Cost inefficiency or just heterogeneity?
An application of stochastic frontier models to the Italian water industry

Graziano ABRATE, Fabrizio ERBETTA, Giovanni FRAQUELLI

Working Paper 3, anno 2008

Fondazione Collegio Carlo Alberto
Via Real Collegio, 30
10024 - Moncalieri (TO)
Tel: 011 670 5250
Fax: 011 6705089
info@hermesricerche.it
http://www.hermesricerche.it

Cost inefficiency or just heterogeneity? An application of stochastic frontier models to the Italian water industry

Graziano Abrate
University of Piemonte Orientale, Faculty of Economics, Via Perrone 18, 28100 Novara
HERMES, Via Real Collegio 30, 10024 Moncalieri (TO)

Fabrizio Erbetta*
University of Piemonte Orientale, Faculty of Economics, Via Perrone 18, 28100 Novara
Ceris-CNR and HERMES, Via Real Collegio 30, 10024 Moncalieri (TO)

Giovanni Fraquelli
University of Piemonte Orientale, Faculty of Economics, Via Perrone 18, 28100 Novara
Ceris-CNR and HERMES, Via Real Collegio 30, 10024 Moncalieri (TO)

Abstract

The question of correctly benchmarking regulated firms operating in different environmental conditions has been extensively debated in the literature. One major problem is the treatment of unobserved heterogeneity and its possible interconnection with structural (persistent) inefficiency. The peculiarity of the reformed Italian water industry, which is based on local authorities defining accurate budget plans over a long period of time, provides a suitable field to test the performance of several frontier models incorporating different specifications for observed and unobserved heterogeneity and efficiency estimates. The results can also shed some light on the consequence of decentralizing efficiency improvements to local authorities and on the potential need to centralize benchmarking activity.

Keywords: Stochastic cost frontier, heterogeneity, cost inefficiency, water industry

JEL Classification C23, D24, G18, L95

* Corresponding Author: Fabrizio Erbetta, University of Piemonte Orientale, Faculty of Economics, via Perrone 18, 28100 Novara, Italy. Tel +39-(0)321-375-440, fax +39-(0)321-375-405, email: fabrizio.erbetta@eco.unipmn.it

1. Introduction

In the last three decades, the theory of regulation has emphasized the relationship between incentive-based mechanisms and efficiency. Regulation schemes based on price-cap and yardstick competition can promote cost reduction and enhance efficiency of regulated monopolies, aiming to replicating the beneficial effects of competition even where it is absent. These high-powered incentive schemes have been largely implemented by policy-makers all around the world to regulate network industries.

The efficacy of these regulatory schemes crucially depends on the actual capability of the regulator to carry out benchmarking activity, that is, to properly compare the current performance of a utility with a performance reference involving similar companies (Shleifer 1985). However, a major problem concerning benchmarking is that utilities serving different areas are likely to face highly differentiated environmental contexts. Neglecting firms' heterogeneity may lead to inaccurate assessments by the regulator because of the effects of either unfavorable or favorable environments on companies' costs.

Identification of a regulatory model able to take into account company heterogeneity when implementing comparative competition models appears to be an important feature in the water industry (Sawkins 1995). OFWAT, the water and sewerage industry regulator in the UK, introduced a price-cap regime following privatization in 1989, whereby price reviews are defined in a flexible way by applying minimum efficiency improvement rates for each company and claiming to be consistent with the firm-specific operating environment (OFWAT 1999, 2004).

In Italy, the current regulation system, introduced in 1994, is based on local regulatory authorities entitled to determine final customer tariffs, to plan and monitor the capital investment programs and the quality levels in the territory they administer. In practice, each agency is required to establish a long-term economic and financial plan of the integrated water service, which then becomes the basic instrument to fix tariffs to the operating companies. Although the investments are remunerated at a given rate, in the same way as standard ROR regulation, the tariffs are fixed ex-ante for a long period based on the local authorities' budget plans. Moreover, local authorities are supposed to provide incentives to recover efficiency based on the comparison of their budget plans with a given benchmarking formula defined at national level. This mechanism more closely resembles a price-cap, at least as far as operating costs are concerned. The effectiveness of such a regulatory system crucially relies on the idea that, under the constraint of a pre-defined benchmarking formula, local authorities would behave virtuously by incorporating adequate rates of efficiency improvements in their plans. On the one hand, it is possible to argue that local authorities may have better knowledge of the specific operating environment, on the other, however, one might contend that such a decentralized regulatory method could make it more difficult to correctly discern between pure environmental effects and inefficiency. Furthermore, local authorities seem to be more exposed to the risk of regulatory capture than a single national authority. Local agencies could, for example, be tempted to accommodate the public interest of water utilities when defining revenues in order to avoid conflicts. The effect that will prevail is certainly not clear a priori.

Given this premise, the Italian case provides a particularly suitable field to study the sensitivity of benchmarking analysis to alternative hypotheses on treatment of firms' heterogeneity. We gathered information from 46 local regulatory plans that typically unfolded over a 20-30 year period, providing rich panel data that also includes information on time-invariant environmental factors (i.e. *observed* heterogeneity). The empirical analysis subsequently proposed in this paper contributes to literature in two key ways. From a methodological point of view, we provide a comparison of several alternative cost frontier models, presenting the impact on the efficiency estimates of their different hypotheses on the error term and showing the effects of either including or excluding specific regressors in the models that account for the observed heterogeneity. In particular, the aim is to discuss which part of the cost differences for each model are attributable to the environmental condition rather than to efficiency, stressing the difficulty of distinguishing between heterogeneity and persistent inefficiency. From a policy point of view, interesting considerations come to light from the analysis of the time-trend of the efficiency component and the evaluation of the incentives to efficiency improvements intrinsically included in plans by local regulators. The results also shed some light on the potential role of ex-post benchmarking by a single national Authority.

The paper is organized as follows. Section 2 depicts the Italian regulatory system as far as territorial organization of integrated water services and price determination are concerned. Section 3 surveys the empirical studies dealing with the problem of measuring efficiency in the water industry. In section 4, we focus on the econometric problem of separating heterogeneity from inefficiency, comparing the theoretical assumptions of various frontier models proposed in the literature. Section 5 presents the specification of the cost function frontier model along with the data and variables used. Estimation results are shown in Section 6, while Section 7 summarizes and concludes.

2. Regulation of the integrated water services in Italy

The water supply system can be divided into three functions: production (abstraction and distribution), sewage collection and sewage disposal. Distribution involves the construction and maintenance of plants including wells, pumps and storage facilities as well as the delivery of water to household and non-household customers using the distribution network. Sewage collection conveys wastewater to treatment plants through pipelines. Finally, sewage disposal processes conveyed wastewater and releases purified water into the environment.

The Italian water industry was highly fragmented. It was composed of around 6,000 actors – almost all directly owned by local public authorities (provinces or municipalities) – that typically operated the distribution function.[1] The average population served by each distribution company was around 9,000 inhabitants (Fabbri and Fraquelli 2000). If, in addition, we consider that the 200 largest firms served around half the population, then the under-sizing of the remaining firms appears even more serious. This situation not only generated inefficiency but also had negative effects on the service quality.

The reorganization of the Italian water industry, which began in 1994 but has not yet been completed, was aimed at favoring new investments and improving both scale and managerial efficiency, attributing functions at national and local level (see Table 1).[2] In line with the declared goal of efficiency improvement, the core element of the regulatory reform consisted in implementing a full-cost pricing principle, defining the costs of the service including the cost of new investments (depreciation and a rate of return on capital investments), wherein unitary operating cost is defined according to a

[1] The other functions, i.e. conveyance and treatment of wastewater, were generally directly operated by local municipalities.

[2] The current regulatory scheme was introduced by the law no. 36/94, whereas the general criterion to determine water tariffs is contained in the Decree 1/8/96 (the so-called *Metodo Tariffario Normalizzato*, MTN). For a detailed chronology of the process of institutional change of the Italian water industry see Goria and Lugaresi (2004).

capping rule. [3] In order to accomplish these tasks, the national territory was divided into 90 Optimal Territorial Areas (henceforth ATOs, *Ambiti Territoriali Ottimali*), based on both hydrographical and political-administrative criteria, entitled to locally administer the integrated water service. Within each ATO, water services must extend beyond the municipal level while integrated service must be assigned to a unique operating company, thus facilitating the attainment of cost benefits due to economies of scale and scope. Each ATO is then subject to the surveillance of local regulatory authorities (the so-called *Autorità di Ambito*), entitled to define infrastructural investment plans and, consequently, to schedule the costs of the service from a long term perspective (usually 20 to 30 years) including both operating and capital costs.

Table 1 Aims and regulatory tasks at national and local level

Aim	National Authority	Local Authorities
Quality	Promote investments establishing a certain rate of return	Define the required investments at the ATO level and monitor their realization
Scale and scope economies	Define the number and sizing of ATOs (*Ambiti Territoriali Ottimali*)	Define economic and financial budget plans, considering the aggregation of the existing productive structures
Managerial efficiency	Give benchmarking guidelines by means of an ex-ante parametric formula	Establish the efficiency improvement rate on the operating cost (and, ultimately, the tariff)

In practice, the model works as follows. The national authority gives a pre-specified parametric formula, used to define a benchmark for the operating cost of each ATO.[4] Thereafter, the local authorities are required to compare their own operating planned costs with the above mentioned benchmark and, accordingly, fix the efficiency improvement rates.[5] Such regulatory arrangement should balance the need to place price

[3] In the reformulation of the price determination system proposed in 2002, but still not enforced, also depreciation and return on capital components should be included in the capping mechanism and therefore subject to the regulatory assessment concerning efficiency improvement rates.

[4] Each function composing the integrated water system has its own formula. The modeled operating costs are then added up in order to obtain the cost of the integrated service. These formulas are not reported here. For a description of modeled costs in water distribution, *see* Antonioli and Filippini (2001).

[5] In more detail, if planned operating costs exceed modeled costs augmented by 20% in a certain year, they must be reduced by at least 2% (based on the planned operating cost of the year immediately preceding); if planned operating costs are inferior to modeled costs augmented by 20%, they must be reduced by at least 1%; finally, if planned operating costs are inferior to modeled costs, they must be reduced by at least 0.5%.

limits that may challenge inefficiency with the one to secure that each company is able to finance their investment programs and carry out their functions without any deterioration of quality (Muraro 2008).[6]

Although the benchmarking criteria are established at national level, the responsibility of defining effective incentives to reduce operating costs is delegated to the local level. Indeed, planned costs and efficiency improvement rates are defined at the same time, with no ex-post benchmarking over local authorities' budget plans. Thus, the model relies on the idea that, under the constraint of a pre-defined benchmarking formula, local authorities will behave virtuously incorporating adequate rates of efficiency improvements and also taking into account the need for case-by-case adaptation to the environmental conditions.

In this paper, we carry out a benchmarking exercise of the local authorities' budget plans in order to evaluate their ability to achieve efficiency improvements. We believe this case study fits interestingly into the debate on the treatment of heterogeneity in benchmarking analysis and the contraposition between centralized and local tasks in pursuing efficiency improvements.

3. Efficiency analysis in the water sector: literature review.

Amongst empirical studies regarding the water industry, two main topics can be found: the analysis of cost characteristics such as economies of scale, scope and/or density and the analysis of cost efficiency. Here we mainly concentrate on the second issue, given that our objective is to evaluate the extent to which Italian local authorities actually succeeded in encompassing cost inefficiency recovery rates when designing their long-term budget plans, in accordance with the declared aims of the national regulatory scheme.

In one of the first studies on water utilities, Bhattacharyia et al. (1995) used a short-run translog cost frontier function to estimate the inefficiency cost of publicly and privately-owned urban US water distributors. Firm-specific variables were included to account for cross-sectional variations in the variable cost and the cost share equations. These variables embraced service quality, system losses, type of input water source and ownership structure. Other unobserved effects entered into the model through additive

[6] The average water loss rate in Italy is between 35-40% while the services of wastewater conveyance and treatment have coverage rates equal to 84% and 70.1% respectively (Utilitatis 2006). The latter two, however, are likely to be overestimated given that they are computed on a basis that does not represent the total population and "equivalent inhabitants". Such deficiencies explain the extensive investment programs the water industry is expected to realize in the future.

one-sided distributed error components. The relevance of firm heterogeneity has also been emphasized by Ashton (2000). Analyzing firm-specific cost efficiency conditions of the UK privatized water and sewerage companies, this study found a moderate dispersion of average inefficiency, indicative of both diversity in operating environments and managerial practices.

Saal et al. (2007) analyzed the productivity of the UK water and sewerage industry using a stochastic input distance function. Productivity growth is decomposed into technical change, efficiency change and scale efficiency change. This study shows that while technical change occurred as a consequence of privatization (occurred in 1989), efficiency improvements did not come about due to excessively lax regulatory price control. The impact of regulation was also examined by Aubert and Reynaud (2005) observing a sample of US water utilities operating in the State of Wisconsin. The particular Wisconsin regulation system, based on the simultaneous presence of price-cap and rate-of-return schemes in the same region at the same time, allowed the authors to compare the effects of the two different regulatory regimes. Using a stochastic cost frontier approach (where the inefficiency error term is modeled as a regulatory type function) they concluded that the most efficient utilities are those operating under a rate-of-return regime[7] and subject to extensive information gathering by the regulator. This emphasizes the importance of the availability of extensive information to establish forceful benchmarking.

Garcia and Thomas (2001) examined the production structure of French municipal water distributors located in the Bordeaux region. They estimated a system of variable cost function and input cost shares using a Generalized Method of Moments (GMM) procedure adapted to panel data. A set of firm-specific characteristics of the served area was accounted for, including the number of metered connections and the number of local communities in the service area. Delivered water and network losses were jointly considered as outputs, which allowed the authors to obtain positive measures of economies of scope between water distribution and network losses.[8]

More recently, Filippini et al. (2008) compared several long-run translog cost frontier models to evaluate the cost efficiency of a sample of Slovenian water distribution

---

[7] A rate-of-return regulation basically consists in letting the firms freely choose their price under the constraint that return on capital should be fair but below a pre-specified level. This method allows prices to increase to cover costs, and in this way, is expected to provide fewer incentives to pursue cost efficiency.

[8] This result indicates that increasing water production while keeping the rate of network losses constant is a preferable option to keeping production constant and improving network efficiency by means of more frequent repairs when facing increases in water demand.

utilities. In order to determine whether unobserved heterogeneity among firms significantly influenced efficiency results, conventional random effects (RE) panel data models were compared with the more recently developed 'true' fixed-effect model (Greene 2005a, b). The latter extends previous panel data models by including an additional time-invariant error term to account for unobserved heterogeneity. The results highlight that while conventional RE panel data models seem to overestimate cost inefficiency, since the inefficiency estimates include all time-invariant firm-specific characteristics (improperly labeled as inefficiency because they are out of managers' control), the 'true'-fixed effects model seems to underestimate cost inefficiency since all time-invariant factors, including the time-invariant structural inefficiency, are purged out of the inefficiency term and treated as heterogeneity. As for the production structure, delivered water was used as output while the costs of labor, capital and other materials were included on the input side. Output characteristics (namely, number of customers and size of service area) and several dummy variables representing the percentage of network losses and the types of water source, entered the function in order to control for observed heterogeneity.

An alternative method to the stochastic approaches for modeling technology and assessing efficiency is *Data Envelopment Analysis* (DEA). This approach is widely used by regulators all around the world to regulate water services. A comprehensive description of the use of DEA for regulatory purposes is provided in Thanassoulis (2000a, b), where DEA methodology was adopted with the aim of estimating potential cost savings of the distribution function of the UK water utilities. In these studies, several output and network characteristics, that might have an effect upon firms' costs, were directly introduced in the DEA-based linear programming problems. The variables were number of customers, length of distribution network and the annual number of burst pipelines. Other applications of DEA can be found in Tupper and Resende (2004) referring to the Brazilian water and sewerage industry and in Coelli and Walding (2006) referring to the use of this methodology to set the efficiency improvement rates (the so-call *X*-factor in the price-cap regulation formula) for the Australian water industry. In particular, the former study provides a second stage correction of the DEA efficiency scores in order to account for regional operational heterogeneity. Finally, De Witte and Marques (2007 and forthcoming) provide an international benchmarking of drinking water industries using DEA-based bootstrapping algorithms and a metafrontier approach. In these studies applications to account for environmental variables in the efficiency measurement are also presented[9]

Overall, the studies highlight the relevance of controlling cost efficiency for variations in environmental characteristics faced by different utilities in their daily operations. This paper tries to account for this issue by comparing stochastic cost frontier models that differently disentangle firm heterogeneity and cost inefficiency, using panel data from budget plans established by local Italian regulation authorities. The details concerning properties, advantages and disadvantages of the models compared in this study are presented in the next Section.

4. The econometric model

It is well known that, for regulatory purposes, various techniques for estimating and comparing efficiency levels can be used to provide regulated firms with the incentive to improve their performance and move towards a best-practice. However, if the efficiency scores are obtained without properly taking into account the heterogeneity in the environmental conditions, the regulator may end up providing misleading incentives and undesirable effects. The econometric problem is how to isolate the managerial inefficiency component from the random noise and the firm-specific environmental conditions that may affect cost performance. In general, a variable cost frontier model may be written as follows:

$$C_{it} = \alpha_0 + c(Y_{it}, P_{it}, K_{it}, Z_i, t) + \varepsilon_{it} \;, \quad \text{with } \varepsilon_{it} = \alpha_i + \lambda_i + u_{it} + v_{it} \qquad (1)$$

where i (i = 1,…I) denotes firm and t (t = 1,…, T) time, C is the variable operating cost, Y is the vector of outputs, P is the vector of variable input prices, K is the quasi-fixed input, Z is a vector of *observed* firm-specific environmental conditions (invariant over time) and t is the time trend which reflects technological change.

The *unobserved* firm-specific characteristics are represented by two distinct components: α_i is the environmental component which reflects cost differences intrinsically due to the territorial area where the firm operates, while λ_i can be interpreted as a component of persistent inefficiency, due to the (low or high) innate

[9] DEA is a deterministic approach in the sense that the total deviation from the estimated frontier is interpreted as inefficiency. Simar and Wilson (2000, 2007) introduced a bootstrapping model to tackle this problem, thus making statistical inference possible.

aptitude of firm managers (or, more generally, the innate ability of the firm shareholders to select good managers from the market). In terms of policy, it is worth stressing the possibility of separating these two components: while firms operating in adverse environmental conditions should not be penalized by the incentive schemes, structural inefficiency, on the contrary, can and should be deterred by means of regulation. For example, managers' entrenchment or, also, the persistence of managers' selection procedures not primarily driven by efficiency purposes are likely to systematically harm efficiency. In such cases, incentive regulation schemes may lead to the removal of these structural sources of inefficiency.[10] The other two error terms in Eq. 1 are respectively the time-varying (non-negative) component of inefficiency (u_{it}) and the conventional random noise ($v_{it} \sim i.i.d.\ N(0, \sigma^2)$).

Ideally, a regulator would need to insert in the price-cap formulation an efficiency recovery factor strictly connected to a measure including both time-varying and persistent components of inefficiency. In our case the benchmarking exercise concerns authorities' budget plans that are the basis for the tariff calculation (see Section 2). Therefore, such a measure represents the level of inefficiency implicitly allowed by the regulatory authorities of a certain ATO *i* during each period *t*. Moreover, the study of the dynamics of these levels of inefficiency would allow understanding the power of incentives introduced in order to recover efficiency.

Unfortunately, a complete solution for the econometric estimation of the model in Eq. 1 is not available, and further assumptions need to be added to the various components described above. A particularly clear analysis of the hypotheses of various frontier models is given by Filippini et al. (2008). In the present study, the alternative assumptions of seven different frontier models were discussed and Table 2 provides a summary.

One first option is given by the classical frontier model proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) (MODEL I, POOLED). In this case, the problem of unobserved heterogeneity is not directly handled (the terms α_i and λ_i are implicitly assumed to be equal to 0), and the model can be estimated using the maximum likelihood (ML) method following an apposite distributional assumption made on the

[10] It could be argued that in the long run persistent inefficiency should not be an issue since the managerial board will certainly change. Nevertheless, if managers are selected not only according to financial motivations but also in order to satisfy other stakeholders' goals (for example, political), then this may be another source of persistent inefficiency.

one-sided and non-negative inefficiency term u_{it}.[11] The crucial hypothesis of the model is that all the sources of firm heterogeneity can be observed and included in the term Z_i, otherwise, both the efficiency estimates and the parameters of the cost function may be potentially biased.

Table 2 Alternative estimators for Model (1) and relative assumptions

Model	Hypothesis	Cost inefficiency estimation (w_{it})
MODEL I – POOLED	α_i=0; λ_i=0 $u_{it} \sim i.i.d.\ N^+(0, \sigma_u^2)$	$E(u_{it} \mid \hat{\varepsilon}_{it})$
MODEL II – RE-GLS	α_i=0; u_{it}=0; $\lambda_i \sim i.i.d.\ N^+(0, \sigma_\lambda^2)$	$\hat{\lambda}_i - \min\{\hat{\lambda}_i\}$
MODEL III – RE-ML	α_i=0; u_{it}=0 $\lambda_i \sim i.i.d.\ N^+(0, \sigma_\lambda^2)$	$E(\lambda_i \mid \hat{\varepsilon}_{i1},\ldots,\hat{\varepsilon}_{iT})$
MODEL IV – FE	α_i=0; u_{it}=0 λ_i fixed (group dummies)	$\hat{\lambda}_i - \min\{\hat{\lambda}_i\}$
MODEL V – BC	α_i=0; $\lambda_i \sim i.i.d.\ N^+(0, \sigma_\lambda^2)$ $\exp[-\eta(t-T_i)] = \beta(t)$	$E(\lambda_i\beta(t) \mid \hat{\varepsilon}_{i1},\ldots,\hat{\varepsilon}_{iT})$
MODEL VI – TFE	α_i fixed (group dummies); λ_i =0; $u_{it} \sim i.i.d.\ N^+(0, \sigma_u^2)$	$E(u_{it} \mid \hat{\varepsilon}_{it})$
MODEL VII – TRE	$\alpha_i \sim i.i.d.\ (0, \sigma_\alpha^2)$; λ_i =0 $u_{it} \sim i.i.d.\ N^+(0, \sigma_u^2)$	$E(u_{it} \mid \hat{\varepsilon}_{it})$

An alternative specification is given by the random-effects (RE) model. In this case, the entire inefficiency is forced to be time-invariant, thus u_{it} is assumed equal to 0. Two possible estimation procedures have been suggested. Schmidt and Sickles (1984) applied a GLS procedure to the standard random effects panel specification and derived the inefficiency of each firm as the difference between the firm-specific random effects and the minimum value assumed by λ_i, i.e. the most efficient firm (MODEL II, RE-GLS).[12] Pitt and Lee (1981) estimated the model with the ML method assuming λ_i distributed as half-normal (MODEL III, RE-ML). In these models, *all* unobserved heterogeneity among firms is assumed to be inefficiency (therefore, α_i=0). Thus, RE models impose even stricter assumptions on the inefficiency term with respect to the pooled model. Nevertheless, the main advantage with respect to MODEL I is that the firm-specific effects are explicitly modeled and as such – at least as long as the latter are

[11] Possible alternatives to model u_{it} are the half-normal, exponential, truncated-normal and gamma distribution. In this study we used the half-normal assumption.
[12] It should be noted that in the RE-GLS model, while the term λ_i is symmetrically distributed, the estimate of cost efficiency is always non-negative.

not correlated with the regressors – allow obtaining unbiased and efficient estimates of the cost function parameters.

A similar approach is proposed by Schmidt and Sickles (1984) based on fixed-effects estimation (MODEL IV, FE). Again, α_i and u_{it} are assumed equal to 0. With respect to RE models, this approach allows obtaining unbiased estimates of the cost function parameters even if firm-specific effects are correlated with the regressors. On the other hand, the FE model can produce biased estimates of the inefficiency scores due to the incidental parameter issue, especially if the number of time periods is limited.[13]

In the approach proposed by Battese and Coelli (1992), within a panel data specification, the inefficiency is allowed to vary over time (MODEL V, BC). The model implicitly assumes $\alpha_i = 0$ and the distributional assumption of λ_i is the same as in Pitt and Lee (1981). The Battese and Coelli (1992) model specifies that the inefficiency is modeled as the product of an exponential function of time and non-negative firm-specific random variables. In this model, η is an unknown parameter that defines the trend of the inefficiency over time. This model, however, does not allow firm-specific patterns of temporal change of cost inefficiency.[14]

A specific treatment of the variable α_i is given in Greene (2005a, b), who proposed two specifications named 'true' fixed-effects (MODEL VI, TFE) and 'true' random-effects (MODEL VII, TRE) models to deal with the problem of unobserved heterogeneity.[15] In the first case, the firm-specific effects are modeled with group dummy variables, while in the second case α_i is distributed as a random variable. In both cases, the inefficiency is modeled as time-varying, with u_{it} distributed as a half-normal random variable. The major weakness of these models is that *all* time-invariant unobservable effects, and thus any potential persistent component of inefficiency, are included in the term α_i. In other words, in these cases, the implicit assumption is that $\lambda_i=0$. In addition, in the case of the TFE model, one cannot include among the regressors any time-invariant factor (Z_i). This is not an issue in the TRE model which allows time-invariant factors to be included in the specification of the cost function. This feature may also allow a further interpretation of TRE results, since if it is assumed that all the environmental (time-

[13] Wang and Ho (2007) developed a model, based on first-difference and within-transformation of regressors, to solve the incidental parameters problem.

[14] A positive value of η parameter means that inefficiency (u_{it}) decreases over time while e negative value means that inefficiency (u_{it}) increases over time. The λ_i term could be interpreted as the inefficiency of the firm i when $t = T$, i.e. firm i is observed in the last period of the time series.

[15] Application of these models may be found, for example, in Farsi et al. (2005a, b and 2006a, b) and Filippini et al. (2008)

invariant) heterogeneity were observable and measurable through the Z_i variables, then the firm-specific error term, λ_i, would actually represent the persistent component of cost inefficiency (i.e. α_i=0 and $\lambda_i \geq 0$ with $\lambda_i \sim N^+(0, \sigma_\lambda^2)$).

Once point estimates of cost inefficiency (w_{it}) are obtained (in the ways displayed in Table 2), estimates of cost efficiency, *CE*, for each firm in each period can be derived as:

$$CE_{it} = e^{-\hat{w}_{it}} = C*_{it} / C_{it} \tag{2}$$

where $C*_{it}$ denoted the minimum cost frontier for firm *i* in time *t*. Such a measure is comprised between 0 and 1, equal to one when the firm is on the frontier. In this case, 1-CE_{it} indicates a standardized measure of cost inefficiency.

5. The cost function specification

A translog functional form was chosen for the estimation of the variable cost function frontier, as showed in the following equation[16]:

$$\begin{aligned}\ln C_{it} &= \beta_0 + \sum_{m\in M}\beta_m \ln Y_m + \sum_{r\in R}\beta_r \ln P_r + \beta_K \ln K + \frac{1}{2}\sum_{m\in M}\sum_{n\in M}\beta_{mn}\ln Y_m \ln Y_n + \\ &+ \frac{1}{2}\sum_{r\in R}\sum_{s\in R}\beta_{rs}\ln P_r \ln P_s + \frac{1}{2}\beta_{KK}\ln K^2 + \frac{1}{2}\sum_{m\in M}\sum_{r\in R}\beta_{mr}\ln Y_m \ln P_r + \\ &+ \frac{1}{2}\sum_{m\in M}\beta_{mK}\ln Y_m \ln K + \frac{1}{2}\sum_{r\in R}\beta_{rK}\ln P_r \ln K + \gamma_t t + \sum_{q\in Q}\beta_q Z_q + \varepsilon_{it}\end{aligned} \tag{3}$$

where *M* indicates the output set, *R* the input price set and Q the environmental characteristics set. The output set includes the volume of water input into the distribution systems (Y_V) and the inhabitants served by the depuration service (Y_D). The first output was preferred over the volume of delivered water as the main cost driver for variable operating costs. Indeed, doubts over the reliability of delivered volume predictions were raised, since the latter seemed inconsistent with the tariff dynamic (Abrate and Fraquelli 2007). The second output reflects the degree of coverage of the depuration service within each ATO. A third output concerning the sewage collection should have been, in principle, included here. One possibility was to use the length of the sewage collection pipelines, given the absence in our dataset of any information on the population served. This was however excluded due to the high correlation with Y_D

[16] The planning of the integrated water service over a long period is based on exceptional investment programs that do not allow sustaining the hypothesis that capital is employed at its optimal level. Therefore, the adoption of a long-run cost function would seem an inappropriate option.

(Pearson correlation = 0.98). On the input side, the input price set consists of two variables, the price of labor (P_L) and the price of other variable inputs (P_M). The price of labor was obtained by dividing the cost of labor by the number of employees. The price of other variable inputs was obtained by dividing the non-labor variable costs (essentially costs for materials and services) by the number of inhabitants of the served area.[17] One peculiarity of our study is that the data was planned in real terms and therefore, did not require deflation. In addition, input prices were kept constant over time, in accordance with the hypothesis of the authorities. [18]

As for the capital variable (K), very detailed data was available on the monetary investment requirements for each year of the plan and the amount of depreciation. A permanent inventory method was therefore applied. The principal difficulty concerned the construction of a value of the initial stock of capital defined coherently for all the ATOs. From the analysis of the plans and after directly contacting several authorities, we decided to follow a procedure based on the valorization of the networks (both distribution and sewerage network) at current construction prices, adjusting the resulted value on the basis of the average age of the infrastructures.[19]

Finally, the vector Z contains a series of time-invariant environmental factors that can be supposed to affect the level of operating costs. In particular, we decided to build a series of binary variables based on:

- service area extension (Z_1 = 1 if the extension exceeds the sample mean; 0 otherwise);
- morphology of the territory in terms of percentage of highlands (Z_2 = 1 if the percentage of highlands exceeds the sample mean; 0 otherwise);
- number of municipalities (Z_3 = 1 if their number exceeds the sample mean; 0 otherwise)
- geographical location (Z_4 = 1 if the ATO is located in the Centre-North, 0 if it is located in the South).

[17] Two alternative measures of P_M were constructed, using as denominators respectively the length of the network and the volume of delivered water. Main results are not affected by the choice of the P_M variable.

[18] Given the difficulty of forecasting the trend of prices in the future, all the authorities analyzed made their plans in real terms. In addition, they all kept the labor price constant over the whole length of the financial plans. Coherently with this approach, therefore, the price of other variable inputs were computed for each ATO only for the first year of the plan and then left constant over time.

[19] This procedure was applied in those ATOs that presented more detailed data on the value of other plants and infrastructures. It turned out that the estimated value of the two networks (distribution and sewerage) was always around the 80-90 per cent of the whole value of the capital. This allowed us to confidently proceed with this simplification, reconstructing the value of the stock with relatively few data (the price of a new kilometer of network, the average age of the network and the kilometers of network at year one), which was requested directly to the ATOs when not available from the plans.

Clearly, in the fixed effects models (MODEL IV and MODEL VI), these variables (as well as all the regressors containing prices) were removed due to perfect collinearity issues.

Some descriptive statistics of the variables used for the empirical analysis are presented in Table 3. The dataset refers to 46 authorities' budget plans with variable duration (from a minimum of 12 years to a maximum of 30 years), leading to 1,115 total observations. In the average ATO, the volume of water input is around 90 million cubic meters, while the depuration service supplies around 630,000 inhabitants, with operating costs amounting to more than 50 million Euro. As for the environmental variables, 29% of the observations are characterized by an extension of the service area above the sample mean, whereas slightly fewer than 40% of the observations operate in territories with proportions of highland and urban concentrations above the sample mean. Finally, 41% of the observations are located in the Centre-North area. For the empirical analysis, the variables were standardized over their geometric mean. Moreover, in order to impose the homogeneity of degree one in inputs, the normalization of cost and prices over the price of materials was applied.

Table 3 Summary statistics

Variable	Mean	Std. dev.	Min	Max
C *(000 €)*	51,315	52,549	4,824	276,054
Y_V *(000 m^3)*	90,671	99,079	9703	479,200
Y_D	632,354	693,992	70,754	4,220,000
P_L *(€ / Employee)*	39,096	3,676	28,754	44,887
P_M *(€ /Inhabitant)*	48.01	11.50	25.20	78.70
K *(000 €)*	793,110	696,575	90,528	6,058,642
Z_1	0.29	0.45	0	1
Z_2	0.37	0.48	0	1
Z_3	0.39	0.49	0	1
Z_4	0.41	0.49	0	1

6. Results

The estimated parameters from the different specifications of the stochastic cost frontier function described in Eq. 3 are presented in Table 4. The parameters β_V and β_D are positive and generally highly significant, implying that operating costs are increasing in outputs. Worthy of notice is that both coefficients remain fairly stable across all frontier models except in the pooled model, which is the only model that does not account for individual effects. This suggests that the cost function coefficients can be biased if heterogeneity is not explicitly modeled. The parameter β_V, the output elasticity at the sample mean for water distribution, is considerably larger than β_D, the output elasticity at the sample mean for wastewater disposal service, thus indicating that the most important cost driver is the physical water input into the pipelines. This may be explained by a major incidence of fixed costs in the wastewater treatment phase.

The β_L parameter associated with the price of labor is positive and significant, thus meaning that an increase in the price of labor, leaving the amount of output unchanged, produces an increase in variable costs.[20]

The estimated coefficient for capital (β_K) is always negative and statistically different from zero. This is consistent with the economic theory and signals that an increase in capital endowment allows firms to reduce variable costs.

The time trend is positive (except in MODEL I and VII) even though the magnitude of the γ_t parameter is very low (and significantly different from zero in MODEL IV-V-VI). The stability of the technological change seems consistent with the nature of the dataset. The regulators' planning of activities characterizing the water and sewerage industry is based on the technological features existing at the initial time of the planned period. Therefore, the absence of any technological progress/regression does not seem surprising.

As for the environmental factors, which, as already noted, represent sources of observed heterogeneity, the signs of the dummy variables are, in general, as expected and comparable among models.[21] Consistently with many studies (see, for instance, Ashton 2000; Garcia and Thomas 2001; Filippini et al. 2008), operating costs, in general, depend positively and significantly upon the extension of the service area (β_1) and the number of municipalities (β_3). The percentage of highlands (β_2) influences costs negatively and significantly, thus indicating that higher expected costs for maintenance in highland areas are probably offset by the proximity to the water sources.[22] Likewise, the geographical dummy shows a negative and statistically significant sign, thus

[20] The concavity in input prices is one of the theoretical properties of the cost function. Using the estimated parameters β_L and β_{LL} it is possible to immediately compute the second partial derivative of cost with respect to the input price at the sample mean. In this point concavity is satisfied for models I, II, III and V. In models IV and VI input prices were removed because of multicollinearity problems. Thus, differentiability of the cost function with respect to input prices is not applicable in these cases.

[21] Other dummy variables, such as type of water sources (i.e. boreholes or surface sources) and percentage of network losses in the initial year were tested, but they did not prove statistically significant.

[22] It could also be argued that an operating environment characterized by higher percentage of highlands would imply higher expenses for capital infrastructures due to the major difficulty of realizing water facilities. Such an effect, however, is not accounted for in our short-run variable cost function.

denoting a structural shortfall in southern Italy, with respect to northern Italy, which might be attributed to the different status of the network and other capital facilities. This highlights the high penalization suffered by the southern area in terms of major maintenance and intervention costs.

Table 4 Estimated parameters of the stochastic cost frontier function

Var.	Par.	MODEL I *POOLED*	MODEL II *RE-GLS*	MODEL III *RE-ML*	MODEL IV *FE*	MODEL V *BC*	MODEL VI *TFE*	MODEL VII *TRE*
$\ln Y_V$	β_V	0.323*** (14.00)	0.757*** (29.57)	0.749*** (26.27)	0.660*** (19.27)	0.778*** (33.58)	0.830*** (38.3)	0.617*** (99.70)
$\ln Y_D$	β_D	0.660*** (20.41)	0.062*** (3.72)	0.051*** (3.08)	0.018 (1.04)	0.076*** (4.56)	0.031* (1.80)	0.048*** (7.73)
$\ln P_L$	β_L	0.388*** (16.48)	0.352*** (3.48)	0.283** (2.31)		0.175*** (3.29)		0.112*** (15.34)
$\ln K$	β_K	-0.096*** (-4.18)	-0.050*** (-3.84)	-0.052*** (-4.13)	-0.059*** (-4.58)	-0.066*** (-5.06)	-0.033** (-2.45)	-0.131*** (-2.54)
$(\ln Y_V)^2$	β_{VV}	-0.018 (-0.37)	-0.020 (-0.83)	-0.048** (-2.01)	-0.133*** (-4.71)	-0.039 (-1.42)	-0.059** (-2.28)	-0.112*** (-10.16)
$(\ln Y_D)^2$	β_{DD}	-0.201*** (-3.81)	-0.099*** (-4.62)	-0.106*** (-5.17)	-0.128*** (-6.05)	-0.060*** (-2.72)	-0.103*** (-4.93)	-0.073*** (-5.61)
$(\ln P_L)^2$	β_{LL}	-0.113 (-1.57)	0.017 (0.06)	-0.161 (-0.82)		0.066 (0.61)		0.112*** (4.67)
$(\ln K)^2$	β_{KK}	0.167*** (4.28)	-0.081*** (-5.50)	-0.078*** (-5.53)	-0.087*** (-6.09)	-0.096*** (-6.64)	-0.077*** (-5.29)	-0.044*** (-7.24)
$\ln Y_V \ln Y_D$	β_{VD}	0.254*** (2.58)	0.043 (1.28)	0.054* (1.68)	0.075** (2.23)	-0.037 (-0.97)	0.044 (1.31)	0.014 (0.67)
$\ln Y_V \ln P_L$	β_{VL}	-0.091 (-1.12)	-0.494*** (-4.90)	-0.943*** (-8.35)		-0.671*** (-5.40)		0.305*** (19.74)
$\ln Y_V \ln K$	β_{VK}	-0.090 (-1.39)	0.043* (1.73)	0.046* (1.91)	0.077*** (3.25)	0.090*** (3.51)	0.040* (1.68)	-0.017* (-1.85)
$\ln Y_D \ln P_L$	β_{DL}	0.177 (1.42)	-0.064 (-1.46)	-0.057 (-1.36)		-0.016 (-0.39)		-0.025 (-1.15)
$\ln Y_D \ln K$	β_{DK}	-0.091 (-1.39)	0.102*** (4.27)	0.099*** (4.35)	0.094*** (4.08)	0.093*** (4.03)	0.101*** (4.34)	0.087*** (7.82)
$\ln P_L \ln K$	β_{LK}	0.238*** (3.08)	0.106* (1.91)	0.084* (1.60)		0.044 (0.84)		0.276*** (13.96)
t	γ_t	-0.008*** (-11.79)	0.0002 (0.80)	0.0004 (1.41)	0.0008*** (2.60)	0.001*** (3.53)	0.001** (2.25)	-0.0004 (-0.347)
Z_1	β_1	0.145*** (8.13)	0.248*** (3.04)	0.193 (1.44)		0.406*** (13.29)		0.063*** (12.47)
Z_2	β_2	-0.020 (-1.50)	-0.142** (-2.25)	-0.200*** (-2.51)		-0.358*** (-7.93)		-0.316*** (-7.65)
Z_3	β_3	-0.053*** (-3.17)	0.012 (0.18)	0.232*** (2.45)		0.068 (1.37)		0.214*** (3.98)
Z_4	β_4	-0.042*** (-3.43)	-0.054 (-0.98)	-0.123** (-2.23)		-0.056** (-2.11)		-0.413*** (-8.58)
Constant	β_0	-0.056*** (-2.80)	-0.066 (-1.32)	-0.466*** (-6.24)	0.080*** (4.84)	-0.538*** (-15.11)		-0.168*** (-35.77)
σ_v		0.111		0.012		0.011	0.032	0.032
σ_u		0.165		0.254		0.197	0.026	0.034
LogL		550.88		2002.70		2011.24	2134.08	1905.57

Notes: t-statistics in brackets
*** Statistically significant at 1%; ** statistically significant at 5%; * statistically significant at 10%

Focusing on the main object of our study, we first derived the cost efficiency estimates from a reduced model excluding dummy variables.[23] We then analyzed the modification of the cost efficiency estimates when including – in the extended model – the set of environmental dummy variables in order to infer some conclusion on the role of inefficiency or unobserved heterogeneity in explaining excess cost.

Table 5 presents the cost efficiency estimates provided by the different models (without dummy variables) using Eq. 2. Average cost efficiency ranges from 0.505 to 0.578 for MODEL II to IV.[24] Such approaches model inefficiency by means of time-invariant effects that may confound the excess cost due to managerial responsibility with that associated with external factors outside managers' control. Moreover, in the case of the FE model, which registers the lowest level of efficiency (0.505), the price effect too (which are time-invariant in our database) cannot be separated from the inefficiency term because of multicollinearity issues. As previously mentioned, these models have been criticized as they risk underestimating cost efficiency. MODEL V, the Battese and Coelli 1992 model, differs from previous RE panel data models as it relaxes the restriction on time-invariant cost inefficiency by introducing a temporal trend in the inefficiency term, following the exponential formula reported in Table 2. The estimated η-parameter, that defines the efficiency change over time, is equal to 0.00072 (p-value = 0.183), corresponding to a very low annual growth rate of cost efficiency of around 0.07% over an average time period of 24 years. Thus, local authorities did not incorporate any efficiency improvement in the operating costs in their plans. This result is in contrast with one of the declared goals of the regulatory reform (see Section 2), and contributes to engendering high tariffs for consumers.

The last two columns in Table 5 provide the efficiency estimates using the 'true' fixed effects and 'true' random effects models, which introduce individual time-invariant effects in order to elicit all unobserved heterogeneity separately from a time-varying inefficiency term. As already noted, these models yield efficiency estimates that can be interpreted as higher-bound, since their drawback is that persistent inefficiency cannot be disentangled from other environmental firm-specific effects. Indeed, the average cost efficiency rises to 0.979, with minimum values at 0.92 and 0.93 in the TRE and TFE

[23] The estimated parameters of the translog cost function frontier without considering dummy variables are very similar to those shown in Table 4 and thus are omitted here. However, results are available upon request to the authors.

[24] MODEL I (pooled model) provides a much higher average value, equal to 0.870. Differently from panel models, in this case, panel structure is not accounted for and, therefore, time-invariant effects given by structural inefficiency and unobserved heterogeneity are not properly modeled.

models, respectively. This evidence, which is consistent with the stable efficiency trend noted above, highlights either that the differentials in cost performance across ATOs are almost entirely due to environmental conditions (thus meaning that managerial inefficiency is not an issue in the Italian water industry), or that managerial inefficiency has a persistent nature which makes it inextricably confounded in the heterogeneity term. While it seems implausible to attribute the massive differences between efficiency estimates solely to the environmental configuration, we are more prone to accept the assertion that most cost inefficiency allowed in the local authorities' budget plans is of structural type.

Table 5 Cost efficiency estimates

	MODEL I *POOLED*	MODEL II *RE-GLS*	MODEL III *RE-ML*	MODEL IV *FE*	MODEL V *BC*	MODEL VI *TFE*	MODEL VII *TRE*
Mean	0.870	0.578	0.567	0.505	0.567	0.979	0.979
Std- dev	0.063	0.144	0.187	0.189	0.184	0.007	0.007
Min	0.560	0.312	0.195	0.172	0.206	0.929	0.919
Max	0.972	1	0.992	1	0.991	0.994	0.994

The results from the reduced model have been compared with those obtained from the extended model in Table 6. The idea is testing the sensitivity of the efficiency estimates when including among the covariates – without claiming it to be exhaustive – a set of dummy variables (the vector Z) able to capture the effect of several environmental factors. In theory, should all the environmental sources of heterogeneity be observable and included in Z, the α_i-term would be equal to 0 and this would allow correctly estimating the structural cost inefficiency. However, even if it is not possible to account for any source of heterogeneity in practice, it could be argued that the higher the number of environmental variables observed, the lower the impact of α_i-term on operating costs. Moreover, in order to provide a better comparison of the results across MODEL I to VII, we derived for all cases an overall cost performance score due to the impact of all the *unobservable* terms (α_i, λ_i and u_{it}) excluding the random noise. The overall score was then split into a time-varying (u_{it}) and a time-invariant component (α_i and λ_i).[25] Notice that the estimated value of the time-invariant component can be affected by both

[25] More precisely, in the specific cases of TFE and TRE the time-invariant cost performance is given by $e^{-[\hat{\xi}_i - \min\{\hat{\xi}_i\}]}$, where $\xi_i = \alpha_i + \lambda_i$. Similarly, the overall cost performance is given by $e^{-\hat{u}_{it} - [\hat{\xi}_i - \min\{\hat{\xi}_i\}]}$.

inefficiency and environmental effects, and its interpretation is in fact different across models. Furthermore, the set of dummy variables is expected to modify, to some extent, the estimation of such time-invariant (inefficiency or heterogeneity) cost components.

In the pooled model, the panel dimension of the dataset is not taken into account and the efficiency levels remain fairly stable with and without dummy variables. In the conventional RE models (Model II and III) – where a time-varying term is not explicitly accounted for and the time-invariant term is entirely interpreted as cost inefficiency – the average time-invariant cost performance estimated when environmental dummy variables are accounted for (columns *b*), appears sensibly higher than the corresponding average values estimated excluding dummy variables (columns *a*). Similar results are also obtained in the BC model, where the inclusion of dummy variables raises the time-invariant average performance to 0.743.[26] These results may be explained by the fact that the time-invariant error terms (λ_i) partly include environmental effects that would be better labeled as heterogeneity rather that inefficiency. When removing such environmental effects (or at least part of them) the efficiency rises.

Table 6 Composition of overall cost performance with and without environmental dummy variables

	Model I *Pooled*		Model II *RE-GLS*		Model III *RE-ML*		Model IV *FE*	Model V *BC*		Model VI *TFE*	Model VII *TRE*	
	(a)	*(b)*	*(a)*	*(b)*	*(a)*	*(b)*	*(a)*	*(a)*	*(b)*	*(a)*	*(a)*	*(b)*
Time-varying cost perform.										0.979	0.979	0.973
Time-invar. cost perform.			0.578	0.641	0.567	0.714	0.505	0.576	0.743	0.568	0.507	0.605
Overall cost perform.	0.870	0.877	0.578	0.641	0.567	0.714	0.505	0.567	0.735	0.558	0.499	0.599

(a) estimated values from models *without* environmental dummy variables
(b) estimated values from models *with* environmental dummy variables (it was not possible to perform estimations for Model IV and Model VI given the time-invariant nature of the environmental covariates)

In the case of the TRE model, the introduction of a set of dummy variables does not change the estimated time-varying cost performance (time-varying efficiency). This seems correct as the inclusion of various sources of observed heterogeneity into the cost

[26] The Battese and Coelli 1992's model is quite different from the other conventional RE models as the time-invariant (λ_i) term corresponds to the cost inefficiency in the *last* year of the observed period. The model then computes the year-by-year inefficiency using the exponential formula described in Table 2. In such a way, it is possible to compute overall cost performance measures both with and without dummy variables. The average values are slightly lower than the corresponding time-invariant terms. This is due to the mild positive efficiency trend observed over the whole period.

function is only expected to influence the time-invariant term. In the TRE model, the overall (and undistinguished) time-invariant term $\alpha_i+\lambda_i$ is interpreted as excess cost due to unobserved heterogeneity. In this formulation, however, we point out that such an assumption may lead to wrong conclusions if cost inefficiency is mostly persistent. As can be seen in Table 6, when including the set of dummy variables, the average time-invariant cost performance passes from 0.499 to 0.599. This result suggests that in the extended random effects model the dummy variables explain, at least partly, the unobserved heterogeneity. Therefore, their inclusion reduces the α_i component and contributes to better isolate the λ_i term. Since the performance score increases by 0.1, but it still remains quite low[27], we argue that a not negligible part of the unobserved time-invariant cost differences across ATOs may be reasonably attributed to persistent inefficiency. In any case, the results of our analysis cast some doubt on the ability by local authorities to undertake proper actions to promote efficiency and reduce structural gaps, suggesting the opportunity of an ex-post benchmarking control by a centralized national authority.

7. Conclusions

In this study, we analyzed the cost efficiency encompassed in local regulators' long-term financial plans concerning water and sewerage functions in Italy by estimating several cost frontier models. Specifically, the available dataset is particularly suitable to provide a comprehensive comparison of different panel data models in order to infer some conclusions on the sensitivity of results to alternative specifications of (observed and unobserved) heterogeneity. We compared traditional cost frontier models based on random and fixed effects approaches with the more recently proposed 'true' fixed and random effects models. While the first category assumes that all cost differences due to (unobserved) time-invariant factors are interpreted as inefficiency, the second category introduces a time-varying inefficiency (one-sided) error component and includes all (unobserved) heterogeneity due to environmental specificities into a separate term.

Results indicate that average cost efficiency differs substantially across models, in accordance with the underlying assumptions. The estimated values range from 0.50-0.60 for standard panel models to 0.98 for the TRE and TFE models. Reasonably, these

[27] In general, we acknowledge that potential omissions could have occurred here. On the other hand, the large excess cost remaining after adjusting for dummy variables makes it implausible to think it might be explained uniquely by omitted environmental factors.

values provide, respectively, the lower and higher bound for the "true" managerial efficiency that will be closer to the lower or the higher bound depending on the magnitude of persistent inefficiency.

We attempted to account for the influence of observed heterogeneity, modeled with a series of environmental dummies, in order to elicit more accurate estimates of cost efficiency. As expected, the inclusion of firm-specific characteristics leads to a reduction of cost inefficiency in traditional panel models (II-III-V), while it implies a reduction of performance gaps due to unobserved heterogeneity in the TRE model. Nevertheless, apart from the different interpretation in terms of heterogeneity or inefficiency, the impact of time-invariant effects on the average cost performance remains substantial.

From a policy point of view, our findings lead to conjecture that the largest part of managerial inefficiency is of a structural nature and local authorities may be unable to disentangle it from environmental effects. In addition, the observation of the stable efficiency trend resulting from the BC model indicates that local authorities' plans do not incorporate proper incentives to improve efficiency on the operating costs, in contrast with one of the declared goal of the reform. Thus, the role of centralized benchmarking activity seems crucial in order to provide high-powered incentives to improve managerial efficiency.

References

Abrate G, Fraquelli G (2007) Domanda e tariffe nel settore idrico italiano: un’analisi di sensitività sui Piani d’Ambito. Economia Pubblica 3/4: 119-132

Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econ 6(1): 21–37

Antonioli B, Filippini M (2001) The use of a variable cost function in the regulation of the Italian water industry. Util Policy 10:181–187

Ashton JK (2000) Cost efficiency in the UK water and sewerage industry. Appl Econ Lett 7:455–458.

Aubert C, Reynaud A (2005) The impact of regulation on cost efficiency: an empirical analysis of Wisconsin water utilities. J Prod Anal 23: 383–409

Battese GE, Coelli TJ (1992) Frontier production function, technical efficiency and panel data: with application to paddy farmers in India. J Prod Anal 3:153–169

Bhattacharyia A, Harris TR, Narayanan R, Raffiee K (1995) Specification and estimation of the effect of ownership on the economic efficiency of the water utilities. Reg Sci Urban Econ 25: 759–784

Coelli TJ, Walding S (2006) Performance measurement in the Australian water supply industry: a preliminary analysis. In: Coelli TJ, Lawrence D (ed) Performance measurement and regulation of network utilities. Edward Elgar Publishing, Cheltenham, pp 29–61.

De Witte K, Marques RC (2007) Designing incentives in local public utilities, an international comparison of the drinking water sector. CES Discussion Paper 07.32

De Witte K, Marques RC (forthcoming) Capturing the environment, a metafrontier approach to the drinking water sector. Int Trans Oper Res

Fabbri P, Fraquelli G (2000) Costs and structure of technology in the Italian water industry. Empirica 27:65–82

Farsi M, Filippini M, Greene W (2005a) Efficiency measurement in network industries: application to the Swiss railway companies. J Regul Econ 28(1):69-90

Farsi M, Filippini M, Kuenzle M (2005b) Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes. Appl Econ 37(18):2127-2141

Farsi M, Filippini M, Kuenzle M (2006a) Cost efficiency in regional bus companies: an application of alternative stochastic frontier models. J Transp Econ Policy 40(1):95-118

Farsi M, Filippini M, Greene W (2006b) Application of panel data models in benchmarking analysis of the electricity distribution sector. Ann Public Coop Econ 77(3): 271-290

Filippini M, Hrovatin N, Zorić J (2008) Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods. J Prod Anal 29:169–182

Garcia S, Thomas A (2001) The structure of municipal water supply costs: application to a panel of French local communities. J Prod Anal 16:5–29

Goria A, Lugaresi N (2004) The evolution of the water regime in Italy. In: Kissling-Näf I, Kuks S (ed) The evolution of national water regimes in Europe, Kluwer Academic Publishers, Dordrecht, pp 265–291

Greene WH (2005a) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. J Econ 126:269–303

Greene WH (2005b) Fixed and random effects in stochastic frontier models. J Prod Anal 23(1):7–32

Meeusen W, van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Int Econ Rev 18(2):435–444

Muraro G (2008) Water services and water policy in Italy. In: Clini C, Musu I, Gullino ML (ed) Sustainable development and environmental management, Springer, Dordrecht, pp 65–80

OFWAT (1999) Final determinations: future water and sewerage charges 2000-05. Available via OFWAT website. http://www.ofwat.gov.uk/aptrix/ofwat/publish.nsf/Content/finaldeterminations

OFWAT (2004) Future water and sewerage charges 2005-10 - Final determinations. Available via OFWAT website. http://www.ofwat.gov.uk/aptrix/ofwat/publish.nsf/Content/pr04fd

Pitt M, Lee LF (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. J Dev Econ 9:43–64

Saal DS, Parker D, Weyman-Jones T (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. J Prod Anal 28:127–139

Sawkins JW (1995) Yardstick competition in the English and Welsh water industry. Util Policy 5: 27–36

Schmidt P, Sickles RC (1984) Production frontiers and panel data. J Bus Econ Stat 2(4):367–374

Shleifer A (1985) A theory of yardstick competition. Rand J Econ 16(3):319–327

Simar L, Wilson P (2000) Statistical inference in nonparametric frontier models: the state of the art. J Prod Anal 13(1): 49–78

Simar L, Wilson P (2007) Estimation and inference in two-stage, semi-parametric models of production processes. J Econ, 136(1):31–64

Thanassoulis E (2000a) The use of data envelopment analysis in the regulation of UK water utilities: water distribution. Eur J Op Res 126:436–453

Thanassoulis E (2000b) DEA and its use in the regulation of water companies. Eur J Op Res 127:1–13

Tupper HC and Resende M (2004) Efficiency and regulatory issues in the Brazilian water and sewerage sector: an Empirical study. Util Policy 12:29–40

Utilitatis – pro acqua energia ambiente (2006) Blue Book. I dati sul servizio idrico integrato in Italia. Roma, Italy

Wang HJ, Ho CW (2007) Estimating panel stochastic frontier models with fixed effects by model transformation. Paper presented at the EWEPA X Conference, Lille 27-30 June, GAPEM & IÉSEG School of Management