Centre for Energy Policy and Economics
Swiss Federal Institutes of Technology

An Analysis of Cost-Efficiency in Swiss Multi-utilities

Mehdi Farsi, Massimo Filippini

CEPE Working Paper No. 60
May 2008

CEPE
Zurichbergstrasse 18 (ZUE E)
CH-8032 Zurich
www.cepe.ethz.ch

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

AN ANALYSIS OF COST EFFICIENCY IN SWISS MULTI-UTILITIES *

Mehdi Farsi Massimo Filippini

mfarsi@ethz.ch mfilippini@ethz.ch

Department of Management, Technology and Economics, ETH Zurich

Zurichbergstr. 18, CH-8032 Zurich, Switzerland

Tel. +41-44-632 06 50, Fax. +41-44-632 10 50

and

Department of Economics, University of Lugano

Abstract

This study presents an empirical analysis of the cost efficiency of a sample of Swiss multi-utilities operating in the distribution of electricity, natural gas and water. The multi-utilities that operate in different sectors are characterized by a strong unobserved heterogeneity. Therefore the measurement of their performance poses an important challenge for the regulators. The purpose of this paper is to study the potential advantages of recently developed panel data stochastic frontier models in the measurement of the level of efficiency for multi-utility companies. These models are estimated for a sample of 34 multi-output utilities operating from 1997 to 2005. The alternative models are compared regarding the cost function slopes and inefficiency estimates. For the inefficiency estimates, the correlation between different models and the effect of econometric specification have been analyzed. The results suggest that the inefficiency estimates are substantially lower when the unobserved firm-specific effects are taken into account.

JEL Classification: C33, D24, L11, L25, L94, L95

Keywords: cost function; efficiency; panel data; multi-output utilities

* This paper is based upon the results of a research project (Farsi and Filippini, 2007) sponsored by Switzerland's State Secretariat for Economic Affairs (SECO). We gratefully acknowledge their financial support but also the extraordinary cooperation of their staff especially Peter Balastèr, throughout the project. We also wish to thank Aurelio Fetz for compiling the data used in this study. The views expressed here do not necessarily reflect the position of any sponsoring agency.

1. Introduction

Along with the recent waves of liberalization and deregulation in network industries throughout Europe, the authorities are increasingly concerned about the productive efficiency of the utilities that, due to their natural monopoly characteristics, are not fully liberalized. In sectors such as power, gas and water distribution, because of the considerable economies of scale, a direct introduction of competition is not optimal. Instead, incentive regulation has been used to ensure (or maximize) the productive efficiency of the locally monopolistic companies. Everywhere in Europe, the traditional regulatory systems are being gradually replaced by incentive regulation schemes. Unlike the traditional contracting systems based on a reasonable rate of return, the incentive contracts are designed to induce incentives for reducing costs and increasing efficiency. Most incentive regulation schemes use benchmarking to evaluate the productive performance of the regulated companies in order to reward/punish them accordingly. Based on their efficiency performance, companies are allowed to retain part of their profits/savings through either differentiated price caps or adjustments in budget or network access fees.

Several OECD countries have already integrated a benchmarking practice in their regulation systems for electricity distribution networks (Farsi, Fetz and Filippini, 2007a; Crouch, 2007). A few countries have also introduced such incentive schemes based on performance in their water industry (Saal et al., 2007; Antonioli and Filippini, 2001).[1] The application of benchmarking methods in the gas sector is probably not as advanced as that observed in the electricity industry. However, the use of incentive schemes based on performance has been proposed in several studies (*cf.* Casarin, 2007). In spite of a relatively common usage in each one of the distribution sectors, the direct application of benchmarking analysis in the regulation of multi-utilities has hardly been explored. This is especially interesting in Switzerland and some other

[1] In Switzerland the distribution utilities are monitored and regulated by cantonal and federal governments. Although Switzerland has not yet implemented any incentive regulation system, the actual debates suggest that the regulators will probably follow similar reforms in the near future.

European countries, where multi-utilities dominate the distribution sectors in electricity, natural gas and water.

To our knowledge there is no reported empirical application of efficiency measurement in the multi-utility sector. This may perhaps be considered in line with arguments in favor of unbundling the multiple-utilities into separate legal entities. In fact, horizontal unbundling is a recurring subject of the public policy debates both in the EU and Switzerland. However, the dominance of multi-utilities in Switzerland is not expected to be affected by the ongoing reforms. According to the observed tendencies in the EU regulatory reforms, the multi-utilities especially those with moderate and small networks (less than 100,000 customers), will remain exempt from unbundling requirements.

Noting the importance of multi-utilities in many countries an important question is whether the benchmarking methods can be applied to multi-utilities as well as single-output distribution utilities. It is often argued that an accounting unbundling is sufficient for applying separate benchmarking analyses to each branch of a multi-utility. However, due to the fact that in certain situations only part of these sectors are regulated with incentive regulation schemes, a company could artificially shift part of the costs to the sector for which the regulation does not foresee a benchmarking process. Similarly, because of the different levels of incentive regulation across various sectors of a single firm, the management might focus their efforts in one sector, thus permit slackness in others. Moreover, extending single-sector benchmarking to multi-utilities requires pooling the data from single-output distributors with the corresponding branches of the multi-utilities. The latter units, benefiting from the economies of scope, are arguably not comparable to specialized firms, thus might bias the efficiency estimates. In these cases, a benchmarking across the entire operation of multi-utilities might be more relevant than separate benchmarking analyses for individual sectors. In many cases, with a mixture of mutli-output utilities and specialized distributors, the two types of analyses can also be combined to assess the potential differences among firms and also across sectors.

The effectiveness of the regulation systems relies upon the accuracy of estimated efficiency levels of individual companies. However, due to a great variety of available methods of efficiency measurement and the observed discrepancy of results across different methods, benchmarking practice requires a methodology to adopt a single model among several legitimate approaches and specifications. This task is particularly complicated in network utilities in which unobserved firm-specific factors might be confounded with inefficiency. Obviously the problem of unobserved heterogeneity is more important in multi-output distributors that operate in several networks, each of which could have different types of cost drivers with specific characteristics.

Unobserved firm-specific heterogeneity can be taken into account with conventional fixed or random effects in a panel data model. In order to distinguish external heterogeneities from cost efficiency, Greene (2005a) proposed a model that integrates an additional stochastic term representing inefficiency in both fixed and random effects models.[2] These models assume that the firm-specific heterogeneity does not change over time but sources of inefficiency vary both across firms and over time. In this paper we use a 'true random-effects' model, which is a random-constant frontier model, obtained by combining a conventional random-effects model with a skewed stochastic term representing inefficiency. The extended model includes separate stochastic terms for latent heterogeneity and inefficiency.

The empirical results reported in the literature obtained from true random effects models suggest that modeling unobserved heterogeneity could significantly decrease the inefficiency estimates.[3] This could lend certain support to the application of benchmarking methods in the regulation of strongly heterogeneous network industries, in which the conventional inefficiency estimates appear to be overstated. Provided that they can sufficiently control for the unobserved heterogeneity across firms, these methods can be used to have a better estimate of cost-inefficiency in the sector or at individual companies.

[2] Kumbhakar (1991) proposed a similar approach using a three-stage estimation procedure. See also Heshmati and Kumbhakar (1994) and Kumbhakar and Hjalmarsson (1995) for two applications

[3] See for instance Greene (2004), Farsi, Filippini and Kuenzle (2005) and Alvarez, Arias and Greene (2004).

The purpose of this paper is to study the potential advantages of these extended models in an application to Switzerland's multi-output utilities. The models are estimated for a sample of 34 companies operating in Switzerland from 1997 to 2005. The alternative models are compared regarding the cost function slopes and inefficiency estimates. For the inefficiency estimates, the correlation between different models and the effect of econometric specification have been analyzed. The results suggest that the inefficiency estimates are substantially lower when the unobserved firm-specific effects are taken into account.

The rest of the paper is organized as follows: Section 2 presents the model specification and the methodology. The data are explained in section 3. Section 4 presents the estimation results and discusses their implications, and section 5 provides the conclusions.

2. Stochastic frontier models for panel data

The methods used for measuring technical, allocative and cost efficiency are commonly referred to as frontier approaches, classified into two main categories of linear programing methods and econometric approaches.[4] The latter group, also known as Stochastic Frontier Analysis (SFA) is easily adaptable to panel data structure and therefore used in this study. In SFA models, first developed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977), the regression residuals are decomposed into a symmetric component representing statistical noise and a skewed term one representing inefficiency.

As opposed to cross-sections, in panel data the repeated observation of the same company over time allows an estimation of unobserved firm-specific factors, which might affect costs but are not under the firm's control. Individual companies operate in different regions with various environmental and network characteristics that are only partially observed. It is crucial for the regulator to disentangle such exogenous heterogeneities from inefficiency estimates. However the distinction between these

[4] Murillo-Zamorano (2004) provides an account of advantages and shortcomings of each group. Other interseting surveys are Coelli et al. (2005), Simar (1992) and Kumbhakar and Lovell (2000).

two unobserved terms requires certain assumptions based on judgment. In early applications of SFA models to panel data (Pitt and Lee, 1981; Schmidt and Sickles, 1984; Battese and Coelli, 1988), the common assumption was that the productive efficiency is a time-invariant characteristic that can be captured by firm-specific effects in a random or fixed effects model.

A general form of a cost frontier based on these models can be written as:

$$ln\ C_{it} = f\,(\boldsymbol{y}_{it}, \boldsymbol{w}_{it}\,) + u_i + v_{it}. \qquad (1),$$

where subscripts i and t denote the firm and the operation year, C is the cost variable usually in logarithms and $\boldsymbol{y}$ and $\boldsymbol{w}$ are respectively vectors of outputs and input factor prices. The time-varying error component v_{it} , typically a normal variable, represents the unobserved heterogeneity and random errors, whereas the time-invariant term u_i is assumed to represent excess costs due to inefficiency. The latter term is considered with different distributions: While Pitt and Lee (1981) adopt a half-normal distribution that is, a normal distribution truncated at zero. Battese and Coelli (1988) extends the model to non-zero truncation points and Schmidt and Sickles (1984) propose two variations in which they relax the distribution assumptions respectively using Generalized Least Squares (GLS) and fixed-effect estimators. In particular, in the latter model, the individual effects u_i can be correlated with the explanatory variables.

In more recent papers the random effects model has been extended to include time-variant inefficiency. Cornwell, Schmidt and Sickles (1990), Kumbhakar (1990), and Battese and Coelli (1992) are the important contributions that consider a time function to account for variation of efficiency. In particular the former paper proposes a flexible function of time with parameters varying among firms. In all these models, however, the unobserved external heterogeneity is suppressed in an *iid* error term across observations. This implies that the cost variations due to factors other than firm's efficiency are randomly assigned to each observation. This could be a restrictive assumption in network industries in which certain external cost drivers specific to environment and/or network complexity remain practically unchanged over fairly long periods of time.

To the extent that environmental factors and network characteristics do not change considerably over time, associating the time-invariant excess costs to external factors rather than inefficiency can be a sensible assumption. On the other hand, improvements in efficiency are usually linked to a dynamic learning process and adaptation to new technologies. Therefore, it can be assumed that inefficiencies are captured by the time-varying excess costs. These assumptions combined with the distribution assumption in line with the original frontier model allow a disentanglement of inefficiencies from firm-specific heterogeneity captured by panel's individual effects.[5]

In fact, the SFA model in its original form (Aigner, Lovell and Schmidt, 1977) can be readily extended to panel data models, by adding a fixed or random effect in the model. Although similar extensions have been proposed by several previous authors,[6] Greene (2005a,b) provides effective numerical solutions for both models with random and fixed effects, which he respectively refers to as "true" fixed and random effects models. Several recent studies such as Greene (2004), Farsi, Filippini and Kuenzle (2005), Alvarez, Arias and Greene (2004) and Tsionas (2002) have followed this line. Some of these models have proved a certain success in a broad range of applications in network industries in that they give more plausible efficiency estimates.[7] These results raise an important question as to what extent the panel-data-adapted models can be used to have a better understanding of the inefficiencies and whether they can provide a reliable basis for benchmarking and incentive regulation systems in industries characterized by strong heterogeneity. This question is especially important in the multi-utility sector, in which the companies operate in multiple networks, entailing several network-specific heterogeneity dimensions.

Greene's (2005a) 'true' cost frontier model can be written as:

[5] There are evidently other feasible econometric specifications that can incorporate these assumptions. A remarkable example is the flexible framework proposed by Sickles (2005).

[6] In particular Kumbhakar (1991) proposed a three-stage estimation procedure to solve the model with time- and firm-specific effects, Polachek and Yoon (1996) estimated a panel data frontier model with firm dummies and Heshmati (1998) used a two-step procedure in a random-effect framework to separate the firm-specific effects from efficiency differences.

[7] See Saal, Parker and Weyman-Jones (2007), Farsi, Filippini and Greene (2006), Farsi, Filippini and Kuenzle (2006) and Farsi, Filippini and Greene (2005) for applications in water distribution, electricity networks, bus transport and railroads respectively.

$$ln\ C_{it} = f\,(\mathbf{y}_{it}, \mathbf{w}_{it}) + \alpha_i + u_{it} + v_{it},. \quad (2)$$

The term (α_i) is a normal *i.i.d.* in random-effects framework, or a constant parameter in fixed-effects approach. u_{it} and v_{it} are respectively a half-normal variable representing inefficiency and a normal random variable that captures the statistical noise. In this study, we used the true random effect model, mainly because the numerical solution of the fixed effects model was cumbersome and did not converge to sensible results for the estimates of inefficiencies and individual intercepts. In order to provide a basis for comparing the results, three other models namely, Pitt and Lee (1981), Battese and Coelli (1992) and a GLS model in line with Schmidt and Sickles (1984) have also been considered. These models will be described in the next section.

3. Data and model specification

The data used in this study includes financial and technical information from a sample of electricity, natural gas and water distribution companies that have operated in Switzerland between 1997 and 2005. The data have been mainly collected from the annual reports. Information on the size of the firm's distribution area is from the "Arealstatistik 2002" published by the Federal Statistical Office and the "Preisüberwacher". The original data set covers about 90 companies covering about 42% of total electricity, 67% of total gas and 22% of total water distribution in Switzerland. That sample includes multi-utility firms as well as specialized companies in electricity, gas and water sectors and several double-output utilities, but excludes companies with more than 10% self-generation of total electricity distribution.

Since the focus of this study is on the horizontal integrated multi-utilities, we focused on a sub-sample of the data used by Farsi, Fetz and Filippini (2008),[8] including-

[8] In that study we analyzed the economies of scope and scale in Swiss multi-utilities using a quadratic cost function without performing a frontier analysis. In contrast with the present study, the estimation of the economies of scope requires data from the integrated multi-utilities as well as specialized distributors. Pooling the data across different types of utilities is not appropriate for a benchmarking analysis that relies on comparing comparable companies.

ing observations from 34 companies. Moreover, as pointed out by Saal and Parker (2006) assuming a similar cost frontier among multi-output companies and specialized utilities is not a realistic assumption and might cause considerable distortion in efficiency estimates and ranking. Because the primary purpose of this analysis is the estimation of cost-efficiency, we did not pool the multi-utilities with specialized companies.

The final sample used in this analysis consists of an unbalanced panel data set including observations from 34 multi-utilities during the nine-year period spanning from 1997 to 2005. The sample represents about 60% of the integrated multi-utilities in Switzerland. According to our estimates based on the available information, the multi-utilities included in the sample cover about half of the national electricity and gas consumption provided by multi-utilities and about a fifth of the water distributed by multi-utilities. Overall, these companies cover approximately 13% of electricity, 38% of gas and 14% of water distribution in the entire country.

The model specification is based on a cost function with three outputs namely, the distributed electricity, gas and water and four input factors that is, labor and capital as well as the electricity and gas inputs. As in Sing (1987) customer density is introduced as a service area characteristic. This variable should capture, at least partially, the cost impact of the heterogeneity of the service area of the companies. In fact, differences in networks and environments influence the production process and therefore the costs. Obviously, the heterogeneity of the service area cannot be summarized into a single variable. However, the available data do not allow for any other environmental or network characteristic that is reasonably independent of the included explanatory variables. Given the risk of multi-collineraity in the translog function, especially in the second-order terms, we preferred to retain a relatively simple specification. Thus, some of these characteristics are inevitably omitted from the cost function specification. As we see later these omitted factors are represented by firm-specific stochastic components in the adopted panel data econometric models.

Assuming that the technology is convex and the firm minimizes cost, the adopted total cost function can be written as:

$$C = C(q^{(1)}, q^{(2)}, q^{(3)}, r, w^{(0)}, w^{(1)}, w^{(2)}, w^{(3)}, D_t), \qquad (3)$$

where C represents total costs; $q^{(1)}$, $q^{(2)}$ and $q^{(3)}$ are respectively the distributed electricity, gas and water during the year, $w^{(0)}, w^{(1)}, w^{(2)}$ and $w^{(3)}$ are respectively the input factor prices for capital and labor services and the purchased electricity and gas; r is the customer density measured by the number of customers divided by the size of the service area measured in square kilometers; and D_t is a vector of year dummies that represent technical change and other year-to-year variations with the first year of the sample (1997) as the omitted category.[9] The technical change is assumed to be neutral with respect to cost minimizing input ratios, that is, it is represented by a cost shift that does not alter the optimal input bundles.

An important implication of the above specification is that the estimated economies of scale are based on the usual assumption (in line with Caves et al., 1981) that any change in the production scale entails a uniform proportional change in all outputs and network characteristics, thus retaining the same ratios in particular the same customer density. This assumption is consistent with many policy applications such as the economic assessment of mergers and acquisitions and the extension of local monopolists to new areas. However, the potential synergies could be understated in other cases such as the assessment of side-by-side competition, where considerable economies might also be achieved by increasing the density, namely the economies of density.[10] Unfortunately, the sample's independent variations in networks and outputs do not seem to be sufficient for a meaningful empirical distinction between the economies of scale and the economies of density. In fact, our preliminary analyses with several alternative specifications particularly, models including the size of the service area and/or the number of customers, indicated certain discrepancy in the signs and statistical significance of output coeffi-

[9] As we will see later our regressions suggest that the time-variation of costs is not linear. These variations can be explained by many unobserved factors (such as changes in labor contracts or seasonal composition of the demand) that change uniformly across companies.

[10] The economies of output (customer) density describe the effects of changes in output (number of customers) keeping all other network characteristics fixed (Caves et al., 1985; 1984). As illustrated in Farsi, Filippini and Kuenzle (2007, 2006), the economies of density are usually greater than the economies of scale.

cients, which can be explained by multicollinearity problems due to the strong correlation of output variables with those characteristics.

The variables for the cost function specification were constructed as follows. Total costs (C) are calculated as the total firm's expenditures in a given year. The outputs $q^{(m)}$ are measured by the total quantity delivered to the customers. The measurement units are GWh for electricity and gas and million cubic meters for water. Input prices are defined as factor expenditures per factor unit. Following Friedlaender and Chiang (1983), we used the residual approach for estimating the capital prices. The residual costs are specified as the company's total costs net of labor expenditures and purchases of electricity and natural gas. Capital price for each network is obtained by dividing the residual costs by the capital stock measured by the network length. The overall capital price ($w^{(0)}$) is then calculated as a weighted average of capital prices for each of the three sectors namely, electricity, natural gas and water. The weights, similar to Fraquelli, Piacenza et al. (2004), are proportional to the share of the residual costs in each sector out of the multi-utlity's total residual costs. Labor price ($w^{(1)}$) is defined as the ratio of annual labor costs to the total number of employees in terms of full time equivalent worker. In a few cases in which the full time equivalent was not available, in order to avoid the underestimation of labor price due to part-time employees, we considered a correction based on the mean labor price values within the same canton. The electricity and gas prices ($w^{(2)}, w^{(3)}$) are defined as the expenditures of purchasing the input factors divided by the amount purchased (in MWh).

Table 1 provides a descriptive summary of the variables included in the model. All the costs and prices are adjusted for inflation using consumer price index and are measured in year 2000 Swiss Francs (CHF). As can be seen in the table, the sample shows a considerable variation in costs and all three outputs.

Table 1: Descriptive statistics (237 observations from 34 companies)

Variable		Min.	Median	Mean	Max.
C	Total costs (CHF Mio.)	11.20	41.10	77.60	503.00
$q^{(1)}$	Electricity distribution (GWh)	38.78	126.89	293.23	2'023.59
$q^{(2)}$	Gas distribution (GWh)	28.82	226.34	512.60	4'294.20
$q^{(3)}$	Water distribution (Mio. m^3)	0.78	2.45	5.28	33.35
r	Customer density (customer/ km^2)	44.35	298.33	387.57	1'554.09
$w^{(0)}$	Capital price (CHF/ km)	11'853	31'167	38'385	234'796
$w^{(1)}$	Labor price (CHF/ employee)	77'789	106'466	107'851	146'816
$w^{(2)}$	Electricity price (CHF/ MWh)	44.6	107.4	105.9	163.5
$w^{(3)}$	Gas price (CHF/ MWh)	16.6	28.4	29.3	63.2

Following Christensen et al. (1973) we use a translog model which is probably the most widely used functional form in empirical studies of cost and production functions.[11] This flexible functional form is a local, second-order approximation to any arbitrary cost function. The approximation point is usually set at the sample mean or median. Here the approximation point has been set at the sample median. Compared to the mean, the median values are less affected by outlier values. The translog form does not impose any restrictions on the elasticity of substitution and allows the economies of scale to vary with the output level. In order to avoid the excessive number of parameters we have considered a homothetic cost function in which the interaction terms between input price variables and output variables are

[11] See Caves et al. (1980) on the advantages of translog form in multiproduct settings and Griffin et al. (1987) for a discussion of the criteria used for the choice of the functional form.

excluded.[12] This brings about another assumption namely that marginal costs particularly cost complementarities and scale elasticities depend only upon the technological characteristics of the production, thus being independent of input prices. This is a valid assumption insofar as the input prices remain in a reasonable range, especially because the potential changes in the shape of the cost function can easily be dominated by other approximations entailed by the functional form.

It is generally assumed that the cost function is the result of cost minimization given input prices and output and should therefore satisfy certain properties. Mainly, this function must be non-decreasing in output and non-decreasing, concave and linearly homogeneous in input prices (Cornes, 1992). We imposed the latter condition by normalization of prices namely, by dividing the costs and all factor prices by one common factor price referred to as numeraire (*cf.* Farsi, Fetz et al., 2007b; Featherstone and Moss, 1994; Jara-Diaz, Martinez-Budria et al., 2003). Here we used the capital price as the numeraire. The remaining conditions can be tested based on the estimation results.

The general econometric specification of the cost function in (3) can be written as:

$$
\begin{aligned}
\ln\left(\frac{C_{it}}{w_{it}^{(0)}}\right) = {} & \sum_m \alpha^m \ln q_{it}^{(m)} + \alpha^r \ln r_{it} + \sum_k \beta^k \ln \frac{w_{it}^{(k)}}{w_{it}^{(0)}} + \frac{1}{2}\sum_m \alpha^{mm}\left(\ln q_{it}^{(m)}\right)^2 \\
& + \sum_{m(m\neq n)} \sum_n \alpha^{mn} \ln q_{it}^{(m)} \ln q_{it}^{(n)} + \frac{1}{2}\alpha^{rr}\left(\ln r_{it}\right)^2 + \sum_m \alpha^{rm} \ln q_{it}^{(m)} \ln r_{it} \\
& + \frac{1}{2}\sum_k \beta^{kk}\left(\ln \frac{w_{it}^{(k)}}{w_{it}^{(0)}}\right)^2 + \sum_{k(k\neq l)}\sum_l \beta^{kl} \ln \frac{w_{it}^{(k)}}{w_{it}^{(0)}} \ln \frac{w_{it}^{(l)}}{w_{it}^{(0)}} \\
& + \sum_t \delta^t D_t + \alpha^0 + \alpha_i + u_{it} + v_{it} \ ,
\end{aligned} \tag{4}
$$

[12] We evaluated the possibility of applying a non-homothetic translog form. However, the relatively large number of parameters created certain numerical problems in some of the econometric models, especially the true random effects model that requires a simulated likelihood maximization method. This is perhaps related to problems due to the model's over-identification and perhaps multicollinearity as suggested by the lack of significance and counter-intuitive signs for some of the main variables.

where subscripts i and t denote the company and year respectively; the parameters $\alpha^m, \beta^k, \alpha^{mn}, \beta^{kl}, \delta^t$ and α^0 ($m,n,k,l = 1,2,3; t = 1998,...,2005$) are the regression coefficients to be estimated; and all second-order parameters α^{nm} and β^{kl}, satisfy the symmetry conditions ($\beta^{kl} = \beta^{lk}; \alpha^{mn} = \alpha^{nm}$); α_i is a firm-specific effect; u_{it} is an asymmetric stochastic component term that captures the time-variant inefficiency and v_{it} is a symmetric term representing random noise and statistical errors.

We consider four variations of the above model. These models are summarized in Table 2. The first model (Model *I*) is a random effects model in line with Schmidt and Sickles (1984). The model is estimated using the Generalized Least Squares (GLS) method. The specification includes a firm-specific random effect α_i, and a random noise term v_{it}, which are both assumed to be identically and independently distributed (*iid*) with any arbitrary distribution. In this model, the inefficiency is assumed to be constant over time, namely the term u_{it} in Equation (4) is set equal to zero. A given company i's inefficiency is considered as the difference between its estimated random effect α_i and that of the firm with the "best performance" namely, the minimum estimated random effect ($\min\{\alpha_i\}$).

The GLS model benefits from certain robustness in that no specific distribution assumption is imposed, except for the usual assumption that the random terms are uncorrelated with the explanatory variables. However, the very construction of this model implies that companies are compared to a single, fully efficient firm that has the lowest observed costs after adjusting for explanatory variables and allowing for random noise. This could be an unrealistic assumption that only one company is completely efficient. Moreover, there is always a probability of wrong identification of a single "best" company because of some firm-specific unobserved factor, in which case the efficiency estimates will be completely distorted. The advantage of imposing a distribution assumption on efficiency attenuates at least partly such seriously misleading outcomes. A commonly used distribution in the literature is the half-normal distribution which is obtained by a zero-mean normal distribution truncated at zero. This distribution assumption that dates back to the original frontier models (Aigner et al., 1977; Meeusen and van der Broek, 1997), implies that full efficiency is the most frequent outcome located at the mode of the distribution.

Table 2: Econometric specifications of the stochastic cost frontier

Stochastic term	*Model I* **GLS** **(Schmidt-Sickles)**	*Model II* **ML** **(Pitt-Lee)**	*Model II* **ML** **(Battese-Coelli)**	*Model IV* **True RE** **(Greene)**
Firm-specific effect α_i	$\alpha_i \sim iid\ (0, \sigma_\alpha^2)$	$\alpha_i \sim N^+(0, \sigma_\alpha^2)$	0	$\alpha_i \sim N(0, \sigma_\alpha^2)$
Time-varying inefficiency u_{it}	0	0	$u_{it} =$ $u_i \exp\{-\eta(t\text{-}T)\}$ $u_i \sim N^+(0, \sigma_u^2)$	$u_{it} \sim N^+(0, \sigma_u^2)$
Random noise v_{it}	$v_{it} \sim iid\ (0, \sigma_v^2)$	$v_{it} \sim N\ (0, \sigma_v^2)$	$v_{it} \sim N(0, \sigma_v^2)$	$v_{it} \sim N(0, \sigma_v^2)$
Inefficiency estimate	$\hat{\alpha}_i - \min\{\hat{\alpha}_i\}$	$E\left[\alpha_i \mid \hat{\omega}_{i1}, \hat{\omega}_{i2}, \ldots\right]$ with $\omega_{it} = \alpha_i + v_{it}$	$E\left[u_{it} \mid \hat{\varepsilon}_{it}\right]$ with $\varepsilon_{it} = u_{it} + v_{it}$	$E\left[u_{it} \mid \hat{r}_{it}\right]$ with $r_{it} = \alpha_i + u_{it} + v_{it}$

The half-normal distribution not only provides a relatively solid benchmark performance observed in a relatively large number of cases, it is also more consistent with the economic theory. In fact the half-normal distribution implies that higher levels of inefficiency have lower incidence. This is aligned with the theory that predicts the prevalence of rational and cost-minimizing behavior and considers the non-optimal performance as sporadic and rare outcomes. Following this assumption in the other three models, we assume a half-normal distribution for inefficiency.

Model *II* is a random effects model proposed by Pitt and Lee (1981). Similar to the first model, the efficiency is assumed to be constant over time (u_{it}=0). As opposed to Model *I* that does not impose any distribution, here the stochastic terms are assumed to follow a composite normal-half-normal distribution: The firm-specific effect α_i that represents (time-invariant) inefficiency, follows a half-normal distribution, and the random noise v_{it} is simply a normal variable with zero mean. This model is estimated using the maximum likelihood method. In line with Kumbhakar and Lovell (2000) we will refer to this model as the maximum likelihood (ML) model. The firm's inefficiency is estimated using the conditional mean of the inefficiency term proposed by Jondrow et al. (1982),[13] that is: $\mathrm{E}\left[\alpha_i \left|\hat{\omega}_{i1}, \hat{\omega}_{i2}, ..., \hat{\omega}_{iT}\right.\right] = \mathrm{E}\left[\alpha_i \left|\bar{\omega}_i\right.\right]$, where the hat symbol ^ is used to indicate the post-estimation predicted value; $\omega_{it} = \alpha_i + v_{it}$; and $\bar{\omega}_i = \frac{1}{T}\sum_{t=1}^{T}\hat{\omega}_{it}$.

The assumption of the firm's inefficiency being constant over time can be relaxed by assuming a parametric form. A commonly used functional form is the exponential decay function proposed by Battese and Coelli (1992). Model *III* is based on one of the specifications proposed by those authors. In this model the inefficiency is defined as $u_{it}=u_i\exp\{-\eta(t\text{-}T)\}$, where u_i is a firm-specific stochastic term, T is the end period and η is a positive constant to be estimated. The adopted functional form implies that a given company i starts with an initial level of inefficiency of $u_{i0}=u_i\exp(\eta T)$, that declines over time with an exponential rate of $\exp(-\eta)$ per period, reaching $u_{iT}=u_i$ at the end of the sample period.[14] This specification, while recognizing individual differences in efficiency, assumes a similar improvement rate for all companies. The firm-specific heterogeneity term α_i in Equation (4), is set equal to zero.[15] This model is also estimated using the maximum likelihood method. The firm's inefficiency is estimated using the conditional mean of the inefficiency term, namely: $\mathrm{E}\left[u_{it} \left|\varepsilon_{it}\right.\right] = \mathrm{E}\left[u_i \left|\hat{\varepsilon}_{i1}, \hat{\varepsilon}_{i2}, ..., \hat{\varepsilon}_{iT}\right.\right]\exp\{-\eta(t-T)\}$, where $\varepsilon_{it}= u_{it}+ v_{it}$.

In both models *I* and *II*, it is assumed that all the unobserved differences across firms that do not vary over time are related to inefficiency. Model *III* relaxes the time-invariance by imposing a deterministic form of evolution that is uniform among all companies. In all three models, all the unobserved differences that cannot be captured by the random noise (v_{it}) are assumed to be due to inefficiency. As we have seen in

[13] See also Greene (2005a).

[14] Note that a more general notation T_i is usually used for the end of sample period (T) that can be specific to company. Here we dropped the subscript for simplicity.

[15] Battese and Coelli (1992, 1995) have proposed variations of this model with different distributions for u_i, including truncated normal distribution. In this study we assume a half-normal distribution.

the previous section this could be a restrictive assumption in network industries especially in multi-utilities, which might entail a considerable cost variation through unobserved factors that vary from one network to another but are more or less constant over time and cannot be changed by the management. This implies that in these cases some of the unobserved heterogeneity, for instance, the complexity of the distribution network that is mainly determined by the topology of the service area, can be identified as inefficiency.

Model *IV* allows for a separate stochastic term that captures the time-invariant unobserved heterogeneity. This model is the 'true random effects' frontier specification proposed by Greene (2005a,b), which extends the original frontier model (Aigner et al., 1977) in a panel data framework with random effects. The stochastic components α_i, u_{it} and v_{it} respectively represent the firm-specific random effect, inefficiency term and random noise: $\alpha_{\mathrm{i}} \sim N(0,\sigma_\alpha^2), v_{it} \sim N(0,\sigma_v^2)$ and $u_{it} \sim N^+(0,\sigma_u^2)$. This model is estimated using Simulated Maximum Likelihood (SML) method. We use quasi-random Halton draws to minimize the potential sensitivity of the results to simulation process. Number of draws has been fixed to 1000. Our sensitivity analysis using several options suggested that the estimation results are not sensitive when the number of draws is higher than a few hundred. The inefficiency is estimated using the (simulated) conditional mean of the inefficiency term (u_{it}) given by $\mathrm{E}\left[u_{it} \middle| \hat{r}_{it}\right]$, where $r_{it}= \alpha_i+u_{it}+v_{it}$ is the regression residual. The above conditional expectation is also calculated by Monte Carlo simulations.[16]

With two heterogeneity terms, Model *IV* is expected to provide a better distinction between inefficiency and other unexplained variations. This advantage is especially important in network industries, in which a significant part of unobserved differences is due to time-invariant factors. All the adopted models assume that the stochastic terms namely, cost-efficiency and unobserved heterogeneity are independent from each other and are both uncorrelated with the explanatory variables included in

[16] See Greene (2005b) for more details. A general discussion of the SML estimation method is also provided by Greene (2007).

the model. There are several methods to relax these assumptions. For instance the correlation between firm-specific effects and explanatory variables can be allowed by Mundlak's specification (Farsi, Filippini and Greene, 2005; Farsi, Filippini and Kuenzle; 2005) or the impact of explanatory variables on efficiency can be modeled by specifying the truncation point of the normal distribution as a function of observed factors (Kumbhakar et al., 1991; Battese and Coelli, 1995) or as a general functional form (Wang and Schmidt, 2002). However, such elaborations can only be achieved through more complicated and often arbitrary assumptions that might compromise the clarity of the original assumptions and make the interpretations more difficult. Moreover, including explanatory variables in several forms in the model specification could cause over-identification and multi-collinearity issues. Such problems could bias the estimated coefficients or lower their accuracy, and eventually cause misleading estimates of cost-efficiency as well as technological characteristics such as the economies of scale. Finally, most of these "refinements" cannot be combined with the true random effects model that provides an already rich structure of the stochastic terms.

4. Empirical results

Table 3 lists the regression results of the cost frontier analysis, using the four alternative models as presented in Equation (4) and Table 2. The estimated coefficients of the first-order terms generally have the expected signs and are statistically significant across all models. Given that all the variables except the dummy variables are in logarithmic form, these coefficients can be directly interpreted as elasticities. The coefficients of first-order output variables represent the cost elasticities with respect to the corresponding outputs at the sample median. These coefficients indicate that the marginal costs of electricity distribution are considerably higher than those of natural gas, which in turn are substantially greater than those of water distribution.

Table 3: Estimation results

	Model I		*Model II*		*Model III*		*Model IV*	
	GLS (Schmidt-Sickles)		**ML (Pitt-Lee)**		**ML (Battese-Coelli)**		**True RE (Greene)**	
α^1 (Electricity output)	0.505 **	(.053)	0.460 **	(.069)	0.418 **	(.063)	0.527 **	(.020)
α^2 (Gas output)	0.317 **	(.032)	0.298 **	(.041)	0.245 **	(.045)	0.258 **	(.012)
α^3 (Water output)	0.092 **	(.039)	0.178 **	(.053)	0.212 **	(.047)	0.146 **	(.015)
α^r (Customer density	0.064 **	(.027)	0.043	(.038)	0.026	(.037)	0.007	(.009)
β^1 (Labor price)	0.242 **	(.057)	0.229 **	(.054)	0.236 **	(.058)	0.201 **	(.027)
β^2 (Electricity price	0.326 **	(.059)	0.317 **	(.051)	0.333 **	(.052)	0.370 **	(.033)
β^3 (Gas price)	0.234 **	(.043)	0.243 **	(.039)	0.223 **	(.038)	0.215 **	(.024)
α^{11}	0.646 **	(.197)	0.368 *	(.221)	0.218	(.193)	0.231 **	(.086)
α^{22}	0.234 **	(.055)	0.154 *	(.080)	0.067	(.071)	0.093 **	(.023)
α^{33}	0.287 **	(.141)	0.042	(.176)	0.186	(.167)	0.089 *	(.052)
α^{rr}	0.019	(.061)	-0.063	(.095)	-0.233 **	(.089)	-0.146 **	(.026)
α^{12}	-0.273 **	(.086)	-0.182 *	(.105)	-0.048	(.091)	-0.099 **	(.041)
α^{13}	-0.327 **	(.149)	-0.124	(.158)	-0.214	(.148)	-0.133 **	(.058)
α^{1r}	-0.215 **	(.070)	-0.220 **	(.097)	0.074	(.104)	-0.119 **	(.030)
α^{23}	-0.002	(.059)	0.049	(.072)	0.051	(.068)	0.037	(.026)
α^{2r}	0.123 **	(.059)	-0.002	(.079)	-0.147 *	(.080)	-0.065 **	(.027)
α^{3r}	0.085 *	(.050)	0.120	(.081)	0.104	(.076)	0.122 **	(.020)
β^{11}	0.419	(.279)	-0.031	(.270)	0.051	(.248)	0.384 **	(.121)
β^{22}	0.695 **	(.205)	0.524 **	(.172)	0.565 **	(.167)	0.758 **	(.110)
β^{33}	-0.243 **	(.120)	-0.291 **	(.106)	-0.278 **	(.110)	-0.217 **	(.108)
β^{12}	-0.701 **	(.221)	-0.419 **	(.197)	-0.460 **	(.189)	-0.724 **	(.102)
β^{13}	0.294 **	(.147)	0.422 **	(.137)	0.386 **	(.136)	0.351 **	(.096)
β^{23}	-0.096	(.135)	-0.154	(.118)	-0.156	(.115)	-0.136	(.092)
δ^{1998}	-0.004	(.019)	-0.005	(.015)	0.011	(.016)	-0.005	(.032)
δ^{1999}	-0.003	(.020)	-0.002	(.016)	0.028	(.019)	-0.005	(.021)
δ^{2000}	-0.015	(.021)	-0.013	(.018)	0.035	(.024)	-0.006	(.025)
δ^{2001}	-0.014	(.023)	-0.015	(.020)	0.049 *	(.029)	-0.012	(.022)
δ^{2002}	-0.037 *	(.021)	-0.036 **	(.018)	0.036	(.030)	-0.040 *	(.022)
δ^{2003}	-0.041 *	(.021)	-0.044 **	(.018)	0.039	(.033)	-0.039 *	(.023)
δ^{2004}	-0.064 **	(.023)	-0.069 **	(.020)	0.032	(.038)	-0.067 **	(.024)
δ^{2005}	-0.059 **	(.026)	-0.065 **	(.023)	0.046	(.043)	-0.073 **	(.022)
α^0	7.164 **	(.029)	6.989 **	(.032)	6.917 **	(.046)	7.120 **	(.019)
σ_α	.053		0.217 **	(.034)			0.114 **	(.005)
σ_u					0.210 **	(.039)	0.081 **	(.030)
σ_v	.054		0.054 **	(.003)	0.052 **	(.003)	0.024 **	(.006)
η					0.048 **	(.015)		
logL	Not Applicable (R^2=0.982)		296.785		299.355		303.786	

** and * refer to 5% and 10% significance levels respectively. Standard errors are given in parentheses.

Approximately, the results suggest that by adding electricity output by 10 percent, the total costs will increase by about 5 percent on average, but the same relative increase in other outputs will raise the company's total costs by about 2.5 to 3 percent for gas and only about 0.9 to 2 percent for water output. These predictions vary slightly across different models. Many of the second-order terms are also statistically significant, implying that the assumption of constant elasticities is unrealistic. The coefficients of the squared output terms (α^{11}, α^{22}, α^{33}) are positive and mostly significant across all models. This suggests that a marginal increase in a given output increases the cost elasticity of that output. Therefore, as expected, the (product-specific) economies of scale are decreasing in output.

As we see in Table 3 the output cross-interaction terms (α^{12}, α^{13}, α^{23}) are mostly negative across the models. In particular, the cross effect between electricity and other two outputs (natural gas and water), is statistically significant. This suggests that the multi-utilities with higher electricity output have a relatively low marginal cost for distributing water and gas. This cost complementarity also applies to companies with high gas or water output, which according to the estimation results, have lower marginal cost for electricity output. The results show however that the cost complementarity between gas and water outputs (as shown by coefficient α^{23}) is not statistically significant. If we interpret this is a zero effect, this result suggests that the marginal cost of distributing gas (water) is not related to the volume of water (gas) output. This is a weak form of cost complementarity, implying that the marginal costs of one output will not increase in the amount of the other output.

As for the effect of customer density, the results show that the first order term is positive but statistically insignificant in most models. This suggests that the effect at the median company is probably not important. However, the mostly negative coefficient of the square term (α^{rr}) suggests that higher densities could have a decreasing effect on costs. At first impression, this can be considered as counter-intuitive because increasing the customer density may be economical in low-density areas, but could create extra costs in congested areas. However, the statistically significant interaction terms between customer density and outputs, suggest that the density has a strongly non-linear effect depending on the output combination across the three services.

For instance the interaction term with electricity output (α^{1r}) is mostly negative and significant, suggesting that the marginal cost of electricity output is lower in networks with higher customer density. This cannot be said for gas and water outputs. Especially the corresponding interaction term for water distribution (α^{3r}) is mostly on the positive side, suggesting that an increase in customer density will increase the marginal cost of water distribution. These results can be related to different costs of network connection for various outputs, and also different amount of extra cables and pipes required for the provision of greater volumes of electricity, gas and water, depending on the actual customer density. For instance, in a dense and crowded area providing more electricity might be handled easier than a considerable increase in gas and water output. Moreover, connection of new customers to electricity networks is probably less costly than that of water and gas distribution networks.

The coefficients of the first-order terms of input prices are an indicator of the share of each factor price at the sample median.[17] Based on the regression results, the shares of labor, electricity and gas inputs respectively amount to about 22, 33 and 23 percent of the total costs. These numbers are comparable to the sample mean of the observed factor shares which is 12, 35 and 17 percent of the company's total costs, respectively for labor, electricity and gas inputs. As we see the share of electricity and gas expenses are quite close the average observed values. The remaining costs have been considered as 'capital' costs that are 36 percent on average, but about 22 percent from the regression results. Therefore in the model, the share of labor costs is overstated compared to that of the residual capital costs.

We explored if the estimated cost functions satisfy the theoretical properties implied by cost-minimization. As shown by the positive coefficients of the first order terms (Table 3), all the estimated cost functions are non-decreasing in output and input prices at the approximation point (sample median). However, our calculations showed that the Hessian matrix defined by the second derivatives of the translog cost

[17] Note that in translog form, any statement about sample points other than the approximation point (here, sample median), should consider the second-order terms in addition to the main effects.

function with respect to log of input prices, is not negative semi-definite. The violation of this necessary condition[18] for concavity might be considered as an indication that the concavity in input prices is not satisfied. This result can be explained by the fact that the multi-utilities are probably not as sensitive to price changes as the textbook economic theory might predict. Theoretically the companies are expected to substitute labor with capital or capital with energy in response to changes in the relative prices. However, in practice these substitutions are not feasible in many cases. For instance if the relative price of electricity to gas increases, the companies cannot substitute electricity input with gas input, because these inputs are mainly determined by the demand side.

In any case, even if we consider the lack of concavity in input prices as an indication that the companies do not fully minimize their costs the estimated cost functions can be useful to study the marginal effects of different factors on costs and also to compare the companies' performance. In such cases, as pointed out by Bös (1986) and Breyer (1987), functions based on cost optimization can still be used as 'behavioral' cost functions and can be helpful in studying the firms' behavior. Moreover, we should keep in mind that we are estimating a cost frontier function, which allow the possibility that some companies do not minimize their costs.

Cost efficiency

The estimates of inefficiency scores obtained from the four models are summarized in Table 4. As expected, compared to all other models, the True RE model's estimates provide generally lower inefficiency. According to this model the multi-utilities have on average about 6 percent excess costs compared to the fully efficient production whereas the other models predict from 18 to 21 percent excess cost on average. The median inefficiency for the True RE model is about 5%, while being about

[18] As pointed out by Diewert and Wales (1987), even with a negative semi-definite Hessian matrix for the translog cost function, the costs might be concave with respect to input prices. So applying such a condition on the coefficient matrix of a translog cost function is too strong for concavity in input prices.

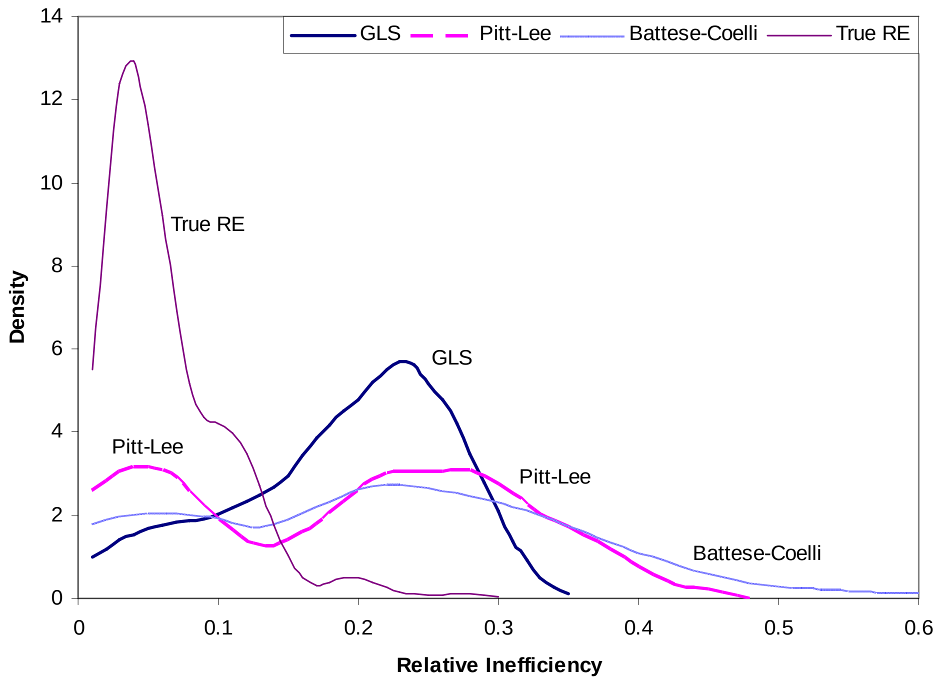
20% for all other models. It should be noted that the True RE model's estimates do not include the persistent inefficiencies that might remain more or less constant over time. To the extent that there are certain sources of inefficiency that result in time-invariant excess costs, the estimates of the True RE model should provide a reasonable lower bound for the companies' inefficiency. On the other hand, in all the three other models, it is assumed that all the time-invariant cost differences due to exogenous heterogeneity are accounted for by the observed explanatory variables included in the model, and whatever remains can be interpreted as inefficiency. Therefore, the overall estimates of inefficiency obtained from these models can be considered as a kind of upper bound for the actual level of inefficiency in the sector.

Table 4: Descriptive summary of inefficiency estimates

	Model I **GLS (Schmidt-Sickles)**	*Model II* **ML (Pitt-Lee)**	*Model III* **ML (Battese-Coelli)**	*Model IV* **True RE (Greene)**
Mean	0.184	0.183	0.216	0.063
Std. Deviation	0.079	0.119	0.143	0.043
Minimum	0.000	0.013	0.014	0.010
1st Quartile	0.144	0.060	0.075	0.031
Median	0.202	0.207	0.214	0.050
3rd Quartile	0.251	0.275	0.303	0.082
Maximum	0.303	0.401	0.699	0.277

The distribution of the inefficiency estimates in the sample is depicted in Figure 1. The distribution densities have been smoothed using Kernel density method. As seen in the figure the extent of inefficiency in the True RE model is considerably narrower than in other models. Moreover, the distribution of the GLS estimates suggest a negative skewness, which contradicts the usual assumption of positive skewness in cost-inefficiencies. Moreover, both Models *II* and *III* indicate a tendency toward a bimodal distribution, which goes against the underlying half-normal distribution assumption in these models. These peculiar patterns might be indicative that the econometric specification of the error term in the first three models could be insufficient to capture the inefficiencies in a coherent way. This can be explained by unobserved cost differences that are not due to inefficiency but to other external factors.

Figure 1: Distribution of inefficiency estimates



In order to explore if the efficiency estimates provide a consistent ranking pattern across different modes, we studied the correlation coefficients between these estimates. Table 5 provides the correlation matrix of inefficiency scores across the four models. The results suggest a high positive correlation among the first three models. There is however a relatively low correlation between each one of these models and the True RE model. The Spearman rank correlation matrix shows slightly lower correlation in general but confirms the above pattern namely low correlation between Model *IV* and the other three models, and high correlation among the latter models. This result suggests that even if we are only interested in efficiency ranking rather than the numerical level of inefficiency, using the inadequate model can give a misleading ordering of individual companies.

Table 5: Pearson correlation matrix between inefficiency estimates

	Model I GLS (Schmidt-Sickles)	*Model II* ML (Pitt-Lee)	*Model III* ML (Battese-Coelli)	*Model IV* True RE (Greene)
I	1	0.863**	0.715**	0.124*
II		1	0.793**	0.140**
III			1	0.128**

** and * refer to 5% and 10% significance levels respectively.

5. Conclusions

This study presents an empirical analysis of cost inefficiency in a sample of Swiss multi-utilities operating in the distribution of electricity, natural gas and water. The issues addressed in the study involve an important question related to the application of benchmarking analysis in incentive regulation schemes for multi-utilities. In general, the benchmarking of multiple-output companies is more complicated than in utilities with a similar output. Multi-utilities that operate in several different sectors, are characterized by a strong unobserved heterogeneity making the measurement of their performance an important challenge for the regulators.

It is shown that the recent methodological developments in the estimation of cost frontier functions using panel data methods can be helpful to achieve more reliable estimates of inefficiency in presence of unobserved and omitted factors. The previous studies have used some of these methods in single-network distributors such as electricity and gas. However to our knowledge there is no reported empirical application in the multi-utility sector. The present analysis serves as a first illustration of the difficulties involved in the estimation of efficiency in multi-network utilities.

Using a translog cost function and several stochastic frontier models this analysis indicates the presence of unexploited global scale economies in the majority of the companies included in the sample. The efficiency estimates are sensitive to the econometric specification of unobserved factors through the model's stochastic components. While highlighting the potential problems in benchmarking multi-utilities, this study shows that adequate panel data models can be used to identify the inefficient companies and determine to certain extent, which part of their excess costs has been persistent and which part has varied over time.

Combining several frontier models also allows two types of inefficiency estimates: a "lower bound" estimate that includes only the transient part of the firm's excess costs assuming that all persistent cost differences are due to unobserved factors rather than poor efficiency performance, and an "upper bound" that associates all the firm-specific unaccounted cost differences to their productive efficiency and neglects the effect of external unobserved factors. Both estimates could be useful for the regulator, as they can use them to identify the companies that are persistently more costly than others and those that have high time-variant inefficiency. The regulator should perform further detailed and possibly case-by-case studies to assess to what extent the excessive costs of the former group can be associated with productive inefficiency and identify the potential external factors and peculiarities that might have caused such excessive costs.

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.

Alvarez, A., C. Arias, and W. H. Greene (2004) 'Accounting for unobservables in production models: management and inefficiency.' Working Paper No. E2004/72, Centro de Estudios Andaluces, Spain.

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.

Battese, G.E. and T. Coelli (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Econometrics*, 20: 325-332.

Battese, George E. and Tim J. Coelli (1992). "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India", *Journal of Productivity Analysis*, 3 (1): 153-169.

Battese, George E. and Tim J. Coelli (1988). "Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data", *Journal of Econometrics*, 38: 387-399.

Bös, D. (1986). *Public Enterprise Economics*, Amsterdam, North-Holland.

Breyer, Friedrich (1987). "The specification of a hospital cost function: A comment on the recent literature", *Journal of Health Economics*, 6 (1987): 147-157.

Casarin, A. A. (2007). 'Efficient industry configurations in downstream gas markets. An empirical assessment', *Energy Economics* 29: 312–328.

Caves, Douglas W., L. R. Christensen, M. W. Tretheway, and R. J. Windle (1985) 'Network effects and the measurement of returns to scale and density for U.S. railroads', in *Analytical Studies in Transpot Economics*, editor. A. F. Daugherty, Cambridge University Press, pp. 97-120.

Caves, D. W., L. R. Christensen, and M. W. Tretheway (1984): "Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ". *Rand Journal of Economics*, 15 (4): 471–489.

Caves, W. C., L. R. Christensen and J. A. Swanson (1981). "Productivity growth, scale economies, and capacity utilization in U.S. railroads, 1955-74", *American Economic Review*, 71 (5): 994-1002.

Caves, D. W., L. R. Christensen and M. W. Tretheway (1980). "Flexible Cost Functions for Multiproduct Firms." *Review of Economics and Statistics* 62(3): 477-81.

Christensen, L. R., D. W. Jorgenson, L. J. Lau (1973): 'Transcendental Logarithmic Production Frontiers', *Review of Economics and Statistics*, 55(1): 28-45.

Coelli, T., D. S. P. Rao, and G. E. Battese (2005) *An Introduction to Efficiency and Productivity Analysis*, 2nd edition, Springer Verlag.

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

Crouch, M. (2007). 'Investment under RPI-X: Practical experience with an incentive compatible approach in the GB electricity distribution sector', *Utilities Policy*, 14: 240-244.

Diewert, W. E. and T. J. Wales (1987): 'Flexible Functional Forms and Global Curvature Conditions', *Econometrica*, 55 (1): 43-68.

Farsi, M., A. Fetz and M. Filippini (2008): 'Economies of scale and scope in multi-utilities', *Energy Journal*, forthcoming.

Farsi, M., A. Fetz and M. Filippini (2007a): 'Benchmarking and Regulation in the Electricity Distribution Sector', In Servizi Pubblici: Nuovo tendenze nella regolametazione nella produzione et nel finanziamento (M. Marrelli, F. Padovano and I. Rizzo, Eds), Franco Angeli, Milano, pp. 159-176.

Farsi, M., A. Fetz and M. Filippini (2007b). "Economies of scale and scope in local public transportation." *Journal of Transport Economics and Policy*, 41 (3): 345-361.

Farsi, M. and M. Filippini (2007). *Cost Efficiency and Scope Economies in Multi-output Utilities in Switzerland*, final report submitted to the State Secretariat for Economic Affairs (SECO), Bern, Switzerand.

Farsi, M., M. Filippini and W. H. Greene (2006), Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector, *Annals of Public and Cooperative Economics*, 77 (3): 271-290.

Farsi, M., M. Filippini and W. Greene (2005). 'Efficiency Measurement in Network Industries: Application to the Swiss Railway Companies', *Journal of Regulatory Economics*, 28 (1): 65-86.

Farsi, M., M. Filippini and M. Kuenzle (2007). 'Cost Efficiency in the Swiss Gas Distribution Sector', *Energy Economics*, 29 (1): 64-78.

Farsi, M., M. Filippini and M. Kuenzle (2006): 'Cost efficiency in regional bus companies: An application of new stochastic frontier models', *Journal of Transport Economics and Policy*, 40(1): 95-118.

Farsi, M., M. Filippini and M. Kuenzle (2005). "Unobserved Heterogeneity in Stochastic Cost Frontier Models: An Application to Swiss Nursing Homes", *Applied Economics*, 37, 2127-2141.

Featherstone, A. M. and C. B. Moss (1994). "Measuring Economies of Scale and Scope in Agricultural Banking." *American Journal of Agricultural Economics* 76(3): 655-61.

Fraquelli, G., M. Piacenza and D. Vannoni (2004). "Scope and Scale Economies in Multi-Utilities: Evidence from Gas, Water and Electricity Combinations." *Applied Economics* 36(18): 2045-57.

Friedlaender, A. F. and W. Chiang (1983). "Productivity Growth in the Regulated Trucking Industry." *Research in Transportation and Economics* 1: 149-84.

Greene, William H. (2007): *Econometric Analysis*, 6th edition, Prentice Hall, Upper Saddle River, New Jersey.

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

Greene, William (2005b). "Fixed and Random Effects in Stochastic Frontier Models", *Journal of Productivity Analysis*, 23 (1): 7-32.

Greene, William (2004). "Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems", *Health Economics* 13: 959-980.

Griffin, Ronald C., J. M. Montgomery and M. E. Rister (1987). "Selecting Functional Form in Production Function Analysis", *Western Journal of Agricultural Economics*, December 1987, pp. 216-227.

Heshmati, Almas (1998). "Efficiency measurement in rotating panel data". *Applied Economics*, 30 (7): 919-930.

Jara-Diaz, S. R., E. Martinez-Budria and F. J. Ramos-Real (2003). "Adapting Productivity Theory to the Quadratic Cost Function: An Application to the Spanish Electric Sector." *Journal of Productivity Analysis* 20(2): 213-29.

Jondrow, J., I. Materov, K. Lovell, and P. Schmidt (1982): 'On the estimation of technical inefficiency in the stochastic frontier production function model' *Journal of Econometrics*, 19, 2/3 (August), 233-238.

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

Kumbhakar, S. C. (1991). "Estimation of technical inefficiency in panel data models with firm- and time-specific effects". *Economics Letters*, 36 (1991): 43-48.

Kumbhakar, S. C. (1990). 'Production frontiers and panel data, and time varying technical inefficiency'. *Journal of Econometrics*, 46(1990): 201-211.

Meeusen, Wim, and Julien van den Broeck (1977). 'Efficiency Estimation From Cobb-Douglas Production Functions With Composed Error', *International Economic Review*, **18** (2), 435-444.

Murillo-Zamorano, Luis R. (2004), Economic Efficiency and Frontier Techniques, *Journal of Economic Surveys*, 18 (1): 33-77.

Pitt, M. and L. Lee (1981). "The measurement and sources of technical inefficiency in Indonesian weaving industry", *Journal of Development Economics*, 9: 43-64.

Polachek, S. and B. Yoon (1996). "Panel estimates of a two-tiered earnings frontier" *Journal of Applied Econometrics*, 11, 169-178.

Saal, D. S., D. Parker and T. Weyman-Jones (2007). ' Determining the contribution of technical change, efficiency change and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000', *Journal of Productivity Analysis*, 28:127–139.

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

Sickles RC. 2005. Panel estimators and the identification of firm-specific efficiency levels in parametric, semiparametric and nonparametric settings. *J Econom* **126** (2): 305-334.

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

Sing, M. (1987). "Are combination Gas and Electric Utilities multiproduct natural Monopolies?" *Review of Economics and Statistics* 69(3): 392-8.

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

Wang, H. and P. Schmidt (2002): "One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels". *Journal of Productivity Analysis*, 18, 129–144.