

COST EFFICIENCY OF SLOVENIAN WATER DISTRIBUTION UTILITIES: AN APPLICATION OF STOCHASTIC FRONTIER METHODS

Massimo Filippini

CEPE, Federal Institute of Technology, ETH Zentrum, Zürich and
Department of Economics, University of Lugano, Switzerland

E-mail: massimo.filippini@lu.unisi.ch

Nevenka Hrovatin

University of Ljubljana, Faculty of Economics, Kardeljeva pl. 17, 1000 Ljubljana,
Slovenia

E-mail: nevenka.hrovatin@ef.uni-lj.si

Jelena Zorić*

University of Ljubljana, Faculty of Economics, Kardeljeva pl. 17, 1000 Ljubljana,
Slovenia

E-mail: jelena.zoric@ef.uni-lj.si

Abstract

This study estimates cost inefficiency and economies of scale of Slovenian water distribution utilities over the 1997-2003 period by employing several different stochastic frontier methods. The results indicate that significant cost inefficiencies are present in the utilities. An introduction of incentive-based price regulation scheme might help resolve this problem. However, the inefficiency scores obtained from different cost frontier models are not found to be robust. The levels of inefficiency estimates as well as the rankings depend on the econometric specification of the model. The established lack of robustness can be at least partly explained by different ability of the models to separate unobserved heterogeneity from inefficiency. Newly proposed true fixed effects model (Greene, 2005a, b) appears to perform better than the conventional panel data models with respect to distinguishing between unobserved heterogeneity and inefficiency. On the other hand, different models produce fairly robust results with respect to estimates of economies of output density, customer density and economies of scale. The optimal size of a company is found to closely corresponds to the sample median. Economies of scale are found in small-sized utilities, while large companies exhibit diseconomies of scale.

Keywords: *SFA, cost frontier function, cost inefficiency, water distribution utilities, price regulation*

* Corresponding author.

1. Introduction

Over the last two decades it has become increasingly important to promote the efficiency and improve the performance of natural monopolies operating within network industries. In this respect, incentive-based regulation schemes appear to be superior to the traditional rate-of-return regulation. The most widely adopted incentive-based regulatory schemes in the water distribution sector involve price cap (RPI-X), revenue cap, and yardstick regulation models. Most of these regulation schemes used in practice are based on *benchmarking* that is, measuring a company's productive efficiency against a reference performance. In benchmarking applications the regulator is generally interested in obtaining a measure of firms' efficiency in order to reward (or punish) companies accordingly. Hence, there is a close link between efficiency measurements and incentive-based price regulation.

In the EU water industry context, the two best-practice regulatory examples are the UK regulator OFWAT where benchmarking combined with a price-cap is in use (OFWAT, 1999, 2004) and the Italian Regulation Authority where benchmarking is combined with rate-of-return regulation (Massarutto, 1999). In Slovenia, the current price regulation of water distribution utilities still closely resembles the traditional rate-of-return regulation. Nonetheless, it should be noted that the rules on price regulation recently issued by the government (i.e., Rules on Price Determination of Obligatory Local Public Utilities for Environment Protection, 2004) envisage the benchmarking of costs and quality combined with the rate-of-return regulation. However, the rules have not yet been put into practice, nor has been the benchmarking method specified.¹

The reliability of efficiency scores is crucial for an effective implementation of incentive-based price regulation. Unfortunately, the evidence from empirical studies shows that the various benchmarking methods – the *parametric* approach and the *non-parametric* approach – often produce different results with respect to firms' efficiency scores and rankings.² A possible explanation of this lack of robustness problem relates to the difficulty of benchmarking methods in accounting for observable and unobservable heterogeneity in environmental and network characteristics across companies.³ This is particularly undesirable if the results are to be used in economic policy-making. Despite extensive research carried out in the field of efficiency measurement, so far there is no general consensus on which approach (*parametric* or *non-parametric*) has been found to perform the best.⁴

¹ To note that in Slovenia a price-cap incentive regulation scheme combined with benchmarking analysis has been already applied to the electricity distribution utilities (AERS, 2004). Use of stochastic frontier benchmarking in estimating cost inefficiency of Slovenian electricity distribution companies is considered in Filippini, Hrovatin and Zorić (2004).

² For example, see Bauer et al. (1998), Estache, Rossi and Ruzzier (2004), Jamasb and Pollitt (2003), Farsi, Filippini and Greene (2005), Farsi and Filippini (2006).

³ See Daraio and Simar (2005) for a discussion on the possibility to take into account environmental (*observed*) heterogeneity in a nonparametric approach and Farsi et. al. (2005) for a discussion in a parametric approach.

⁴ Both approaches have advocates in the scientific community. The purpose of this paper is not to stress the advantages and drawbacks of these two approaches, but to explore if some limitations of conventional stochastic frontier models can be overcome if panel data are available. Particularly, the focus of the paper is on the effect that *unobserved* heterogeneity can have on the inefficiency estimates.

In the following paper, parametric frontier benchmarking methods are used to study performance of the Slovenian water distribution utilities. This is one of the first studies on performance of water distribution companies from Central and Eastern European transition countries.⁵ Several stochastic frontier methods for panel data are used to estimate the cost frontier function for a sample of water distribution companies operating in Slovenia between 1997 and 2003. Stochastic Frontier Analysis (SFA) was originally introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). In subsequent papers, Pitt and Lee (1981) and Schmidt and Sickles (1984) proposed stochastic frontier models for panel data. Over the years, many extensions to the originally proposed stochastic frontier models have been developed.⁶ Since water distribution utilities operate in different regions with different environmental and network characteristics that are only partially observed, it is essential to be able to distinguish between inefficiency and unobserved heterogeneity that influences the costs. Until recently, this issue has been neglected in the empirical work since the conventional stochastic frontier models are unable to make a distinction between these two effects. As a result, unobserved heterogeneity has often been confounded with inefficiency. Since this may have serious financial consequences for regulated firms, it is crucial to be able to explicitly model cost differences that are due to heterogeneity and inefficiency. New developments in the field of stochastic frontier analysis, namely true random and true fixed effects models proposed by Greene (2005a, b) can help us address this issue. These models extend the previous models by adding an additional stochastic error component for the heterogeneity.⁷

In order to find out whether accounting for unobserved heterogeneity in the model significantly influences the results, the cost inefficiency estimates obtained from both conventional panel data models and the newly proposed models are compared. We also analyze the robustness and reliability of obtained cost inefficiency scores and propose how the results from benchmarking analysis could be employed in regulating water prices in Slovenia. In addition, economies of scale and density are estimated and the optimal size of water distribution utilities is ascertained.

The paper is organized as follows. Section 2 shortly reviews studies estimating the cost function of water distribution companies. Section 3 presents the model specification and the methodology employed. The data description is provided in Section 4. Section 5 presents the estimation results and Section 6 concludes the paper.

2. Review of the Relevant Studies

In the literature we can find two types of studies on costs of water distribution companies: (i) studies estimating cost function and economies of output density,

⁵ By becoming a new EU member state on May 1 2004, the transition process of Slovenian economy has been formally completed. Thus, we can now speak of ex-transition countries.

⁶ A good review of different stochastic frontier methods is provided in Kumbhakar and Lovell (2000).

⁷ For example, the true random effects model has been already applied to the network industries by Farsi, Filippini and Greene (2005) and Farsi, Filippini and Kuenzle (2005, 2006). A similar model but with a three-stage estimation procedure has been proposed by Kumbhakar (1991) and Heshmati and Kumbhakar (1994).

customer density and/or economies of scale, and (ii) studies estimating cost frontier function and cost efficiency. In what follows, we provide a short review of the most relevant papers, covering sample description, model specification, functional form, variables included in the cost function, method of estimation and the obtained results. Table 1 summarizes the reviewed studies.

Kim and Clark (1988) examine the multiproduct nature of water supply relative to economies of scale and scope. The data used in the study come from a cross-section of 60 water utilities in the United States for 1973. In estimating the total cost function, two outputs are considered: the amounts of water delivered to residential and non-residential users. To take into account spatial variation of demand, service distance is included. The input prices included in the model are price of labour, price of capital and price of energy. Capacity utilization rate, measured by the load factor of a water system, is also incorporated in the model. The translog multiproduct total cost function is estimated jointly with the cost share equations by using Zellner's iterative seemingly unrelated regression (SUR) method. No significant economies of scale in the utility's overall operation are discovered, where overall economies of scale for the sample mean are estimated to be 0.99. Small utilities exhibit rather marked economies of scale (1.33), while large utilities exhibit moderate diseconomies of scale (0.88). The utilities on the whole are found to enjoy considerable economies of scale for non-residential water supply, but suffer from diseconomies in residential supply. The utilities also experience economies of scope associated with joint production of the two services. The shortcoming of this study is that it does not consider output characteristics in estimating the total cost function. Due to excluded variables such as number of customers and area size, the obtained results may be biased.

Bhattacharyya et al. (1995) uses a stochastic frontier cost function to specify the costs and inefficiency of 221 publicly and privately owned urban water utilities operating in the US in 1992. A translog functional form is employed to estimate variable cost function. Explanatory variables used are the output (total quantity of water sales), the input prices (price of energy, labour and material), the stock of capital and the network variables. Network configuration variables include different types of water sources used, total quality of water produced and total system loss. The error term is composed of a random noise and the cost inefficiency term. Both mean and variance of inefficiency are specified in the model as functions of firm-specific factors. The model is estimated by the two-step estimation procedure.⁸ The public water utilities on average out-perform the private water companies; the estimated mean cost inefficiencies are 9.8% and 18.7%, respectively. Again, the estimated inefficiency scores may be biased since output characteristics are not included in the model. Another important thing to be noticed is that the coefficient of the capital stock has a positive sign. This is quite often in the applied literature, although it contradicts the cost theory.⁹

⁸ This procedure has some serious econometric flaws which results in biased and inconsistent estimators (Wang and Schmidt, 2002).

⁹ Possible explanations of this theoretically implausible sign are provided in Cowing and Holtmann (1983), Guyomard and Vermersch (1989), and Filippini (1996). One of the reasons in this particular case may be also poorly specified capital stock variable.

Antonioli and Filippini (2001) explore economies of scale and density in the Italian water industry. The panel consists of 32 water distribution firms over the 1991-1995 period. The variable cost function is estimated using the OLS and random-effects model. Due to the time-invariance of some explanatory variables, the fixed-effects model was disregarded. To avoid the multicollinearity problem, a Cobb-Douglas functional form is employed. The explanatory variables employed are the amount of water distributed, the price of labour, the number of customers, the length of the pipes, percentage of water losses, the number of water wells (as a proxy for the capital stock), the treatment dummy variable and the time variable to capture the shift in technology. The inclusion of output characteristics in the cost function, allows for the distinction of economies of output density, economies of customer density and economies of scale. Since the returns to scale are estimated to be 0.95, the results based on the random-effects model suggest the presence of weak diseconomies of scale. On the contrary, there exist economies of output and customer density, the estimates being equal to 1.46 and 1.16, respectively. Since the random-effects model is not able to control for unobserved heterogeneity constant over time, the results may be biased.

Table 1: Summary of the findings from the literature review

Author(s) of the paper	Data sample	Model and functional form	Method of estimation / calculation	Estimated economies of scale	Estimated economies of density	Estimated cost efficiency
Kim and Clark (1988)	60 US water utilities in 1973	Translog multi-product TC function	SUR method	0.992 (sample average)	/	/
Bhattacharyya et al. (1995)	221 US water utilities from 1992 survey	Translog VC function	SFA (SUR and two-step estimation)	/	1.246 (E_{OD} , private, SR) 0.932 (E_{OD} , public, SR) ¹	0.901 (average; public more efficient)
Antonioli and Filippini (2001)	32 Italian water utilities between 1991-1995	Log-log VC function	OLS and RE panel data model	0.95 (LR)	1.46 (E_{OD} , LR) 1.16 (E_{CD} , LR)	/
Garcia and Thomas (2001)	55 French water utilities between 1995-1997	Multi-product translog VC function	GMM (IV method), SUR method	1.002 (sample average, LR)	1.21 (E_{OD} , LR) 0.87 (E_{CD} , LR) ²	/

¹ E_{OD} stands for economies of output density, while E_{CD} stands for economies of customer density. SR stands for the short run, while LR stands for the long run.

² In the short run, $E_{OD} = 1.14$ and $E_{CD} = 1.05$.

Garcia and Thomas (2001) examine the cost structure of French municipal water utilities. The sample is composed of 55 water utilities from the Bordeaux region for the years 1995 to 1997. Generalized Method of Moments (GMM) procedure is used to estimate the system of variable cost and input cost shares. Multi-product translog variable cost function is employed. The following explanatory variables are used: the output variables, factor prices, and technical variables. Technical variables used are the number of customers, the number of municipalities supplied (as proxy for area size), and several variables representing the existing capital stock: network length, production capacity, stocking and the pumping capacity. Estimated economies of scope at the variables sample mean are positive (0.237) indicating that there are potential gains in production water losses (undesirable output) jointly with water sold to final customers (desirable output). Possible explanation for this is that costs associated with network repairs and maintenance in order to decrease water losses are higher than costs involved

in satisfying customer demand by simply increasing water production. Furthermore, returns estimates at the sample mean show that in the short run there are economies of output density (1.14) as well as economies of customer density (1.05) present. In the long run, economies of output density are found (1.21), while there are no longer economies of customer density present (0.87). Finally, (the long run) scale economies are estimated to be 1.002.

The reviewed studies mostly focus on the estimation of the economies of scale and not on the measurement of cost inefficiency. With respect to the previous studies in the water distribution sector, the contribution of this study is that it notably improves on stochastic frontier methods used to estimate cost inefficiency and it explicitly recognises the problem of unobserved heterogeneity in measuring the cost and scale inefficiency. Moreover, the results of this empirical analysis could be of interest for the authorities responsible for regulation of the Slovenian water distribution companies.

3. Model Specification and Methodology

The main purpose of water supply utilities is to produce drinking water with sufficient quality from a resource (groundwater or surface water) that may require preliminary treatments to make drinking water wholesome and clean, and to distribute water by continuously adapting supply to daily demand while preserving water quality during its transportation through transmission and distribution network. Water supply utilities typically cover all operations from resource extraction to consumer taps (Garcia and Thomas, 2001). Accordingly, water production process consists of the following activities: (i) water extraction (or production) from groundwater or surface water and treatment of water, (ii) transfer of water through transmission pipelines, (iii) storage of water, (iv) pressurization of water pipelines, and (v) distribution of water to final customers through distribution mains; includes also quality monitoring and metering (Fabbri and Fraquelli, 2000, and Garcia and Thomas, 2001).

The costs of operating a water distribution system are the costs of building and maintaining the water system (wells and springs, pumps, treatment facilities, storage facilities, transmission and distribution pipelines and other facilities), and of measuring and billing water. For the specification of the cost model, we consider a water distribution company which uses three inputs, labour, capital and material, to distribute a single output to a number of customers within its service area of size. The number of customers and the network size can be considered as output characteristic variables. The output characteristics are included as explanatory variables to control for the cost differences that occur merely due to the (observed) heterogeneity of output.

If it is assumed that firms are in long run static equilibrium with respect to all inputs employed and that they minimize total cost, a cost function can be written as:¹⁰

$$C = C(Q, P_L, P_M, P_K, CU, AS, D_{LOSL}, D_{TREAT}, D_S, D_U, T) \quad (1)$$

¹⁰ Over the last few decades we have not witnessed any notable changes in water demand in Slovenia so the demand for water can be considered to have been relatively stable. There is thus no reason to believe that utilities are considerably deviating from the optimal level of capital employed. Therefore, the adoption of a long-run concept seems an appropriate choice.

where C represents total cost and Q is the output represented by the total cubic meters of water delivered. P_L , P_M , and P_K are the price of labour, the price of material and the price of capital, respectively. CU stands for the number of customers served, while AS is the size of the service area. D_{LOSL} is a dummy variable of water losses bearing value 1 if the firm has low water losses, and 0 value otherwise. D_{TREAT} is a dummy variable for water treatment and takes on value 1 if the firm distributes water that has to be treated chemically before distribution and 0 value otherwise. The treatment is necessary in a situation when, from a medical point of view, the quality of the water does not reach a predefined standard and, therefore, it is not suitable for drinking. Water distribution utilities can use different water resources: surface water, underground water or mix of both sources. D_S represents a dummy variable for the use of surface water only and D_U is a dummy variable for underground water only. Groundwater usually implies higher drilling and pumping costs, whereas treatment costs are usually higher with surface water. Finally, T is a time variable, which captures the shift in technology.

Estimation of cost function requires a specification of the functional form. The Cobb-Douglas is not locally flexible functional form, but is widely used in the literature because of its simplicity of application and clearness of interpretation of its parameters. The major limit of the Cobb–Douglas functional form is that the estimated values of the economies of scale and density do not vary with the size of the firms in the sample but are assumed to be constant. Generally, the translog cost function, which is a more flexible functional form, offers an appropriate functional form for answering questions about economies of scale and density. For that reason, the translog functional form is applied.¹¹ However, it should be noted that the translog function is not without shortcomings. Since the translog functional form is a local approximation, the estimation results are reliable close to the approximation point, while its global properties are unsatisfactory. Several studies have noted the problem of fitting a single parametric cost function across companies of widely varying size.¹² The translog form of Eq.(1) can be written in the following way:

¹¹ In fact, general translog specification was tested against Cobb-Douglas and translog with hedonic specification of the output and it has been found that the general translog as specified in Eq.(2) is the preferable functional form.

¹² An alternative to deal with this potential problem would be to use a semi-parametric orthogonal series estimator (the Fourier flexible form), which increases the number of parameters to be estimated and thus requires large samples. Due to the relatively small sample size this functional form had to be disregarded in our case. Similarly, nonparametric alternatives (e.g., kernel regression, local polynomial smoothing) could not be considered. Moreover, the use of nonparametric methods would not enable us to make a distinction between unobserved heterogeneity and inefficiency. For a discussion and application of the abovementioned methods see, for example, McAllister and McManus (1993), Mitchell and Onvural (1996), Wheelock and Wilson (2001) and Wilson and Carey (2004).

$$\begin{aligned}
\ln \frac{C_{it}}{PK_{it}} = & \ln \alpha_{(i)} + \beta_Q \ln Q_{it} + \beta_{CU} \ln CU_{it} + \beta_{AS} \ln AS_{it} + \beta_{PL} \ln \frac{PL_{it}}{PK_{it}} + \beta_{PM} \ln \frac{PM_{it}}{PK_{it}} \\
& + \frac{1}{2} \beta_{Q,Q} \ln Q_{it} \ln Q_{it} + \frac{1}{2} \beta_{CU,CU} \ln CU_{it} \ln CU_{it} + \frac{1}{2} \beta_{AS,AS} \ln AS_{it} \ln AS_{it} \\
& + \beta_{Q,CU} \ln Q_{it} \ln CU_{it} + \beta_{Q,AS} \ln Q_{it} \ln AS_{it} + \beta_{CU,AS} \ln CU_{it} \ln AS_{it} \\
& + \frac{1}{2} \beta_{PL,PL} \ln \frac{PL_{it}}{PK_{it}} \ln \frac{PL_{it}}{PK_{it}} + \frac{1}{2} \beta_{PM,PM} \ln \frac{PM_{it}}{PK_{it}} \ln \frac{PM_{it}}{PK_{it}} + \beta_{PL,PM} \ln \frac{PL_{it}}{PK_{it}} \ln \frac{PM_{it}}{PK_{it}} \\
& + \beta_{PL,Q} \ln \frac{PL_{it}}{PK_{it}} \ln Q_{it} + \beta_{PL,CU} \ln \frac{PL_{it}}{PK_{it}} \ln CU_{it} + \beta_{PL,AS} \ln \frac{PL_{it}}{PK_{it}} \ln AS_{it} \\
& + \beta_{PM,Q} \ln \frac{PM_{it}}{PK_{it}} \ln Q_{it} + \beta_{PM,CU} \ln \frac{PM_{it}}{PK_{it}} \ln CU_{it} + \beta_{PM,AS} \ln \frac{PM_{it}}{PK_{it}} \ln AS_{it} \\
& + \gamma_{LOSL} D_{LOSL} + \gamma_{TREAT} D_{TREAT} + \gamma_S D_S + \gamma_U D_U + \gamma_T T + \varepsilon_{it}, \tag{2}
\end{aligned}$$

with ε_{it} the error term, $i = 1, \dots, N$ and $t = 1, \dots, T_i$. The properties of the cost function are that it is concave and linearly homogeneous in input prices, nondecreasing in input prices and nondecreasing in output. Notice that normalization of cost and input prices by one of the input prices is used to impose linear homogeneity in input prices. Hence, the total cost, the price of labour and the price of material are divided by the price of capital. Other properties remain to be verified after the estimation of the translog cost function is conducted.

The stochastic frontier cost function in Eq.(2) is estimated using four different SFA methods. Table 2 summarises the models used in the analysis. The differences between the various specifications are related to the assumptions imposed on the error term (ε_{it}) introduced in Eq.(2), cost inefficiency and firm-specific effects. **Model I** is a pooled frontier model estimated by maximum likelihood (ML) method as proposed by Aigner, Lovell and Schmidt (1977). Since the focus of the SFA is not on estimating the frontier cost function but rather on the error term, especially the inefficiency component, let us express the cost frontier function in Eq.(2) in the following way:

$$\ln C_{it}^* = \alpha + c(\mathbf{x}_{it}; \boldsymbol{\beta}) + v_{it} + u_{it}, \tag{3}$$

where C^* denotes the normalised costs, \mathbf{x} stands for the vector of explanatory variables, $\boldsymbol{\beta}$ is the vector of coefficients and α is the regression constant. The error term (ε_{it}) in Model I is composed of two parts: a stochastic error (v_{it}), capturing the effect of noise, and a one-sided non-negative disturbance capturing the effect of inefficiency ($u_{it} \geq 0$). To estimate the stochastic cost frontier using the ML method, the following distributional assumptions have to be made: $v_{it} \sim \text{iid } N(0, \sigma_v^2)$, $u_{it} \sim \text{iid } N^+(0, \sigma_u^2)$, and v_{it} and u_{it} are distributed independently of each other and of the regressors. This model is referred to as a Normal-Half Normal Model.¹³ The cost inefficiency is usually expressed in terms of cost inefficiency score:

$$EFF_{it} = \frac{C_{it}}{C_{it}^F} = \exp(\hat{u}_{it}) \tag{4}$$

¹³ Alternative distributional assumptions on u_i can as well be made (e.g., exponential, truncated normal, and gamma distribution).

where C_{it} is the observed total cost and C_{it}^F is the frontier or minimum cost of the i -th firm in time t . Cost inefficiency score of one indicates a firm on the frontier, while non-frontier firms receive scores above one. Alternatively, the cost efficiency score can be calculated as the reciprocal of the cost inefficiency score defined in Eq.(4).

Table 2: Econometric specification of the models employed

<i>Model</i>	<i>Firm-specific component</i>	<i>Random error</i> ε_{it}	<i>Inefficiency</i> u_{it}
Model I Pooled (ML)	None	$\varepsilon_{it} = v_{it} + u_{it}$ $u_{it} \sim \text{iid } N^+(0, \sigma_u^2)$ $v_{it} \sim \text{iid } N(0, \sigma_v^2)$	$E(u_{it} \varepsilon_{it})$
Model II RE (GLS)	$\alpha_i \sim \text{iid}(0, \sigma_\alpha^2)$	$\varepsilon_{it} = v_{it} + \alpha_i$ $\varepsilon_{it} \sim \text{iid}(0, \sigma_\varepsilon^2)$	$u_i = \hat{\alpha}_i - \min_i \{\hat{\alpha}_i\}$
Model III RE (ML)	$u_i \sim \text{iid } N^+(0, \sigma_u^2)$	$\varepsilon_{it} = v_{it} + u_i$ $v_{it} \sim \text{iid } N(0, \sigma_v^2)$ $u_{it} \sim \text{iid } N^+(0, \sigma_u^2)$	$E(u_i \varepsilon_i)$
Model IV TFE (ML)	Fixed (group dummies α_i)	$\varepsilon_{it} = v_{it} + u_{it}$ $u_{it} \sim \text{iid } N^+(0, \sigma_u^2)$ $v_{it} \sim \text{iid } N(0, \sigma_v^2)$	$E(u_{it} \varepsilon_{it})$

Since **Model I** does not assume any firm-specific effects, it does not have the ability to distinguish between cost inefficiency and unobserved heterogeneity of the firms. We therefore turn to the panel data stochastic frontier models and examine how the abovementioned shortcoming of the pooled model is addressed. In **Model II** we consider random-effects (RE) model proposed by Schmidt and Sickles (1984):

$$\ln C_{it}^* = \alpha_o + c(\mathbf{x}_{it}; \boldsymbol{\beta}) + \alpha_i + v_{it} . \quad (5)$$

The model in Eq.(5) is estimated by feasible Generalized Least Squares (GLS) method. If we allow stronger distributional assumptions on the inefficiency term to hold, we can use ML procedure to estimate the RE model, which is done in **Model III**:

$$\ln C_{it}^* = \alpha_o + c(\mathbf{x}_{it}; \boldsymbol{\beta}) + v_{it} + u_i . \quad (6)$$

The latter method was introduced by Pitt and Lee (1981). In both conventional RE models cost inefficiency is assumed to be time-invariant, which can be rather limiting assumption, particularly in long panels. However, it may be a plausible assumption in non-competitive operating environment. From cost inefficiency estimates (α_i or u_i) cost inefficiency scores are also obtained using Eq.(4). The fixed-effect (FE) model is not considered appropriate since its precision relies on the within variation which is very low in our case.¹⁴ Also, time-invariant variables which are often present in the network industries can not be included in the FE model. Nonetheless, the appeal of the FE model as opposed to the GLS estimator is that the former produces unbiased estimates of the

¹⁴ See Cameron and Trivedi (2005) for a discussion on this issue.

regression coefficients even if the firm specific effects are correlated with the regressors. On the other hand, the FE model produces biased inefficiency estimates due to the incidental parameter problem. Based on the above arguments we decided not to employ this model.

The main weakness of Model II and Model III is that they force any time-invariant firm-specific heterogeneity into the same term that is being used to capture the inefficiency (Greene, 2005 b). Consequently, these models do not have the ability to distinguish between time-invariant unobserved heterogeneity and cost inefficiency. Any time-invariant firm-specific effects are treated as inefficiency. By introducing environmental and exogenous factors in the model, one can control for observed heterogeneity. However, not all relevant data are always available and some factors may even be too complex to be properly measurable. This results in unobserved heterogeneity which is beyond the firms' control but may affect their costs significantly. To deal with the unobserved heterogeneity, the alternative 'true' fixed-effects and 'true' random-effects models recently proposed by Greene (2005a, b) are considered. Therefore, in **Model IV** we additionally estimate the stochastic frontier cost function by applying true fixed effects (TFE) model formulated in the following way:

$$\ln C_{it} = \alpha_i + c(\mathbf{x}_{it}; \boldsymbol{\beta}) + v_{it} + u_{it} \quad (7)$$

TFE model treats firm-specific time-invariant *fixed* effects (α_i) and time-varying inefficiency (u_{it}) separately and is therefore able to distinguish between the unobserved heterogeneity and inefficiency. In this way it tries to overcome some limitations of the conventional panel data models.¹⁵ The model is estimated by the 'brute force' maximum likelihood, i.e., by simply creating dummy variables for each firm. The remaining shortcoming of the TFE model is the incidental parameters problem.¹⁶ Also, time-invariant firm characteristics can not be included in the model as explanatory variables. Nevertheless, these effects are viewed as unobserved heterogeneity and are (at least partially) captured by the firm-specific time-invariant term additionally specified by this model.

A final note to be made is that newly proposed TFE and TRE do not fully resolve the problem of distinguishing between unobserved heterogeneity and inefficiency. The problem of these two models is that any time-invariant or persistent component of inefficiency is completely absorbed in the firm-specific constant term. As the conventional FE and RE models tend to overestimate the inefficiency, it may be the case that the TFE and TRE models underestimate it. Thus, the choice of appropriate model is also based on the researcher's belief whether there is some time-invariant

¹⁵ TRE model was as well applied, but the simulated maximum likelihood estimation method did not converge. The true random-effects model (TRE) is specified as: $\ln C_{it} = \alpha + c(\mathbf{x}_{it}; \boldsymbol{\beta}) + \omega_i + v_{it} + u_{it}$. The difference between this formulation and the TFE model is that ω_i is a (time-invariant and firm-specific) *random* effect meant to capture unobserved heterogeneity. A possible explanation why this model did not perform well in our case is that the model specification is too rich for our data and, as a result, some of the error terms degenerate to zero.

¹⁶ Greene (2005b) finds the bias to be small with respect to the estimates of the regression coefficients. For the inefficiency estimates the bias is found to be larger, where the overestimation error of about 20% is reported.

unobserved heterogeneity in the model or whether the inefficiency does not in fact vary over time (Greene, 2005b).

Besides the cost inefficiency, the main objective of this paper is to ascertain the presence of economies of scale and density. In the case of network industries, the output typically possesses several dimensions. Therefore, besides output distributed, several output characteristics such as number of customers, size of service area or length of network can influence the costs. According to Caves et al. (1984) and Roberts (1986), the inclusion of the number of customers and the size of service area in the cost function allows us to distinguish between economies of output density, economies of customer density and economies of size. From estimated cost frontier function in Eq.(2), economies of output density are obtained as follows:

$$E_{OD} = \left(\frac{\partial \ln C}{\partial \ln Q} \right)^{-1}. \quad (8)$$

Economies of output density measure the reaction of costs to an increase in output, holding the number of customers and the size of the service area constant. It also follows that the customer density, defined as a ratio of the number of customers to the area size, is held constant. The existence of economies of output density ($E_{OD} > 1$) implies that the average cost of water distribution utility decreases as a physical output increases.

Furthermore, if average cost decreases as the output and number of customers are proportionally increased then economies of customer density exist ($E_{CD} > 1$). This measure allows us to analyze the existing service area which becomes more densely populated. In addition, it is assumed that, on average, new customers consume as much as the existing ones. Using Eq.(2) economies of customer density are calculated as:

$$E_{CD} = \left(\frac{\partial \ln C}{\partial \ln Q} + \frac{\partial \ln C}{\partial \ln CU} \right)^{-1}. \quad (9)$$

Finally, economies of scale measure the reaction of costs when the output, the number of customers and the area size increase proportionally. It is assumed that customer density and output per customer are held fixed. Economies of scale are obtained from Eq.(2) in the following way:

$$E_S = \left(\frac{\partial \ln C}{\partial \ln Q} + \frac{\partial \ln C}{\partial \ln CU} + \frac{\partial \ln C}{\partial \ln AS} \right)^{-1}. \quad (10)$$

The economies of scale exist when $E_S > 1$. This measure becomes important when analysing whether or not it is beneficial to expand the size of the service area.

4. Data Description

Slovenian water industry is part of the Slovenian communal sector providing public services of supply of drinking water, wastewater treatment, solid waste management

and some other services. The study is based on a panel data set for Slovenian water distribution utilities over the 1997-2003 period. Since water supply utilities are under the responsibility of local communities, the data on their operation are not collected systematically at the national level. Thus, the data had to be gathered via questionnaire issued by the Ministry of the Environment and Spatial Planning. In this way we obtained data on 52 water supply utilities over the 1997-2003 period. The sample is an unbalanced panel consisting of total 332 observations.

Utilities included in the sample supply 153 out of 192 municipalities in Slovenia, that is almost 80% of all municipalities. All Slovenian regions are covered by the utilities in the sample. Only four companies in the sample are not public utilities, but operate as private companies or they have a concession. Utilities differ in terms of size and as well in some environmental conditions. Some utilities also provide other services like wastewater treatment, solid waste disposal etc. Usually, in smaller municipalities all communal activities are joined in a single company, while in larger municipalities communal activities are provided separately by several companies. Since 1997, utilities are obliged to have separate accounts for different regulated activities. This separation of activities is aimed to increase transparency and enable easier monitoring of regulated utilities.

Descriptive statistics of the variables included in the model are presented in Table 3. Total distribution cost (C) equals to operating and capital expenditure. The price of labour (P_L) is equal to average annual wages, estimated as labour expenditures divided by the average number of employees for a given year. The price of capital (P_K) is calculated as the ratio of capital cost and the capital stock, which is approximated by the capacity of pumps measured in litres per second. Capital cost consists of depreciation and interests, where depreciation accounts for most of the capital cost.¹⁷ The price of material (P_M) is obtained by dividing material cost by the length of distribution network in kilometres. Material cost consists of various groups of costs obtained when subtracting capital and labour cost from the total company's cost. Material cost thus includes cost of energy, material and services. All input prices and costs were deflated to 2000 constant Slovenian tolar (SIT) using the producers' price index.

The output (Q) is measured as the amount of water supplied to the final customers expressed in cubic metres. The number of final customers (CU) is the sum of household and non-household customers. The size of service area (AS) is expressed in square kilometres.

Water losses are the difference between the amount of water pumped into the distribution system and the amount of water supplied to the customers. The share of water losses is calculated as the ratio of water losses and the water pumped into the pipes. It is considered that the utility has low water losses if the share of water losses does not exceed the first quartile which equals to 19% of water losses. The variable is included in the model as a dummy variable D_{LOSL} with value 1 if the firm has low water

¹⁷ There are strict rules in place which set limits regarding the purpose of the loan and the amount of the loan that public utilities can take. Interest expenses thus mostly consist of costs related to the fixed capital.

losses, and 0 otherwise.¹⁸ In some cases water needs treatment in order to be suitable for drinking. A dummy variable D_{TREAT} takes on value 1 if the firm distributes water that has to be treated chemically before distribution and 0 value otherwise. Only demanding chemical treatment is taken into the account; simple chemical treatment (disinfection and chlorination) is not considered. Since water distribution utilities can use surface water, underground water or mix of both resources, a type of water resource is as well included in the model. D_S is a dummy variable for the use of surface water only and D_U is a dummy variable for the use of underground water only.

Table 3: Descriptive statistics

Variable description	Variable	Mean	Std. Dev.	Minimum	Maximum
Total annual cost (10^3 SIT) ¹	<i>TOTEX</i>	304,698	538,387	7,208	2,997,534
Price of labour (10^3 SIT/ employee)	<i>PL</i>	3,047.7	397.1	2,131.9	4,162.7
Price of capital (10^3 SIT/ litre per sec.)	<i>PK</i>	449.4	564.9	13.5	1,484.0
Price of material (10^3 SIT/ km of network)	<i>PM</i>	312.0	244.3	46.9	1,412.0
Water supplied (m^3)	<i>Y</i>	2,298,780	3,835,452	106,627	25,507,653
Number of customers	<i>CUST</i>	7,402.1	7,777.4	515.0	43,272.0
Size of service area (km^2)	<i>AREA</i>	336.9	240.0	57.8	949.1
Treatment dummy	<i>D_{TREAT}</i>	0.120	0.326	0	1
Dummy for surface water	<i>D_S</i>	0.199	0.400	0	1
Dummy for underground water	<i>D_U</i>	0.355	0.479	0	1
Dummy for low water losses	<i>D_{LOSL}</i>	0.250	0.434	0	1

¹ The average official exchange rate of Slovenian tolar (SIT) in 2000 was 1 EUR = 205.0316 SIT (Bank of Slovenia, 2001).

5. Parameter Estimates of Cost Frontier Function

The estimation results of the translog cost frontier function of Slovenian water distribution utilities obtained by the four different models are given in Table 4.¹⁹ The expansion point of the translog stochastic frontier cost function specified in Eq.(2) is chosen to be the sample median. Since total cost and all the continuous explanatory variables are in logarithms, the estimated first-order coefficients can be interpreted as cost elasticities evaluated at the sample median. As expected, results show that input prices, output and output characteristics are positive and highly significant across models. Different models, however, lead to different results with respect to coefficient estimates.

¹⁸ In order to avoid a multicollinearity problem we did not include a continuous variable for water losses in the model. The correlation between water losses, output, number of customers and size of the service area is quite high.

¹⁹ The models are estimated using NLogit 3.0.

Table 4: Estimation results of the cost frontier function

Coefficient	Model I Pooled (ML)	Model II RE (GLS)	Model III RE (ML)	Model IV TFE (ML)
$\ln a$	11.570*** (0.036)	11.856*** (0.054)	11.424*** (0.079)	-
c_{PL}	0.579*** (0.024)	0.405*** (0.024)	0.401*** (0.041)	0.521**** (0.030)
c_{PM}	0.180*** (0.022)	0.341*** (0.020)	0.339*** (0.042)	0.193*** (0.028)
b_Q	0.329*** (0.059)	0.289*** (0.063)	0.290*** (0.091)	0.258*** (0.069)
b_{CU}	0.449*** (0.059)	0.471*** (0.071)	0.454*** (0.095)	0.503*** (0.072)
b_{AS}	0.193*** (0.023)	0.201*** (0.040)	0.218*** (0.074)	0.158*** (0.032)
$c_{PL,PL}$	-0.097*** (0.047)	0.034*** (0.037)	0.015*** (0.069)	-0.178*** (0.060)
$c_{PM,PM}$	-0.052*** (0.036)	0.014*** (0.028)	-0.005*** (0.044)	-0.109*** (0.046)
$c_{PL,PM}$	0.144*** (0.037)	0.023*** (0.029)	0.040*** (0.050)	0.222*** (0.050)
$b_{Q,Q}$	0.587*** (0.152)	0.330*** (0.124)	0.248*** (0.239)	0.673*** (0.149)
$b_{CU,CU}$	0.122*** (0.226)	-0.029*** (0.189)	-0.094*** (0.281)	-0.184*** (0.252)
$b_{AS,AS}$	0.195*** (0.055)	0.086*** (0.116)	0.026*** (0.163)	0.287*** (0.078)
$b_{Q,CU}$	-0.432*** (0.177)	-0.209*** (0.139)	-0.149*** (0.228)	-0.350*** (0.183)
$b_{Q,AS}$	0.022*** (0.076)	-0.032*** (0.088)	-0.031*** (0.121)	-0.103*** (0.089)
$b_{CU,AS}$	0.155*** (0.078)	0.127*** (0.092)	0.185*** (0.174)	0.250*** (0.093)
$d_{PL,Q}$	0.096*** (0.080)	0.104*** (0.063)	0.117*** (0.118)	0.089*** (0.089)
$d_{PL,CU}$	-0.056*** (0.078)	-0.137*** (0.066)	-0.157*** (0.119)	-0.060*** (0.093)
$d_{PL,AS}$	-0.065*** (0.037)	0.031*** (0.039)	0.028*** (0.076)	-0.105*** (0.048)
$d_{PM,Q}$	-0.109*** (0.059)	-0.101*** (0.050)	-0.092*** (0.069)	-0.148*** (0.069)
$d_{PM,CU}$	0.094*** (0.065)	0.126*** (0.051)	0.135*** (0.067)	0.093*** (0.079)
$d_{PM,AS}$	0.046*** (0.035)	-0.055*** (0.030)	-0.081*** (0.081)	0.124*** (0.044)
h_T	0.002*** (0.005)	-0.002*** (0.003)	-0.002*** (0.004)	-0.008*** (0.004)
g_S	0.202*** (0.029)	0.097*** (0.069)	0.216*** (0.116)	-
g_U	0.090*** (0.026)	0.040*** (0.059)	0.219*** (0.201)	-
g_{TREAT}	0.120*** (0.037)	0.212*** (0.080)	0.287*** (0.164)	-
g_{LOSL}	-0.156*** (0.027)	-0.037*** (0.020)	-0.020*** (0.022)	-

Notes: standard errors in brackets;
 * – significant at 10%, ** – significant at 5%, *** – significant at 1%, **** – significant at 0.1% (two-sided significance level)

Table 4: Continuation

Coefficient	Model I Pooled (ML)	Model II RE (GLS)	Model III RE (ML)	Model IV TFE (ML)
$\sigma_v (S_v)$	0.0976	0.0712	0.0698	0.1542
$\sigma_u (S_u)$	0.2502	0.1714	0.4282	0.2611
$\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$	0.2686****	-	0.4338***	0.3032****
$\lambda = \sigma_u/\sigma_v$	2.564**** (0.3397)	-	6.137** (3.0475)	1.693**** (0.2079)

Notes: standard errors in brackets;

* – significant at 10%, ** – significant at 5%, *** – significant at 1%, **** – significant at 0.1% (two-sided significance level)

The results of the four models show that the output coefficient (b_Q) is positive and highly significant in all models. It suggests that, on average, a one percent increase in the amount of water supplied will increase the total cost of Slovenian water distribution utilities by 0.26% to 0.33%, depending on the model considered. Similarly, the coefficients of the two output characteristics, the number of customers (b_{CU}) and the size of service area (b_{AS}), are found to be significantly positive. The coefficient of the number of customers varies between 0.45 and 0.50, while the coefficient of the service area size is found to be between 0.16 and 0.22.

The cost frontier function is non-decreasing in input prices since both the labour price coefficient as well as the material price coefficient are positive and highly significant. The concavity in input prices and twice-continuous differentiability of the cost function imply that the Hessian is negative semidefinite. This condition is confirmed at the sample median for Models I and IV. In Models II and III, the respective coefficients are not found to be significant. It should be noted that theoretical cost function satisfying all the required properties can only be derived under the cost minimising behaviour. If this is not the case, it is likely that some properties of the cost function, for example concavity in input prices, will not be satisfied. Therefore, in the presence of inefficiencies in the model there is no reason to assume that by employing the cost function, one can truly arrive at the economic representation of the production possibility set. In such cases the estimated empirical cost function cannot be viewed as the ‘true’ cost function but rather as the ‘behavioural’ cost function (Evans, 1971, and Breyer, 1987).

Time does not seem to have a significant influence on the costs of Slovenian water distribution utilities. By assuming a one-sided hypothesis, only in Model IV are costs found to be significantly decreasing over the analysed period. Based on the results it cannot be concluded that total cost has considerably changed over time. This is largely consistent with non-competitive environment in which the public utilities operate. Also, price regulation is not designed in a way that would stimulate utilities to decrease their costs and operate more efficiently.

Table 5 provides descriptive statistics on the cost inefficiency estimates of Slovenian water distribution utilities obtained from Models I – IV. We can observe some notable differences in the estimated cost inefficiency levels. By employing the pooled stochastic

frontier model (Model I), the average cost inefficiency is estimated to be 22.5%. On the contrary, in the case of the RE panel data stochastic frontier models, the estimated average cost inefficiencies are quite high; the inefficiency amounts to 66.3% in the RE GLS model (Model II) and 50% in the RE ML model (Model III). The relatively high inefficiency levels of the RE models might to some extent be attributed to unobserved firm-specific time-invariant effects. The RE models treat these effects as time-invariant cost inefficiency so the cost inefficiency estimates obtained by these models are most likely overestimated. This is not the case of the pooled model since each observation is treated as independent and, accordingly, the inefficiency is considered to vary across utilities and over time. Further, in the RE models the median values of cost inefficiency are considerably lower compared to the means, indicating that the means are influenced by the extreme values. Finally, the average cost inefficiency based on the true fixed effects model (Model IV) is estimated to be 19.1%. Moreover, the standard deviation of efficiency scores is much lower in this model than in the other models. Lower inefficiency levels in comparison to the other models are expected since the true fixed effects model is able to distinguish unobserved firm-specific fixed effects from inefficiency and is thus able to treat the two effects separately. We therefore believe that the results obtained by the TFE model can be regarded as a good approximation of the general picture of the actual cost inefficiency of Slovenian water distribution utilities.

Table 5: Estimated cost inefficiency scores

Inefficiency score (EFF)	Model I Pooled (ML)	Model II RE (GLS)	Model III RE (ML)	Model IV TFE (ML)
<i>Mean</i>	1.225	1.663	1.500	1.191
<i>Median</i>	1.181	1.556	1.378	1.182
<i>Std. Dev.</i>	0.162	0.376	0.346	0.057
<i>Minimum</i>	1.031	1.000	1.118	1.067
<i>Maximum</i>	1.710	2.690	2.599	1.514

What remains to be tested is whether the models provide similar rankings of the utilities with respect to the cost inefficiency scores. From the regulatory point of view, this issue is considered to be vital. Table 6 provides the pair-wise Pearson correlation coefficients between the cost inefficiency estimates.²⁰ We can observe that, with the exception of Model IV, the correlation between the inefficiency scores resulting from different models is significant, positive and, overall, not particularly high. The correlation is especially high between the inefficiency scores from Model II and Model III (RE panel data models). The correlation between inefficiency scores from Model IV and Model I is significant but quite moderate, whereas the correlation between Model IV and Models II and III is not significantly different from zero.²¹ Again, the reason may be found in the fact that the TFE model treats firm-specific fixed effects (α_i) separately from the

²⁰ The conclusions based on the rank correlation between the inefficiency scores from different models (Spearman correlation coefficients) are very similar to those found in Table 6.

²¹ Similar conclusion can be found in Farsi, Filippini and Greene (2005) and Farsi, Filippini and Kuenzle (2005, 2006) where the efficiency estimates of the true random effects model are compared with conventional panel data stochastic frontier models. The conclusions based on the rank correlation between the inefficiency scores from different models are very similar to those found in Table 6.

inefficiency (u_{it}). As a result, some effects that might be attributed to inefficiency by other models are here captured by the firm-specific effects and thus attributed to firm heterogeneity rather than inefficiency. This may be a plausible reason for the no correlation with the two RE panel data models.

These results show the sensitivity of the stochastic frontier benchmarking methods in our sample. This is not particularly encouraging since the results cannot be considered as reliable, especially if they are to be applied in the price-regulation process. Therefore, the direct use of inefficiency estimates in the regulation of water distribution utilities may be misleading. Nevertheless, some lack of robustness of inefficiency estimates is expected since the various models employ different assumptions regarding cost inefficiency and heterogeneity. We thus cannot expect the results to be completely invariant to these different assumptions.

Whether time-invariant effects belong to unobserved heterogeneity or cost inefficiency is debatable. If there is some time-invariant inefficiency, the inefficiency scores obtained by TFE model could be underestimated. On the other hand, if there is some unobserved time-invariant heterogeneity present the other panel data models treat it as cost inefficiency and thus tend to overestimate it. How we handle time-invariant effects obviously has a large influence on the findings. Ultimately, firm-specific heterogeneity and inefficiency both might contain time-invariant and time-varying elements and there is no perfect way to disentangle them based on the observed data (Greene, 2005a, b). If one wishes to be on the safe side, then the conclusion would be that the TFE sets the lower bound, whereas the two RE models set the upper bound for the cost inefficiency of Slovenian water distribution utilities.

Table 6: Correlation between inefficiency scores (Pearson correlation coefficients)

<i>R</i>	<i>Model I</i> Pooled (ML)	<i>Model II</i> RE (GLS)	<i>Model III</i> RE (ML)	<i>Model IV</i> TFE (ML)
<i>Model I</i>	1	0.667*	0.614*	0.399*
<i>Model II</i>		1	0.932*	0.023
<i>Model III</i>			1	0.027
<i>Model IV</i>				1

Note: * – significant at 0.1% (two-sided significance level)

Estimated economies of output density, customer density and economies of scale for Slovenian water distribution utilities can be found in Table 7. The respective measures for all four models are calculated using Eq.(8), Eq.(9) and Eq.(10), where the input prices are held fixed at their median values. With respect to the amount of water distributed, the number of customers, and the size of service area three types of representative companies are chosen – a first-quartile company (small companies), a median company (medium-sized companies) and a third-quartile company (large companies). Here the results from different models as reported in Table 7 demonstrate far more consistency than in the case of cost-inefficiency scores. All results follow the same pattern and lead us to the same conclusions.

Economies of output density (E_{OD}) are present for all three types of companies with respect to size. Since $E_{OD} > 1$, a 1% increase in cost (C) is associated with a more than 1% increase in the amount of water distributed (Q), holding the number of customers (CU) and the size of the service area (AS) constant. It would therefore be beneficial for water companies if they managed to distribute larger amounts of output to the existing customers within their service areas. E_{OD} are the highest for small utilities, followed by medium-sized utilities and large utilities. The economies of customer density (E_{CD}) are also confirmed for all three different types of companies. A 1% proportional increase in both the output and the number of customers leads to an increase in cost by less than 1% ($E_{CD} > 1$), holding the area size constant. Thus, it would be beneficial for companies if the existing service areas were to become more densely populated or if the companies could manage to get new customers. The highest E_{CD} are reported for the median company.

Table 7: Economies of output density (E_{OD}), customer density (E_{CD}) and scale (E_S)

Economies	Quartile	Model I Pooled (ML)	Model II RE (GLS)	Model III RE (ML)	Model IV TFE (ML)
E_{OD}	1st Quartile	3.099	3.485	3.500	4.605
	Median	3.042	3.455	3.448	3.874
	3rd Quartile	1.846	2.509	2.689	2.029
E_{CD}	1st Quartile	1.214	1.222	1.277	1.109
	Median	1.286	1.316	1.344	1.313
	3rd Quartile	1.182	1.265	1.263	1.208
E_S	1st Quartile	1.289	1.121	1.157	1.311
	Median	1.030	1.040	1.039	1.088
	3rd Quartile	0.816	0.933	0.925	0.846

The economies of scale (E_S) equal the inverse of the percentage change in costs when the output, number of customers and area size increase by 1%. The results show that substantial economies of scale are present in smaller companies ($E_S > 1$). It would be thus rational for the smaller companies to expand their service area or, if possible, for adjacent companies to merge.²² Economies of scale are also present in medium-sized companies, where they are close to one. This is also an indication that the optimal size of Slovenian water distribution utilities is relatively close to the median point of the sample. The median company corresponds to a company with an annual water supply of 1.17 million cubic metres, 5,168 customers and 264 square kilometres of service area size. On the other hand, diseconomies of scale prevail in large companies ($E_S < 1$). Apparently, the largest water distribution utilities in the sample have already exhausted their potential for cost savings resulting from economies of scale and their operations are found to be on the interval where average costs already start to rise.

²² Of course, to come up with realistic estimates of potential gains from merging the companies, it would be necessary to look in more detail geographical position of the companies and areas they operate and to analyse whether it would be feasible to connect two or more water distribution networks. This analysis goes beyond the scope of this study.

6. Conclusions

In the study several different stochastic frontier methods were considered to estimate a translog cost frontier function for the sample of Slovenian water distribution utilities over the 1997-2003 period. The inefficiency scores obtained from the different methods are not found to be robust in their levels and rankings of the companies. A possible explanation for this lack of robustness can be found in the different ability of stochastic frontier methods to account for unobservable heterogeneity. From the methodological point of view the empirical results show that conventional random effects models tend to overestimate cost inefficiency since the inefficiency estimates also contain unobserved heterogeneity. The true fixed effects model recently proposed by Greene (2005a, b) seems to be able to distinguish between unobserved heterogeneity and inefficiency but it may underestimate the inefficiency since all time-invariant effects are treated as unobserved heterogeneity. Therefore, this model seems to contribute, at least partially, to the solution of the problem of time invariant unobserved heterogeneity in the estimation of cost inefficiency. However, the problem is not completely resolved and, as suggested by a referee, the next step in the research would be to explore the possibility to solve this problem using semiparametric and nonparametric methods.

From the policy and regulatory point of view, the lack of robustness of the results suggests that a mechanical use of SFA inefficiency scores results in a price-setting process is not recommended. It is thus recommended to use the benchmarking results only as a starting point for providing information about the range in which the inefficiency scores can be located. Finally, with respect to economies of scale and density the results are more consistent. The estimated economies of scale close to one for the sample median point indicate that medium-sized utilities closely correspond to the optimal size of water distribution utilities in Slovenia. Large utilities are found to operate at levels where diseconomies of scale are already present, while smaller utilities should be interested in expanding their service areas since this would lead to a decrease in average operating costs. Economies of output density and customer density are confirmed for all three different types of utilities with respect to the size of the operation. Therefore, franchised monopolies, rather than side-by-side competition, seem to be the most efficient form of production organization in the water distribution industry.

7. Bibliography

- AERS – Energy Agency of the Republic of Slovenia (2004). Report on the Energy Sector 2003. Maribor.
- Aigner, D.J., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics* 6(1), 21-37.
- Antonioli, B. and Filippini, M. (2001). The use of a variable cost function in the regulation of the Italian water industry. *Utilities Policy* 10, 181-187.
- Bank of Slovenia (2001). Monthly Bulletin 8(1). January 2001, Ljubljana.

- Bauer, P., Berger, A., Ferrier, G., Humphrey, D. (1998). Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. *Journal of Economics and Business* **50**, 85-114.
- Bhattacharyya, A., Harris, T.R., Narayanan, R., Raffiee, K. (1995). Specification and estimation of the effect of ownership on the economic efficiency of the water utilities. *Regional Science and Urban Economics* **25**, 759-784.
- Breyer, F. (1987). The specification of a hospital cost function: a comment on the recent literature. *Journal of Health Economics* **6(2)**, 147-157.
- Cameron, A.C. and Trivedi, P.K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Caves, D.W., Christensen, L.R., Tretheway, J.A. (1984). Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ. *Rand Journal of Economics* **15(4)**, 471-489.
- Cowing, T. and Holtmann A.G. (1983). Multiproduct Short-Run Hospital Cost Functions: Empirical Evidence and Policy Implications from Cross-Section Data. *Southern Economic Journal* **49(3)**, 637-653.
- Daraio, C. and Simar, L. (2005). Introducing Environmental Variables in Nonparametric Frontier Models: a Probabilistic Approach. *Journal of Productivity Analysis* **24**, 93-121.
- Estache, A., Rossi, M.A., Ruzzier, C.A. (2004). The Case of International Coordination of Electricity Regulation: Evidence from Measurement of Efficiency in South America. *Journal of Regulatory Economics* **25(3)**, 271-295.
- Evans, R.G. (1971). Behavioural Cost Functions for Hospitals. *The Canadian Journal of Economics* **4(2)**, 198-215.
- Fabbri, P. and Fraquelli, G. (2000). Costs and Structure of Technology in the Italian Water Industry. *Empirica* **27**, 65-82.
- Farsi, M. and Filippini, M. (2006). An analysis of efficiency and productivity in Swiss hospitals. *Swiss Journal of Economics and Statistics* **142(1)**, 1-37.
- Farsi, M., Filippini, M. and Greene, W. (2005). Efficiency Measurement in Network Industries: Application to the Swiss Railway Companies. *Journal of Regulatory Economics* **28(1)**, 69-90.
- Farsi, M., Fillipini, M. and Kuenzle, M. (2005). Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. *Applied Economics* **37(18)**, 2127-2141.
- Farsi, M., Fillipini, M. and Kuenzle, M. (2006). Cost Efficiency in Regional Bus Companies: An Application of Alternative Stochastic Frontier Models. *Journal of Transport Economics and Policy* **40(1)**, 95-118.

- Filippini, M. (1996). Economies of scale and utilization in the Swiss electric power distribution industry. *Applied Economics* **28**, 543– 550.
- Filippini, M., Hrovatin, N. and Zorić, J. (2004), Efficiency and Regulation of Slovenian Electricity Distribution Companies. *Energy Policy* **32(3)**, 335-344.
- Garcia, S. and Thomas, A. (2001). The Structure of Municipal Water Supply Costs: Application to a Panel of French Local Communities. *Journal of Productivity Analysis* **16**, 5-29.
- Greene, W.H. (2005a). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics* **126**, 269-303.
- Greene, W.H. (2005b). Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis* **23(1)**, 7-32 .
- Guyomard, H. and Vermersch, D. (1989). Derivation of long-run factor demands from short-run responses. *Agricultural Economics* **3**, 213–230.
- Heshmati, A. and Kumbhakar, S.C. (1994). Farm heterogeneity and technical efficiency: Some results from Swedish dairy farms. *Journal of Productivity Analysis* **5(1)**, 45-61.
- Jamasb, T. and Pollitt, M. (2003). International benchmarking and regulation: an application to European electricity distribution utilities. *Energy Policy* **31(15)**, 1609-1622.
- Kim, H.Y. and Clark, R.M. (1988). Economies of Scale and Scope in Water Supply. *Regional Science and Urban Economics* **18**, 479-502.
- Kumbhakar, S.C. (1991). Estimation of technical inefficiency in panel data models with firm- and time-specific effects. *Economics Letters* **36(1)**, 43-48.
- Kumbhakar, S. C. and Lovell, C.A.K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.
- Massarutto, A. (1999). Water management and water prices in Italy. Paper presented at the conference Pricing Water: Economics, Environment & Society, Sintra, Portugal, 6–7 September 1999. European Commission DG XI, Brussels and Instituto da Agua, Portugal.
- McAllister, P.H. and McManus, D. (1993). Resolving the scale efficiency puzzle in banking. *Journal of Banking and Finance* **17**, 389–405.
- Mitchell, K. and Onvural, N.M. (1996). Economies of scale and scope at large commercial banks: evidence from the Fourier flexible functional form. *Journal of Money, Credit, and Banking* **28**, 178–199.
- Meeusen, W. and van den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International Economic Review* **18(2)**, 435-444.

- OFWAT (1999). Final Determinations: Future water and sewerage charges 2000-05. 1998 Periodic Review. Birmingham, UK.
- OFWAT (2004). Future water and sewerage charges 2005-10: Final determinations. 2004 Periodic Review. Birmingham, UK.
- Pitt, M. and Lee, L.F. (1981). The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics* **9**, 43-64.
- Roberts, M.J. (1986). Economies of Density and Size in the Production and Delivery of Electric Power. *Land Economics* **62(4)**, 378-387.
- Rules on Price Determination of Obligatory Local Public Utilities for Environment Protection (2004) [Pravilnik o oblikovanju cen storitev obveznih občinskih gospodarskih javnih služb varstva okolja]. Official Gazette of the Republic of Slovenia 128/2004, Ljubljana.
- Schmidt, P. and Sickles, R.C. (1984). Production Frontiers and Panel Data. *Journal of Business and Economic Statistics* **2(4)**, 367-374.
- Wang, H. and Schmidt, P. (2002). One-Step and Two-Step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels. *Journal of Productivity Analysis* **18**, 129-144.
- Wheelock, D.C. and Wilson, P.W. (2001). New evidence on returns to scale and product mix among US commercial banks. *Journal of Monetary Economics* **47**, 653–674.
- Wilson, P.W. and Carey, K. (2004). Nonparametric analysis of returns to scale in the US hospital industry. *Journal of Applied Econometrics* **19**, 505–524.