Enforcement of Yardstick Contracts & Consistency in Performance Rankings: An Application to the England and Wales Regulated Water Industry*

Aude LE LANNIER†

ADIS-GRJM, University of Paris XI-Sud

GREGOR, IAE, University of Paris I Panthéon-Sorbonne

Crigen, GDF-SUEZ

Preliminary Version - 2009

1 Introduction

The UK Government has appointed Professor M. Cave to lead a review of competition and innovation in water markets in order to propose changes to the current legislative and regulatory frameworks. The final objective of the Cave review is to "*deliver benefits to both business and household customers and increase the efficiency of water use*" in this sector.[1] The first section of the Cave review deals with the way to improve competition in water and sewerage industries, which are regulated by the Water Service Regulation Authority (Ofwat).

In privatizing the Regional Water Authorities of England and Wales in November 1989, the government faced the classic regulatory problem: how to prevent the management of privately owned natural monopolies exploiting their market power, through higher prices and reduced quality, in the long run. The solution was innovative. An independent regulator was created, Ofwat, who has the power to limit price rises with a price cap, used in conjunction with a system of comparative or

*L'auteur remercie Stéphane Saussier, Eshien Chong et Miguel Amaral pour leurs conseils et commentaires ainsi que les membres de la Chaire EPPP (Economies des Partenariats Publics-Privés de l'IAE - Paris 1. L'auteur est seul responsable pour toute erreur ou omission.

†e-mail: audelelannier@yahoo.fr

[1]See the Cave review of competition and innovation in water markets: a call for evidence [2008].

yardstick regulation.

As noted by Sawkins [1995], yardstick competition had clearly a place in the 1989 regulatory settlement and the Director General of Water Services, Ian Byatt, confirmed the advantages of this mechanism.[2] This approach has the potential of breaking the direct link between the regulated firm's costs and revenues, by linking the firm's costs to those of comparator(s). Once prices are set, if firms manage to deliver service at a lower average cost than that assumed by the regulator, they keep the resulting benefits. The regulator can thus provide firms with the incentive to increase their efficiency and then return part of the realized cost savings to the customers through a subsequent reduction in prices. Then, high-powered incentives similar to those of a competitive market should be created (see Shleifer [1985]). Provided that a regulator could identify at least one comparator, this benchmarking or yardstick competition became a tool to overcome the information asymmetry which is inherent in regulation.

Today, yardstick competition is widely used by regulators worldwide and has led to various forms of modeling comparative efficiency. However, concerning the incentive regulation used by Ofwat, the Cave review notes that "*a number of stakeholders have suggested that the incentives in the regime may not fully support the demands of consumers, or the environment; and encourage innovation.*" The Cave review raises the question whether the comparative competition could be reformed to better deliver the needs of consumers and the environment without further actual competition. More precisely, the question, in a regulatory context, is to challenge benchmarking methods and specifications of regulated inputs and outputs currently used by Ofwat and to determine whether the regulator should make more use of comparisons between companies' performance.

Shleifer [1985] recognized that some conditions should be met for benchmarking to be efficient and reliable. First, regulators would need to protect against the dangers of "collusive manipulation" by the regulated firms. Second, there would have to be

[2] "*I [Byatt] shall compare the performance of the 39 appointed companies and use the examples of the best to set standards for the others to introduce an element of comparative competition. Such comparisons will cover differences in operating cost, capital cost, levels of service and customer care. There will be allowances for differences such as geographical conditions which are outside the control of efficient managements. These comparisons will help me achieve a better deal for all water customers in England and Wales.*"

a shadow firm (i.e. a comparator) that has cost structures similar to that of the regulated firm's. If cost structures were heterogeneous then the results might be unreliable. Finally, the regulator should be able to commit to disregarding firms' complaints and would, in extremes, have to be prepared to let inefficient companies go bankrupt.[3] To overcome the lack of a perfect shadow firm, Shleifer [1985] noted that multivariate regression models should be developed, which account for differences between firms costs that are outside firm's control, such as network and customer density for instance.

The use of four main empirical methodologies has emerged over time to implement a benchmarking of firm's performances: (1) stochastic analysis of cost and production functions, involving OLS (Ordinary Least Squares), COLS (Corrected Ordinary Least Squares) and, more recently, SFA (Stochastic Frontier Analysis); (2) mathematical modeling, using DEA techniques (Data Envelopment Analysis); (3) productivity and unit cost indices (total factor productivity measures and descriptive statistics); and (4) engineering-based models (see Coelli *et al.* [1998, 2003]). The advantages of benchmarking have been extensively highlighted in the literature, even if there is no consensus on the best method to use. For instance Berg and Lin [2007], Corton and Berg [2009], Botasso and Conti [2003] and Cubbin and Tzanidakis [1998] compare different benchmarking methods and find either that the methods yield similar rankings or underline that firms' efficiency analysis is actually sensible to the empirical methodology employed and to the variables included into the benchmarking process.

However, in yardstick competition applications, the regulator is often interested in obtaining a measure of firms' efficiency in order to reward (or punish) companies accordingly. Hence, there is a close link between efficiency measurements and incentive-based price regulation. If different benchmarking specifications lead to significantly different rankings, this may create tensions and disputes between the regulator and the regulated companies. Indeed, the companies may attempt to take advantage from the incompleteness of the method used, by trying to renege on the incentive scheme implemented by the regulator, thus creating enforcement difficulties of yardstick mechanisms. Berg and Lin [2007] note that "*the regulatory commission reviews studies and establishes performance incentives to achieve policy objectives.*

[3] But as noted by Bös [1991], yardstick competition can be criticized precisely because the threat of bankruptcy may not be credible where firms supply essential services.

Without confidence in the scores and relative rankings, those responsible for creating incentives will not risk their credibility by instituting rewards or applying penalties. Regulators will be unwilling to apply incentives based on performance unless they are very confident that the rankings can survive challenges." Sage [1999] analyzes the implementation of benchmarking by Ofwat and argues that the incompleteness of the method is regularly subject to claims from companies who feel disadvantaged by the variables included in the model and by the non-inclusion of an "essential element of their specific situation". Similarly, Sawkins [1995] explains, in the case of the benchmarking performed by Ofwat, that the regulated companies often challenge the results provided by the regulator.[4] In the same way, Dassler *et al.* [2006] underline the difficulties encountered by Ofwat when applying its benchmarking: "*But even here its use has been controversial with the regulated companies questioning both the models used and their results.*" These enforcement difficulties also occur in other activity sectors, like the gas and electricity sectors, as well as the hospital sector, that are regulated by benchmarking and yardstick competition. For instance, Burns *et al.* [2006] show that Ofgem (the English energy regulator) has encountered difficulties in implementing its benchmarking methods in the context of a yardstick regulation. They show that it is due to a lack of credibility of the benchmarking method used, and more generally to the credibility of the regulatory body. Similarly, Hesseling and Sari [2006], members of the Netherlands Competition Authority (NMa) note that "*The main formal accountability of DTe*[5] *is to the Dutch courts. If a stakeholder does not agree with a decision of DTe, he can take recourse to the specialized court (CBB). Over the past few years, this has proven a very popular tool for both energy companies and end users' organizations. On average, DTe has won about half of these cases." "In the first regulatory period, DTe applied an input-oriented DEA-based benchmark [...]. However, the system was defeated in court.*". Moreover, in the American hospital sector, the *CPB Netherlands Bureau for Economic Policy Analysis* [2000] underlines that renegotiation often occurs concerning the benchmark definition.[6]

[4]"*While both sides [Ofwat and the UK water utilities] admitted that the technical issue of accounting for company heterogeneity had not been completely resolved, companies were more unhappy than the Regulator with the current means of comparison. And while some of this may be quite justifiable, the research did not reveal the extent to which these arguments were part of the process of strategic manipulation: whether or not this was a conscious or unconscious decision by those interviewed.*"

[5]DTe is the Dutch regulatory authority for electricity and gas sectors.

[6]"*A costly element is the recurrent renegotiation [...]. Pooling hospitals with higher than average costs may lobby for such renegotiation.*"

Regulators can use a number of alternative methodologies and data for comparing firms' efficiency, but these approaches need to be robust to be accepted by stakeholders. Indeed, the reliability of efficiency scores is crucial for an effective implementation of yardstick competition. However, it is well known, as showing in empirical studies, that the various benchmarking methods (generally parametric vs non-parametric approaches) often produce different results with respect to firms' efficiency scores and rankings. A possible explanation of this lack of robustness problem could relate to the difficulty of benchmarking methods in finding a consensus on which explanatory variables should be taken into account, how heterogeneity in environmental and network characteristics across companies can be controlled and how quality indicators should be included in the models. This lack of robustness is particularly undesirable if the results are to be used in economic policy-making. As argued by Berg and Lin [2007], "*If the criterion of consistency is not met, these groups [the stakeholders] cannot be confident that the relative performance indicators are meaningful. However, when alternative methodologies yield broadly similar rankings, stakeholders are less likely to engage in acrimonious high-stakes disputes.*"

In this paper, we wonder how the recurrent firms-led renegotiations (i.e. enforcement difficulties) of yardstick competition and benchmarking methods can be explained. Since water distribution utilities operate in different regions with different environmental and network characteristics, it is essential to be able to distinguish between inefficiency and data noises. Therefore, Stochastic Frontier Analysis (SFA) method is used to study relative performances of the UK water distribution utilities. Indeed, the main strength of the stochastic frontier approach is that it deals with stochastic noise. The need for imposing an explicit functional form for the underlying technology and an explicit distributional assumption for the inefficiency term is the main weakness of the stochastic frontier. To limit this weakness, different specifications of stochastic cost functions are tested, in order to check the consistency of rankings provided by this method. The dataset consists of an unbalanced panel of 22 company observations on operating costs (explained variable), input prices (price of labor and other variable inputs), physical outputs (volume of water delivered and number of properties connected), environmental variables (population density, water losses, proportion of water delivered to non-households) and quality indicators (measuring the drinking water quality, the service quality and the "technical" performance of companies), observed over the period 2002-2008. Ofwat has been using benchmarking mechanisms and yardstick competition for several years. It's an experimented and well informed regulator (a lot of data are available). In that sense, UK water and sewerage regulation is often considered as a model of benchmarking implementation. However, as noted before, both the UK government and the regulated companies are now challenging the Ofwat regulation. As far I know, this study is the first to test different specifications of SFA method in the UK water and sewerage industry. This empirical strategy aims to evaluate the consistency of UK water utility performance rankings. In this context, we wonder whether the enforcement difficulties of benchmarking schemes could be due to a lack of information for the regulator or these problems are inherent to the benchmarking tools used by regulators.

The results first indicate that environmental and quality variables should not be ignored when evaluating water utilities performances. Environmental variables should be included directly into the cost function as regressors to control for (observable) heterogeneity between companies[7], whereas some quality variables provide an adequate explanation of technical efficiency and should be included in the inefficiency component.

Second, the study shows that there are some potential difficulties with this approach of evaluating performances. Indeed, performances results and rankings are significantly different, depending on the data retained for the evaluation made by a SFA mechanism. More precisely, the rankings significantly diverge according to the way to include environmental and quality variables into the benchmarking model. The higher consistency in rankings is obtained when physical outputs and environmental variables are directly included as regressors and no quality indicator is taken into account. In this case, 86.36% of the results enable to obtain differences lower than 5 ranks. When introducing quality variables as additional outputs, the models fail to determine a same best and worse performing company. However, the correlation between rankings is still very high and 72.73% of the results enable to limit the differences in rankings (differences lower than 5 ranks) and there is no case where a firm has a difference of ranks higher than 10. Moreover, the results demonstrate that some quality variables should be included in the inefficiency component for the

[7] By introducing environmental and exogenous factors in the model, one can control for observed heterogeneity. However, as argued by Greene [2005a,2005b] and Filippini *et al.* [2008], not all relevant data are always available and some factors may even be too complex to be properly measurable. This could result in unobserved heterogeneity which is beyond the firms' control but may affect their costs significantly.

model specification to be reliable. However, in this case, the correlation between rankings becomes very low and in 13.64% of the cases, differences in rankings are higher than 10.
Note that this study does not aim to make an exhaustive analysis of water industry performance. Hence, the example developed in this paper should be taken to be an illustration of the incompleteness of the benchmarking methods, even in an "ideal" case where a lot of data are available (i.e. the case of UK water and sewerage utilities regulation). Therefore, the lack of consistency, and indirectly the enforcement difficulties, don't seem to be due to a lack of information for the regulator. On the contrary, the results indicate that, when adding more variables into the SFA specification (more precisely, when quality variables are added), the consistency in rankings is reduced. This issue may explain the fact that companies regulated by yardstick mechanisms often try to renege on the regulatory decisions based on the results of benchmarking. Finally, this paper contains discussion on the degree to which relative performances evaluations and yardstick competition should be used in a "light-handed" manner in regulatory policies. Indeed, if rankings or scores are to be used in regulatory proceedings, great care must be taken to avoid unduly penalizing utilities, since the direct use of inefficiency estimates in the regulation of water distribution utilities may be misleading.

The paper is organized as follows. Section 2 shortly reviews studies estimating the cost function of water distribution companies and the UK experience of benchmarking in this sector. The data description is provided in Section 3 and Section 4 presents the model specifications and the methodologies employed. In order to check the robustness of our results, our empirical strategy is to estimate different models, with different specifications in terms of physical output variables, environmental variables and quality indicators. Moreover, we test different assumptions concerning the cost inefficiency and the evolution of performances over the time period considered in order to analyze the robustness and reliability of obtained cost inefficiency scores and rankings. The estimation results are given in Section 5 and Section 6 concludes the paper.

2 Benchmarking Water Utilities & The UK Experience

2.1 Past Studies on Benchmarking in Water Sectors

There is a growing importance of benchmarking of water utilities and various empirical studies analyze the impact of different variables on firms' efficiency.[8] For instance, some papers recently have specifically focused on assessing efficiency in water utilities explicitly accounting for quality. As far as I know, the first paper that did take quality into account when measuring water utilities' performance was that by Saal and Parker [2000,2001]. More recently, Antonioli and Filippini [2001], Estache and Rossi [2002], Tupper and Resende [2004], Lin [2005], Saal and Parker [2006] and Saal *et al.* [2007], Berg and Lin [2007], Bouscasse, Destandau and Garcia [2008], Picazo-Tadeo *et al.* [2008] also develop the analysis of quality variables in benchmarking in different countries.

Recent studies have focused on the relative performance of public and private water utilities in various countries.[9] For instance, Estache and Kouassi [2002], Estache and Rossi [2002] and Bouscasse *et al.* [2008] analyze the effects of ownership on utility performance and find significant differences between private and public water utilities: the private operators are more efficient than public operators. For their part, Kirkpatrick *et al.* [2006] and Seroa da Motta and Moreira [2006] find no significant evidence of difference between ownership types.

The relevance of firm heterogeneity has also been emphasized by several authors, like for instance Bhattacharyya *et al.* [1995], Ashton [2000], Tupper and Resende [2004] and Filippini *et al.* [2008].

Other authors analyze the impact of regulation on firms' efficiency (see for instance Aubert and Reynaud [2005] and Saal *et al.* [2007]) and the economies of scale (see for instance Saal and Parker [2005], Garcia *et al.* [2007] and Filippini *et al.* [2008]).

Although the advantages of benchmarking have been extensively highlighted in the literature, there is still no consensus on the best method to use. For instance, Berg and Lin [2007] compare the results provided by a SFA and a DEA-based benchmarking in the case of water utilities in Peru and find that these methods yield similar

[8] See von Hirschhausen *et al.* [2008] for a recent and detailed survey on benchmarking in water utilities.

[9] See Estache *et al.* [2004] for a comprehensive literature review. They survey productivity and efficiency literatures in infrastructure industries in developing countries.

rankings. In addition, the techniques have comparable success for identifying the best and worse performing utilities. Similarly, Corton and Berg [2009] show that SFA and DEA methods provide the same best performer when comparing the water utilities in 6 countries in the Central American region.
However, some authors have underlined the fact that firms' efficiency analysis is actually sensible to the empirical methodology employed and to the variables included into the benchmarking process (See for instance Bauer *et al.* [1998], Botasso and Conti [2003], Jamasb and Pollitt [2003], Farsi, Filippini and Greene [2005] and Farsi and Filippini [2006]). Cubbin and Tzanidakis [1998] find that the rankings based on the OLS model are different from the ones based on DEA models when comparing the UK water utilities.[10]
Hence, different methodologies could provide different rankings. But this problem may also appear within one method, depending on the specifications of the model. Despite the importance of this issue, only a few papers examine the sensitivity of efficiency and rankings based on different specifications within one benchmarking method. For instance, Cubbin and Tzanidakis [1998] explain that the scores and rankings may change, depending on the variables used in a DEA-based benchmarking.[11] They also highlight the fact that the rankings may be altered by the nature of the returns to scale in a DEA model.[12]

With respect to the previous studies in the water distribution sector, the methodological contribution of this paper is that it analyzes in detail the results provided by different specifications of SFA method, depending on the variables generally included into this type of model. Moreover, the results of this empirical analysis could be of interest in a regulatory context, since the study analyzes the consistency between the rankings provided by the different specifications. Hence, this paper contains discussion on the degree to which relative performances evaluations and yardstick competition should be used in a "light-handed" manner in regulatory policies.

[10]"*Under ideal conditions both RA (Regression Analysis) and DEA should be able to identify the true company efficiency and produce similar results. This is not, however, the case in practice since application of these two techniques on the same data set often produces strikingly different efficiency results, particularly in the small samples which are prevalent in regulated industries.*"

[11]"*Equally worrying, however, are the changes in scores and rankings when we alter the specification of the DEA model. With two more variables [...], some companies showed dramatic improvements in their DEA scores [...].*"

[12]"*Another example of this feature is that in DEA with constant returns there are three companies appearing as fully efficient while in DEA with variable returns 13 out of 29 companies appear as fully efficient.*"

2.2 The UK Experience of Benchmarking

Efficiency analysis and benchmarking was first applied to the price reviews of the UK water industry.[13] After the Regional Water Authorities of England and Wales were privatized in November 1989, the successors Water and Sewerage Companies (10 WASCs) and Water Only Companies (12 WOCs) faced a new regulatory regime that was designed to promote economic efficiency while simultaneously improving drinking water and environmental quality.

Ofwat placed considerable emphasis on yardstick competition in price cap setting, and it is clear that the system of regulation comes closest in practice to Shleifer's [1985] benchmarking model. Ofwat assess the relative efficiency of the water and sewerage companies in order to set price limits for the local monopolies. In the absence of a competitive market, Ofwat also use comparative competition to drive the companies to become more efficient. The regulator monitors and publishes the companies' progress each year. At each price review, Ofwat use its assessments to derive efficiency factors to include in price limits and to identify which companies qualify for enhanced future incentives. Hence, Ofwat has two objectives for comparing firms' performances:

1. The sunshine regulation leads to publish the rankings of firms' performances and distribute the information regarding the relative performances of the regulated firms. If a firm appears less efficient than the others, this mechanism will have a negative effect on the firm's reputation. The distribution of the comparison results will act as an indirect competition pressure, exerted above all by stakeholders (users, medias, politicians, NGOs...).

2. The comparisons are used to set the price cap during the price reviews. The regulator calculates the companies' costs, taking into account a number of indicators to capture some of the heterogeneity of operating conditions.

These price reviews are conducted by Ofwat every 5 years (1994, 1999, 2004, 2009). The approach used by Ofwat in the 1994 review was described in detail by Thanassoulis [2000a,2000b]. Ofwat applied Data Envelopment Analysis and the water activities were identified by "distribution", "resources & treat" and "business activities". The outputs chosen were the number of connections, the length of main and the volume of water delivered (measured and delivered) with the operating expenditures as

[13] See for instance Sawkins [1995], Cubbin [2005], Allan [2006], Dassler *et al.* [2006], von Hirschhausen *et al.* [2008] for a description of the Ofwat regulatory policy.

an input. The efficiency results were then compared to regression results, and entered into the price determination with the exact usage being confidential to Ofwat. Therefore, price caps are not automatically determined by Ofwat (Stern [2005]). Although Ofwat still used econometric techniques for the determination of price caps in 1999, DEA was no longer employed (Dassler *et al.* [2006], von Hirschhausen *et al.* [2008]). For the 2004 price review, Ofwat commissioned studies comparing the results of OLS with DEA and SFA. At the time this paper was written, Ofwat had not clearly announced its plans of using either DEA or SFA in the 2009 price review. The first review of water and sewerage charges after privatization came in 1994. Ofwat maintained that there were difficulties in calculating total expenditure due to problems with data collection and making comparisons between companies serving different geographical areas (Dassler *et al.* [2006]). Therefore, operating costs and capital costs were modeled separately, as were water supply and sewerage services.

In parallel, each company must report to the regulator some information regarding its performance in terms of quality of service. The "*Report on levels of service for the water industry in England and Wales*" evaluates the firms' performance for 8 quality standards: (1) Reliability of the water distribution system (DG2, DG3); (2) Water resources management by the company (DG4); (3) Reliability and quality of the sewerage network (DG5); (4) Reactivity of each company concerning its customers' complaints (DG6, DG7, DG8, DG9).

If Ofwat argues that it uses companies' costs and operating conditions to assess their relative efficiency, the regulator specifies that high costs do not always indicate inefficiency, as a high cost company may be operating in a particularly unfavorable environment. Similarly, low costs do not necessarily point to high efficiency, while rising costs do not automatically indicate that a company is becoming less efficient. For example, many companies face increasing costs for operating new treatment works to meet higher quality standards. Hence, in both the operating expenditure and capital maintenance models, Ofwat has reduced the modeling residuals (the difference between actual costs and the costs predicted by the models) in order to take some account of possible errors in the data and in the statistical process. In other words, Ofwat recognizes the importance of specific and environmental conditions when comparing firms' performances. Thus, the benchmark company is not always the company at the efficiency frontier. It needs to satisfy a number of criteria, including size, to make it suitable for comparison with the rest of the industry.

The water and sewerage industry regulation is thus a particularly interesting case. First, Ofwat is the first regulator to have explicitly developed yardstick competition and benchmarking in his regulatory mechanism. Second, Ofwat is an experimented regulator who has collected a lot of information about this industry. Moreover, this information is available from the Ofwat website. However, as introduced above, enforcement difficulties and recurrent firms-led renegotiation occur in this sector. Therefore, testing different specifications of SFA method on the UK water industry enables to wonder whether the potential enforcement difficulties of benchmarking and yardstick mechanisms come from informational problems for the regulator or from the incompletness of the method used.

3 The Data

The dataset used in this study consists of an unbalanced panel of 22 company observations on both Water And Sewerage Companies (10 *WASCs*) and Water Only Companies (12 *WOCs*) observed over the period 2002-2008. The source of data essentially comes from the "*June Returns for the Water and Sewerage industries in England and Wales*" published by Ofwat each year. In the empirical application we focus only on the water service and we do not consider the sewerage one.

The dependent variable. The dependent variable is **the operating costs (*Opex*)**, like for Cubbin and Tzanidakis [1998] and Thanassoulis [2000] in the case of water companies in England and Wales, Corton [2003] for water companies in Peru, Tupper and Resende [2004] for water and sewerage companies in Brazil and Corton and Berg [2009] for Central American water utilities. The operating costs in the Ofwat database include energy, employment and material costs, but exclude the costs of providing third party services. It also excludes atypical and exceptional costs, such as restructuring costs.

The explanatory variables. The physical outputs or explanatory variables are:

1. The volume of water delivered (*Vol*);

2. The number of properties connected for water supply only (*Prop*).

Several papers have highlighted that improvement in assessing efficiency of water utilities can be accomplished if both the volume of water delivered and the number of connected properties are considered as outputs (Thanassoulis [2000], Garcia

and Thomas [2001], Saal and Parker [2006], Picazo-Tadeo *et al.* [2007], Corton and Berg [2009]). Saal and Parker [2006] argue that this specification is relevant because the characteristics of outputs associated with the volume of water delivered to existing customers are rather different from those required for the provision of new connections. Moreover, both outputs have substantially different marginal costs.
The volume of water delivered (*Vol*) is a conventional measure of the water production activity and is represented, in the Ofwat database, by the total volume of water delivered and billed to households and non-households.
The number of properties connected for water supply only (*Prop*) includes the total number of household and non-household water-connected properties. As noted by Saal and Reid [2004], previous researches (see for instance Antonioli and Filippini [2001] and Garcia and Thomas [2001]) have suggested that because of the cost of maintaining network connections, the number of connected properties is an important determinant of water industry costs. According to Erbetta and Cave [2007], this specification is a proxy for the scale of the distribution activity.
We expect that the higher these two variables, the higher the operating costs.[14]

Environmental Variables. The efficiency of a firm could be affected by exogenous conditions that are not under the direct control of managers. Environmental variables have been included because they may influence the technology under which water utilities operate and may account for exogenous differences in operating environment experienced by each firm (see Bhattacharyya *et al.* [1995], Garcia and Thomas [2001] among others). These variables enable to take account of the impact of the different characteristics of the network and of the area where the service is provided, thus controlling for heterogeneity among firms[15]:

1. The population density (PD);
2. The percentage of water losses ($Loss$);
3. The proportion of water delivered to non-households ($NonH$).

[14] Following Ofwat, we also tested the length of mains as an output, like Cubbin and Tzanidakis [1998], Thanassoulis [2000] and Corton [2003] for instance. Thanassoulis [2000] argued the length of mains reflects the geographical dispersion of connections. For Berg and Lin [2007], this variable is an indicator of capital. However, in our dataset, the lenght of mains is highly correlated to the volume of water delivered. Because the volume of water delivered is commonly used as the main output in water utilities, the lenght of mains indicator is dropped from the tests.

[15] The environmental variables used are consistent with many of the mentioned empirical studies. See for instance Erbetta and Cave [2007].

As noted by Saal and Reid [2004], an extensive literature has included measures of the density of operations as an important determinant of water industry costs (see for instance Bhattacharyya *et al.* [1995], Cubbin and Tzanidakis [1998], Antonioli and Fillippini [2001], Estache and Rossi [2002]). Therefore, the water service density or, in other words, **The population density (*PD*)** is included in our specification and is defined as population per kilometer of water main (i.e. the ratio between the population provided with water and the length of mains). For Erbetta and Cave [2007], providing service to a more concentrated population is, generally, cheaper than providing a dispersed population. The idea is that more dispersions of the network, more frequent maintenance and more energy are needed. However, as argued by Botasso and Conti [2003], the population density may have ambiguous effects on cost inefficiency since, one the one hand, it may be more expensive too service dispersed customers, but in the other hand, a higher density may create congestion problems.

The percentage of water losses (*Loss*) with respect to the total volume of water delivered (Distribution losses/Volume of water delivered) is a general proxy for the operational condition of the distribution network. According to Erbetta and Cave [2007], "*a higher proportion of losses implies more critical conditions of the network, thus a higher input use is expected*". However, other studies use the water losses to take account for deficiencies in either operational or commercial practices. Indeed, as argued by Corton and Berg [2009], water losses may reflect a cost trade off between increasing water production and repairing network leaks to keep up with water demand. Hence, the idea is that, to satisfy demand, managers may find it more costly to repair leaks and to control water losses than to increase water production. For Garcia and Thomas [2001], water network losses are considered as a non-desirable output produced jointly with the service of water delivery. For their part, Coelli *et al.* [2003] regard water losses as an indicator of the technical quality of service. In this study, we assume, like Erbetta and Cave [2007], that the percentage of water losses reflects an environmental and exogenous variable, which might control for heterogeneity between firms.

The proportion of water delivered to non-households (*NonH*) is the share of water delivered to non-households customers on total water delivered. It reflects the cost savings associated with supplies to larger customers. It is a proxy for the importance of large (industrial) users. The idea is that a higher proportion of large users is expected to reduce cost inefficiency because it is cheaper to distribute the same amount of water to a few large users than to an high number of small customers (see Cubbin and Tzanidakis [1998], Botasso and Conti [2003] and Corton [2003]).

Finally, we add two dummy variables:

1. A "regulatory dummy" (*Reg*);
2. An "activity dummy" (*Act*).

Indeed, an other type of exogenous variable may be considered through a policy variable that relates to regulatory policy and more specifically, the change in the economic environment that occurs after the introduction of new regulatory constraints. Because there has been one price control (in 2004) during the period under observation, we introduce in the model **a first dummy variable (*Reg*)**, which assumes a value of 1 for the three years after 2004-2005. The inclusion of this dummy variable allows a test of whether, after this price review, the performance of the companies was significantly greater. Finally, we add a variable reflecting the activity of companies as **another dummy variable (*Act*)**, which assumes a value of 1 for the 10 *WASCs* and 0 for the 12 *WOCs*. This dummy should pick up technology differences existing between the *WASCs* and the *WOCs*.

Quality Indicators. In addition to traditional measures of technical efficiency, service quality is a performance indicator that warrants attention, since one important characteristic of water companies is that they must comply with quality standards. In the water industry, the variables representing quality might differ considerably from one country to another. In some developing countries, service coverage, service continuity or the percentage of water receiving chemical treatment are adequate variables to measure water quality (see for instance Lin [2005] and Berg and Lin [2007] in the case of Peru, Corton and Berg [2009] for the Central American water utilities). In contrast, in developed countries where water services cover nearly all the population and water quality reaches higher standards, alternative measures of quality are required (see for instance Alegre *et al.* [2006], Picazo-Tadeo *et al.* [2007] and Bouscasse *et al.* [2008]). However, this paper does not aim to discuss the reliability of different variables to measure quality in water utilities, but rather to analyze whether the introduction of quality indicators changes the results of efficiency measurement. For that purpose, we retain the following quality variables:

1. Concerning the drinking water quality, we retain the percentage mean zonal compliance with drinking water Regulations (*Drink*).

2. To measure the service quality provided to customers, we use:

    - A measure of the perception by users of the offered quality of services on the basis of the customers' complaints (*NonComp*);
    - The percentage of written complaints dealt with within 10 working days (*Written*);
    - The Call handling satisfaction indicator (*Call*).

3. To measure the "technical" performance of companies, we use:

    - The percentage of properties receiving standard water pressure (*SPress*);
    - An indicator reflecting the performance in terms of water supply interruptions (*NInterrupt*).

The Drinking water quality variable (*Drink*) reflects the percentage mean zonal compliance with samples taken according to the current *Drinking Water Quality Regulations* during the calendar year (see for instance Saal and Parker [2000]). These variables have been collected on the *Drinking Water Inspectorate*'s website.[16]
To measure the service quality, we use three proxies reflecting the "perception by users of the quality of services provided" (see for instance Bouscasse *et al.* [2008]). We use the number of customers' complaints (like Corton and Berg [2009] and Bouscasse *et al.* [2008], among others), which includes the total number of written and telephone customers' complaints for various reasons. This variable is available in the annual *June Returns* of each company and on the Ofwat's website. On the basis of this variable, we calculate **the share of customers who do not complain during the year (*NonComp*)**. A small number of complaints indicates a higher quality of service. Therefore, we expect that the higher the share of non-complaining customers, the higher the operating costs.
The Ofwat regulation includes different quality standards, in particular the DG7 and DG9 indicators: **The percentage of written complaints dealt with within 10 working days (*Written*)** and **the call handling quality (*Call*)**.[17] We expect that the higher these variables, the higher the service quality provided by the company and therefore, the higher the operating costs.
The last two quality variables are also included in the Ofwat's quality standards (DG2 and DG3 quality indicators) and represent a measure of the reliability of the

[16] http://www.dwi.gov.uk
[17] It represents an annual satisfaction score generated by 4 waves of customer surveys.

distribution network. We use the percentage of properties receiving low water pressure, which is available on the *June Returns*. It represents the part of the properties in the company's area of water supply which, at the end of the year, has received and is likely to continue to receive a pressure or flow below the reference level. As noted by Saal and Reid [2004], "*Improvements in water pressure require substantial expenditure on leakage control and improved system design and management. Moreover, improved pressure was an important quality parameter pursued by Ofwat in the years following privatization.*" On the basis of this variable, we calculate **the percentage of properties receiving standard water pressure (*SPress*)**.
We also use the water supply interruptions indicator of the *June Returns*, which reflects the percentage of properties in the company's area affected by unplanned and unwarned supply interruptions (greater than 6 hours, 12 hours and 24 hours). On this basis, we calculate **the percentage of properties that is not affected by unplanned and unwarned supply interruptions (*NInterrupt*)**. The higher these two variables, the higher the reliability of the water distribution network. Hence, we expect that the higher these variables, the higher the operating costs.

The variables are summarized in Tables 1 and 2.

4 The SFA Model Specifications

The setting of the X factor in a price cap regulation is always the subject of debate. This issue has encouraged some regulators to consider the use of industry benchmarks in the setting of X factors.[18] This generally involves the calculation of firm-level measures of relative efficiency, using a method such as OLS, DEA[19] or SFA.[20] These methods have the advantage that they provide greater incentives for efficiency improvements. However, it is often difficult to capture all aspects of a particular businesses' operating environment in a single model.

OLS methods are well known and easy to implement, however they could be criticized since they require the specification of a functional form for the production

[18] See Coelli *et al.* [1998] for details regarding these various methods and their relative merits.

[19] See Charnes, Cooper and Rhodes [1978]. A comprehensive description of the use of DEA for regulatory purposes is provided in Thanassoulis [2000]. See also Charnes *et al.* [1994] and Cooper *et al.* [2004] for a survey of DEA applications.

[20] See Kumbhakar and Lovell [2000] for an exhaustive analysis on stochastic frontier and an overview of various SFA methodologies.

Table 1: Variables in SFA

Name	Description	Role in SFA	Sign
Opex	Operating expenditure	Dependant	
	Physical outputs	Explanatory	
Vol	Volume of water delivered		+
Prop	Connected properties		+
	Quality indicators	Explanatory or inefficiency variables	
Drink	Mean zonal compliance with drinking quality water standards		+
NonComp	Customers who do not complain		+
Written	Written complaints dealt with within 10 days		+
Call	Call handling quality		+
SPress	Properties receiving standard pressure		+
NInterrupt	Properties not affected by unplanned interruptions		+
	Environmental variables	Control variables or inefficiency variables	
PD	Population density		+/-
Loss	Percentage of water losses		+
NonH	Water delivered to non-households		-
	Dummies	Control variables or inefficiency variables	
Reg	Regulatory variable		-
Act	Activity variable		-

Operating expenditures in millions of GB £
Volume of water delivered in megaliters per day
Properties connected for water only in number
Mean zonal compliance with drinking quality standards in percentage
Share of customers who do not complain during the year in percentage
Written complaints dealt with within 10 working days in percentage
Call handling quality is a score
Properties receiving standard water pressure in percentage
Properties not affected by unplanned and unwarned supply interruptions in percentage
Population density reflects the population provided with water per kilometer of network
Water losses is the proportion of losses to the total volume of water delivered in megaliters per day
Proportion of water delivered to non-households in megaliters per day

Table 2: Sample summary statistics
(22 companies in 2002-2008)

Variables	**Mean**	**Standard deviation**	**Min**	**Max**
$Opex$	78402.43	6925.60	5204	353799
Inputs price				
P_L	12.02	21.06	1.50	129.21
P_M	46.09	26.26	21.94	315.84
Physical outputs				
Vol	562.29	576.06	24.56	2179.44
$Prop$	286203.4	288855.6	59525	1262225
Quality indicators				
$Drink$	99.91	0.10	98.96	100
$NonComp$	96.27	5.09	74.75	99.97
$Written$	98.70	5.49	64.79	100
$Call$	4.49	0.15	4.10	4.81
$SPress$	96.67	4.18	62.00	100
$NInterrupt$	99.51	1.63	81.55	100
Environmental variables				
PD	68.83	15.52	46.27	110.32
$Loss$	16.72	5.06	6.78	32.26
$NonH$	28.55	5.76	17.27	49.84
Dummies				
Reg	0.50	0.50	0	1
Act	0.45	0.50	0	1

technology, like SFA. Moreover OLS methods provide information on average performance rather than frontier performance. DEA and SFA methods address this latter problem by building efficiency frontier. As noted by Dassler *et al.* [2006], Ofwat has commissioned a comparison of different benchmarking methods since 2000. DEA and SFA were compared with OLS results and were seen as credible alternatives to OLS regression.

DEA does not require any specification of the functional form of the production relationship but develops a frontier relating inputs to outputs. SFA constructs the efficiency frontier with a sophisticated economic specification of the production relationship. Its advantage is that the approach attempts to account for the effects of noise in the data (data errors and omitted variables). With both DEA and regression analysis, all deviation is attributed to inefficiency. Moreover, standard statistical tests can be used with SFA to test hypotheses on model specification and significance of the variables included in the model. The limitation of SFA is that, in accounting for noise, it assumes that random shock and inefficiency display a specific distribution. As argued by Filippini *et al.* [2008], since water distribution utilities operate in different regions with different environmental and network characteristics that are only partially observed, it is essential to be able to distinguish between inefficiency and data noises. Therefore, we adopt in this study a SFA approach, originally proposed by Aigner, Lovell and Schmidt [1977] and Meeusen and Van den Broeck [1977], assuming that deviations from the best practice frontier might be due to both inefficiency and other random factors.

4.1 The stochastic cost frontier

This study examines cost efficiency by statistically estimating cost relationships given a level of output produced. SFA constructs an efficiency frontier with a sophisticated economic specification of cost relationship. The cost frontier is chosen because it can accommodate multiple outputs easily. A cost frontier shows costs as a function of the level of outputs and the prices of inputs. Conceptually, the minimum cost function defines a frontier showing costs technically possible associated with various levels of inputs and control variables.[21]

The stochastic cost frontier can be expressed as:

$$C_i = C(y_i, p_i, q_i, e_i, \beta) \times exp\{\epsilon_i\} \quad i = 1, ..., N \tag{1}$$

Where:

- C_i is the cost of the company i;
- y_i is a vector of output quantities of the company i;
- p_i is an input price vector faced by firm i. The input prices include the price of labor and other variable inputs (essentially power and materials and services). The price of labor (P_L) is equal to total employment cost[22] divided by total number of employees. The price of other inputs (P_M) is equal to the difference between total inputs costs and labor cost, divided by total volume of water delivered;
- q_i is a vector of quality level for company i;

[21] A production frontier reveals technical relationships between inputs and outputs of firms and represents an alternative when cost frontiers can not be calculated due to lack of data. The estimated output is the maximum possible output for given inputs of an individual firm.

[22] The employment costs are the sum of the total costs of "non-manual and manual manpower" which are directly attributable to each water service (e.g the gross salaries and wages of all employees within the water activity, including payments resulting from bonus and profit-related payment schemes, employer's National Insurance contributions, superannuation, unfunded pension liabilities, private health insurance…). The employment cost is not readily available and had to be compiled using information in the companies' regulatory and/or group annual accounts.

- e_i is a set of contextual factors reflecting the environment of firm i;
- β is a vector of technology parameters to be estimated.

With:
$y \in \{Vol, Prop\}$; $p \in \{P_L, P_M\}$;
$q \in \{Drink, NonComp, Written, Call, SPress, NInterrupt\}$;
$e \in \{PD, Loss, NonH, Act, Reg\}$.

Equation (1) is the cost frontier common to all water utilities, which determines the minimum operating cost achievable for a given set of outputs, input prices, control variables and quality level.

4.2 Taking the panel data into account

In early applications of SFA models to panel data (Pitt and Lee, [1981], Schmidt and Sickles [1984] and Battese and Coelli [1988]), the common assumption was that the productive efficiency is a time-invariant characteristic, i.e. the inefficiency is assumed to be constant over time. This can be rather limiting assumption, particularly in long panels. However, it may be a plausible assumption in non-competitive operating environment. More recent papers (see for instance Battese and Coelli [1992]) proposed a time-variant model to deal with the SFA panel data, in order to account for variation of efficiency. However, Coelli *et al.* [2003] argue that a time varying efficiency model restricts the technical efficiency of all firms since they follow the same trend direction (either all increasing or all decreasing over time), and it is unlikely to be valid in many instances.
In the case of time-varying model, according to the recommendations of Coelli *et al.* [2003] and Estache *et al.* [2004], a time trend is added to the cost function to capture the technical change.[23]
Therefore, equation (1) can be expressed as:

$$C_{it} = C(y_{it}, p_{it}, q_{it}, e_{it}, t, \beta) \times exp\,\{\epsilon_{it}\} \tag{2}$$

In order to check the robustness of our results, these two stochastic frontier methods for panel data are used in this study.

[23] See also Lin [2005].

4.3 Modeling inefficiency effects

In the stochastic cost frontier, the error term ϵ_i can be decomposed in two parts in order to separate noise from inefficiency in the model, with:

$$\epsilon_{it} = v_{it} + u_{it} \tag{3}$$

Where u_{it} is a positive one-sided disturbance that captures the effect of firm- and time-specific cost inefficiency. Therefore, u_{it} is associated with the inefficiency of the level of operating costs in the UK water companies, given the levels of outputs and input prices. In other words, this non-negative error component reflects the inability if a firm i at the observation t to attain the potential minimum operating cost defined by the stochastic frontier (2). v_i accounts for measurement errors in the operating costs, for the effects of unspecified explanatory variables in the model and for other company-specific random factors.
To separate noise from inefficiency in the model, typically one of a number of possible distributional assumptions on ϵ_i is made. Different distribution models are tested in this study in order to reduce the impact of choosing a specific distribution function arbitrarily. The main distributional assumptions retained in the existing literature are the following:

1. First assumption: an half-normal model.
   The idiosyncratic error term v_i is independently and identically distributed as $v_i \sim N(0, \sigma_v^2)$, while $u_i \sim iid\ N^+(0, \sigma_u^2)$. v_i and u_i are distributed independently of each other and of other regressors.

2. Second assumption: a truncated normal model.
   The idiosyncratic error term v_i is still independently and identically distributed as $v_i \sim N(0, \sigma_v^2)$, while $u_i \sim iid\ N^+(\mu, \sigma_u^2)$.

Battese and Coelli [1995] proposed a conditional mean efficiency model based on a truncated normal model in order to identify some of the reasons for differences in predicted efficiencies among firms in an industry. The model can be expressed as:

$$\mu_i = z_i \delta \tag{4}$$

μ_i is the mean parameter of the truncated normal distribution (of the truncated normal model). z_i is a vector of variables, which may influence the efficiency of a company. δ is a vector of parameters to be estimated.

The vectors of parameters δ and β are estimated by the Maximum likelihood method, as the associated parameters $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\gamma = \sigma_u^2/\sigma^2$. The parameter γ must lie between 0 and 1 and provides a useful indication of the relative contributions of u_{it} and v_{it} to ϵ_{it}. As $\gamma \to 0$ the symmetric noise component v_{it} dominates the one-sided cost inefficiency term u_{it} in determining the variation of global residual ϵ_{it}. When γ moves toward 1, the relative effect of u_i is more important.

Concerning, the environmental variables, Coelli *et al.* [1999] suggest that the literature offers two alternative approaches to the inclusion of these variables. One assumes that environmental variables influence the shape of the technology and that these factors should be included directly into the cost functions as regressors. For the other approach, environmental factors directly influence the degree of technical efficiency and should be included in equation (4).
Similarly, Lin [2005] and Bouscasse *et al.* [2008] consider the quality variables either as environmental variables that may influence the efficiency of a firm (quality variables are therefore included in equation (4)) or as additional outputs of the cost function.[24] On this issue, different approaches have been followed in the literature on water industry cost structure. For example, Bhattacharyya *et al.* [1995] include, among others, a dummy variable for typology of sources in the frontier and a proxy for the presence of industrial users in the inefficiency term. The same author, in another paper, includes the sources dummy in the inefficiency term.
In order to check the robustness of our results, our empirical strategy is to estimate different models, where the environmental and quality variables alternatively enter as additional variables in the cost frontier or as variables explaining cost inefficiency.

4.4 The translog cost function

Estimation of cost function requires a specification of the functional form. The most frequent are the translog (transcendental logarithmic) function and the Cobb-Douglas function (Kumbhakar and Lovell [2000]). The translog cost function is generally recognized to be a more flexible functional form than the Cobb-Douglas one (see Chambers [1988], among others).[25] For that reason, the translog functional form is applied in this study.

[24]Quality variables could be regarded as additional outputs since the fulfillment of quality programs is usually expensive.

[25]The translog function is a flexible functional form which approximates any twice differentiable function without imposing any a priori restrictions on the production technology.

Quality and Environmental variables in the cost function. If we assume that quality and environmental variables are additional variables in the cost frontier (additional outputs for the quality variables, control variables for environmental variables), the translog form of (2) can be written in the following way:

$$\begin{aligned}
\ln C_{it} &= \beta_0 + \textstyle\sum_m \beta_m \ln p_{mit} + \sum_j \beta_j \ln y_{jit} + \sum_k \beta_k \ln q_{kit} + \sum_l \beta_l \ln e_{lit} \\
&+\tfrac{1}{2} \textstyle\sum_m \sum_{m'} \beta_{mm'} \ln p_{mit} \ln p_{m'it} + \tfrac{1}{2} \sum_j \sum_{j'} \beta_{jj'} \ln y_{jit} \ln y_{j'it} \\
&+\tfrac{1}{2} \textstyle\sum_k \sum_{k'} \beta_{kk'} \ln q_{kit} \ln q_{k'it} + \sum_m \sum_j \beta_{mj} \ln p_{mit} \ln y_{jit} \\
&+ \textstyle\sum_m \sum_k \beta_{mk} \ln p_{mit} \ln q_{kit} + \sum_j \sum_k \beta_{jk} \ln y_{jit} \ln q_{kit} \\
&+ d_1 Act + d_2 Reg + \alpha T + v_{it} + u_{it}
\end{aligned} \tag{5}$$

With $(\beta_0, \beta_m, \beta_j, \beta_k, \beta_l, \beta_{mm'}, \beta_{jj'}, \beta_{kk'}, \beta_{mj}, \beta_{mk}, \beta_{jk}, d_1, d_2, \alpha)$ the parameters to be estimated.

T is a time trend variable used to capture the time varying effect which is common to all the utilities in a specific time period, like for instance technology change or policy change (Berg and Lin [2007]). Therefore, the time variable is interpreted as a proxy for technological changes but not for changes in technical efficiency conditions, which are embodied in the one-sided distributed error component.

The properties of the cost function are that it is concave, symmetric and linearly homogeneous in input prices[26], nondecreasing in input prices and nondecreasing in output.

The explanation of cost inefficiencies. As noted before, the model proposed by Battese and Coelli [1995] defines the inefficiency effects u_{it} as non-negative random variables assumed to be a function of a set of firm-specific explanatory variables which may vary over time, z_{it}, and an unknown vector of coefficients, δ, associated with the z_{it}. The explanatory variables in the inefficiency model are expected to include any factors that help explain the extent to which the variable cost observations exceed the corresponding stochastic frontier cost values.

[26] The symmetric restrictions required for this are: $\beta_{mm'} = \beta_{m'm}$, $\beta_{jj'} = \beta_{j'j}$, $\beta_{kk'} = \beta_{k'k}$, $\beta_{mj} = \beta_{jm}$, $\beta_{mk} = \beta_{km}$, $\beta_{jk} = \beta_{kj}$. Following Jorgenson [1986], Carrington *et al.* [2002] and Estache *et al.* [2004], homogeneity can be imposed by normalizing the dependent variable and factor prices with the price of one of the inputs: we normalized for the price of labor.

The inefficiency effect incorporated in the composed error term ϵ_{it} of the general stochastic frontier model (equations (2) and (3)) could be specified by the equation:

$$u_{it} = \delta_0 + \delta z_{it} + w_{it} \tag{6}$$

u_{it} reflects the inability of firm i at the observation t to attain the potential minimum cost defined by the stochastic frontier.
If we assume that quality and environmental variables are variables explaining inefficiency, equation (6) can be expressed as:

$$u_{it} = \delta_0 + \sum_{\phi} \delta_{\phi} q_{\phi it} + \sum_{\psi} \delta_{\psi} e_{\psi it} + \delta_{\tau} T_{it} + w_{it} \tag{7}$$

where the ϕ subscript on δ and q_{it} indexes quality variables ($\phi = Drink, NonComp, Written, Call, SPress, NInterrupt$); the ψ subscript on δ and e_{it} indexes environmental and dummy variables ($\psi = PD, Loss, NonH, Reg, Act$); w_{it} is a random variable making the inefficiency effect stochastic.
The Battese and Coelli's model enables us to include both firm-specific and time effects in the specification of inefficiency model. Therefore, when analyzing the time-variant model, we also incorporate in equation (7) a time variable T_{it} indicating the year of the observation involved. It specifies that inefficiency may change linearly with respect to time according to the sign of the associated parameter, δ_{τ}.

The cost inefficiency is usually expressed in terms of cost inefficiency score:

$$EFF_{it} = \frac{C_{it}}{C_{it}^F} = \exp\{u_{it}\}$$

where C_{it} is the observed operating cost; C_{it}^F is the frontier or minimum cost of the i-th firm in time t. EFF_{it} takes a value between one (when $u_{it} = 0$) and infinity (when $u_{it} \to \infty$). A cost inefficiency score of 1 indicates a firm on the frontier, while non-frontier firms receive scores above 1.
When equation (7) is assumed, the overall cost inefficiency for the firm i at the t-th observation is defined by:

$$EFF_{it} = \exp\{u_{it}\} = \exp\{\delta_0 + \delta z_{it} + w_{it}\} \tag{8}$$

When quality and contextual variables are assumed to explain cost inefficiencies, the translog form of (2) can be written in the following way:

$$\ln C_{it} = \beta_0 + \sum_m \beta_m \ln p_{mit} + \sum_j \beta_j \ln y_{jit} + \frac{1}{2} \sum_m \sum_{m'} \beta_{mm'} \ln p_{mit} \ln p_{m'it}$$
$$+ \frac{1}{2} \sum_j \sum_{j'} \beta_{jj'} \ln y_{jit} \ln y_{j'it} + \sum_m \sum_j \beta_{mj} \ln p_{mit} \ln y_{jit} + v_{it} + u_{it} \tag{9}$$

Firms providing a higher quality level are expected to be more efficient than firms producing a lower quality level with the same costs. Similarly, it is assumed here that environmental factors directly influence the degree of technical efficiency. Therefore, firms with favorable exogenous conditions (better characteristics of the network and of the area) are expected to be more efficient than firms producing in worse environments.

5 The SFA Empirical Results

The computer program FRONTIER 4.1 (developed by Coelli [1996]) is used to obtain Maximum likelihood estimates of the parameters of the cost function and the inefficiency component. The program can accommodate panel data; time-varying and invariant efficiencies; half-normal and truncated normal distributions, which have been proposed in the literature.

The stochastic frontier cost function is estimated using four different SFA specifications. Tables 3, 10, 20, 26 summarize the specifications tested. The differences between the various specifications are related to the environmental and quality variables used, the assumptions imposed on the error term, cost inefficiency and the evolution of productivity over time.

We label Model A a specification of equation (5) where physical outputs are introduced as explanatory variables of operating costs. However, in many cases the ranking obtained may be an unreliable indicator of the "true performance" of the companies. It could be due to factors not under the immediate control of the managers. Therefore, we also introduced variables controlling for heterogeneity between utilities. We alternatively defined the vector of environmental variables as $e_i = [Act, PD, NonH]$ or $e_i = [Act, Reg, PD, NonH, Loss]$. In Model A_{EFF}, these environmental variables are included in the inefficiency component. In order to take quality performances into account, quality indicators are included, first as additional explanatory variables (in Model B), second as variables explaining cost inefficiency, i.e. the quality variables are included in the inefficiency component (in Model B_{EFF}).

5.1 Model A: Outputs and Environmental tests results

First, we estimate the cost frontier model with physical outputs and environmental variables. The different specifications of Model A are summarized in Table 3. The correlation coefficients between the physical outputs and between the environmental variables are given in Tables 4 and 5. The literature offers two alternative approaches to the inclusion of environmental variables. One assumes that these variables influence the shape of the technology and that these factors should be included directly into the cost functions as regressors (section 5.1.1). For the other approach, environmental factors directly influence the degree of technical efficiency and should be included in equation (4), i.e. in the error term (section 5.1.2).

Table 3: Model A - The different specifications

Models	Explanatory variables	Dummy	Environmental variables	Distribution	Time assumption
A_1	P_m, Vol, $Prop$	Act	PD, $NonH$	Half-normal	Invariant
A_2	"	Act, Reg	PD, $NonH$, $Loss$	$u_i \sim iid\ N^+(0, \sigma_u^2)$	"
	"			$v_i \sim iid\ N(0, \sigma_v^2)$	"
A'_1	P_m, Vol, $Prop$	Act	PD, $NonH$	Truncated-normal	Invariant
A'_2	"	Act, Reg	PD, $NonH$, $Loss$	$u_i \sim iid\ N^+(\mu, \sigma_u^2)$	"
	"			$v_i \sim iid\ N(0, \sigma_v^2)$	"
A_{t1}	P_m, Vol, $Prop$	Act	PD, $NonH$, T	Half-normal	Variant
A_{t2}	"	Act, Reg	PD, $NonH$, $Loss$, T	$u_i \sim iid\ N^+(0, \sigma_u^2)$	"
	"			$v_i \sim iid\ N(0, \sigma_v^2)$	"
A'_{t1}	P_m, Vol, $Prop$	Act	PD, $NonH$, T	Truncated-normal	Variant
A'_{t2}	"	Act, Reg	PD, $NonH$, $Loss$, T	$u_i \sim iid\ N^+(\mu, \sigma_u^2)$	"
	"			$v_i \sim iid\ N(0, \sigma_v^2)$	"

Table 4: Correlation between the physical outputs

	$\ln P_m$	$\ln Vol$	$\ln Prop$
$\ln P_m$	1.000		
$\ln Vol$	0.133	1.000	
$\ln Prop$	-0.150	0.162	1.000

Table 5: Correlation between the environmental variables

	$\ln PD$	$\ln Loss$	$\ln NonH$
$\ln PD$	1.000		
$\ln Loss$	0.059	1.000	
$\ln NonH$	-0.351	-0.003	1.000

5.1.1 Environmental variables as regressors to control for heterogeneity between utilities

The efficiency of a firm could be affected by exogenous conditions that are not under the direct control of managers. The inclusion of environmental variables enables to take account of the impact of the different characteristics of the network and of the area where the service is provided, thus controlling for heterogeneity among firms. Since different operating conditions may have serious financial consequences for regulated firms, it is crucial to be able to explicitly model cost differences that are due to heterogeneity and inefficiency. In order to find out whether accounting for heterogeneity in the model significantly influences the results, the cost inefficiency estimates obtained from the inclusion of different environmental variables are compared. We also analyze the robustness and reliability of obtained rankings. The estimation results of the translog cost frontier function of UK water distribution utilities obtained by the eight specifications of the output & environmental model are given in Table 6 (Appendix 1) in the case of an half-normal distribution and a time-invariant model. The other results are available from the author.

With the eight specifications, the coefficients of Vol are surprisingly insignificant. However, the coefficients of the number of properties ($Prop$) are significantly positive, that is consistent with the economic theory: more outputs leads to higher costs. Moreover, the input price coefficients are positive and highly significant with all specifications, i.e. whatever the assumptions retained and the environmental variables included in the model. Concerning the environmental variables, the coefficients on PD are negative and significant[27] at he 0.10 level. It suggests that a high population density in a area decreases the operating costs of a company. However, the results indicate no significant impact of the variable $NonH$ on operating costs, except for models A_1 and A'_1, suggesting no significant cost savings associated with supplies to larger customers. The coefficients on $Loss$ are either significantly negative in time-invariant models or insignificant in time-varying models, reflecting an

[27]Except for the model A_{t2}, where the coefficient on PD is not significant.

ambiguous impact of water losses on operating costs, as noted in section 3. The positive and significant coefficients on the dummy *Act* surprisingly reflects cost advantages for the $WOSCs$. For the English and Welsh water and sewerage industry, empirical evidence on the existence of scope economies between water and sewerage services is mixed (See Hunt and Lynk [1995], Saal and Parker [2001] and Botasso and Conti [2003]). Turning to the regulatory dummy, it should be note that *Reg* has the expected (negative) sign and is significant. This result could be explained by the change in the regulatory policy that took place in 2004. Finally, time varying intercepts are never significant with these specifications, which suggests no significant technological change over this time period.

What remains to be tested is whether the different specifications provide similar rankings of the companies with respect to the cost inefficiency scores. Tables 7 and 8 of Appendix 1 show respectively the differences and the correlation between the obtained rankings. When comparing the results on the basis of the environmental variables included as regressors in the model and the assumptions retained (for the error term distribution and the evolution of performance over time), we can note that the correlation between the obtained rankings is very high (between 0.916 and 0.992) and that there is no strong difference between the rankings. Indeed, the higher difference is for firm 13 (7^{th} with model $A_1'/14^{th}$ with model A_{t2}). Therefore, the introduction of environmental variables as regressors in the cost function enables to obtain relatively similar rankings of the companies with respect to the cost inefficiency scores. In addition, the different specifications have relatively comparable success for identifying the best and worse performing utilities (Firms 8 and 19 alternatively rank 1^{st} or 2^{nd}, whereas firm 5 is always the last in the ranking). Table 9 shows the distribution of differences in rankings provided by Model A and indicates that 86.36% of results provide differences in rankings lower than 5 ranks. There is no case in which a firm has a variation of ranking higher than 10.

Table 9: Model A - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	19	86.36
$5 < Diff \leq 10$	3	13.64
$10 < Diff \leq 15$	0	0
$Diff > 15$	0	0
Total	22	100

Environmental variables may also be assumed to directly influence the degree of cost inefficiency. Therefore, some authors consider that these factors should be included in the inefficiency term. Section 5.1.2 analyzes this case and compare the results obtained with the situation where environmental variables are included as regressors.

5.1.2 Environmental variables to explain cost inefficiency

In order to check the robustness of our results, our empirical strategy is to estimate different models, where the "environmental" variables *Act*, *Reg*, PD, *Loss* and $NonH$ alternatively enter as additional variables in the cost frontier or as variables explaining inefficiency. Moreover, in order to analyze the time pattern of inefficiency we have included a trend variable in the inefficiency component in all models. We label Model A_{EFF1} a specification of equations (7) and (9) where the activity and the policy dummies enter directly in the cost function, whereas the vector $e = [PD, Loss, NonH]$ are included in the inefficiency component. In Model A_{EFF2}, vector $e = [PD, Loss, NonH]$ is dropped from the inefficiency term and replaced by the dummies. Finally, for Model A_{EFF3}, all environmental variables and dummy variables are included in the inefficiency component.[28] Table 10 summarizes the three different models tested.

Table 10: Model A_{EFF} - The different specifications

Models	Explanatory var.	Environmental var.	Inefficiency var.
A_{EFF1}	P_m, Vol, *Prop*	*Act*, *Reg*	PD, *Loss*, $NonH$, T
A_{EFF2}	"	PD, *Loss*, $NonH$	*Act*, *Reg*, T
A_{EFF3}	"	None	*Act*, *Reg*, PD, *Loss*, $NonH$, T

The results are given in Appendix 1, Table 11. The volume of water delivered (Vol) is significantly positive only in Model A_{EFF3}, and insignificant with other models. Since cost and explanatory variables are in logarithms, the estimated first-order coefficients can be interpreted as cost elasticities evaluated at the sample median. Indeed, in a translog function, there are varying elasticities. Therefore, we have to calculate the elasticities for the "average" network. The calculation of elasticities for the input price, the volume of water delivered, the number of connected properties are calculated in Appendix 1 and Table 12 in Appendix 1 indicates a cost elasticity equal to 2.487 in Model A_{EFF3}. In addition, cost elasticities with respect to the number of connections (*Prop*) show that this variable significantly raise operating cost in Models 2 (2.547) and 3 (2.487). The input price has also a positive and

[28]Other specifications have also been tested, without significant changes in results. These other tests are available from the author.

highly significant impact on operating costs with all specifications, with cost elasticities of 0.063 in Model A_{EFF2} and 0.036 in Model A_{EFF3}. Turning to environmental variables, the activity dummy (*Act*) is positive and highly significant, when it is included directly as a regressor (Model A_{EFF1}). However, when this variable is included in the inefficiency component, its coefficient is either not significant (Model A_{EFF2}) or significantly negative (Model A_{EFF3}), reflecting an advantage in term of cost efficiency for the 10 $WASCs$. The coefficient on policy dummy (*Reg*) is significantly negative, both when it is added as a regressor (Model A_{EFF1}) and a variable explaining cost inefficiency (Model A_{EFF3}). However, its coefficient is insignificant in Model A_{EFF2}. Moreover, the results indicate that a higher population density (PD) decreases cost inefficiency when included as an explanatory variable for inefficiency. A higher proportion of large (industrial) consumers ($NonH$) and a higher percentage of losses (*Loss*) is found to decrease operating costs only in Model A_{EFF1}, reflecting the fact that a higher proportion of losses could imply deficiencies in operational or commercial practices, rather than a more critical situation that involves higher costs for the company. Finally, time varying intercepts are still never significant with these specifications, which suggests no significant technological change over this time period.

When comparing the way to include environmental variables into the inefficiency component (Tables 13 and 14 in Appendix 1), we observe that the correlation of rankings is now very low (between 0.078 and 0.284) and the models fail to identify the same best and worse performing company. 31.8% of the cases provide differences in rankings higher than 10 ranks with a maximum of 19 ranks for firms 14 and 20. Only 18.2% of the cases provide differences lower than 5 ranks. The distribution of differences in rankings in given in Table 15.

Table 15: Model A_{EFF} - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	4	18.18
$5 < Diff \leq 10$	11	50
$10 < Diff \leq 15$	5	22.73
$Diff > 15$	2	9.09
Total	22	100

Moreover, when comparing models A and A_{EFF} (Tables 16 and 17 in Appendix 1), the correlation between the rankings is slightly higher, but only one case provides differences lower than 5 ranks. Table 18 shows the distribution of differences in rankings when comparings these two models.

Table 18: Models A & A_{EFF} - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	1	4.55
$5 < Diff \leq 10$	11	50
$10 < Diff \leq 15$	6	27.27
$Diff > 15$	4	18.18
Total	22	100

Almost half the cases (45.45%) gives differences in rankings higher than 10 ranks with a maximum of 21 ranks for firm 8. Here again, the models fail to identify the same best and worse performing company. These results show the sensitivity of the stochastic frontier benchmarking methods in our sample. This is not particularly encouraging since the results cannot be considered as reliable.

However, to check whether or not the inclusion of environmental variables in the explanation of cost inefficiency is reliable, a likelihood ratio (LR) test must be conducted. Therefore, turning to the one-side error component u_{it}, we test the null hypothesis that coefficients of the environmental variables are equal to zero simultaneously, e.g. $H_0 = \gamma = \beta_{PD} = \beta_{Loss} = \beta_{NonH} = \beta_{Act} = \beta_{Reg} = \beta_T = 0$ for model A_{EFF3}. The test statistic is distributed as a chi-square (χ^2) random variable with degrees of freedom equal to the number of restrictions involved (See Kodde and Palm [1986]). This method enables to test the significance of the γ parameter. If the null hypothesis is accepted, this would indicate that $\sigma_u^2 = 0$. Therefore, the u_{it} term should be removed from the model.

The LR tests suggest that the inclusion of environmental variables in the explanation of cost inefficiency is unreliable, as the hypothesis that all parameters are jointly zero can be accepted for the three models. Indeed, LR test results reported in Table 19 leads to accept the null hypothesis at the 0.001 level. Therefore, environmental variables should be removed from the inefficiency term and included as regressors. It can explain the wide differences in rankings provide by this kind of specification, since Model A_{EFF} is misleading.

In addition, one important characteristic of water companies is that they must comply with quality standards. In section 5.2, stochastic cost frontier models are used to explore the impact of quality variables on the firm efficiency evaluation.

Table 19: Likelihood-ratio test for environmental variables as variables explaining cost inefficiency

	Null hypothesis H_0 : no inefficiency	Likelihood ratio	χ^2 statistic	Decision
A_{EFF1}	$\gamma = \beta_{PD} = \beta_{Loss} = \beta_{NonH} = \beta_T = 0$	14.739	$\chi^2_6(0.001) = 21.666$	Accept H_0
A_{EFF2}	$\gamma = \beta_{Act} = \beta_{Reg} = \beta_T = 0$	7.580	$\chi^2_5(0.001) = 19.696$	Accept H_0
A_{EFF3}	$\gamma = \beta_{PD} = \beta_{Loss} = \beta_{NonH} = \beta_{Act} = \beta_{Reg} = \beta_T = 0$	19.619	$\chi^2_8(0.001) = 25.370$	Accept H_0

5.2 Model B: The quality variables

As noted before, the quality variables can be considered either as additional outputs of the cost function (section 5.2.1) or as environmental variables that may influence the efficiency of a firm (quality variables are therefore included in the inefficiency component) (section 5.2.2).

5.2.1 Quality variables as additional outputs

Drinking water, service and technical quality are other potential outputs (alongside physical outputs) since a firm can always lower its costs by reducing the quality "offered". Incorporating quality variables into the benchmarking is crucial in order to have a good understanding of utility performance. In order to find out whether accounting for quality level in the model significantly influences the results, the ranking obtained from the inclusion of different quality variables are compared.

In Models 1 (i.e. B_1 B_1' B_{t1} and B_{t1}' depending on the assumption on the error term and the evolution of performance over time), we introduce the variable $Drink$ as an additional output, alongside the volume of water delivered and the number of connections, to proxy the drinking water quality level. We label Models 2 a specification where the quality of service is included as an output, with the vector $q = [NonComp, Written, Call]$ which takes account for the customer's complaints and the reactivity of the company to deal with these complaints. Models 3 includes proxies of "technical" quality as outputs, with the vector $q = [SPress, NInterrupt]$ to take account for the pressure level and the number of interruptions on the network. Finally, all the quality variables are included as outputs in Models 4.
Table 20 summarizes the different specifications tested[29] and Table 21 shows the correlation coefficients between the quality variables included in the model.

[29]Other specifications have also been tested but it does not significantly change the obtained rankings. These other tests are available from the author.

Table 20: Model B - The different specifications

		Time-Invariant models			
Models	Physical outputs	Quality outputs	Dummy	Environmental variables	Distrib.
B_1	P_m, Vol, $Prop$	$Drink$	Act, Reg	PD, $Loss$, $NonH$	Half-
B_2	”	$NonComp$, $Written$, $Call$	”	”	Normal
B_3	”	$SPress$, $NInterrupt$	”	”	”
B_4	”	$Drink$, $NonComp$, $Written$, $Call$, $Spress$, $NInterrupt$	”	”	”
B'_1	P_m, Vol, $Prop$	$Drink$	Act, Reg	PD, $Loss$, $NonH$	Truncated-
B'_2	”	$NonComp$, $Written$, $Call$	”	”	Normal
B'_3	”	$SPress$, $NInterrupt$	”	”	”
B'_4	”	$Drink$, $NonComp$, $Written$, $Call$, $Spress$, $NInterrupt$	”	”	”
		Time-Variant models			
B_{t1}	P_m, Vol, $Prop$	$Drink$	Act, Reg	PD, $Loss$, $NonH$, T	Half-
B_{t2}	”	$NonComp$, $Written$, $Call$	”	”	Normal
B_{t3}	”	$SPress$, $NInterrupt$	”	”	”
B_{t4}	”	$Drink$, $NonComp$, $Written$, $Written$, $Call$, $Spress$, $NInterrupt$	”	”	”
B'_{t1}	P_m, Vol, $Prop$	$Drink$	Act, Reg	PD, $Loss$, $NonH$, T	Truncated-
B'_{t2}	”	$NonComp$, $Written$, $Call$	”	”	Normal
B'_{t3}	”	$SPress$, $NInterrupt$	”	”	”
B'_{t4}	”	$Drink$, $NonComp$, $Written$, $Written$, $Call$, $Spress$, $NInterrupt$	”	”	”

Table 21: Correlation between the quality variables

	$\ln Drink$	$\ln NonComp$	$\ln Written$	$\ln Call$	$\ln SPress$	$\ln NInterrupt$
$\ln Drink$	1.000					
$\ln NonComp$	0.115	1.000				
$\ln Written$	-0.068	0.148	1.000			
$\ln Call$	0.135	0.075	0.206	1.000		
$\ln SPress$	0.158	0.214	0.028	0.004	1.000	
$\ln NInterrupt$	0.766	0.070	-0.026	0.112	0.038	1.000

The estimation results are given in Appendix 2, Table 22, in the case of a truncated-normal distribution and a time-variant model. Other results are available from the author.

The input price has a positive and highly significant impact on operating costs with all specifications. The volume of water delivered (Vol) is significantly positive only in Models 4, where all quality variables are included as outputs, and insignificant in other models. In addition, cost elasticities with respect to the number of connections (*Prop*) show that this variable significantly raise operating cost in Models 1, Models 2 and 4, with, for instance, elasticities between 0.110 and 0.940 in the case of truncated-normal distributions and time-varying specifications.
Turning to quality outputs, the drinking water quality (*Drink*) has a significant and positive impact on operating costs in Model 1 (with cost elasticities between 0.819 and 0.906 depending on the specifications retained), and is insignificant in Model 4. Therefore, in Model 1 a higher compliance with drinking water quality standards leads, in average, to higher costs for the company. The proxies for the customer's complaints ($NonComp$) and the pressure level ($SPress$) are never significant. Finally the proxy for the number of interruption on the network ($NInterrupt$) is always significantly positive: a low number of interruptions reflects higher operating costs for the company.
The activity dummy (*Act*) is positive and highly significant with all specifications, like in the output models, whereas the policy dummy (*Reg*) has the expected (negative) sign, reflecting higher performances after 2004, i.e. after the price review. Moreover, the results indicate that a higher population density (PD) decreases costs, except for Models 3 where this variable is not significant. A higher percentage of losses (*Loss*) is found to decrease significantly the operational costs in Models 4 whatever the specifications, and in Models 1, 2, 3 when a time-invariant model is assumed. Therefore, a higher proportion of losses may reflects an indicator of the technical quality of service rather than a proxy for the operational condition of the

distribution network: a higher proportion of losses does not imply more critical conditions of the network in these tests. Finally, the proportion of large consumers ($NonH$) and the time coefficients are never significant.

Tables 23 and 24 in Appendix 2 show respectively the consistency and the correlation between the ranking when quality variables are introduced as additional outputs. There are still few differences when comparing the rankings on the basis of the distribution assumption and time assumption. Indeed, with the Half-Normal vs Truncated-Normal models, the correlation of rankings is very high[30] and only firm 13 obtains really different rankings ($12th$ with model $B_{t4}/5^{th}$ with model B'_{t4}). With the Time-invariant vs Time-variant models, the correlation between the rankings is also very high[31], and only firm 13 obtains really different rankings ($7th$ with model $B_{t3}/14^{th}$ with model B'_{t3}). However, when comparing the results on the basis of the quality outputs considered in the model, although the correlation is high, between 0.847 and 0.980, some firms have really different rankings. For instances, firm 10 ranks 4^{th} in models B_4 and B'_4 and 13^{th} in models B_2 and B'_2. Firm 13 ranks 5^{th} in model B'_{t4} and 14^{th} in model B'_{t3}. Moreover, when introducing quality variables as additional outputs, the models still fail to identify the same best and worse performing company. However, with these specifications, 72.73% of the cases now allow to obtain differences in rankings lower than 5 ranks and there is no difference higher than 9 ranks as noted in Table 25.

Table 25: Model B - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	16	72.73
$5 < Diff \leq 10$	6	27.27
$10 < Diff \leq 15$	0	0
$Diff > 15$	0	0
Total	22	100

Like in the case of environmental variables, quality factors may be assumed to directly influence the degree of cost inefficiency, i.e. they may be included in the inefficiency term rather than directly in the cost function. Section 5.2.2 analyzes this case and compares the results obtained with the situation where quality variables are included as regressors.

[30] Between 0.918 (model B_{t4} vs B'_{t4}) and 0.993 (model B_{t2} vs B'_{t2}).
[31] between 0.932 (model B_{t3} vs B'_{t3}) and 0.996 (model B_{t4} vs B'_{t4}).

5.2.2 Quality variables to explain cost inefficiency

In order to check the robustness of our results, we estimate different models, where the drinking water quality ($Drink$), the quality of service ($NonComp$, $Written$, $Call$) and the technical quality ($SPress$, $NInterrupt$) alternatively enter either as additional variables in the cost frontier or as variables explaining cost inefficiency. Moreover, in order to analyze the time pattern of inefficiency we include a trend variable in the inefficiency component in all models.

The technical quality is included in the inefficiency component in Model B_{EFF1}, the quality of service is included in Model B_{EFF2} and the water quality in Model B_{EFF3}. We label Model B_{EFF4} (resp. B_{EFF5} or B_{EFF6}) a specification where the technical quality (resp. the quality of service or the water quality) is included in the inefficiency component and the other quality variables are directly used as regressors, alongside the physical outputs (Vol and $Prop$). In Model B_{EFF7}, all quality variables are assumed to explain cost inefficiency, and finally, in B_{EFF8}, environmental variables are added to the inefficiency component, alongside all quality variables. Table 26 summarizes the eight different models tested.

Table 26: Model B_{EFF} - The different specifications

Models	Explanatory variables	Quality outputs	Environmental variables	Inefficiency variables
B_{EFF1}	P_m, Vol, $Prop$	None	Act, Reg, PD $Loss$, $NonH$	$SPress$, $NInterrupt$, T
B_{EFF2}	”	None	”	$NonComp$, $Written$, $Call$, T
B_{EFF3}	”	None	”	$Drink$, T
B_{EFF4}	”	$Drink$, $NonComp$ $Written$, $Call$	”	$SPress$, $NInterrupt$, T
B_{EFF5}	”	$Drink$, $SPress$ $NInterrupt$	”	$NonComp$, $Written$, $Call$, T
B_{EFF6}	”	$NonComp$, $Written$ $Call$, $SPress$ $NInterrupt$	”	$Drink$, T
B_{EFF7}	”	None	”	$Drink$, $NonComp$, $Written$ $Call$, $SPress$ $NInterrupt$,T
B_{EFF8}	”	None	None	Act, Reg, PD $Loss$, $NonH$, $Drink$ $NonComp$, $Written$ $Call$, $SPress$ $NInterrupt$, T

The results are given in Appendix 2, Table 27, for Models B_{EFF1}, B_{EFF5} and B_{EFF8}. Indeed, LR test results reported in Table 28 lead to reject the null hypothesis at the 0.001 level only for these three tests. The results for the other specifications are available from the author.

Table 28: Likelihood-ratio test for quality variables as variables explaining cost inefficiency

Test	Null hypothesis	Likelihood ratio	χ^2 statistic	Decision
B_{EFF1}	H_0 : no inefficiency $H_0 = \gamma = \beta_{SPress} = \beta_{NInterrupt} = \beta_T = 0$	38.780	$\chi^2_5(0.001) = 19.696$	Reject H_0
B_{EFF2}	$H_0 = \gamma = \beta_{NonComp} = \beta_{Written} = \beta_{Call} = \beta_T = 0$	2.932	$\chi^2_6(0.001) = 21.666$	Accept H_0
B_{EFF3}	$H_0 = \gamma = \beta_{Drink} = \beta_T = 0$	2.450	$\chi^2_4(0.001) = 17.612$	Accept H_0
B_{EFF4}	$H_0 = \gamma = \beta_{SPress} = \beta_{NInterrupt} = \beta_T = 0$	5.511	$\chi^2_5(0.001) = 19.696$	Accept H_0
B_{EFF5}	$H_0 = \gamma = \beta_{NonComp} = \beta_{Written} = \beta_{Call} = \beta_T = 0$	40.827	$\chi^2_6(0.001) = 21.666$	Reject H_0
B_{EFF6}	$H_0 = \gamma = \beta_{Drink} = \beta_T = 0$	14.597	$\chi^2_4(0.001) = 17.612$	Accept H_0
B_{EFF7}	$H_0 = \beta_{Drink} = \beta_{NonComp} = \beta_{Written} = \beta_{Call} = \beta_{SPress} = \beta_{NInterrupt} = \beta_T = 0$	8.711	$\chi^2_9(0.001) = 27.133$	Accept H_0
B_{EFF8}	$H_0 = \beta_{Act} = \beta_{Reg} = \beta_{PD} = \beta_{Loss} = \beta_{NonH} = \beta_{Drink} = \beta_{NonComp} = \beta_{Written} = \beta_{Call} = \beta_{SPress} = \beta_{NInterrupt} = \beta_T = 0$	37.533	$\chi^2_{14}(0.001) = 35.425$	Reject H_0

As expected, the input price and the physical outputs have positive and significant impact on operating costs, except for Model B_{EFF8} where the coefficient on input price is not significant. Cost elasticities for the number of properties connected are respectively equal to 0.964, 0.786 and 0.833 for Models B_{EFF1}, B_{EFF5} and B_{EFF8}.

Concerning the environmental variables, the population density (PD) has a significantly negative impact on costs when it is assumed to control for heterogeneity between companies (Models B_{EFF1} and B_{EFF5}) but it is insignificant as variable explaining cost inefficiencies, whereas the coefficient on the percentage of distribution losses ($Loss$) is always significantly negative. Indeed, in Model B_{EFF8}, the percentage of distribution losses has a significantly negative impact on cost inefficiency: the higher the losses, the lower cost inefficiency. This result does not confirm the results of Erbetta and Cave [2007] who consider that "*a higher proportion of losses implies more critical conditions of the network, thus a higher input use is expected*". This result suggests instead that water losses may reflect that managers may find it more costly to repair leaks and to control water losses than to increase water production. The proportion of non-household customers ($NonH$) has a negative impact on costs, only when added as a regressor: it reflects cost savings associated with supplies to larger customers in this case. However, this variable seems to have no direct impact on cost inefficiency since its coefficient is not significant in Model B_{EFF8}. Moreover, the coefficient on the activity dummy (Act) is always significantly negative when added as a regressor, suggesting an advantage for the ten $WASCs$ due to the combination of water and sewerage activities. However, Model B_{EFF8} shows that the activity dummy has a positive and significant impact on cost inefficiency, reflecting here an advantage for the $WOCs$. These results reflect the current debate concerning the existence of scope economies between water and sewerage services. Finally, the policy dummy (Reg) still has a negative impact on operating costs when added as an environmental variable, but it is insignificant to explain cost inefficiencies.

Turning to quality factors, when technical quality variables are included as regressors in Model B_{EFF5}, only the proxy for the number of interruptions on the network ($NInterrupt$) is significant. However, the two variables that proxy the techniqual quality (($NInterrupt$ and $SPress$) are significant in Models B_{EFF1} and B_{EFF8}, reflecting a significant and negative impact on cost inefficiency. This result shows that a higher quality of infrastructure and network increases the efficiency of a company. The quality of service, estimated by the customer's satifaction and the reactivity of companies ($NonComp$, $Written$ and $Call$) has a significant and negative impact on cost inefficiency when the service quality is the only measure of quality level (Model B_{EFF5}). It suggests here that a higher user's satisfaction and a higher reactivity to complaints increase the company efficiency. However, when other quality factors are included in the inefficiency component (Model B_{EFF8}), the service quality becomes no significant to explain cost inefficiencies. The impact on operating costs of the drinking water quality ($Drink$) is not significant when this variable is assumed as an output (Model B_{EFF5}) but is highly significant to explain cost inefficiency (Model B_{EFF8}): a better drinking water quality reflects a higher cost efficiency of a firm. Finally, in Models B_{EFF1} and B_{EFF5} estimate results now show the existence of a decreasing path of cost inefficiency over the sample period, as the coefficient related to trend is significantly negative, but this trend is not significant in Model B_{EFF8}.

Moreover, SFA recognizes that some of the distance from the best practice frontier is attributable to random shocks or statistical noise in the data. The Maximum likelihood estimated parameters γ are 0.999 in Models B_{EFF1} and B_{EFF5} and 0.643 in Model B_{EFF8}. These results imply that deviations from the best practice frontier are mostly entirely due to inefficiency in the two first models. Inefficiency is also the

main resource of deviations in Model B_{EFF8}, but here, random shocks still explain more than 35% of deviations.
Since the null hypothesis is rejected by the LR test in these three models, our results show that these specifications provide an adequate explanation of the inefficiency sources in the UK water and sewerage sector.
Therefore, a first adequate specification consider that the technical quality is a good explanation of cost inefficiency, when other quality variables are ignored (Model B_{EFF1}). However, Model B_{EFF5} indicates that an adequate specification is to include the number of interruptions on the network as a proxy for the technical quality directly as an additional output and to introduce service quality variables to explain cost inefficiencies. Finally, with Model B_{EFF8}, the inclusion of the two technical quality variables alongside the drinking water quality in the inefficiency component provides an adequate specification for UK water utilities cost efficiency.

Hence, as regards these results, it seems difficult to decide what a proper specification of UK water utilities cost inefficiencies, since the results depend on the variables included directly in the cost function and in the inefficiency term. However, these three tests show that omitting these quality indicators would wrongly reduce the efficiency scores attributed to good quality services. Moreover, these results differ from those found with Model B, when quality variables are assumed as additional outputs. It underlines the interest of comparing the two specifications and it shows how some factors can directly impact the technology structure and other influence the performance of services (see Kumbhakar and Lovell [2000]).

Turning to the rankings obtained with these models, Tables 29 and 30 in Appendix 2 show the consistency and the correlation between the rankings when quality variables are assumed to explain cost inefficiencies. The correlation coefficients are relatively low and equal to 0.434 (Model B_{EFF5} vs B_{EFF8}) 0.579 (Model B_{EFF1} vs B_{EFF8}) and 0.803 (Model B_{EFF1} vs B_{EFF5}) and differences exist between the rankings (up to 13 ranks for firm 11). Moreover, the models still fail to identify the same best performing company, which is different with the three models, and the worse performer is the same only with models B_{EFF1} and B_{EFF5}. As noted in Table 31, in this case 22.7% of the cases provide differences in rankings higher than 10 ranks and 36.36% give differences lower than 5 ranks.
What remains to be tested is whether the different specifications of quality variables (as regressors or as variables in the inefficiency component) provide similar rankings

Table 31: Model B_{EFF} - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	8	36.36
$5 < Diff \leq 10$	11	50
$10 < Diff \leq 15$	3	13.64
$Diff > 15$	0	0
Total	22	100

of the companies with respect to the cost inefficiency scores. Table 32 in Appendix 2 indicates the consistency between the different rankings. Table 33 in Appendix 2 indicates that the correlation coefficients between the rankings obtained in Models B and B_{EFF} are very low. When comparing the efficiency scores with models B'_{t4} and B_{EFF8}, where all quality variables are included either as regressors or in the inefficiency component, the correlation is equal to 0.447 and the rankings vary between 0 and 12. For instance firm 8 ranks 1^{st} with Model B'_{t4} and 13^{th} with Model B_{EFF8}; firm 18 ranks alternatively 18^{th} and 6^{th}. When the technical quality variables are included either as regressors (Model B'_{t3}) or in the inefficiency term (Model B_{EFF1}), the correlation between the rankings is 0.535 with differences varying between 1 and 17.[32] Finally, in Model B'_{t2} service quality variables were included as regressors, while in Model B_{EFF5} these variables are assumed to explain cost inefficiency and other quality variables are included as regressors. When comparing these two specifications, the correlation between the rankings is still very low (0.399) and the rankings vary between 0 and 17.[33] More generally, only 18.2% of the cases provide differences lower than 5 ranks and 54.5% give differences higher than 10 ranks as noted in Table 34. The differences in rankings are between 3 and 18 such that the different specifications fail to determine a same best and worse performing company.

Table 34: Models B & B_{EFF} - Distribution of differences in rankings

Differences in rankings	Number of cases	%
$Diff \leq 5$	4	18.18
$5 < Diff \leq 10$	8	36.36
$10 < Diff \leq 15$	9	40.91
$Diff > 15$	1	4.55
Total	22	100

[32]Firm 18 ranks 21^{st} with Model B'_{t3} and 4^{th} with Model B_{EFF1}.
[33]Firm 18 ranks 20^{th} with Model B'_{t2} and 3^{d} with Model B_{EFF5}.

6 Conclusion

Benchmarking has become a very important tool especially for comparing the relative performance of different companies within a sector, informing citizens and providing information to regulatory bodies, helping them improve incentives. In this study, SFA model is used to explore the impact of the inclusion of different variables and different assumptions on the firms efficiency evaluation. In addition to traditional measures of efficiency which include physical outputs, environmental variables are used to control for heterogeneity in environmental and network characteristics across companies. Moreover, since quality is a performance indicator that warrants attention when evaluating utilities performances, the study takes different aspects of quality into account. This paper presents different specifications of SFA model to illustrate how these additional indicators can be incorporated into benchmarking studies and how performance rankings might be affected.

The results indicate that, in the case of UK water distribution utilities, environmental factors influence the shape of the technology and hence these factors should be included directly into the cost function as regressors, rather than in the inefficiency component. Moreover, the results show that some aspects of quality directly influence the degree of technical efficiency and hence should be included in the inefficiency term, rather than as additional outputs.

In addition, this study examines how the introduction of different variables affects performance comparisons across utilities. From the methodological point of view, different distribution and time assumptions are tested in the study in order to reduce the impact of choosing a specific distribution function and time consideration arbitrarily. The empirical results show that these different assumptions do not have a significant impact on the rankings resulting from the benchmarking. However, the inefficiency scores obtained from specifications of the cost function are not found to be robust. The levels of inefficiency estimates as well as the rankings depend on which variables are introduced into the model and on the way these variables are included. The smallest differences in rankings are provided by Model A, when physical outputs and environmental variables are directly included as regressors and no quality indicator is taken into account. In this case, 86.36% of the results enable to obtain differences lower than 5 ranks. When introducing quality variables as additional outputs (Model B), the models fail to determine a same best and worse performing company. However, the correlation between rankings is still very high and 72.73% of the results enable to limit the differences in rankings (differences lower than 5 ranks). Like with Model A, there is no case where a firm has a difference of ranks higher than 10. Moreover, the results demonstrate that some quality variables should be included in the inefficiency component for the model specification to be reliable. However, in this case, the correlation between rankings becomes very low and in 13.64% of the cases, differences in rankings are higher than 10.

What can we conclude from these results and what are the implications for regulators who decide to implement benchmarking methods and yardstick competition? Rankings help inform the public, directing attention toward poorly performing utilities and providing information to policy-makers and regulators regarding deviations from best practice. The reliability of efficiency scores is crucial for an effective implementation of incentive-based price regulation. However, the evidence from empirical studies shows that the various benchmarking methods (parametric vs non-parametric approaches) often produce different results with respect to firms' efficiency scores and rankings. This study demonstrates that this difficulty also exists when using different specifications within one method: although the ranking correlation is often high within the models, the rankings can change dramatically for specific utilities. These results show the sensitivity of the stochastic frontier benchmarking methods in our sample, even in an "ideal" case where a lot of data are available. Therefore, the lack of consistency, and indirectly the enforcement difficulties, don't seem to be due to a lack of information for the regulator. On the contrary, the results indicate that, when adding more variables into the SFA specification (more precisely, when quality variables are added), the consistency in rankings is reduced. This issue may explain the fact that companies regulated by yardstick mechanisms often try to renege on the regulatory decisions based on the results of benchmarking. In this context, firms-led renegotiation would be partly justified by the incompleteness of the method used. As a consequence, if rankings or scores are to be used in regulatory proceedings, great care must be taken to avoid unduly penalizing utilities, since the direct use of inefficiency estimates in the regulation of water distribution utilities may be misleading. The lack of consistency of the results suggests that a mechanical use of SFA inefficiency scores in a price-setting process is not necessarily recommended.

In a policy and regulatory point of view, different recommendations may be proposed. First, one can conclude that benchmarking results should only be used as a starting point for providing information about the range in which the inefficiency

scores can be located, but should not result in the implementation of penalties for relative bad performing utilities. In the same way, the consistency problem could lead to privilege a "softer" form of yardstick competition, i.e. the sunshine regulation which does not lead to implement mechanisms for rewarding or penalizing poorly performing firms, other than "naming and shaming" the latter.
Indeed, a "strict" form of yardstick competition denotes the use of benchmarking results to determine maximum prices or revenues. This type of application could lead (and often leads) to tensions and disputes between the regulator and regulated companies. These enforcement difficulties are due to the incompleteness of the benchmarking method used and its challenge by operators. However, it does not mean that companies refuse any application of benchmarking, since the implementation of softer forms of yardstick competition, such as sunshine regulation, is less subject to discussion by the regulated firms. For instance, De Witte and Saal [2008] studied the application of yardstick competition in the Dutch water sector. They explain that the regulator initially applied a voluntary sunshine regulatory model in which firms committed themselves to publicize their performance. Then, the Dutch government proposed a new law to implement a stricter form of yardstick competition to the sector. However, as noted by De Witte and Saal [2008], "*it seems that the uncertainty relatively to the regulatory model undermined the willingness to participate in the voluntary benchmark. Whereas in 1997 and 2000, respectively, 78 and 71 percent of the companies participated, in 2003 this decreased to only two thirds of the utilities. Although all companies are officially in favour of benchmarking, in their annual accounts some companies commented on the imprecise methodology (e.g. measuring costs per* m^3 *or per connection could deliver significant different results).*"

However, this study shows that the different approaches are not consistent in identifying "best" or "worse" firms (except with Model A where no quality indicators are included). Therefore, even a softer form of yardstick competition may be difficult to implement. Indeed, as noted by Estache *et al.* [2004], a "mild" form of benchmarking regulation can be relied on when different models are consistent in identifying best and/or worse performing companies, even if the ranking consistency test fails.
As regard these two difficulties (lack of consistency in efficiency rankings and in identifying best or worse firms), we can wonder whether yardstick mechanism could be efficient in terms of incentives provided to regulated companies and information collected by regulators since even "mild" forms of comparative regulation could be unreliable. In this context, Le Lannier [2009] shows that taking account of enforce-

ment difficulties of yardstick competition does not prevent the implementation of this incentive scheme, but requires an adaptation of the contractual design. Indeed, consistency problems (i.e. enforcement difficulties) should lead the regulatory bodies to choose more "flexible" yardstick contracts to take account of the possibility of firms-led renegotiation. With this type of contractual design the regulators will have the opportunity to adapt the regulatory contract to future contingences or forecasting errors. Therefore, a flexible yardstick contract would enable to partly overcome the incompletness of benchmarking methods.

References

AIGNER D., LOVELL C.A.K., SCHMIDT P. [1977], Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1): 21-37.

ALEGRE H., BAPTISTA J.M., CABRERA E., CUBILLO F., DUARTE P., HIRNER W., MERKEL W., PARENA R. [2006], Performance Indicators for Water Supply Services, second ed. International Water Association Publishing, London.

ALLAN G. [2006], Assessing capital efficiency in the water and sewerage industry in England and Wales: Ofwat's approach. *Utilities Policy*, 14(4): 224-233.

ANTONIOLI B., FILIPPINI M. [2001], The Use of Variable Cost Function in the Regulation of the Italian Water Industry. *Utilities Policy*, 10(3-4): 181-187.

ASHTON J.K. [2000], Cost Efficiency in the UK's Water and Sewerage Industry. *Applied Economics Letters*, 7: 455-458.

AUBERT C., REYNAUD A. [2005], The impact of regulation on cost efficiency: an empirical analysis of Wisconsin water utilities. *Journal of Productivity Analysis*, 23: 383409.

BATTESE G.E., COELLI T.J. [1988], Prediction of Firm-Level Technical Efficiencies With a Generalised Frontier Production Function and Panel Data. *Journal of Econometrics*, 38: 387-399.

BATTESE G.E., COELLI T.J. [1992], Frontier production functions, technical efficiency, and panel data: with application to paddy farmers in India. In Gulledge T.R., Lovell C.A.K. (Eds.), International Applications of Productivity and Efficiency Analysis. Kluwer Academic Publishers, Boston.

BATTESE G.E., COELLI T.J. [1995], A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20(2): 325-332.

BAUER P., BERGER A., FERRIER G., HUMPHREY D. [1998], Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Efficiency Methods. *Journal of Economics and Business*, 50: 85-114.

BERG S.V., LIN C. [2007], Consistency in Performance Rankings: The Peru Water Sector. *Journal of Applied Economics* 40(6): 93-805.

BHATTACHARYYA A., HARRIS T.R., NARAYANAN R., RAFFIEE K. [1995], Specification and Estimation of the Effect of Ownership on the Economic Efficiency of the Water Utilities. *Regional Science and Urban Economics*, 25(6): 759-784.

BOS D. [1991], Privatization: A Theoretical Treatment. Clarendon Press, Oxford.

BOTASSO A., CONTI M. [2003], Cost Inefficiency in the English and Welsh Water industry: An Heteroskedastic Stochastic Cost Frontier Approach. Economics Discussion papers 573, University of Essex, Department of Economics.

BOUSCASSE H., DESTANDAU F., GARCIA S. [2008], Analyse économétrique des coûts des services d'eau potable et qualité des prestations offertes aux usagers. *Revue d'économie industrielle*, 122: 7-26.

BURNS P., JENKINS C., MIKKERS M., RIECHMANN C. [2006a], The role of the policy framework for the effectiveness of benchmarking in regulatory proceedings. In Performance Measurement And Regulation Of Network Utilities, Edited by Tim Coelli, Centre for Efficiency and Productivity Analysis School of Economics, University of Queensland, Australia and Denis Lawrence, Meyrick and Associates, Australia.

CARRINGTON R., COELLI T.J., GROOM E. [2002], International Benchmarking of Monopoly Price Regulation: The Case of Australian Gas Distribution. *Journal of Regulatory Economics*, 21(2): 191-216.

Cave review of competition and innovation in water markets: a call for evidence [2008]. www.defra.gov.uk/environment/water/industry/cavereview.

CHAMBERS R.G. [1988], Applied Production Economics: A Dual Approach. Cambridge University Press, UK.

CHARNES A., COOPER W.W., RHODES E. [1978], Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2: 429-444.

CHARNES A., COOPER W.W., LEWIN A.Y., SEIFORD L.M. [1994], Data Envelopment Analysis: Theory, Methodology and Application. Boston: Kluwer Academic Publishers.

COELLI T.J. [1996], A guide to frontier version 4.1 : A computer program for frontier production function estimation. Centre for Efficiency and Productivity Analysis Working Paper 96/07, Department of Econometrics, University of New England, Armidale.

COELLI T.J., RAO D.S.P., BATTESE G. [1998], An Introduction to Efficiency and Productivity Analysis. Kluwer, Dordrecht.

COELLI T.J., PERELMAN S., ROMANO E. [1999], Accounting for environmental influences in stochastic frontier models: with application to international airlines. *Journal of Productivity Analysis*, 11: 251-273.

COELLI T.J., ESTACHE A., PERELMAN S., TRUJILLO L. [2003], A Primer on Efficiency Measurement for Utilities and Transport Regulators. World Bank, Washington, DC.

COOPER W.W., SEIFORD L.M., ZHU J. [2004], Handbook on Data Envelopment Analysis. Boston: Kluwer Academic Publishers.

CORTON M.L. [2003], Benchmarking in the Latin American water sector: the case of Peru. *Utilities Policy*, 11: 133-142.

CORTON M.L., BERG S.V. [2009], Benchmarking Central American water utilities. *Utilities Policy*, 17(3-4): 267-275.

CPB Netherlands Bureau for Economic Policy Analysis [2000], Yardstick competition: Theory, design, and practice. Working Paper n°133.

CUBBIN J., TZANIDAKIS G. [1998], Regression versus Data Envelopment Analysis for Efficiency Measurement: An Application to the England and Wales Regulated Water Industry. *Utilities Policy*, 7(2), 75-85.

CUBBIN J. [2005], Efficiency in the water industry. *Utilities Policy*, 13(4): 289-293.

DASSLER T., PARKER D., SAAL D.S. [2006], Methods and Trends of Performance Benchmarking in UK Utility Regulation. *Utilities Policy*, 14(3): 166-174.

ERBETTA F., CAVE M. [2007], Regulation and Efficiency Incentives: Evidence from the England and Wales Water and Sewerage Industry. *Review of Network Economics*, 6(4): 425-452.

ESTACHE A., KOUASSI E. [2002], Sector Organization, Governance and the Inefficiency of African Water Utilities. World Bank Policy Research Working Paper 2890.

ESTACHE A., ROSSI M. [2002], How Different is the Efficiency of Public and Private Water Companies in Asia. *World Bank Economic Review*, 16(1): 139-148.

ESTACHE A., ROSSI M., RUZZIER C.A. [2004], The Case for International Coordination of Electricity Regulation: Evidence from the Measurement of Efficiency in South America. *Journal of Regulatory Economics*, 25(3): 271-95.

FARSI M., FILIPPINI M., GREENE W. [2005], Efficiency Measurement in Network Industries: Application to the Swiss Railway Companies. *Journal of Regulatory Economics*, 28(1): 69-90.

FARSI M., FILIPPINI M. [2006], An analysis of efficiency and productivity in Swiss hospitals. *Swiss Journal of Economics and Statistics*, 142(1): 1-37.

FILIPPINI M., HROVATIN N., ZORIC J. [2008], Cost efficiency and economies of scale of slovenian water distribution utilities: an application of panel data stochastic frontier methods. *Journal of Productivity Analysis*, 29(2): 169-182.

GARCIA S., THOMAS A. [2001], The Structure of Municipal Water Supply Costs: Application to a Panel of French Local Communities. *Journal of Productivity Analysis*, 16(1): 5-29.

GARCIA S., MOREAUX M., REYNAUD A. [2007], Measuring Economies of Vertical Integration in Network Industries: An Application to the Water Sector. *International Journal of Industrial Organization*, 25(4): 791-820.
GREENE W.H. [2005a], Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126: 269-303.
GREENE W.H. [2005b], Fixed and Random Effects in Stochastic Frontier Models. *Journal of Productivity Analysis*, 23(1): 7-32.
HESSELING D., SARI M. [2006], The introduction of quality regulation of electricity distribution in The Netherland. NMa document, published in the European Energy Law Report III.
HUNT L.C., LYNK E.L. [1995], Privatization and efficiency in the UK water industry: an empirical analysis. *Oxford Bulletin of Economics and Statistics*, 57: 371-388.
JAMASB T., POLLITT M. [2003], International benchmarking and regulation: an application to European electricity distribution utilities. *Energy Policy*, 31(15): 1609-1622.
JORGENSON D.W. [1986], Econometric Methods for Modeling Producer Behavior. In Z. Griliches and M.D. Intriligator eds, Handbook of Econometrics, vol 3, Amsterdam, Elsevier, 1841-1915.
KIRKPATRICK C., PARKER D., ZHANG Y.F. [2006], State versus Private Sector Provision of Water Services in Africa: An Empirical Analysis. *The World Bank Economic Review*, 20(1): 143-163.
KODDE D., PALM F. [1986], Wald criteria for jointly testing equality and inequality restrictions. *Econometrica*, 54(5): 1243-1248.
KUMBHAKAR S.C. and LOVELL C.A.K. [2000], Stochastic Frontier Analysis. Cambridge University Press.
LIN C. [2007], Service Quality and Prospects for Benchmarking: Evidence from the Peru Water Sector. *Utilities Policy*, 13(3): 230-239.
MEEUSEN W., VAN DEN BROECK J. [1977], Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2): 435-444.
PICAZO-TADEO A.J., SAEZ-FERNANDEZ F.J., GONZALEZ-GOMEZ F. [2008], Does Service Quality Matter in Measuring the Performance of Water Utilities? *Utilities Policy*, 16(1): 30-38.
PITT M., LEE L.F. [1981], The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics*, 9: 43-64.
SAAL D.S., PARKER D. [2000], The Impact of Privatization and Regulation on

the Water and Sewerage Industry in England and Wales: A Translog Cost Function Model. *Managerial and Decision Economics*, 21(6): 253-268.

SAAL D.S., PARKER D. [2001], Productivity and Price Performance in the Privatized Water and Sewerage Companies of England and Wales. *Journal of Regulatory Economics*, 20(1): 6190.

SAAL D.S., PARKER D. [2005], Assessing the Performance of Water Operations in the English and Welsh Water industry: A Panel Input Distance Function Approach. Aston Business School Working Paper RP0502.

SAAL D.S., PARKER D. [2006], Assessing the Performance of Water Operations in the English and Welsh Water Industry: A Lesson in the Implications of Inappropriately Assuming a Common Frontier. In Coelli T., Lawrence D. (editors): Performance Measurement and Regulation of Network Utilities. Edward Elgar Publishing, Cheltenham.

SAAL D.S., 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. *Journal of Productivity Analysis*, 28(1): 127-139.

SAAL D.S., REID D. [2004], Estimating Opex Productivity Growth in English and Welsh Water and Sewerage Companies: 1993-2003. Aston Business School Working Paper RP0434.

SAGE E. [1999], La concurrence par comparaison : théorie et appplications. Une proposition pour le secteur de l'eau en France. Thèse de doctorat, Université Paris-Dauphine.

SAWKINS J.W. [1995], Yardstick competition in the English and Welsh water industry. *Utilities Policy*, 5(1): 27-36.

SCHMIDT P., SICKLES R.C. [1984], Production Frontiers and Panel Data. *Journal of Business and Economic Statistics*, 2(4): 367-374.

SEROA DA MOTTA R., MOREIRA A.R.B. [2006], Efficiency and regulation in the Sanitation Sector in Brazil. *Utilities Policy*, 14(3): 185-195.

SHLEIFER A. [1985], A Theory of Yardstick competition. *RAND Journal of Economics*, 16(3): 319-327.

STERN J. [2005], UK Regulatory Price Reviews and the Role of Efficiency Estimates. *Utilities Policy* , 13(4): 273-278.

THANASSOULIS E. [2000a], DEA and its use in the Regulation of Water Companies. *European Journal of Operational Research*, 127(1): 1-13.

THANASSOULIS E. [2000b], The Use of Data Envelopment Analysis in the Regu-

lation of UK Water Utilities: Water Distribution. *European Journal of Operational Research*, 126(2): 436-453.
TUPPER H.C., RESENDE M. [2004], Efficiency and Regulatory Issues in the Brazilian Water and Sewage Sector: An Empirical Study. *Utilities Policy*, 12(1): 29-40.
VON HIRSCHHAUSEN C., CULLMANN A., WALTER M., WAND R., ZSCHILLE M. [2009], Quo vadis efficiency analysis of water distribution? A comparative literature review. *Utilities Policy*, 17(3-4): 225-232.

Appendix

Appendix 1 - Physical Outputs & Environmental Variables

Table 6: Model A - Estimation results of the cost frontier:
Half-normal distribution and time-invariant model

Variables	B_1	B_2
Constant	16.531**	14.690**
	(7.386)	*(6.934)*
$\ln P_m$	1.588**	2.389***
	(0.962)	*(0.905)*
$\ln Vol$	-0.339	-0.201
	(0.615)	*(0.636)*
$\ln Prop$	1.846*	1.789*
	(1.168)	*(1.128)*
$(\ln P_m)^2$	-0.104***	-0.125***
	(0.025)	*(0.024)*
$(\ln Vol)^2$	0.088**	0.107**
	(0.045)	*(0.047)*
$(\ln Prop)^2$	-0.101**	-0.107**
	(0.055)	*(0.057)*
$\ln P_m \ln Vol$	0.220***	0.164***
	(0.045)	*(0.045)*
$\ln P_m \ln Prop$	-0.164**	-0.200***
	(0.075)	*(0.070)*
$\ln Vol \ln Prop$	-0.029	-0.051
	(0.082)	*(0.088)*
Act	0.976***	1.089***
	(0.175)	*(0.179)*
$\ln PD$	-0.360*	-0.344*
	(0.249)	*(0.278)*
$\ln NonH$	-0.210**	-0.117
	(0.120)	*(0.112)*
$\ln Loss$	-	-0.121*
		(0.089)
Reg	-	-0.101***
		(0.026)
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.068***	0.069***
	(0.020)	*(0.022)*
$\gamma = \sigma_u^2/\sigma^2$	0.708***	0.751***
	(0.091)	*(0.083)*
Log Likelihood	51.400	59.175

Notes: standard errors in brackets;
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

The elasticities for the input price, the volume of water delivered, the number of connected properties are calculated as follows:

$$\epsilon_i = \frac{\partial \ln Opex}{\partial \ln Y_i} = \beta_i + 2\beta_{ii} \ln \widetilde{x}_i + \sum_{j \neq i} \beta_{ij} \ln \widetilde{X}_j$$

Table 7: Model A - Consistency in efficiency ranking

ID	Firms	Ranking		
		Best	Worst	Difference
1	Anglian Water services Ltd	8	14	6
2	Northumbrian Water Ltd	8	12	4
3	Severn Trent Water Ltd	15	15	0
4	South West Water Ltd	10	13	3
5	Southern Water Services Ltd	22	22	0
6	Thames Water Utilities Ltd	6	7	1
7	United Utilities Water plc	20	21	1
8	Dwr Cymru Cyfyngedig	1	2	1
9	Wessex Water Services Ltd	12	14	2
10	Yorkshire Water Services Ltd	4	9	5
11	Bournemouth & West Hampshire Water plc	3	9	6
12	Bristol Water plc	16	17	1
13	Cambridge Water Company plc	7	14	7
14	Dee Valley Water plc	19	20	1
15	Veolia Water South East Ltd	4	5	1
16	Mid Kent Water plc	18	18	0
17	Portsmouth Water plc	16	17	1
18	South East Water Ltd	19	21	2
19	South Staffordshire Water plc	1	2	1
20	Sutton & East Surrey Water plc	3	8	5
21	Veolia Water East Ltd	9	11	2
22	Veolia Water Central Ltd	6	11	5
			Min Diff.	0
			Max Diff.	7

Table 8: Model A - Correlation between the rankings

	B_1	B_2	B'_1	B'_2	B_{t1}	B_{t2}	B'_{t1}	B'_{t2}
B_1	1.000							
B_2	0.953	1.000						
B'_1	0.991	0.927	1.0000					
B'_2	0.951	0.985	0.933	1.000				
B_{t1}	0.960	0.969	0.932	0.948	1.000			
B_{t2}	0.945	0.988	0.916	0.970	0.975	1.000		
B'_{t1}	0.953	0.965	0.930	0.950	0.992	0.976	1.000	
B'_{t2}	0.952	0.984	0.926	0.984	0.972	0.992	0.980	1.000

Table 11: Model A_{EFF} - Estimation results of the cost frontier

Model A_{EFF1}

Variables	Estimated parameter	Standard dev.
Constant	2.079	31.576
$\ln P_m$	3.143***	1.021
$\ln Vol$	0.226	0.355
$\ln Prop$	-0.398	0.843
$(\ln P_m)^2$	-0.086***	0.025
$(\ln Vol)^2$	0.129***	0.037
$(\ln Prop)^2$	0.070**	0.036
$\ln P_m \ln Vol$	0.159***	0.053
$\ln P_m \ln Prop$	-0.262***	0.076
$\ln Vol \ln Prop$	-0.106**	0.057
Act	1.005***	0.159
Reg	-0.084*	0.062
Constant	2.328	30.996
$\ln PD$	-0.464***	0.150
$\ln Loss$	-0.156**	0.089
$\ln NonH$	-0.216**	0.104
T	0.002	0.018
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.033***	0.004
$\gamma = \sigma_u^2/\sigma^2$	0.499	9.716
Log Likelihood	37.443	

Model A_{EFF2}

Variables	Estimated parameter	Standard dev.
Constant	25.724***	0.990
$\ln P_m$	-1.783**	0.986
$\ln Vol$	-0.674	0.906
$\ln Prop$	-2.489***	0.537
$(\ln P_m)^2$	-0.058	0.115
$(\ln Vol)^2$	-0.015	0.141
$(\ln Prop)^2$	0.061*	0.044
$\ln P_m \ln Vol$	0.283*	0.196
$\ln P_m \ln Prop$	0.087	0.078
$\ln Vol \ln Prop$	0.112*	0.084
$\ln PD$	-0.296	0.889
$\ln Loss$	0.077	0.951
$\ln NonH$	-0.470	0.987
Constant	-0.044	0.395
Act	0.120	0.420
Reg	0.030	0.686
T	0.009	0.123
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.043***	0.004
$\gamma = \sigma_u^2/\sigma^2$	0.000***	0.000
Log Likelihood	23.900	

$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 11: Continuation

Model A_{EFF3}		
Variables	Estimated parameter	Standard dev.
Constant	23.037***	0.988
$\ln P_m$	-1.655**	0.938
$\ln Vol$	-0.953*	0.665
$\ln Prop$	-2.001***	0.339
$(\ln P_m)^2$	-0.068***	0.026
$(\ln Vol)^2$	-0.053***	0.019
$(\ln Prop)^2$	0.023	0.024
$\ln P_m \ln Vol$	0.264***	0.058
$\ln P_m \ln Prop$	0.089*	0.068
$\ln Vol \ln Prop$	0.177***	0.055
Constant	0.495**	0.253
Act	-0.151***	0.037
Reg	-0.200**	0.101
$\ln PD$	-0.416***	0.110
$\ln Loss$	-0.006	0.033
$\ln NonH$	0.069	0.84
T	-0.009	0.023
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.043***	0.006
$\gamma = \sigma_u^2/\sigma^2$	0.001***	0.000
Log Likelihood	19.917	

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

with $Y_i \in \{P_m, Vol, Prop\}$

For instance, the elasticity of the input price is as follows:

$$\epsilon_{P_m} = \frac{\partial \ln Opex}{\partial \ln P_m} = \beta_{P_m} + 2\beta_{P_m,P_m} \ln \widetilde{P_m} + \beta_{P_m,Vol} \ln \widetilde{Vol} + \beta_{P_m,Prop} \ln \widetilde{Prop}$$

Table 12: Cost elasticities - Model A_{EFF}

Explanatory variables	ϵ_i	
	A_{EFF2}	A_{EFF3}
P_m	0.063	0.036
$Prop$	1.241	0.555
Vol	2.547	2.487

Table 13: Model A_{EFF} - Consistency in efficiency ranking

ID	Firms	Ranking			Difference
		A_{EFF1}	A_{EFF2}	A_{EFF3}	
1	Anglian Water services Ltd	15	16	8	8
2	Northumbrian Water Ltd	6	20	6	14
3	Severn Trent Water Ltd	13	18	16	5
4	South West Water Ltd	18	17	17	1
5	Southern Water Services Ltd	19	22	7	15
6	Thames Water Utilities Ltd	1	14	18	17
7	United Utilities Water plc	12	21	19	9
8	Dwr Cymru Cyfyngedig	16	15	22	7
9	Wessex Water Services Ltd	17	19	9	10
10	Yorkshire Water Services Ltd	4	13	14	10
11	Bournemouth & West Hampshire Water plc	3	6	1	5
12	Bristol Water plc	11	8	10	3
13	Cambridge Water Company plc	14	7	5	9
14	Dee Valley Water plc	21	11	2	19
15	Veolia Water South East Ltd	9	4	3	6
16	Mid Kent Water plc	20	5	13	15
17	Portsmouth Water plc	8	12	4	8
18	South East Water Ltd	22	10	21	12
19	South Staffordshire Water plc	2	3	15	13
20	Sutton & East Surrey Water plc	10	1	20	19
21	Veolia Water East Ltd	5	9	11	6
22	Veolia Water Central Ltd	7	2	1	10
				Min Diff.	1
				Max Diff.	19

Table 14: Model A_{EFF} - Correlation between the rankings

	A_{EFF1}	A_{EFF2}	A_{EFF3}
A_{EFF1}	1.000		
A_{EFF2}	0.284	1.000	
A_{EFF3}	0.085	0.078	1.0000

Table 16: Models A & A_{EFF} - Consistency in efficiency ranking

ID	Firms	Ranking		
		Best	Worse	Diff.
1	Anglian Water services Ltd	8	16	8
2	Northumbrian Water Ltd	6	20	14
3	Severn Trent Water Ltd	13	18	5
4	South West Water Ltd	10	18	8
5	Southern Water Services Ltd	7	22	15
6	Thames Water Utilities Ltd	1	18	17
7	United Utilities Water plc	12	21	9
8	Dwr Cymru Cyfyngedig	1	22	21
9	Wessex Water Services Ltd	9	19	10
10	Yorkshire Water Services Ltd	4	14	10
11	Bournemouth & West Hampshire Water plc	1	9	8
12	Bristol Water plc	8	17	9
13	Cambridge Water Company plc	5	14	9
14	Dee Valley Water plc	2	21	19
15	Veolia Water South East Ltd	3	9	6
16	Mid Kent Water plc	5	20	15
17	Portsmouth Water plc	4	17	13
18	South East Water Ltd	10	22	12
19	South Staffordshire Water plc	1	15	14
20	Sutton & East Surrey Water plc	1	20	19
21	Veolia Water East Ltd	5	11	6
22	Veolia Water Central Ltd	2	12	10
			Min Diff.	5
			Max Diff.	21

Table 17: Models A & A_{EFF} - Correlation between the rankings

	A_{EFF1}	A_{EFF2}	A_{EFF3}	A'_{t1}	A'_{t2}
A_{EFF1}	1.000				
A_{EFF2}	0.284	1.000			
A_{EFF3}	0.085	0.078	1.0000		
A'_{t1}	0.6330	0.411	-0.113	1.000	
A'_{t2}	0.633	0.436	-0.071	0.992	1.000

Appendix 2 - Quality Variables

Table 22: Model B - Estimation results of the cost frontier: Truncated-normal distribution and time-variant model

Variables	B'_{t1}	B'_{t2}	B'_{t3}	B'_{t4}
Constant	14.980**	13.953*	31.219***	-107.304***
	(7.936)	*(10.193)*	*(12.826)*	*(6.985)*
$\ln P_m$	2.175***	2.315***	2.570***	1.327**
	(0.881)	*(0.877)*	*(0.888)*	*(0.748)*
$\ln Vol$	-0.290	-0.271	-0.558	0.990**
	(0.552)	*(0.592)*	*(0.615)*	*(0.542)*
$\ln Prop$	-2.046**	-1.961**	-1.209	-2.809***
	(0.937)	*(0.929)*	*(0.984)*	*(0.821)*
$\ln Drink$	-8.265**	-	-	-1.239
	(4.175)			*(2.997)*
$\ln NonComp$	-	3.420	-	1.703
		(4.188)		*(2.937)*
$\ln Written$	-	0.005	-	4.761**
		(3.350)		*(2.185)*
$\ln Call$	-	3.332*	-	5.220**
		(2.441)		*(2.893)*
$\ln SPress$	-	-	-0.939	-8.180
			(7.985)	*(6.031)*
$\ln NInterrupt$	-	-	1.505**	1.176***
			(5.226)	*(4.084)*
Act	1.083***	1.050***	1.085***	0.898***
	(0.155)	*(0.162)*	*(0.155)*	*(0.141)*
$\ln PD$	-0.360*	-0.358*	-0.156	-0.427***
	(0.251)	*(0.261)*	*(0.347)*	*(0.209)*
$\ln Loss$	-0.109	-0.107	-0.099	-0.151***
	(0.088)	*(0.086)*	*(0.086)*	*(0.075)*
$\ln NonH$	-0.112	-0.087	-0.119	-0.029
	(0.105)	*(0.106)*	*(0.106)*	*(0.090)*
Reg	-0.126**	-0.165**	-0.162**	0.095
	(0.062)	*(0.075)*	*(0.075)*	*(0.086)*
T	0.012	0.17	0.012	0.003
	(0.016)	*(0.018)*	*(0.019)*	*(0.019)*
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.146*	0.138*	0.114**	0.158***
	(0.113)	*(0.092)*	*(0.060)*	*(0.051)*
$\gamma = \sigma_u^2/\sigma^2$	0.890***	0.897***	0.870***	0.950***
	(0.098)	*(0.080)*	*(0.076)*	*(0.019)*
Log Likelihood	63.622	71.132	68.715	107.145

Notes: standard errors in brackets;
$*p < 0.10$; $**p < 0.05$; $***p < 0.01$

Table 23: Model B - Consistency in efficiency ranking

ID	Firms	Ranking		
		Best	Worse	Difference
1	Anglian Water services Ltd	6	13	7
2	Northumbrian Water Ltd	6	11	5
3	Severn Trent Water Ltd	15	21	6
4	South West Water Ltd	9	14	5
5	Southern Water Services Ltd	21	22	1
6	Thames Water Utilities Ltd	4	9	5
7	United Utilities Water plc	20	22	2
8	Dwr Cymru Cyfyngedig	1	4	3
9	Wessex Water Services Ltd	11	14	3
10	Yorkshire Water Services Ltd	4	13	9
11	Bournemouth & West Hampshire Water plc	2	8	6
12	Bristol Water plc	15	18	3
13	Cambridge Water Company plc	5	14	9
14	Dee Valley Water plc	16	19	3
15	Veolia Water South East Ltd	1	5	4
16	Mid Kent Water plc	14	19	5
17	Portsmouth Water plc	16	19	3
18	South East Water Ltd	18	22	4
19	South Staffordshire Water plc	1	4	3
20	Sutton & East Surrey Water plc	3	8	5
21	Veolia Water East Ltd	10	16	6
22	Veolia Water Central Ltd	6	10	4
			Min Diff.	1
			Max Diff.	9

Table 24: Model B - Correlation between the rankings

	Half-Normal vs Truncated-Normal distribution							
	B_1	B_2	B_3	B_4	B'_1	B'_2	B'_3	B'_4
B_1	1.000							
B_2	0.962	1.000						
B_3	0.987	0.968	1.0000					
B_4	0.935	0.900	0.946	1.000				
B'_1	0.981	0.961	0.977	0.932	1.000			
B'_2	0.944	0.988	0.956	0.909	0.966	1.000		
B'_3	0.962	0.983	0.977	0.920	0.974	0.980	1.000	
B'_4	0.871	0.837	0.888	0.959	0.890	0.853	0.887	1.000
	B_{t1}	B_{t2}	B_{t3}	B_{t4}	B'_{t1}	B'_{t2}	B'_{t3}	B'_{t4}
B_{t1}	1.000							
B_{t2}	0.969	1.000						
B_{t3}	0.980	0.966	1.000					
B_{t4}	0.949	0.939	0.965	1.000				
B'_{t1}	0.988	0.983	0.979	0.960	1.000			
B'_{t2}	0.948	0.993	0.959	0.940	0.970	1.000		
B'_{t3}	0.975	0.961	0.993	0.970	0.980	0.957	1.000	
B'_{t4}	0.840	0.839	0.850	0.918	0.863	0.847	0.866	1.000
	Time-Invariant vs Time-Variant assumption							
	B_1	B_2	B_3	B_4	B_{t1}	B_{t2}	B_{t3}	B_{t4}
B_1	1.000							
B_2	0.962	1.000						
B_3	0.987	0.968	1.000					
B_4	0.935	0.900	0.946	1.000				
B_{t1}	0.986	0.940	0.972	0.919	1.000			
B_{t2}	0.972	0.983	0.974	0.920	0.969	1.000		
B_{t3}	0.969	0.931	0.975	0.931	0.980	0.966	1.000	
B_{t4}	0.945	0.908	0.966	0.965	0.949	0.939	0.965	1.000
	B'_1	B'_2	B'_3	B'_4	B'_{t1}	B'_{t2}	B'_{t3}	B'_{t4}
B'_1	1.000							
B'_2	0.966	1.000						
B'_3	0.974	0.980	1.000					
B'_4	0.890	0.853	0.887	1.000				
B'_{t1}	0.967	0.939	0.949	0.852	1.000			
B'_{t2}	0.943	0.972	0.965	0.838	0.970	1.000		
B'_{t3}	0.941	0.913	0.932	0.856	0.980	0.957	1.000	
B'_{t4}	0.895	0.857	0.889	0.996	0.863	0.847	0.866	1.000

Table 27: Model B_{EFF} - Estimation results of the cost frontier

Model B_{EFF1}

Variables	Estimated parameter	Standard dev.
Constant	21.327***	1.059
$\ln P_m$	1.230**	0.694
$\ln Vol$	4.323***	0.462
$\ln Prop$	-4.111***	0.262
$\ln PD$	-0.708***	0.197
$\ln Loss$	-0.105***	0.034
$\ln NonH$	-3.614*	0.074
Act	-0.598***	0.221
Reg	-0.134***	0.021
Constant	-9.895***	1.714
$\ln SPress$	-0.415*	0.287
$\ln NInterrupt$	-2.204***	0.370
T	-0.052***	0.013
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.032***	0.003
$\gamma = \sigma_u^2/\sigma^2$	0.999***	0.000
Log Likelihood	108.032	

Model B_{EFF5}

Variables	Estimated parameter	Standard dev.
Constant	55.038***	0.998
$\ln P_m$	1.321***	0.329
$\ln Vol$	5.607***	0.404
$\ln Prop$	-4.009***	0.275
$\ln Drink$	-0.515	0.918
$\ln SPress$	0.614	0.500
$\ln NInterrupt$	1.798***	0.818
$\ln PD$	-0.595***	0.043
$\ln Loss$	-0.024***	0.034
$\ln NonH$	-0.137***	0.034
Act	-0.412***	0.066
Reg	-0.085***	0.008
Constant	0.400	0.957
$\ln NonComp$	-2.272***	0.442
$\ln Written$	-0.352*	0.239
$\ln Call$	-0.889**	0.535
T	-0.027**	0.012
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.041***	0.001
$\gamma = \sigma_u^2/\sigma^2$	0.999***	0.000
Log Likelihood	125.967	

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

Table 27: Continuation

Model B_{EFF8}		
Variables	Estimated parameter	Standard dev.
Constant	9.131***	1.480
$\ln P_m$	0.806	0.774
$\ln Vol$	7.018***	0.735
$\ln Prop$	-3.446***	0.304
Constant	-1.419	1.160
Act	0.379***	0.117
Reg	0.104	0.085
$\ln PD$	0.162	0.130
$\ln Loss$	-0.205**	0.108
$\ln NonH$	-0.136	0.133
$\ln Drink$	-4.964***	1.618
$\ln NonComp$	0.063	0.402
$\ln Written$	0.045	0.286
$\ln Call$	0.706	0.675
$\ln SPress$	-0.612**	0.335
$\ln NInterrupt$	-4.725***	1.806
T	0.005	0.023
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.018***	0.003
$\gamma = \sigma_u^2/\sigma^2$	0.643***	0.178
Log Likelihood	91.904	

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

Table 29: Model B_{EFF} - Consistency in efficiency ranking

ID	Firms	B_{EFF1}	B_{EFF5}	B_{EFF8}	Diff.
1	Anglian Water services Ltd	14	11	18	7
2	Northumbrian Water Ltd	6	9	16	10
3	Severn Trent Water Ltd	11	14	21	10
4	South West Water Ltd	21	19	20	2
5	Southern Water Services Ltd	19	20	22	3
6	Thames Water Utilities Ltd	17	18	17	1
7	United Utilities Water plc	22	22	19	3
8	Dwr Cymru Cyfyngedig	13	5	13	8
9	Wessex Water Services Ltd	10	8	15	7
10	Yorkshire Water Services Ltd	2	2	14	12
11	Bournemouth & West Hampshire Water plc	9	17	4	13
12	Bristol Water plc	15	16	5	11
13	Cambridge Water Company plc	5	7	1	6
14	Dee Valley Water plc	18	13	10	8
15	Veolia Water South East Ltd	1	4	2	3
16	Mid Kent Water plc	16	12	8	8
17	Portsmouth Water plc	20	21	12	9
18	South East Water Ltd	4	3	6	3
19	South Staffordshire Water plc	3	6	3	3
20	Sutton & East Surrey Water plc	8	1	9	8
21	Veolia Water East Ltd	7	15	7	8
22	Veolia Water Central Ltd	12	10	11	2
		Min Difference			1
		Max Difference			13

Table 30: Model B_{EFF} - Correlation between the rankings

	B_{EFF1}	B_{EFF5}	B_{EFF8}
B_{EFF1}	1.000		
B_{EFF5}	0.803	1.000	
B_{EFF8}	0.579	0.434	1.0000

Table 32: Models B & B_{EFF} - Consistency in efficiency ranking

ID	Firms	Ranking		
		Best	Worse	Diff.
1	Anglian Water services Ltd	6	18	12
2	Northumbrian Water Ltd	6	16	10
3	Severn Trent Water Ltd	11	21	10
4	South West Water Ltd	10	21	11
5	Southern Water Services Ltd	19	22	3
6	Thames Water Utilities Ltd	5	18	13
7	United Utilities Water plc	19	22	3
8	Dwr Cymru Cyfyngedig	1	13	12
9	Wessex Water Services Ltd	8	15	7
10	Yorkshire Water Services Ltd	2	14	12
11	Bournemouth & West Hampshire Water plc	2	17	15
12	Bristol Water plc	5	18	13
13	Cambridge Water Company plc	1	14	13
14	Dee Valley Water plc	10	19	9
15	Veolia Water South East Ltd	1	4	3
16	Mid Kent Water plc	8	19	11
17	Portsmouth Water plc	12	21	9
18	South East Water Ltd	3	21	18
19	South Staffordshire Water plc	1	6	5
20	Sutton & East Surrey Water plc	1	9	8
21	Veolia Water East Ltd	7	16	9
22	Veolia Water Central Ltd	6	12	6
			Min Diff.	3
			Max Diff.	18

Table 33: Models B versus B_{EFF} - Correlation between the rankings

	B'_{t2}	B'_{t3}	B'_{t4}
B_{EFF1}	-	0.535	-
B_{EFF5}	0.399	-	-
B_{EFF8}	-	-	0.447