**, 23-43.

Schleifer, Andrei (1985) ‘A Theory of Yardstick Competition’, *Rand Journal of Economics*, **16** (3): 319-327.

Schmidt, P., and R.E. Sickles, (1984) ‘Production Frontiers and Panel Data’, *Journal of Business and Economic Statistics*, **2**, 367-374.

Simar, Leopold, and P. W. Wilson (2000) ‘Statistical Inference in Nonparametric Frontier Models: The State of the Art’, *Journal of Productivity Analysis*, **13**, 49-78

Simar, Leopold (1992) ‘Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-parametric and Semi-parametric Methods with Bootstapping’, *The Journal of Productivity Analysis*, **3**, 171-203.

Shuttleworth, Graham (2003) ‘Firm-Specific Productive Efficiency: A Response’ *The Electricity Journal*, **16** (3), 42-50.

Street, Andrew (2003) ‘How Much Confidence Should We Place in Efficiency Estimates?’ *Health Economics*, **12** (11): 895-907.

Thompson, H.G. (1997) ‘Cost Efficiency in Power Procurement and Delivery Service in the Electric Utility Industry’, Land Economics, **73**, 287-296.

Tsionas, Efhtymios G. (2002) ‘Stochastic Frontier Models With Random Coefficients’ *Journal of Applied Econometrics*, **17**, 127-147.

Wagstaff, A. (1989) ‘Estimating Efficiency in the Hospital Sector: a Comparison of Three Statistical Cost Frontier Models’, *Applied Economics*, **21**, 659-672.

Wild, Jörg (2001) *Deregulierung und Regulierung der Elektrizitätsverteilung, Eine Mikro-ökonomische Analyse mit Empirischer Anwendung für die Schweiz*, PhD dissertation, Swiss Federal Institute of Technology, Zurich, Switzerland.