An Analysis of Inefficiency of Big Urban Water Utilities in Latin-America

Documento de Discusión CIUP
DD1413

Presentado en la Conferencia Anual del Consorcio de Universidades 2014

José Luis Bonifaz
Reyk Itakura**

Universidad del Pacífico
Centro de Investigación

ABSTRACT

This study uses a stochastic frontier parametrical approach to analyze the inefficiency of firms in the water industry between 1999 and 2010. For this purpose, an unbalanced panel of data from 12 firms from all over Latin America was used. One of the main findings of the study is that companies from the private sector outperform those from the public sector over time. Another conclusion is that there are no economies of scale or density, considering the actual size of the average sector. Finally, inefficiency shows a positive correlation with the firms' size and with the length of the network.

1. Introduction

For several years now, finding a way to measure the performance of firms in the water industry has been one of the main concerns of regulatory bodies around the world. This sector has received very special attention, not only because of water's importance in sustaining life, but also because of its configuration as a natural monopoly. The most relevant studies about performance have assessed the firms' inefficiency and the way it affects costs, production and, consequently, the customers' welfare. Because of its market structure, the regulation and the incentives mechanism ruling the sector are topics of frequent discussion. In particular, the number of privately and publicly-owned entities and the degree of regulatory intervention in pricing schemes and quality standards are especially important (Worthington 2010).

Besides the ownership nature of the utilities within the sector, some other topics such as reform characteristics and implementation, have received considerable interest throughout the literature. Regarding this, reform can influence in two ways: one of them is whether or not benefits arise from a larger participation of the private sector in the industry by expanding the access and improving management within the sector. The second way consists of the incentives scheme designed to increase the level of competitiveness. Specifically, whether or not the regulatory framework established is the best way to promote cost efficiency, network expansion sustainability and service quality, among all the other options available (price caps, rate of return or a hybrid regulation (Lin 2005). Other topics include the optimum size of the service area, the volume of production or firm's size. These are important aspects to be considered because they allow verifying if the firms are able to benefit from economies of scale or economies of population density. Finally, the usefulness itself of a benchmarking exercise to recognize the best practices in various companies that face different regulatory and geographical characteristics among countries is also frequently discussed.

In this respect, some recent studies as Sabbioni (2006) and Lin (2005) have made efforts to address the sources of these differences in order to make a plausible comparison. These sources are often referred to in literature as environmental variables; by ignoring them, non-plausible conclusions may arise. These studies introduce them in the analysis to control for heterogeneity factors (i.e. total water delivered/firm size, population density, production or distributional technology, resource availability, corruption, regional characteristics, etc.).

Together, these issues have pushed an increasing number of researches to try to find the best method to measure inefficiency. These studies recognize that the efficiency levels rest on variables within firms´ control and its value is affected by regulatory and environmental factors surrounding the firms. For this purpose, three different types of efficiency were distinguished: scale, allocative and technical. The first one presents itself when a firm´s size is optimum for profit maximization. The second one, allocative efficiency, refers to a situation where a firm combines inputs in the right shares allowing it to minimize production costs. Allocative efficiency is particularly important when firms have to deal with price distortions that may restrict their ability to minimize costs. Finally, technical efficiency occurs when a firm is producing the maximum possible output with the input combination chosen. This type of efficiency might be oriented in two different directions: it may follow an input orientation or an output orientation. Basically, this depends on the particular output or input target the firm pursues (Alvarez 2001). These measures of efficiency have allowed a more detailed and feasible benchmarking exercise and improve policy making through the implementation of yardstick regulation forms.

The aim of this paper is to analyze, using a stochastic frontier approach, the determinants of economic inefficiency after the reforms that were implemented in Latin America over the last decades in some specific countries. This is to be done by measuring the performance of water utilities that share similar characteristics and play a highly essential role within their countries as benchmarks for other smaller or less relevant firms in the same sector. Hence, the paper ranks the firms being studied according to the impact the new regulation had in their efficiency. In this respect, it is important to note that the regulation process carried out has received theoretical disapproval in some of the countries being studied. The cases of Peru and Brazil have been studied by Lin (2005) and Da Silva e Souza (2007), respectively.

The benchmark system implemented in Peru since 1999, promoted by the World Bank, gave no significant improvements in utilities' performance. This may be explained by the fact that, since the scheme lacked variables, some relationships and costs were not taken into account. For this reason, the regulatory system in Peru was designed to be output-oriented, ignoring a more appropriate, relative to costs, input orientation (Lin 2005). On the other hand, although the participation of the private sector in Brazil has increase over the last decade, the larger and main water companies are still publicly owned. Since it was implemented in the absence of well-defined regulatory policies, the privatization process did not lead to better results. As noted by Tupper and Resende (2004) the tariff policy in Brazil follows a rate of return regulation, which might not represent good incentives neither for public or private companies. The lack of incentives due to this scheme, compared to price-cap regulation could lead a firm to maximize profits without being efficient (Laffont and Tirole 1993). In both cases, the reforms had difficulties in their earlier steps which, as the study proves, did not improve significantly over time.

The rest of the paper is organized as follows. Section 2 shortly reviews studies estimating efficiency by stochastic frontiers for the water distribution companies and provides a brief description of the evolution of this procedure over time. Section 3 presents the model specification and the methodology that was applied. The data description is provided in Section 4. The estimation results are presented in Section 5 and Section 6 shows the study's conclusions.

2. Review of Relevant Studies

The following section provides a short review of the most relevant existing studies that estimate cost functions to determine the firms' level of efficiency and address the presence of economies of scale or economies of output/customer density. Both topics are discussed in the present study. *Table 1* presents a summary of these studies, including the methodology, sample description, functional form and variables in the model, as well as the determinants of the supposed inefficiency in each case.

Filippini et al (2008) studies the inefficiency in Slovenian water distribution utilities between 1997 and 2003. This paper concluded that regulation schemes based on incentives (price caps, revenue caps and yardstick regulation) appear to be superior to the traditional rate of return regulation. It also showed by testing different model specifications that the different results that were obtained were accounting unobservable heterogeneity in the environmental and network characteristics. For this reason, it is essential to distinguish between unobserved heterogeneity and inefficiency that influences costs (Greene 2005). The study deals with the consistency problem and proposes how the results from benchmarking analysis could solve this problem. Regarding the variables, the paper employs total annual costs as the dependent variable, total cubic meters of water delivered, labour and capital prices, prices of materials and customers as inputs. Heterogeneity controls are also included in the cost function: size of area, level of losses, whether or not water is chemically treated, technology changes and whether the water is obtained from subterraneous or superfluous sources. One of the findings that the traditional random effects models tend to overestimate inefficiency and that the true fixed effects model proposed by Greene (2005), although it distinguishes between inefficiency and heterogeneity, may slightly underestimate inefficiency since all the time invariant effects are treated as unobserved heterogeneity. Despite this fact, the true fixed effects model seems to perform better with respect to signs and significances.

On the other hand, Da Silva e Souza et al (2007) discusses the privatization process of over 342 public and private enterprises of the water industry in Brazil, during 2002-2004. The study finds that public companies are efficient over the period under analysis. However, their efficiency is declining, even though the overall efficiency in the sector is improving over time. Souza et al (2005) shows that there is no evidence to support that private and public firms are significantly different in terms of their levels of efficiency. In addition to the lack of adequate incentives, privatization has occurred in very isolated areas and a high percentage of the water coverage and sanitation services are still controlled by the government (Basic Sanitation State Companies). The model specification employed used average costs as the dependent variable, total volume of water produced as output, inputs such as labour and capital price. The inefficiency was considered to be related to population density, percentage of water treated, regions and the ownership type. The estimation also shows that both, translog and Cobb Douglas functional forms, lead to the same results in terms of effects and signs (these result is not common in the literature). The inefficiency components in this study do not differ from those heterogeneity factors as noted in the previous study leading to the associated estimations to be biased.

Other authors like Sauer and Frohberg (2007) focus on the problem of size using a sample of rural water suppliers in East and West Germany during 2000-2001. The methodology employed is based on a demand system with a flexible cost function for input variables; it applies a modified symmetric generalized McFadden functional form. This methodology enables the decomposition of the inefficiency term into input specific effects (made by using the cost function and the input demand equations through Sheppard's Lemma, hence avoiding the "Greene problem"[1]). This study is a response to the ongoing debate on the need and feasibility of the liberalization on the water sector in Germany during 2000. It concludes that the disintegration of utility services seems to be a positive option for an efficient restructuring of the sector since no evidence was found to support the claim that vertical integration of sewerage and water services reduces costs. The model specification considers operational costs as dependent variable, total water production as output; labour, chemicals and energy per m3 as inputs and prices of chemicals and energy as input prices. Finally, inefficiency effects seem to be associated with firms' capital (equity), water intake percentage, distribution lines' net length and number of connections.

Similar to our work, Fraquelli and Moiso (2005) used a sample of 18 utilities in ATOs (Optimum Size of Territorial Service, an administrative division created by the regulating entity) for 30 years. Following the methodology proposed by Battesse and Coelli (1995), a translog cost function was used in an unbalanced data setting. The main purpose of the study is to address the optimal size of the firms and the impact of economies of scale and to determine whether the tariff planning was efficient. The analysis suggests that the disintegration could benefit the overall efficiency in the sector. The model in this case employs total delivered water as output; labor, capital and energy costs as inputs; some other controls as network length, level of water loss and a time trend that accounts for technological changes. Regarding inefficiency, although population density is used, the study argues that other variables would perform well (number of large industrial users, level of pumping costs, percentage of water pumped from rivers, etc.).

For the Peruvian case, Lin (2005) studies how the introduction of quality variables affects performance comparisons across utilities in the water distribution service. For this purpose it compares the inefficiency measures obtained from different specifications of stochastic frontier models for a data set of Peruvian water utilities during 1996-2001. It also uses Battesse and Coelli (1995) methodology for the frontier estimation. As noted by Giannakis et al (2005), cost efficient firms do not necessarily exhibit high service quality in the sense that they could reduce costs by reducing the quality of the services being provided.. It is also showed that quality improvements lead to positive effects in the sector's productivity. The study concludes then that regulatory evaluations should incorporate quality controls instead of using only cost measurements. The model specification employs volume of water billed as output; wages, workers as inputs and capital stock; price of capital (annual capital outlays divided by the capital stock) as input price and a time trend. Some other controls are added to measure the inefficiency level related to the quality of service (positive chlorine tests, service coverage, service continuity, difference between water sold and water supplied).

Finally, Aubert and Reynaud (2005) study a panel data set with 211 water utilities observed from 1998 to 2000. The main finding is that water utilities´ efficiency scores can be partly explained by the regulatory framework in place. Consequently comparing regulated utilities in different states and, therefore, in different regulators jurisdiction, increases the risk of unobservable characteristics related to policy effects biasing the results. It describes the main regulation schemes employed by the sector regulators as follows: The rate of return regulation (or cost plus), which was originally thought to reduce monopoly effects, does not lead the firm's incentives to pursue cost efficiency since no price restrictions exist. On the other hand, the price cap regulation is a high powered incentive scheme because it induces

[1]"The Greene problem" consists on the difficulty of specifying an estimable input system that incorporates both technical and allocative inefficiency as well as the satisfactory derivation of the relationship between allocative inefficiency in the input demand equations and in the cost function (See e.g. Kumbhakar, 1997).

firms to produce efficiently: since the price is set by the regulators, firms can only reduce costs to increase their profits. The model specification uses volume of water sold and customers as output; energy price and labor price as input prices; capital as input and control variables for technical characteristics (Dummy W: utilities that purchase water from another utility, Dummy S: utilities use surface water and Depth: Average depth of pumping wells).

These and other relevant studies that use a similar approach (respect to cost functions and the stochastic frontier method) are summarized in *Table 1*.

Table 1: Summary of Relevant Studies

Author(s)	Methodology	Sample	Specification	Ineffiency Determinants
LynK (1993)	SFA	10 water companies, United Kingdom, 1979-1988.	Dependent: Annual operating costs. Output: water supply, sewerage service. Inputs: Labour costs. Controls: Time trend, regions, private ownership, envirionmental services.	None.
Bhattacharyya, Harris, Narayanan and Raffiee (1995)	SFA	190 public and 31 private utilities, US, 1992	Dependent: Variable costs. Output: Volume of water. Input: Energy, labour, materials, stock of capital. Controls: Water input produced or available for delivery, water input source (surface, ground, both), system losses, age of distribution pipelines, number of emergency breakdowns, length of distribution pipeline, customer type (residential or commercial).	Private ownership, Size, breakdowns.
Antonioli and Filippini (2001)	SFA	32 water distribution firms, Italy, 1991-1995.	Dependent: Variable costs. Output: Total water distributed. Input: Price of labour, capital. Controls: size (length of pipes), extent (number of customers), losses, chemical treatment, time trend.	None
Estache and Rossi (2002)	SFA	50 water companies, 29 Asia-Pacific countries, 1995	Dependent: Operational costs. Output: Number of customers, daily production. Input: Number of connections. Controls: Population density in area served, percentage of water from surface sources, number of hours of water availability per day, percentage of metered connections, qualitative treatment variables (chlorination, desalination).	Ownership
Bottasso and Conti (2003)	SFA	177 water utilities, England and Wales, 1995-2001	Dependent: Operating costs less current cost depreciation and infrastructure renewal charge. Output: Total water delivered. Input: Capital (repalcement costs of net tangible assets), labour costs	Length of mains, average pumping head, river sources, non-households customers, population density, water introduced in mains as proxy of size.
Aubert and Renaud (2005)	SFA	211 water utilities, Wisconsin, 1998-2000.	Dependent: Variable costs. Output: Volume of water sold, customers. Input: Energy price, labour price, capital (average net base rate of water utility). Controls: Dummy utilities that purchase water from another utility, dummy utilities use surface water, average depth of pumping wells.	Dummy for each regulatory regime, time trend.
Lin (2005)	SFA	Water distribution companies, Peru, 1996-2001.	Dependent: Total costs. Output: Volume of water billed, Inputs: Wages, number of employees per 1000 conn., price of capital, capital stock.	Quality service: positive chlorine tests, service coverage, service continuity, difference between water sold and water supplied), time trend.
Fraquelli and Moiso (2005)	SFA	18 ATOs (regional territories), Italy, 1975-2005	Dependent: Total costs. Output: Total delivered water. Input: labour, capital, energy costs. Controls: Network length, level of losses, time trend.	Number of residents to network length.
Kirkpatrick, Parker and Zhang (2006)	SFA and DEA	110 water utilities, Africa, 2000.	Dependent: Operating and maintenance expenditure. Output: Water delivered. Input: Labour price, material price of water distributed. Controls: number of water treatment works.	Ownership.
Sabbioni (2006)	SFA	280 water utilities, Brazil, 2002.	Dependent: Totol costs. Output: Volume of water produce, population served, number connections. Input: Wages.	Water purchased from another company, household consumption, regions.
Da silva e Souza, Faria and Moreira (2007)	SFA	324 public and 18 enterprises, Brazil, 2002-2004.	Dependent: Average costs. Output: Volume of water produced. Input: Price of capital, price of labour.	Population density, water treated, dummy for regions, state/private.
Sauer and Frohberg (2007)	SFA	Rural water suppliers, East and West Germany, 2000-2001.	Dependent: Operational costs. Output: Water output. Input: labour, energy, chemicals, wages, price of energy, price of chemicals.	Level of capital (equity), water intake percentage, net length, number of connections.
Filippini, Hrovatin and Zoric (2008)	SFA	52 water utilities, Slovenia, 1997-2003.	Dependent: Total annual cost. Output: Total water delivered, network size. Input: Labour, capital, materials. Controls: time trend.	Level of losses, water chemically treated before distribution, subterraneous water sources, superflous water.

3. Methodology and Model Specification

Methodology

The present study attempts to identify the determinants of the inefficiency within the water sector for Latin America by a one-step procedure. For this purpose, a cost function frontier was modeled where firms faced a minimization problem. Following this method, the level of inefficiency was obtained as the residual component (different from the error term and different from heterogeneity factors as discussed in the literature). At the same time, this residual was modeled to find the determinants of inefficiency in the water distribution sector in Latin America. Investigations that follow the frontier approach are particularly accurate because, unlike other methods[2], they do not assume that all firms are efficient (Coelli, et al. 1998). These studies try to estimate a frontier from the best practices observed.

Some aspects of the estimation technique require special attention[3]. As is presented by Söderberg (2007), it can be noted that Stochastic Frontier Analysis (SFA), the parametric approach, rests on two basic assumptions: that utilities have a common underlying cost function and that the inefficiency term can be approximated through a pre-specified distribution (in addition to the standard random noise and considering time variations in this term). This way, the inefficiency term can be separated from the data noise and the parametric tests can be carried out. On the other hand, stochastic frontier models are exposed to three potential problems (Schmidt and Sickles, 1984). First of all, estimations of the inefficiency, though unbiased, are inconsistent (variance never approaches to cero as the sample grows). Second, it is necessary to assume some distributional form in order to estimate the model and separate the inefficiency term from the error. Finally, it may not be accurate to suppose that some regressors are independent from the inefficiency term because if a firm knows its level of inefficiency, it could alter its choice of inputs.

Data envelopment Analysis (DEA), the non-parametric approach, constructs its frontier from the best-practice operation assuming that data is noise free. Despite more giving the analysis more flexibility, DEA is very sensitive to the number of included inputs and outputs (Bonifaz and Rodriguez, 2001), the input and output variables selected and data errors that could distort the outcome (Irastorza, 2003). For example if the number of observations grew, the method would lose the ability to differentiate between efficient and inefficient firms. One important problem with DEA is that it does not assume a distribution for the error term. Thus, estimations will not have statistical properties and hypotheses testing are impossible to implement. Furthermore the results can be very sensitive to the number of observations, variables or the presence of outlier, which might bias the efficiency estimates (Yunos and Hawdon, 1997).

Since being able to carry out hypothesis tests on our results is crucial, and the fact that some variables are measured differently by each enterprise cannot be ruled out (which would make the outcomes biased), the SFA approach was chosen in order to carry out this study. Moreover, recent studies have increased the attractiveness of benchmarks based on stochastic frontiers (Söderberg, 2007).

Inefficiency measurements based on Stochastic Frontier Analysis have evolved over the time. They were originally proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and Van den Broeck (1977).

[2]Productivity indexes, production functions, etc.
[3]For estimating a frontier both, parametric and non-parametric approaches can be used.

These authors consider that deviations from the frontier are explained by inefficiency but are also related to other variables. Later, Pitt and Lee (1981) and Schmidt and Sickles (1984) introduced the analysis to panel data. Some other studies tried to explain inefficiency as a function of specific determinants with a two-step approach. The difficulty in this method is that it has a consistency problem with the a priori assumptions about the error term. Kumbhakar et al. (1991), Huang and Liu (1994) and Battesse and Coelli (1995) proposed a one-step approach to model the inefficiency simultaneously distinguishing two components in the error, one of them is associated with a pre-truncated distribution and estimated as a function of exogenous variables related to the firm's inefficiency. Reifschneider and Stevenson (1991), Caudill and Ford (1993), Caudill et al. (1995), and Hadri (1999) modeled these exogenous variables not in the inefficiency error mean but in its variance.

In the present paper, we estimate a stochastic cost frontier model by using the True Fixed Effects Model first proposed by Greene (2005). The True Fixed Effects Model separates all time-invariant sources of heterogeneity from the inefficiency term and hence distinguishes them from other random factors affecting firms' costs minimization. This model tries to solve the limitations of the basic Fixed Effects and Random Effects models by simply introducing dummy variables for each firm and allowing the inefficiency component to vary over time, defined as a random variable that depends on random factors. As the basic Fixed Effects Model, if time-invariant regressors are included in the model specification the model will lead to the estimation of biased parameters. Even if this point is overcome, the model will capture all unobserved time-invariant components of inefficiency within the firm specific constant term (the dummy variable) which may lead to the underestimation of the inefficiency levels.

In addition, the applied methodology will also depend on the frontier specification, which might employ a production function (which shows the quantities produced according to the inputs used) or a cost function (which reflects the total cost as a function of the product and input prices). For the purpose of this paper, we have chosen a cost function specification because, in the water distribution sector (same as transport, electricity and other public services); supply must match demand levels permanently. Following the previous idea, as mentioned by Kumbhakar (1991), since cost functions summarize all the economically relevant aspects of technology, it is common to use models based on costs instead of production functions. The choice of either of them depends on the behavioral assumptions of the producers and availability of data. The cost specification is often used in regulated sectors where firms' outputs are assumed to be exogenous, the same as input prices. Since production is out of firms' control it is reasonable to assume they minimize costs adjusting their inputs usage.

These and other aspects have to be considered in the above election. Sabbioni (2006) organizes them in five points: 1) Operating environment (related with the institutional framework; if the firm is required to produce at a specific level, the production function will not be appropriate), 2) Endogeneity of input quantities, 3) Data limitations (physical quantities vs. prices), 4) Output definition (using a production function, one has to choose among various proxies that might be included as regressors in a cost function), 5)Technology and quality specification are often difficult to include in production functions. According to all these factors it is more appropriate to estimate a cost function in this case.

As mentioned before, inefficiency measures are associated with different sources in the production process. Since we estimate a cost function frontier, the inefficiency measures are related to both, technical and allocative inefficiency; following to Battesse and Coelli (1995), we impose allocative efficiency over time even though it might be a strong assumption. In this respect, we consider all the firms included in the sample as efficiency benchmarks for their countries or regions. Hence, inefficiency must be better explained by technical and firm's specific factors than by particular input price distortions.

Theoretical Model

Since in the water sector the output is assumed to be given for the firms, we estimate a cost function[4] frontier model. The function estimated is the solution to the minimization problem: $\min_{x_i e^{-\mu_i}} w_i^* x_i e^{-\mu_{it}}$ subject to $q_i = f(x_i e^{-\mu_{it}})$. The μ_{it} term represents the nature of the technical inefficiency effects over the production function. As noted, a positive value of this term will increase the firm's need of inputs for a predefined level of output. This procedure will result in the following form:

$$C_i = \sum_j w_{j,i} x_{j,i}$$

In terms of the methodology applied here, this general expression is rewritten for the frontier estimation in the panel data context. Then we have:

$$C_{it} = \exp(\alpha_i + x_{it}\beta + v_{it} + u_{it})$$

Where C_{it} denotes costs at the $t-\text{th}$ observation$(t = 1,2,\ldots,T)$ for the $i-\text{th}$firm$(i = 1,2,\ldots,N)$. This expression is for the case in which the variables have been stated in logarithms where the $\alpha_i$constant is used to control for firm heterogeneity fixed effects (different from the inefficiency component), x_{it} is a $1 \text{ x } k$ ($k = m$ outputs $+$ g imputs $+$ h other *exogenous variables*[5]) vector of values of known functions of inputs, outputs and other explanatory variables associated with the $i-\text{th}$ firm at the $t-\text{th}$ period and β is a$k \text{ x } 1$ vector of unknown parameters to be estimated.

The error term here has two components, v_{it} , that represents a white noise and is assumed to be an iid $N(0, \sigma_v^2)$, independently distributed of the inefficiency termu_{it}, which isa non-negative random variable. Here we assume that the u_{it}term is distributed as a truncated normal (to impose non-negative values) with mean $z_{it}\delta$ and variance σ_u^2.Following Huang and Liu (1994) and Battesse and Coelli (1995), we modeled the inefficiency term mean as a function of other factors under firm's control. Thus z_{it} is a $1 \text{ x } m$ vector of explanatory variables associated with inefficiency of firms over time and δ is a $m \text{ x } 1$ vector of unknown parameters estimated by the model.

Since the underlying production technology is unknown we employ a more flexible functional form for any quadratic approximation of the real function without imposing any restrictions over elasticities (Chambers 1988).Not imposing restrictions over production technology will allow the analysis to measure the level of economies of scale and economies of production density. In the literature, the translog function is one of the most adopted functional forms in the cost frontier context for its flexibility[6]; however as pointed by Conrad and Jorgensen (1997) most empirical studies fail to satisfy the appropriate curvature conditions for the whole range of observations. The translog function must satisfy some regularity conditions according to the cost theory: it must be linearly homogeneous in factor prices, non-decreasing in factor prices and output, symmetric and concave. Some of these conditions

[4]Some other reasons have been exposed previously for this election.
[5]These variables are related to environmental and geographical characteristics.
[6]Although it is more flexible than the more simple Cobb-Douglas functional form, is very common to suffer multicolinearity problems when using translog functions.

might be implemented either by normalizing costs and inputs prices to one of them or by assuming that cross-products related to prices have non-marginal effects.

Econometric Specification

According to the former discussion we restate the previous expressions for a translog functional form, thus the model estimated for the i – th firm and the t – th period is as follows:

$$lnC_{it} = \beta_0 + D_i + \beta_y ln\, y_{it} + \beta_w{}' lnw_{it} + \frac{1}{2}\ln w'_{it} A \ln w_{it} + \frac{1}{2}\ln y_{it}{}^2 + \ln w'_{it} C \ln y_{it}$$

$$+ \beta_f{}' \ln f_{it} + v_{it} + \mu_{it}$$

Where lnC_{it} represents costs expressed in logarithms, D_i accounts for unobserved time-invariant firm specific heterogeneity effects, $ln\, y_{it}$ is the logarithm of the output variable, lnw_{it} is the vector of j inputs in logarithms, and $\ln f_{it}$ is a vector of k exogenous variables that control other environmental and technological factors out of firms' control. Finally, v_{it} and μ_{it} represent, as mentioned before, the error term of the model. We have assumed that matrixes A and C are symmetric since it cannot be imposed in the estimation procedure applied here. Additionally, we have not considered second differentiable effects for the k exogenous variables because they are not related to the original cost minimization problem and, therefore, are not involved functionally in the production technology. Linear homogeneity is imposed by assuming all parameters related to cross-products and squared terms of prices to be cero. This condition can also be attained by normalizing costs and other input prices to one input price.

For the cost variable (COST) we used the operational costs of firms excluding salaries and depreciation related to administrative expenditures. The variable output (CONN) has been approximated by the number of connections of potable water in the service area. The number of connections must have a positive effect on the costs. The actual production (in liters per day) is taken within a density ratio against the population size of the service area: a lower ratio will indicate a densely populated area of distribution and, consequently, less infrastructure expenditures for firms, then is expected to have a negative impact over operation costs[7]; this ratio (DENSITY) and a time trend (TREND) are included as exogenous variables affecting the production technology accounting for environmental characteristics and technical change over time, respectively. For input variables we considered the number of full time employees (EMPL) and the total km. of water network length (LEN). We use the Gross Domestic Product per capita (GDP) as a proxy of salaries or labour price; we did not have information available for capital prices/costs for each firm as used by other studies. In the literature, it is a common practice to employ capital expenditures divided by water production (or other output variable). Unfortunately that information is often not available or is highly inaccurate.

As explained above, the inefficiency term has two components:

v_{it} is a random error assumed to be i.i.d.$N(0, \sigma_v^2)$ and independent from μ_{it}. The μ_{it} term is a non-negative random error obtained from truncating a normal distribution with a mean that can be expressed as a function of some other random variables. This term indicates the deviation degree,

[7] Other density ratios have been used through literature. Franquelli and Moiso (2005) used a density variable obtained by dividing number of customer by the network length. Da Silva e Souza et al (2007) and Estache and Rossi (2002) both use population density for this purpose.

presumably due to inefficiency, for each observation. Following Battesse and Coelli (1995) approach for panel data which extended the model of Kumbhakar, Ghosh and McGukin (1991) and Reifschneider and Stevenson (1991), inefficiency effects are defined by firm-specific variables. Thus we have:

$$\mu_{it} = Z_{it}\delta + e_{it}$$

Here $Z_{it}\delta$ is the mean of the inefficiency term. Notice that it varies over time depending on those firm specific factors, which is different from the more traditional methods that assume the inefficiency term or its mean to be constant in each period. Several studies have argued about the possible variables determining inefficiency, even though, a consensus has not been reached.. In the present paper we proposed the next expression to model inefficiency considering that it should be affected only by variables under firms' control and not by environmental effects (that influenced costs directly) as other studies have shown. Hence, the inefficiency term is specified as:

$$\mu_{it} = \delta_0 + \delta_1 OWNERSHIP + \delta_2 WCOV + \delta_3 SCOV + \delta_4 TREAT + \delta_5 LOSSES + \delta_6 METER + e_{it}$$

Where *OWNERSHIP* is a dummy variable that takes the value of one when the firm is private and cero otherwise (in particular, the variable takes a cero value when the firm is public). This variable is expected to reduce cost inefficiency since private firms might have better management practices and a more effective incentives scheme. *WCOV* and *SCOV* are the percentage level of water and sewerage coverage, respectively; a higher coverage will lead to economies of scale (if they are present), therefore increasing the input usage efficiency. *TREAT* is the percentage level of wastewater that received at least primary treatment after collection; if wastewater were to receive treatment, it would prevent water sources from being corrupt and hence reduce treatment expenses before distribution, which might increase efficiency. *LOSSES* accounts for the percentage level of non-revenue water; a higher level of losses would indicate bigger outlays from illegal connections or a greater waste from leakage due to infrastructure problems. This effect should increase water needs from final customers leading the firm to increase water production. *METER* represents the metering level percentage; a higher metering level will have a positive effect in reducing losses from administrative issues (then this variable might suffer from multicolinearity problems with LOSSES). Finally e_{it} is a random variable accounting for stochastic effects on inefficiency. Then the cost inefficiency for the i – th firm and for the t – th period is measure by:

$$Cost\ Inefficiency = \exp(\mu_{it}) = \exp(Z_{it}\delta + e_{it})$$

The inefficiency measure for each firm is then a non-negative random variable which depends on other explanatory factors that control effects for firm-specific characteristics. It can be observed that the inefficiency mean will vary over time, having independent but not identical distributions for each firm at a specific time.

For the estimation procedure applied we have considered an unbalanced panel data of Latin American water distribution companies for the period 1999-2010. Of these, 9 were public and the other 3 firms were considered private[8]. The panel structure is shown in *Appendix 1*. All the information employed in

[8]Public companies: CAGESE (Ceara), COMPESA (Pernambuco), EMBASA (Bahia), CEDAE (Rio de Janeiro), CASAN (Santa Catarina), CORSAN (Rio Grande do Sul), SANEAGO (Goiania) from Brazil, SEDAPAL from Peru and EAAB from

this study was taken from the data base of the corresponding regulatory authority[9] each country and from the World Bank water benchmarking network data (IBNET)[10]. *Table 2* shows the average of each relevant variable by firm for the available period analyzed in each case. A summary and the statistical description of each variable are presented in *Table 3*.

Table 2: Average of Relevant Variables by Firm

Firm	Operational Costs (US$)	Number of Connections	Number of Employees	Distribution Network of Water (Km)	GDP per Cápita (US$)	Production per Cápita (l/person/day)	Water Coverage (%)	Sewerage Coverage (%)	Wastewater Treated (%)	Non-revenue Water (%)	Metering Level (%)
SEDAPAL	194,732,228	1,259,551	2,062	10,123	2,821	282	88.1%	83.7%	9.0%	39.9%	65.6%
AGUAS ANDINAS	154,973,432	1,388,563	1,167	10,946	7,008	207	100.0%	98.2%	54.0%	29.0%	94.9%
AGUAS ARGENTINAS	201,098,414	1,549,736	2,925	16,256	5,762	576	82.4%	62.1%	10.4%	35.4%	24.5%
E.A.A.B.	170,700,553	1,574,768	1,679	8,446	4,122	184	99.8%	97.5%	0.0%	36.7%	97.7%
O.S.E.	125,776,412	776,524	4,508	12,287	6,234	276	94.1%	31.6%	70.2%	52.8%	96.8%
CAGESE / CE	225,783,255	1,194,739	1,256	8,440	4,604	199	81.2%	53.0%	100.0%	33.4%	93.9%
COMPESA / PE	444,891,381	1,413,258	3,381	10,267	4,604	230	86.1%	60.5%	87.0%	56.8%	64.1%
EMBASA / BA	516,110,007	2,090,788	3,523	19,529	4,604	198	91.2%	26.5%	24.7%	34.7%	87.2%
CEDAE / RJ	2,081,165,497	1,821,905	7,512	16,315	4,604	500	85.7%	67.3%	72.0%	52.4%	62.5%
CASAN / SC	217,933,925	806,196	2,352	13,134	4,604	250	85.4%	54.3%	96.4%	40.1%	91.2%
CORSAN / RS	552,388,173	1,568,383	4,412	22,158	4,604	243	91.6%	49.9%	78.5%	42.3%	66.7%
SEANEAGO /GO	286,276,090	1,076,217	3,493	17,343	4,604	190	88.6%	55.3%	52.6%	35.4%	96.5%
General Average	**432,706,783**	**1,374,574**	**3,190**	**13,843**	**4,914**	**278**	**89.5%**	**61.5%**	**64.3%**	**40.6%**	**78.5%**

Table 3: Summary and Statistical Description of Variables

Variables	Description	Average	Standard Deviation	Maximun	Minimun	Expected Sign
COST	Operational Costs (US$)	432,706,783	593,069,651	4,466,905,410	79,910,000	
Output						
CONN	Number of Connections	1,374,574	412,912	2,513,663	675,217	(+)
Inputs						
EMPL	Number of Employees	3,190	1,754	8,416	1,006	(+)
LEN	Distribution Network of Water (Km)	13,843	4,459	23,924	6,675	(+)
Input Prices						
GDP	GDP per Cápita (US$)	4,914	2,118	10,167	2,044	(+)
Controls						
DENS	Production per Cápita (l/person/day)	278	125	622	179	(-)
TIME	Time Trend	-	-	12	1	(-)
Inefficiency						
OWNERSHIP	Ownership Type (Private or Public)	-	-	1	0	(-)
WCOV	Water Coverage (%)	89.54%	9.34%	100.00%	62.00%	(-)
SCOV	Sewerage Coverage (%)	61.50%	33.10%	100.00%	8.00%	(-)
TREAT	Wastewater Treated (%)	64.26%	42.23%	100.00%	0.00%	(-)
LOSSES	Non-revenue Water (%)	40.63%	9.90%	66.00%	23.00%	(+)
METER	Metering Level (%)	78.54%	21.74%	100.00%	23.61%	(-)

Colombia. Private companies: AGUAS ANDINAS from, Chile; AGUAS ARGENTINAS/AySA (until 2003) from Argentina and OSE from Uruguay (from 2006 is a public-private).

[9]SUNASS for Peru, SNISS for Brazil, SISS for Chile, SUI for Colombia, ERAS for Argentina and URSEA/OSE for Uruguay.

[10]International Benchmarking Network for Water and Sanitation Utilities. http://www.ib-net.org/.

4. Results and Estimates

The cost frontier is estimated by using the Maximum Likelihood algorithm. This allowed the parameters of both the stochastic frontier and the mean of inefficiency term to be estimated simultaneously[11]. *Table 4* shows the results.

Table 4: Cost Frontier and Inefficiency Model Estimates

Variables	Coeficient	Standar Error	Probability
Constant	-218.8178	0.9967	0.0000
D1	0.1579	0.1268	0.2159
D2	-1.1988	0.1144	0.0000
D3	-0.5179	0.2742	0.0619
D4	-0.0158	0.1420	0.9117
D8	0.5773	0.2358	0.0161
D9	2.2124	0.2792	0.0000
D11	0.9419	0.1960	0.0000
D12	0.3859	0.1566	0.0155
CONN	20.2405	0.6858	0.0000
EMPL	6.6489	0.9616	0.0000
LEN	12.4665	1.0362	0.0000
CONN^2	-1.0021	0.2344	0.0000
EMPL^2	0.1717	0.1698	0.3145
LEN^2	-0.0987	0.5972	0.8691
CONN*EMPL	0.6135	0.1878	0.0015
CONN*LEN	0.3858	0.7158	0.5911
EMPL*LEN	-1.9728	0.4290	0.0000
GDP	0.8957	0.0780	0.0000
DENS	-0.0281	0.2782	0.9199
TREND	-0.0234	0.0120	0.0535

Variables	Coeficient	Standar Error	Probability
Constant	0.0575	0.6279	0.9272
OWNERSHIP	-0.9895	0.2111	0.0000
WCOV	-0.9389	0.3989	0.0206
SCOV	0.0963	0.1261	0.4468
TREAT	0.2171	0.1229	0.0803
LOSSES	2.5138	0.5358	0.0000
METER	-0.2956	0.4405	0.5038

Variance Estimates	Coeficient	Standar Error	Probability
sigma-squared	0.0469	0.0122	0.0002
gamma	0.6234	0.1212	0.0000

Number of Firms	12
Number of Periods	12
Total Observations	119

Since all the variables were expressed in logarithms (except for dummies, percentages and the time trend), the coefficients derived from the model represent cost elasticities evaluated at an average point. The overall coefficient estimates obtained for the cost frontier were as expected: results proved that the number of connections has a positive effect on operational costs. In other words, an increase in the number of connections will push costs to a higher level. The coefficients for all the inputs and input prices were significant and had the expected sign. In both cases, the expected coefficients were used as control variables. On the other hand, the higher the density ratio, the lower the operational cost, as it becomes less expensive for firms to supply water and sewerage services for with a higher number of people by reducing the infrastructure extent and maintenance costs. However, the test showed that the density variable was not statistically significant enough. The time trend had a negative and significant effect, which could be interpreted as real cost reduction due to technical change over time.

Second order parameters were estimated but did not show significance for all the effects. The second order coefficient for the production variable was significant and negative, showing that its effect, though positive, is marginally decreasing. Another interesting cross-effect was found the production's effect on costs depended significantly on the number of employees, namely, the higher the number of employees, the greater the effect of production scales on costs. Also, the cross-effect of the network

[11]The model is estimated by using Frontier 4.1 program (developed by Tim Coelli). www.uq.edu.au/economics/cepa/software.htm.

length and number of employees was negative, which may suggest that a degree of complementarity between in the inputs selection. From this, we can infer that an input will depend negatively on the other input's level. From this finding, an optimal cost level might be reached by increasing both the number of employees and the network size.

For the inefficiency term, the majority of explanatory variables were significant. The first remark regarding inefficiency is about the constant term. Although it resulted not to be significant, it cannot be affirmed that none of the time invariant sources of inefficiency where found because the firm identification dummies included were expected to absorb all constant heterogeneity factors (including those affecting inefficiency). As it can be noted from *Table 4* most of the dummy variables were significant[12], hence an overestimation of these coefficients is possible due to model specification. As noted in Greene (2005) there is not a perfect way to avoid overestimation or underestimation of the inefficiency component through model specification because both, the inefficiency and the heterogeneity have time-invariant effects. Second, the ownership variable accounting for differences on inefficiency scores between public and private firms resulted significant. The better private management hypothesis is confirmed for the case studied here; this might be due to the lagged effects of reform over the sector in Latin American countries during the last decade.

Third, the water coverage coefficient was negative and significant which is interpreted as evidence for the existence of economies of scale in the firms' service areas. Due to the high density of the areas served by the sample firms, a coverage expansion might take advantage of this fact leading to a reduction in the cost per unit. The sewerage coverage coefficient was found not to be significant. However, this variable had a positive coefficient, which might call for an unsuccessful vertical integration of these two activities within firms. An interesting finding from the inefficiency model was that the variable treatment (accounting for wastewater treated percentage) presented a positive effect on inefficiency measures. This might be a significant indicator of the previous idea concerning vertical integration of water supply and sewerage services. Variable losses, as expected, had a positive sign, and then it will lead to increase inefficiency by rising water production from its optimal volume to supply a given demand. Its positive sign is explained not only by physical losses due to inadequate pipeline maintenance but also by poor management and billing policies (this incentives illegal connections increasing inefficiency levels). Finally, the metering level had a negative effect on inefficiency but resulted not to be significant.

[12] For estimation purposes we excluded 4 dummy variables from the model specification in order to avoid severe multicolinearity problems.

To validate the use of a translog cost function and the model specification, we tested alternative specifications with different hypothesis. The contrasts are made by comparing the generalized likelihood ratio resulting from the unrestricted and restricted versions with a critical value. The critical values employed correspond to the tables presented in Kodde and Palm (1986)[13]. The results are shown below.

Table 5: Critical values

Null Hypothesis	Restrictions	Likelihood Function	LR Test	Critical Value*
$Ho: Cobb - Douglas$	6	31.43	23.02	11.91
$Ho: \mu_{it} = \delta_0 + e_{it}$	6	13.80	58.27	11.91
$Ho: \mu_{it} = e_{it}$	7	13.28	59.32	13.401
$Ho: D_i = \emptyset$	8	-34.69	155.25	14.853

The first hypothesis consists in comparing the Cobb-Douglas specification against the more flexible translog functional form. The result obtained was the rejection of the hypothesis that the Cobb-Douglas specification has a similar fit in terms of the Likelihood function. The second test resulted in the rejection of the hypothesis that assumed the inefficiency term was not a function of the explanatory variables proposed by the study. The third hypothesis posed that inefficiency had a half normal distribution as opposed to a truncated normal. However, the test showed evidence to reject this assumption. Finally, the last test was made to justify the time-invariant heterogeneity difference among firms, but the hypothesis of non-existence of these effects was also rejected. An additional test may be noticed from the estimation table: the variance parameter gamma was found significant indicating that inefficiency is a stochastic variable and that deviation from the estimated frontier is due to other factors besides random errors.

The average inefficiency score considering all firms for the whole available period is 1.32 which indicates that in average Latin American companies in the water sector experienced an increase in costs of 32% due to inefficiency.. The inefficiency ranking obtained from the estimation is presented graphically in *Figure 1.* It can be observed that the most efficient company is AGUAS ANDINAS from Chile, followed by AGUAS ARGENTINAS[14], and in third place, OSE (OBRAS SANITARIAS DEL ESTADO) from Uruguay. On the other hand, the company presenting the most inefficiency is COMPESA from Brazil. Both AGUAS ARGENTINAS and AGUAS ANDINAS are privately owned, while COMPESA and CEDAE (second to last regarding efficiency) are publicly owned. This result is coherent with the estimation described above and might indicate that the privatization process over the last two decades in some countries in the sample had real efficiency improvements or, at least, that private companies outperform state companies throughout the period under analysis. *Figure 2*, where the average inefficiency score (by category) of all firms classified as public over time are shown.

[13] According Coelli, Prasada and Battese (1998) the number of degrees of freedom used for choosing the critical value equals the number of constraints posed by the null hypothesis.

[14] The name of this company was changed to AGUAS Y SANEAMIENTO ARGENTINO in 2006.

Figure 1: Inefficiency Rankings by Average Inefficiency by Firm

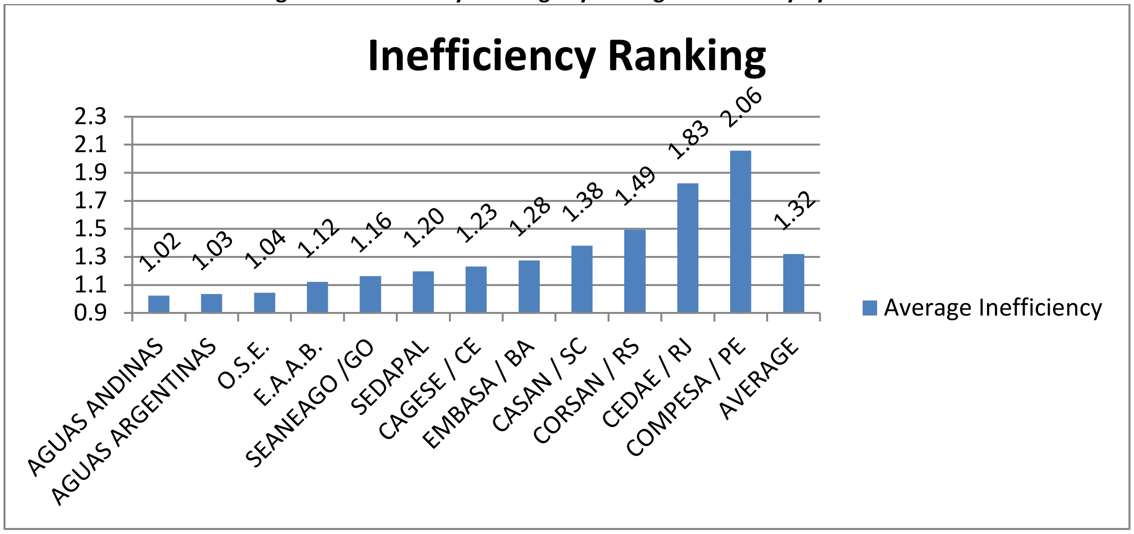
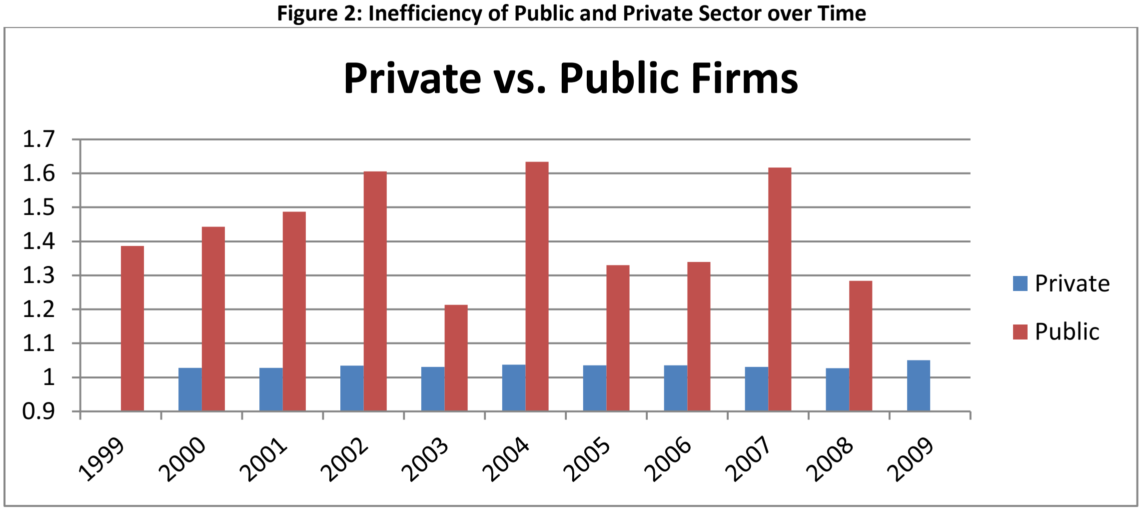
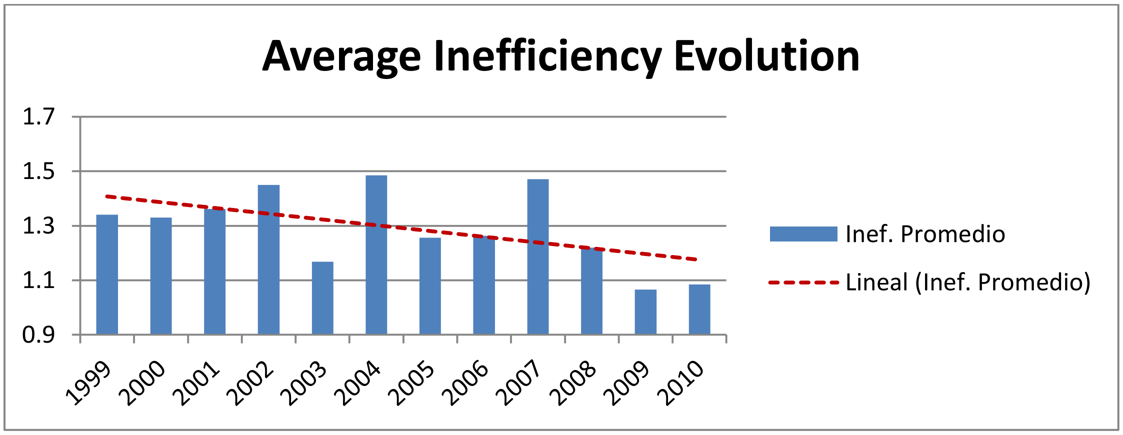


Figure 2: Inefficiency of Public and Private Sector over Time



Another important finding derived from the inefficiency measures is the overall time evolution of the average score. Figure 3 shows this aspect clearly. As is evident, the inefficiency is decreasing over time, which might indicate a progressive improvement for the sector through all the studied firms in Latin America. Decomposition of this evolution by firms can be found in Appendix 2.

Figure 3: Inefficiency Evolution over Time



After analyzing the firms' inefficiency component in costs, the second purpose of this study is to verify the existence of economies of scale and economies of output density. Economies of scale reflect the degree in which costs increase due to an increment in output, keeping all other variables constant. However, as noticed in Fraquelli and Moiso (2005) the output may change with the output characteristics (partly represented by inputs). By including the network length in the cost frontier estimation we allow for the distinction of economies of output density and economies of scale (Filippini 2001). Here we followed Fraquelli and Moiso (2005) by measuring economies of output density as the inverse of output cost elasticity and economies of scale as the inverse of the sum of the elasticity with respect to outputs and the network length. If these measurements are higher than one, it would indicate the presence of economies and therefore cost savings from increases in either outputs or network length.

Then the economies of scale and output density measures for the model estimated here are 0.049 and 0.031, respectively. These scores suggest that, at the current size, the water distribution companies studied are, on average, unable to take advantage of cost savings by increasing output or the network coverage.

Another important finding related to scale effects might be seen graphically in *Figure 4* and *Figure 5*, where firms are presented in ascendant order according to their number of connections and network length. Firms' inefficiency seems to increase with size (measured either by connections or network length). This result could be related to the fact that there was no evidence of economies and hence costs increase more than proportionally when the number of connections or network size grow, presumably because of efficiency losses. This is an indicator then that firms operate at a critical size that do not contribute to efficiency improvements.

Figure 4: Inefficiency by number of Connections (Firms ordered ascending)

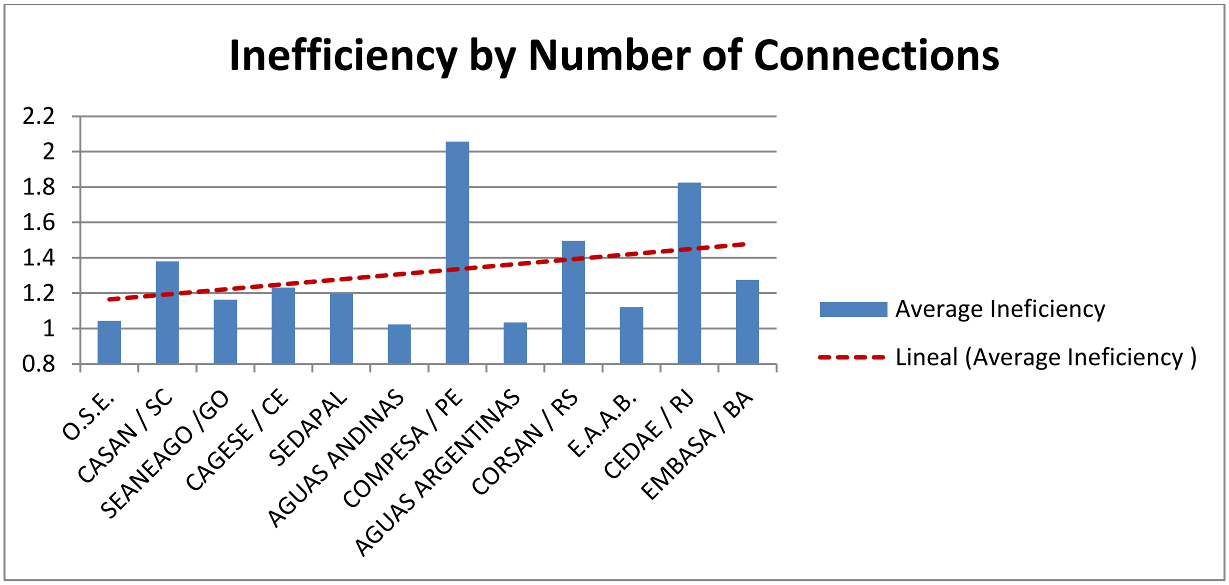
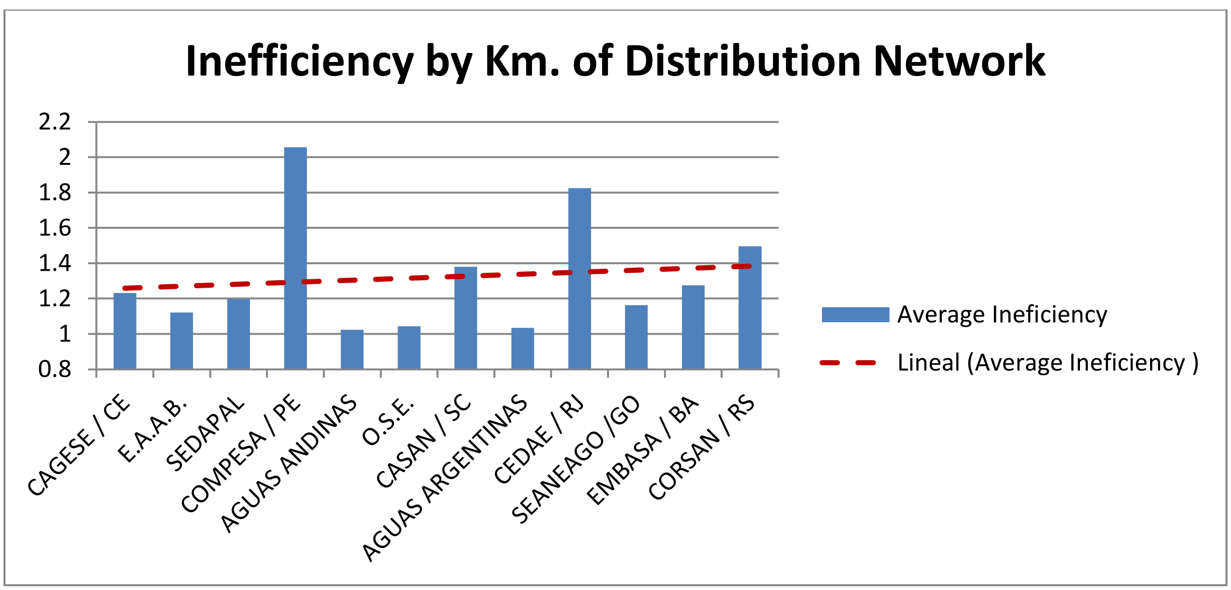


Figure 5: Inefficiency by Network Length (Firms ordered ascending)



5. Conclusions

The main findings of this study concern the level of inefficiency among firms and whether or not economies of scale exist in the water distribution sector for Latin America within the sample analyzed. First, inefficiency scores showed, on average, a declining trend over time. However, this conclusion has to be carefully applied for specific categories within the whole group. For private companies the inefficiency measures were more stable than those for public units; scores for public sector exhibited a more unstable evolution, even though these findings may reflect the presence of distorted observations.. The overall inefficiency for private companies was significantly lower than that for their public pares. This can be interpreted as the consequence of a successful reform and the privatization process over the last decades in some specific countries. Even though the sample used was not significantly large, the results are representative enough to be generalized for other water industries among Latin America.

Regarding economies of scale and output density, the analysis derived from the estimation proved the absence of potential cost savings by increasing the firms' size or network length. At the current size, the overall water sector is not likely to be efficient due to administrative negative effects. This finding might be also related with some evidence from the estimation suggesting that a successful vertical integration of water supply and sewerage services has not yet been achieved. The inefficiency scores showed a positive correlation with size measures, the higher the output and network size, the greater the inefficiency levels. Potential gains could be obtained by reducing the firms' size, especially for those in the public sector.

Finally, the methodology applied allows a more accurate benchmarking exercise, considering environmental and output characteristics differences and isolates inefficiency effects by including firm idiosyncratic factors in this term. Hence an appropriate distinction between heterogeneity sources and inefficiency associated variables is done, which might be a useful tool in order to implement a better yardstick indicator for regulatory purposes. Since this analysis was made for a mixed sample including companies from different countries, a practical application of the results found is hardly possible However, it might allow to recognize best practices from the sample studied and the implementation of a sort of competition by comparison with a more informed customer.

References

- Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977). *Formulation and Estimation of Stochastic Frontier Production Function Models.*Journal of Econometrics 6(1), 21-37.
- Alvarez, A. (2001). La medición de la eficiencia y la productividad. Ediciones Piramide.
- Antoniolli, B. and M. Fillipini (2001). *The use of variable cost function in the regulation of the Italian water industry.* Utilities Policy 10(3-4), 181-187.
- Aubert, C. and A. Reynaud (2005). *The impact of regulation in cost efficiency: An empirical analysis of Wisconsin water utilities.* Journal of Productivity Analysis 23(3), 383-409.
- Battese, G., and Coelli, T. (1995). *A model for technical inefficiency effects in a stochastic frontier production function for panel data.* Empirical Economics 20(2), 325-332.
- Bhattacharyya, A., E. Parker and K. Raffie (1995). *An examination of the effect of ownership on the relative efficiency of public and private water utilities.*LandEconomics 70, 197-209.
- Bonifaz, J. L., y Rodríguez, M. (2001). *Distribución Eléctrica en el Perú: Regulación y Eficiencia.* Consorcio de Investigación Económica y Social, Centro de Investigación de la Universidad del Pacífico (CIUP).
- Bottasso, A. and M. Conti (2003). *Cost Inefficiency in the English and Welsh Water Industry: AnHeteroskedastic Stochastic Cost Frontier Approach*. DIEM, University di Genova.
- Caudill, S.B. and J.M. Ford (1993). *Biases in Frontier Estimation Due to Heteroskedasticity.* Economic Letters 41, 17-20.
- Caudill, S.B. J.M. Ford and D.M. Gropper (1995). *Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroskedasticity*. Journal of Business and Economic Statistics 13(1), 105-11.
- Chambers, R. G. (1988). *Applied Production Analysis: A dual approach*. Cambridge University Press.
- Coelli, T., Prasada R., y Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers.
- Conrad K. and D.W. Jorgensen (1977). *Tests of a model of production for the Federal Republic of Germany*. EuropeanEconomicReview 10, 51-57.
- Da Silva e Souza, G., R. Coelho de Faria and T. B. Moreira (2005). *Estimating the relative efficiency of Brazilian publicly and privately owned water utilities: A stochastic frontier approach*. Journal of the American Water Resources Association (JAWRA).
- Da Silva e Souza, G., R. Coelho de Faria and T. B. Moreira (2007).*Estimating the relative efficiency of Brazilian publicly and privately owned utilities: a stochastic cost frontier approach.* Journal of the American Water Resources Association 43 (5), 1237 - 1244.
- Estache, A., and M. A. Rossi (2002). *How different is the efficiency of public and private water companies in Asia?* The World Bank Economic Review 16(1), 39-148.
- Filippini, M., N. Hrovatin and J. Zoric (2008). *Cost efficiency in Slovenian water distribution utilities: An application of stochastic frontier methods.* Journal of Productivity Analysis 29(2), 169-182.
- Franquelli, G. and V. Moiso (2005). *Cost Efficiency and Scale Economies in the Italian Water Industry, Higher Education and Research on Mobility Regulation and the Economies of Local Services (HERMES)*, Working Papers 8.
- Giannakisa, D., T. Jamasbb and M. Pollit (2005). *Benchmarking and Incentive regulation of quality of service: an application to the UK electricity distribution networks*. Energy Policy 33, 2256-2271.

- Greene, W.H. (2005). *Reconsidering heterogeneity in panel data estimators of the stochastic frontier model.* Journal of Econometrics 126, 269-303.
- Hadri K. (1999). *Estimation of a doubly heteroskedastic Stochastic Frontier Cost Function*. Journal of Business and Economic Statistics 17, 359-63.
- Huang, C.J. and J.T. Liu (1994) *Estimation of a Non-neutral Stochastic Frontier Production Function*. Journal of Productivity Analysis 5, 171-180.
- Irastorza, V. (2003). *Benchmarking for distribution utilities: a problematic approach to defining efficiency*. Electricity Journal (Dec.), 30-38.
- Kirkpatrick, C., D. Parker and Y. F. Zhang (2006). *An empirical analysis of state and private sector provision of Water Services in Africa.* The World Bank Economic Review 20(1), 143-163.
- Kodde, D., y Palm, F. (1986). *Wald Criteria for Jointly Testing Equality and Inequality Restrictions.* Econometrica 54(5), 1243-1248.
- Kumbhakar S. (1997). *Modeling inefficiencies in a Translog Cost Function and Share Equations: An Exact Relationship.* Journal of Econometrics. 76, 351-356.
- Kumbhakar S., Ghosh S., y McGuckin J. (1991). *A generalized production frontier approach for estimating determinants of inefficiency in US dairy farms*. Journal of Business and Economic Statistics 9, 279-286.
- Laffont, J. and J. Tirole (1993). *A theory of incentives in procurement and regulation*. MIT press.
- Lin, C. (2005). *Service Quality and Prospects for Benchmarking: Evidence from Peru Water Sector*. Utilities Policy 13(3), 230-239.
- Lynk, E. L. (1993). *Privatization, joint production and the comparative efficiencies of private and public ownership: The UK water industry case*. Fiscal Studies 14(2), 98-116.
- Meeusen, W. and J. Van den Broeck (1977). *Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error.* International Economic Review 18(2), 435-444.
- Pitt, M. and Lee, L-F. (1981). *The measurement and sources of technical inefficiency in the Indonesian weaving industry.* Journal of Development Economics 9(1), 43-64.
- Reifschneider, D. and R. Stevenson (1991). *Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency.* International Economic Review 32(3), 715-23.
- Sabbioni, G. (2006). *Efficiency in the Brazilian Sanitation Sector*. Utilities Policy. Public Utility Research Center. University of Florida, Research Papers 7(8).
- Sauer, J. and K. Frohberg (2007). *Allocative Efficiency of Rural Water Supply: A Globally Flexible SGM Cost Frontier*. Journal of Productivity Analysis 27, 31-40.
- Schmidt, P., and Sickles R. (1984). *Production Frontiers and Panel Data*. Journal of Business & Economics Statistics 2(4), 367-374.
- Söderberg, M. (2007). *An exploration of factors influencing cost inefficiency in Swedish electricity distribution using stochastic frontier analysis*. Gothenburg School of Business, Economics and Law.
- Tupper, H. C. and M. Resende (2004). *Efficiency and Regulatory Issues in the Brazilian Water and Sewerage Sector: An Empirical Study.* Utilities Policy 12, 29-40.
- Worthington, A. (2010). *A review of frontier approaches to efficiency and productivity measurement in urban water utilities*. Griffith University Business School. Discussion Paper 2010-10.
- Yunos, J. and Hawdon, D. (1997). *The inefficiency on the national electricity board in Malaysia: an intercountry comparison using DEA.* Energy Economics 19.

Appendixes

Appendix 1: Panel Structure (Observations Available)

Observations	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	TOTAL
SEDAPAL	0	1	1	1	1	1	1	1	1	1	0	0	9
AGUAS ANDINAS	1	1	1	1	1	1	1	1	1	1	1	1	12
AGUAS ARGENTINAS	0	1	1	1	1	1	1	1	1	1	1	0	10
E.A.A.B.	0	0	0	0	1	1	1	1	1	1	1	1	8
O.S.E.	0	1	1	1	1	1	1	1	1	1	1	0	10
CAGESE / CE	1	1	1	1	1	1	1	1	1	1	0	0	10
COMPESA / PE	1	1	1	1	1	1	1	1	1	1	0	0	10
EMBASA / BA	1	1	1	1	1	1	1	1	1	1	0	0	10
CEDAE / RJ	1	1	1	1	1	1	1	1	1	1	0	0	10
CASAN / SC	1	1	1	1	1	1	1	1	1	1	0	0	10
CORSAN / RS	1	1	1	1	1	1	1	1	1	1	0	0	10
SEANEAGO /GO	1	1	1	1	1	1	1	1	1	1	0	0	10
TOTAL	8	11	11	11	12	12	12	12	12	12	4	2	119

Appendix 2: Firms' Inefficiency Evolution over Time