

Regulatory Price Performance, Excess Cost Indexes and Profitability: How Effective is Price Cap Regulation in the Water Industry?

By
**ALEXANDROS MAZIOTIS, DAVID SAAL AND EMMANUEL
THANASSOULIS**

RP 0920

Aston Business School
Aston University
Birmingham
B4 7ET, UK
E-mail: maziotia@aston.ac.uk
E-mail: D.S.Saal@aston.ac.uk
E-mail: E.Thanassoulis@aston.ac.uk

September 2009

ISBN No: 978-1-85449-760-4

Aston Academy for Research in Management is the administrative centre for all research activities at Aston Business School. The School comprises more than 70 academic staff organised into thematic research groups along with a Doctoral Programme of more than 50 research students. Research is carried out in all of the major areas of business studies and a number of specialist fields. For further information contact:

The Research Director, Aston Business School, Aston University,
Birmingham B4 7ET

Telephone No: (0121) 204 3000 Fax No: (0121) 204 3326 <http://www.abs.aston.ac.uk/>

Aston Business School Research Papers are published by the Institute to bring the results of research in progress to a wider audience and to facilitate discussion. They will normally be published in a revised form subsequently and the agreement of the authors should be obtained before referring to its contents in other published works.

Abstract

In this study we apply an index number approach to allow for cross sectional comparisons of relative profitability, productivity and price performance of the regulated Water and Sewerage companies (WaSCs) in England and Wales during the years 1991-2008. In order to better analyse the impact of regulation on WaSC performance, we decompose actual economic profits into spatial multilateral Fisher productivity (TFP), the inverse of which is demonstrated to be a regulatory excess cost index that measures the deviation of a firm's actual costs from benchmark costs, and a newly developed regulatory total price performance (TPP) index, which measures the excess of regulated revenues relative to benchmark costs. The results indicate that during the years 1991-2000 price caps were "weak" as prices were high enough for the firms to achieve economic profits despite their low productivity levels. However, after 2001 prices became "catch up promoting" as they required less productive companies to eliminate at least some excess costs in order to eliminate economic losses. When compared to alternative methodologies such as DEA and SFA, our index number based approach has the further advantage of allowing meaningful comparative performance measurement even if the number of available observations is extremely limited. We therefore suggest that our approach should be of great interest, not only to those researchers interested in evaluating the effectiveness of regulation, but also to those researchers more focused on developing effective comparative performance techniques, even if sample sizes are limited.

Keywords: Profits, productivity, price performance, index numbers, regulation, water industry

1. Introduction¹

The Water and Sewerage companies (WaSCs) in England and Wales were privatized as natural monopolies and thus they had strong incentives for monopoly pricing and weak incentives for operating efficiently. A regulatory body, the Office of Water Services (Ofwat) was set up in order to incentivize firms to achieve both productive and allocative efficiency. The method of regulation in UK water and sewerage sector is price cap regulation and is designed to give firms incentives to increase profits by reducing costs by eliminating the potential to manipulate prices and is preferred to rate of return regulation, which potentially leads to overcapitalization (Averch-Johnson, 1962).

In this paper, we demonstrate that the effectiveness of a price cap scheme; e.g. whether it encourages regulated firms to achieve efficiency in production as well as appropriate allocatively efficient pricing; can be evaluated by determining the relationship between productivity, price performance and profits across firms. Moreover, the underlying index number techniques also allow for comparative performance measurement assessment even in cases where the number of observations is extremely limited. Thus, when compared to alternative methodologies such as DEA and SFA, which require a relatively large number of observations to specify an efficient frontier, index number techniques provide a considerable advantage. Previous studies that illustrated the relationship between profits, productivity and price performance using index number techniques include Water and Street (1998), Han and Hughes (1999), Saal & Parker (2001), Salerian (2003), and Diewert & Lawrence (2006).

Our approach could be naively seen as a minor development of Saal & Parker (2001), which employed a methodology that only allowed the measurement of firm-specific profitability, productivity and price performance indices over time. Thus, our study instead employs a cross sectional (spatial) index number technique to measure differences in the level of productivity, price performance and profitability across firms (relative comparative performance). However, moving from the consideration of firm-specific indices to a spatial approach firstly allows comparative performance assessment, which makes the approach directly applicable by regulators

¹ The authors would like to express their gratitude for the support of the Economic and Social Science Research Council as well as the Office of Water Services (Ofwat), and note that the usual disclaimer applies.

in setting price caps. Secondly, and more significantly, it also allows the development of the theoretically consistent model of price cap regulation presented in this paper. As a result, the approach employed in this paper facilitates an analysis of whether price caps are consistent with the achievement of productive and allocative efficiency, which was simply not possible with the methodology employed by Saal and Parker (2001).

The key theoretical contribution that is allowed by the spatial orientation employed in this paper is the decomposition of a firm's actual economic profitability into two sources: a spatial multilateral Fisher productivity index (TFP) and a newly developed regulatory total price performance (TPP) index. The former is calculated using theoretically consistent relative productivity comparisons across companies in any given year (multilateral spatial comparisons) after assuming the most productive company is the base or benchmark firm. Moreover, we demonstrate that the inverse of a spatial multilateral TFP index can be interpreted as a regulatory excess costs index, which measures the excess of a firm's actual costs relative to benchmark costs. The regulatory TPP index is derived as a function of this regulatory excess cost index and the actual economic profitability index, and measures the excess of regulated revenues relative to benchmark costs. As such, it provides a direct measure of how tight price caps are, measured by the proportional deviation between allowed revenues and benchmark costs. Further consideration of the theoretical relationship between actual economic profitability, regulatory excess costs and regulatory price performance allows a characterisation of the power of regulatory price caps, and we illustrate this by offering an analysis of changes in the estimated power of price cap regulation in the English and Welsh water industry over the period 1991-2008.

This paper unfolds as follows. Section 2 discusses the potential application of index number techniques for measuring actual profitability, relative productivity and price performance and its adaption under price cap regulation. Section 3, then considers the methodology necessary to empirically apply this approach in a multilateral setting. The next section provides a discussion of data employed, and the following section details the empirical results. Section 6 then offers some conclusions, as well as suggesting a potential extension of the model.

2. Relative Profitability, Productivity and Price Performance

A firm's economic performance is commonly measured by its economic profitability (π). However, changes in profitability can be decomposed into changes in productivity and price performance. Total factor productivity (TFP) captures changes in performance attributable to increased physical production of outputs relative to inputs. In contrast, total price performance (TPP) captures the impact of changes in output prices relative to input prices. Comparing changes in TFP and TPP therefore allows determination of whether profit change is primarily explained by improvements in productivity or is simply attributable to an increase in output prices relative to input prices that has improved the firm's price mark up relative to actual costs.

Saal & Parker (2001) demonstrates an index number approach to decompose a firm's economic profitability change into TFP change and TPP change. For any given firm, this methodology allows identification of the relative contributions of productivity and price performance to observed profit change and the paper illustrates how changes in regulatory policy influenced both the productivity change and price performance change of regulated water and sewerage companies (WaSCs) in England and Wales (E&W). Nevertheless, while this methodological approach has the strong advantage of allowing the decomposition of profit change even if data is only available for a single firm, it only allows comparison of cross firm differences in the rate of change of TFP, TPP and profitability. Thus, the lack of any cross sectional link between firms' indices makes it impossible to measure differences in the level of TFP, TPP and profitability across firms. The implication of this limitation is highlighted if one notes that Saal & Parker (2001) considers an industry subject to price cap regulation in which prices are set using a comparative yardstick regime that measures firm performance levels relative to other regulated firms, but it does not in fact provide a methodology that allows for measurement of such performance differences. This paper therefore proposes a spatial alternative to Saal & Parker (2001) that allows for measurement of a firm's TFP, TPP and profit performance relative to its peers at any give time.

In this section we first illustrate spatial indices of economic profitability and their decomposition in any given year of our sample and how we can employ spatial indices of productivity and price performance under an ideal incentive regulation

regime. After this illustration, Section 3 will tackle the thornier issue of applying this concept in an empirical multilateral setting.

Spatial Profitability, Productivity, and Price Performance and its Adaptation to Price Cap Regulation

In this section we consider the relationship between profits, productivity and price performance for firm i relative to a base firm b at time t which we call a spatial index, thereby adopting the terminology employed in the price index literature (Hill, 2004). As a result of its definition, these indices only directly measure differences in performance in the spatial dimension (between firms) at any given time.

We define the economic profits of the base firm b at time t , $\Pi_{b,t}$, as a ratio of its total revenues, $R_{b,t}$ and total costs, $C_{b,t}$, at time t . Thus, the total revenues of the base firm b at period t are defined as $R_{b,t} = P_{b,t} \times Y_{b,t}$, where $P_{b,t}$ and $Y_{b,t}$ respectively present the output price index and the aggregate output index of the base firm b at time t . Its total costs at year t , $C_{b,t}$, are defined as $C_{b,t} = W_{b,t} \times X_{b,t}$, where $W_{b,t}$ and $X_{b,t}$ denotes the input price index and the aggregate input index respectively of the base firm at time t . Similarly, we can define the economic profitability of any firm i at time t , $\Pi_{i,t}$ as a ratio of its total revenues, $R_{i,t}$ and its total costs, $C_{i,t}$. We can thus define and decompose a spatial economic profitability index for any firm i relative to the base firm b at time t , $\pi_{b,t}^S$ as follows:

$$\pi_{i,t}^S = \frac{\Pi_{i,t}}{\Pi_{b,t}} = \frac{\frac{R_{i,t}}{C_{i,t}}}{\frac{R_{b,t}}{C_{b,t}}} = \frac{\frac{P_{i,t} Y_{i,t}}{W_{i,t} X_{i,t}}}{\frac{P_{b,t} Y_{b,t}}{W_{b,t} X_{b,t}}} = \frac{TFP_{i,t}}{TFP_{b,t}} \times \frac{TPP_{i,t}}{TPP_{b,t}} = \frac{Y_{i,t}}{X_{i,t}} \times \frac{P_{i,t}}{W_{i,t}} \times \frac{Y_{b,t}}{X_{b,t}} \times \frac{P_{b,t}}{W_{b,t}} = \frac{Y_{i,t}^S}{X_{i,t}^S} \times \frac{P_{i,t}^S}{W_{i,t}^S} = TFP_{i,t}^S \times TPP_{i,t}^S \quad (1)$$

Thus, at time t , a spatial economic profitability index, $\pi_{i,t}^S$ can be expressed as a function of an index of spatial total factor productivity for firm i relative to the base firm b , $TFP_{i,t}^S$ and a spatial index of total price performance between firm i and the base firm b , $TPP_{i,t}^S$. As $TFP_{i,t}^S = Y_{i,t}^S / X_{i,t}^S$ and $TPP_{i,t}^S = P_{i,t}^S / W_{i,t}^S$ these indices can be further decomposed as functions of the spatial output ($Y_{i,t}^S = Y_{i,t} / Y_{b,t}$), input ($X_{i,t}^S = X_{i,t} / X_{b,t}$), output price ($P_{i,t}^S = P_{i,t} / P_{b,t}$) and input price ($W_{i,t}^S = W_{i,t} / W_{b,t}$)

indices. This decomposition of spatial profitability highlights that, at any given time, observed differences in profitability between firms can be explained by differences in productivity, differences in price performance, or differences in both.

By definition spatial indices estimate firm i 's performance relative to any potential base firm b . However, one of the goals of an ideal incentive regulation system is to incentivize firms to improve their productivity so as to catch up to the productivity levels achieved by the frontier or most productive firm. Therefore, a natural candidate for the base firm is the firm that has the highest relative level of productivity at time t . This is because $TFP_{i,t}^S$ then becomes an easily interpretable index of any firm's TFP as a proportion of the best observed productivity level. We therefore henceforward define the base firm (b) as the firm with the highest productivity observed at time t .

However, allocative efficiency is also a goal of an ideal incentive regulation regime, as output prices should in principle be just sufficient to cover the efficient economic costs of production. Stated differently, if output prices are allocatively efficient, the most productive regulated companies should achieve a normal rate of return, or equivalently economic profits should be equal to zero. In contrast, less productive firms should in theory have output prices that would allow them to achieve a normal rate of return only if they achieved the productivity levels achieved by the most productive firm. We therefore choose to adapt our definition of spatial TPP so rather than being based relative to the actual output price index of the base firm ($P_{b,t}$), it is instead based to the output prices that are consistent with zero economic profits for the base firm ($P_{b,t}^*$).

Mathematically, this is easily accomplished by first noting that if the base (highest productivity) firm at time t makes zero economic profits, then

$$\Pi_{b,t} = \frac{R_{b,t}}{C_{b,t}} = \frac{P_{b,t} Y_{b,t}}{W_{b,t} X_{b,t}} = 1. \text{ By simply rearranging this expression, and after assuming}$$

that input prices are exogenous, we obtain the following expression for the optimal output price for the best practice firm, which is consistent with the regulatory goal of achieving allocative efficiency:

$$P_{b,t}^* = W_{b,t} \frac{X_{b,t}}{Y_{b,t}} \quad (2)$$

$P_{b,t}^*$ is an obvious restatement of the condition that for a firm achieving the highest observed productivity, prices should be set such that total revenues are equivalent to economic costs, thereby making economic profits equal to zero. Moreover, it is illustrative to note that in the case of a single output, this is consistent with $P_{b,t}^*$ being equal to average long run total costs.

The simple substitution of $P_{b,t}^*$ for $P_{b,t}$ in $TPP_{i,t}^S$ completes the construction of a “regulatory” TPP index ($TPP_{i,t}^R$), which indicates the contribution to profitability that can be attributed to deviation of firm i 's output prices from those that would be consistent with achieving the same productivity as the base firm, and also realizing zero economic profits:

$$TPP_{i,t}^R = \frac{\frac{P_{i,t}}{P_{b,t}^*}}{\frac{W_{i,t}}{W_{b,t}}} = \frac{\left[\frac{P_{i,t}}{W_{i,t}} \right]}{\left[\frac{X_{b,t}}{Y_{b,t}} \right]} = \frac{P_{i,t} Y_{b,t}}{W_{i,t} X_{b,t}} \quad (3)$$

As suggested by the term after the first equality, $TPP_{i,t}^R$ indicates the deviation of firm i 's output price from those that would be appropriate given its input prices and the assumption that it achieved the same productivity as firm b . It should also be clear that increases (decreases) in $TPP_{i,t}^R$ can be interpreted as loosening (tightening) of regulatory price caps, because this reflects an increase (decrease) in allowed revenues relative to benchmark costs. We can begin to further characterize regulation if we focus on the term after the second equality, if $TPP_{i,t}^R = 1$, $P_{i,t} = P_{i,t}^* = W_{i,t} X_{b,t} / Y_{b,t}$ and firm i will achieve a normal rate of return if it achieved the productivity level of the base firm. If $TPP_{i,t}^R < 1$ then $P_{i,t} < P_{i,t}^*$, thereby suggesting that the regulatory price $P_{i,t}$ has been set “low” as the firm would make an economic loss even if it achieved the productivity levels of the best firm. In contrast, if $TPP_{i,t}^R > 1$, then $P_{i,t} > P_{i,t}^*$, thereby

suggesting that the regulatory price $P_{i,t}$ has been set “high” as the firm would make an economic profit if it achieved the productivity levels of the best firm.

Given the regulatory definition of $TPP_{i,t}^R$, multiplying it by $TFP_{i,t}^S$, no longer results in the spatial measure of profitability ($\pi_{i,t}^S$) detailed in (1). As (4) demonstrates the product of $TPP_{i,t}^R$ and $TFP_{i,t}^S$ has the advantage that it results in the direct measure of actual firm specific economic profitability $\Pi_{i,t}$ previously defined above:

$$\Pi_{i,t} = \frac{R_{i,t}}{C_{i,t}} = \frac{P_{i,t} Y_{i,t}}{W_{i,t} X_{i,t}} = \left[\frac{\frac{P_{i,t}}{W_{i,t}}}{\frac{X_{b,t}}{Y_{b,t}}} \right] \left[\frac{\frac{Y_{i,t}}{X_{i,t}}}{\frac{Y_{b,t}}{X_{b,t}}} \right] = TPP_{i,t}^R TFP_{i,t}^S \quad (4)$$

Interpretation of (4) also demonstrates several useful implications from a regulatory perspective. Focusing on the base firm and given our assumptions, $TFP_{b,t}^S = 1$ thereby revealing the rather obvious finding that the base firm has achieved the regulatory target of achieving best practice TFP. As a result, for the base firm any economic profits (losses) result in $\Pi_{b,t} > 1$ ($\Pi_{b,t} < 1$) and imply that $TPP_{b,t}^R > 1$ ($TPP_{b,t}^R < 1$). Thus, from a regulatory perspective, for the base firm, economic profits (losses) can only result from “inappropriately” high (low) output prices such that $P_{b,t} > P_{b,t}^*$ ($P_{b,t} < P_{b,t}^*$). e.g. for the most productive firm, economic profits or losses can only be attributed to regulatory output prices that are not consistent with zero economic profits, and as a result revenues exceed economic costs. We would note that, in a regulatory context where regulators set price caps that include both “catch-up” and “continuing improvement factors”, it is more than plausible that a regulator could set $P_{b,t} < P_{b,t}^*$ in order to incentive the base firm to further improve its underlying TFP in the future.

For any other firm i , if $TPP_{i,t}^R = 1$, $P_{i,t} = P_{i,t}^*$ and the regulator has set prices so as to fully incentivize the firm to catch up to the productivity of the base firm. As a result, $\Pi_{i,t} = TFP_{i,t}^S$ and the firm’s profitability index, will deviate from 1 in exact proportion to its spatial TFP. In this situation the firm will be making an economic loss because of its below par TFP performance. Moreover, it could be argued that such losses are appropriate as the firm’s revenues have been set equal to an

appropriate benchmark cost associated with the best TFP performance, but the firm's costs exceed these costs by the proportion $1/TFP_{i,t}^S$. Therefore, these economic losses are consistent with setting regulated prices so as to incentivize the firm to fully close the productivity gap between it and the base firm.

When $P_{i,t} \neq P_{i,t}^*$, output prices, and hence $\Pi_{i,t}$ are not consistent with the spatial productivity benchmark $TFP_{i,t}^S$. If, $P_{i,t} < P_{i,t}^*$ then $TPP_{i,t}^R < 1$ and as a result $\Pi_{i,t} < TFP_{i,t}^S$. In this situation regulated revenues are below the benchmark economic costs implied by $TFP_{i,t}^S$, and the economic losses of the firm are partially explained by low output prices that imply the firm would need to not only catch up to but also exceed the base firm's productivity level, if it wished to eliminate its economic losses. Therefore, if $TPP_{i,t}^R < 1$, this could suggest evidence of "powerful" price caps, and/or price caps that are designed to stimulate both catch-up and continuing improvements in TFP.

In contrast, if $P_{i,t} > P_{i,t}^*$ then $TPP_{i,t}^R > 1$ and as a result $\Pi_{i,t} > TFP_{i,t}^S$, this is consistent with regulatory prices not having been set to incentivize a firm to fully close its productivity gap with the base firm. This could result for a variety of reasons. One potential reason is the common regulatory practice of setting price caps in a manner that allows required catch up productivity gains to be accomplished over several years rather than immediately. However, this situation could also be taken as evidence of "weak" regulation that does not fully penalize unproductive firms and cause them to suffer economic losses unless they improve their productivity. As UK regulators, for example, have a duty to maintain the financial viability of regulated companies as well as to improve their productivity, this could even be justified by regulators on the grounds that less productive firms would go bust if tough price caps were set and they were unable to sufficiently improve their productivity performance.

As the above paragraph illustrates, there are plausible and potentially appropriate reasons why regulators may choose to set $P_{i,t} > P_{i,t}^*$. Moreover, even if $P_{i,t} > P_{i,t}^*$, the regulator may still have set output prices in a manner that is designed to better incentivize laggard firms to improve their productivity performance, even if they are not required to close the full productivity gap with the base firm. It is therefore worthwhile to carefully define several critical values of $TPP_{i,t}^R$ that can be

used to characterise the power of price cap regulation. As discussed above, if $TPP_{i,t}^R = 1$ a firm could only eliminate its economic losses by fully catching up to the base firm's TFP level. Therefore, if $TPP_{i,t}^R < 1$, we can characterize price regulation as “powerful” as $P_{i,t}$ is such that a firm could only eliminate its economic losses by catching up to and then exceeding the base firm's TFP.

In contrast, if $TPP_{i,t}^R > 1$ regulation is somewhat dampened in its effect as economic losses can be eliminated without improving TFP to the level of the base firm. However, there is still a clear distinction between “catch-up promoting” regulatory price caps which still retain some incentives to improve productivity, and “weak” price caps, which allow a laggard firm to potentially earn economic profits regardless of whether its productivity is improved. Thus, if $TPP_{i,t}^R > 1$, but $TPP_{i,t}^R < 1/TFP_{i,t}^S$, price caps are “catch-up promoting” as they require some, but not full catch up in TFP to eliminate economic losses. In contrast, if $TPP_{i,t}^R > 1/TFP_{i,t}^S$, price caps are “weak” as prices are high enough for the firm to achieve economic profits despite its low productivity levels, thereby suggesting relatively weak incentives for the firm to improve its productivity

We finally note that it is worthwhile to define a regulatory excess cost index $E_{i,t}^R = 1/TFP_{i,t}^S$, which given the assumption of exogenous input prices, provides an index of the excess of a firm's costs relative to those that would be achieved if it achieved the productivity benchmark. If $TPP_{i,t}^R > 1$, but $TPP_{i,t}^R < E_{i,t}^R$, then the revenues achieved by a firm i when its output prices exceed optimal prices are lower than the “excess costs” relative to the benchmark costs resulting in economic losses ($\Pi_{i,t} < 1$). However, if $TPP_{i,t}^R > 1$, but $TPP_{i,t}^R > E_{i,t}^R$, then the revenues achieved by a firm i when its output prices exceed optimal prices are greater than the “excess costs” relative to the benchmark costs resulting in economic profits ($\Pi_{i,t} > 1$). Thus, it should be clear that if $TPP_{i,t}^R < E_{i,t}^R$ then the regulator has set prices that require laggard firms to improve their productivity/eliminate their excess costs, if they wish to return to economic profitability.

In sum, our discussion highlights that if both productive and allocative efficiency are the goals of price cap regulation, any firm in a regulated industry that has the productivity of the “best practice firm” should in principle have output prices

that result in zero economic profits, while less productive firms should have output prices that would result in economic losses unless they improve their productivity. Therefore, the regulator needs to take into account the impact of both TFP and TPP in the overall performance of the companies when setting price caps. Moreover, at a theoretical level, systematic deviation of $TPP_{i,t}^R$ from a value of 1 can be seen as evidence of deviation from the goal of setting prices that are consistent with a strict interpretation of both the incentive and allocative efficiency based justification for price cap regulation. Similarly, if $TPP_{i,t}^R > E_{i,t}^R$ regulated output prices are high enough to violate a looser interpretation of appropriate incentive regulation, which requires only partial productivity catch up to achieve economic profitability. Given this theoretical discussion, our next section therefore discusses a methodological approach that allows the development of these ideas in an empirical application.

3. Multilateral Productivity, Price Performance and Profitability Computations

In this section, we employ a multilateral Fisher index approach to measure profitability, productivity and price performance across companies at any given year (multilateral spatial comparisons). When the price and quantities across different companies are compared, it is important that such comparisons are undertaken for every pair of companies being considered (multilateral comparisons). However, in order to achieve consistency between all the pairs of comparisons we need to derive multilateral indexes that fulfill the property of transitivity. Internal consistency (transitivity) implies that a direct comparison between two firms gives the same result when comparing indirectly these two firms through a third firm.

Bilateral Fisher output and input indexes between two firms i and j in the case of m outputs and n inputs are respectively, $Y_{i,j}^F$ and $X_{i,j}^F$ where:

$$Y_{i,j}^F = \left[\frac{\sum_{j=1}^M P_j^m Y_i^m}{\sum_{m=1}^M P_j^m Y_j^m} \times \frac{\sum_{m=1}^M P_i^m Y_i^m}{\sum_{m=1}^M P_i^m Y_j^m} \right]^{\frac{1}{2}} \quad X_{i,j}^F = \left[\frac{\sum_{j=1}^N W_j^n X_i^n}{\sum_{n=1}^N W_j^n X_j^n} \times \frac{\sum_{n=1}^N W_i^n X_i^n}{\sum_{n=1}^N W_i^n X_j^n} \right]^{\frac{1}{2}} \quad (5)$$

Y_i^m and Y_j^m denote the quantities for the m th output for firms i and j respectively, whereas X_i^n and X_j^n present the quantities for the n th inputs for firms i and j respectively. Moreover, P_i^m and P_j^m are prices for the m th output, while W_i^n and W_j^n denote input prices. The Fisher output and input indexes measure firm i 's output and input as a proportion of firm j and are the geometric means of Laspeyers and Paasche output and input indexes. For instance, Laspeyers output and input indexes use company j 's prices to weight quantity changes, whereas Paasche output and input indexes use firm i 's prices to weight quantity changes. The bilateral Fisher productivity index can then be constructed as a ratio of the Fisher output index relative to the Fisher input index:

$$TFP_{i,j}^F = \frac{Y_{i,j}^F}{X_{i,j}^F} \quad (6)$$

The above formula is a binary comparison that can be applied directly when we are only interested in making comparisons between two firms. However, when we are interested in making meaningful comparisons between more than two firms, the multilateral nature of spatial comparisons creates some difficulties, which arise from the fact that more than two firms are compared at the same time. Firstly, the number of comparisons may be quite large depending on the number of companies that we have in our sample so the calculation of productivity index can be quite difficult. Secondly, we need consistent comparisons between all firms such that the relative comparisons between any two firms are consistent with other comparisons (transitivity).

Following standard practice, the process of calculating a transitive Fisher output ($Y_{i,j}^F$) and input ($X_{i,j}^F$) indices begins by calculating all the possible binary comparisons, $i, j = 1, \dots, I$ where I is the total number of companies, and results in the following $I \times I$ matrices of binary comparisons:

$$\begin{bmatrix} Y_{1,1}^F & Y_{1,2}^F & \dots & Y_{1,I}^F \\ Y_{2,1}^F & Y_{2,2}^F & \dots & Y_{2,I}^F \\ \dots & \dots & \dots & \dots \\ Y_{I,1}^F & Y_{I,2}^F & \dots & Y_{I,I}^F \end{bmatrix} \quad \begin{bmatrix} X_{1,1}^F & X_{1,2}^F & \dots & X_{1,I}^F \\ X_{2,1}^F & X_{2,2}^F & \dots & X_{2,I}^F \\ \dots & \dots & \dots & \dots \\ X_{I,1}^F & X_{I,2}^F & \dots & X_{I,I}^F \end{bmatrix} \quad (7)$$

These binary Fisher indices can be converted into multilateral consistent transitive indices by applying the EKS method developed by Elteto-Koves (1964) and Szulc (1964) to derive transitive Fisher indices (see Caves, Christensen and Diewert (1982a), Diewert and Lawrence (2006), Caves et al (1981), Ball et al (2001) for a discussion on multilateral transitive indices). We therefore derive transitive Fisher output and input indices using the EKS method, which is equivalent to taking the geometric mean of the I possible direct and indirect (through any possible 3rd firm k) binary Fisher comparisons of firms i and j . The resulting Fisher output and input indices, Y_{ij}^S and X_{ij}^S therefore fulfill the transitivity property:

$$Y_{ij}^S = \prod_{k=1}^I [Y_{ik}^F \times Y_{kj}^F]^{\frac{1}{I}} \quad X_{ij}^S = \prod_{k=1}^I [X_{ik}^F \times X_{kj}^F]^{\frac{1}{I}} \quad (8)$$

Adopting the terminology of the price index literature (Hill, 2004) we refer to these multilateral output and input indices as spatial indices, as they provide spatially consistent measures across all firms.

The spatial total factor productivity Fisher index for a firm i relative to firm j , $TFP_{i,j}^S$, can then be constructed as a ratio of the spatial Fisher output index relative to spatial Fisher input index:

$$TFP_{ij}^S = \frac{Y_{ij}^S}{X_{ij}^S} \quad (9)$$

However, one can also derive fully equivalent transitive Fisher productivity indices using the EKS method by directly taking the geometric mean of all I possible direct and indirect (through any possible 3rd firm k) binary Fisher productivity comparisons of firms i and j :

$$TFP_{ij}^S = \prod_{k=1}^I [TFP_{ik}^F \times TFP_{kj}^F]^{-1} \quad (10)$$

The resulting index fulfills the transitivity property since it is derived using the EKS method, so any direct comparison between two firms i and j is the same with an indirect comparison between these two firms with a third firm k :

$$TFP_{i,j}^S = TFP_{i,k}^S \times TFP_{k,j}^S \quad \forall i, j \quad (11)$$

While we can generate the $I \times I$ possible transitive spatial output, input and productivity indexes between all firms, transitivity also implies that all meaningful information with regard to relative productivity is available in a subset of only I of these indices. Thus, if we arbitrarily choose one firm as a base firm and set $j = b$, then each spatial measure, is a measure of firm i relative to the chosen base firm and we can also simplify notation such that $TFP_{i,b}^S = TFP_i^S$, $Y_{i,b}^S = Y_i^S$, $X_{i,b}^S = X_i^S$. Therefore, productivity relative to the base firm's productivity can be expressed as:

$$TFP_i^S = \frac{Y_i^S}{X_i^S} \quad (12)$$

However, this simplification comes at no loss of generality as another spatial productivity measure between any given firms can simply be calculated as $TFP_{i,j}^S = TFP_i^S / TFP_j^S$. Similarly, $Y_{i,j}^S = Y_i^S / Y_j^S$ and $X_{i,j}^S = X_i^S / X_j^S$.

If spatial comparisons are available for each of T time periods indexed by t , and we assume the same base firm in all years, we can define the spatial productivity of firm i relative to firm b at time t as:

$$TFP_{i,t}^S = \frac{Y_{i,t}^S}{X_{i,t}^S} \quad (13)$$

These $I \times T$ measures then form the elements of a complete set of spatial comparisons indicating the productivity, output and input of firm i relative to the base firm at time t , and can be succinctly illustrated in the matrices:

$$\begin{aligned}
 TFP^S &= \begin{bmatrix} TFP_{1,1}^S & TFP_{1,2}^S & \dots & TFP_{1,T}^S \\ TFP_{2,1}^S & TFP_{2,2}^S & \dots & TFP_{2,T}^S \\ \dots & \dots & \dots & \dots \\ TFP_{I,1}^S & TFP_{I,2}^S & \dots & TFP_{I,T}^S \end{bmatrix} & Y^S &= \begin{bmatrix} Y_{1,1}^S & Y_{1,2}^S & \dots & Y_{1,T}^S \\ Y_{2,1}^S & Y_{2,2}^S & \dots & Y_{2,T}^S \\ \dots & \dots & \dots & \dots \\ Y_{I,1}^S & Y_{I,2}^S & \dots & Y_{I,T}^S \end{bmatrix} \\
 X^S &= \begin{bmatrix} X_{1,1}^S & X_{1,2}^S & \dots & X_{1,T}^S \\ X_{2,1}^S & X_{2,2}^S & \dots & X_{2,T}^S \\ \dots & \dots & \dots & \dots \\ X_{I,1}^S & X_{I,2}^S & \dots & X_{I,T}^S \end{bmatrix} & & (14)
 \end{aligned}$$

Moreover, we report $I \times T$ measures of a regulatory excess cost index for any firm i at time t as the inverse of the spatial productivity, $E_{i,t}^R = 1/TFP_{i,t}^S$, which given the assumption of exogenous input prices, provides an index of the excess of a firm's costs relative to those that would be achieved if it achieved the productivity benchmark. The set of the $I \times T$ measures of the regulatory excess cost index can be illustrated in the following matrix:

$$E^R = \begin{bmatrix} E_{1,1}^R & E_{1,2}^R & \dots & E_{1,T}^R \\ E_{2,1}^R & E_{2,2}^R & \dots & E_{2,T}^R \\ \dots & \dots & \dots & \dots \\ E_{I,1}^R & E_{I,2}^R & \dots & E_{I,T}^R \end{bmatrix} \quad (15)$$

We now turn our discussion to the construction of the spatial total price performance index, $(TPP_{i,t}^S)$ and then the regulatory total price performance index, $(TPP_{i,t}^R)$. In (1) we defined the spatial total price performance of any firm i relative to the base firm as a ratio of output prices to input prices relative to the base firm. Since we defined the spatial TFP index as the productivity index of any firm relative to the best productive firm, we similarly define the spatial TPP index as the price performance index of any firm relative to the price performance of the most productive firm. To accomplish

this, we firstly express turnover of a firm i relative to the base firm as $R_{i,t}^S = R_{i,t}/R_{b,t}$. The spatially consistent aggregate output price index, $(P_{i,t}^S)$ is then calculated as $P_{i,t}^S = R_{i,t}^S/Y_{i,t}^S$. Similarly, we express nominal economic costs of a firm i relative to the base firm as $C_{i,t}^S = C_{i,t}/C_{b,t}$. The spatially consistent aggregate input price index, $(W_{i,t}^S)$ is then calculated as $W_{i,t}^S = C_{i,t}^S/X_{i,t}^S$. Finally, a spatially consistent TPP index of any firm i relative to the base firm at any given time t , $(TPP_{i,t}^S)$ can be obtained as:

$$TPP_{i,t}^S = \frac{\frac{R_{i,t}^S}{Y_{i,t}^S}}{\frac{C_{i,t}^S}{X_{i,t}^S}} = \frac{P_{i,t}^S}{W_{i,t}^S} \quad (16)$$

By rearranging (4) an estimate of $TPP_{i,t}^R$ can be obtained as a function of firm specific economic profitability, $\Pi_{i,t}$, and the spatially consistent regulatory excess cost index, $E_{i,t}^R = 1/TFP_{i,t}^S$:

$$TPP_{i,t}^R = \Pi_{i,t} E_{i,t}^R \quad (17)$$

Recall that $TPP_{i,t}^R$ measures the proportional deviation of output prices from those that are consistent with zero economic profits if the firm eliminated its excess regulatory costs. As discussed above, if $TPP_{i,t}^R < 1$ we can characterize price regulation as “powerful” in the sense of requiring full catch up to the base firm and further productivity improvement to regain economic profitability. From (17) we can see that such “powerful” regulation has the empirically observable requirement that $\Pi_{i,t} < 1/E_{i,t}^R$. Similarly, our discussion above revealed that regulatory prices are “catch up promoting” if $TPP_{i,t}^R < E_{i,t}^R$, which from (17) requires that $\Pi_{i,t} < 1$. E.g. firms are required to at least partially eliminate their regulatory excess costs if they wish to regain economic profitability. Thus, it should be clear that the relatively straight forward comparison of $\Pi_{i,t}$, $E_{i,t}^R$, and $TPP_{i,t}^R$ can provide extremely relevant information with regard to the relative power of regulatory price caps.

Before turning to our empirical application, we must simultaneously highlight both the strength, as well as a potential pitfall, of our index number based methodology. Given that any model of company performance under regulation is only valuable if it can be empirically implemented, our methodology has the distinct advantage of potentially allowing for theoretically consistent cross sectional comparisons of relative productivity, profitability, and price performance, in samples with as few as 2 observations. Thus, when compared to econometric and DEA based approaches to performance measurement, this is a distinct advantage. However, when compared to other methodologies such as SFA and DEA. Our index number methodology does not allow us to as readily take into account differences in operating characteristics that may affect relative measures of productivity or price performance. Nevertheless, given that profitability is not influenced by these characteristics, and if differences in operating characteristics are relatively small, the methodology should be robust enough to accurately characterize trends in regulatory performance over time. Moreover, as we will see below, even when we control for substantial cross sectional and inter temporal variation in the quality of water and sewerage services in England and Wales, our underlying conclusions with regard to the implied power of regulatory price caps is not affected, even though our estimates of underlying productivity catch up are substantially different.

4. Data and the Impact of Quality Adjustment

Our model includes separate outputs for water and sewerage services, and the three inputs, capital, labor and other inputs. The data covered are for the period 1991-2008 for a balanced panel of 10 Water and Sewerage companies (WaSCs). Water connected properties and sewerage connected properties are the proxies for water and sewerage output and are drawn from the companies' regulatory returns to Ofwat. Water and sewage output prices were calculated as the ratio of the appropriate turnover in nominal terms, as available in Ofwat's regulatory returns, to measured output, thereby allowing construction of binary Fisher Output indices. These binary output indices then formed the basis of constructing fully spatially consistent output indices with the EKS method. Finally, spatially consistent aggregate quality-unadjusted output price indices were constructed as the ratio of relative aggregate turnover in nominal terms to this spatial aggregate quality-unadjusted output index, as discussed above.

Our physical capital stock measure is based on the inflation adjusted Modern Equivalent Asset (MEA) estimates of the replacement cost of physical assets contained in the companies' regulatory accounts. However, as periodic revaluations of these replacement cost values could create arbitrary changes in our measure of physical capital, we cannot directly employ these accounting based measures. Instead, we accept the year ending 2006 MEA valuations as our base value, and use net investment in real terms to update this series for earlier and later years. Real net investment is therefore taken as the sum of disposals, additions, investments and depreciation, as deflated by the Construction Output Price Index (COPI). Following Ofwat's approach, we averaged the resulting year ending and year beginning estimates to provide a more accurate estimate of the average physical capital stock available to the companies in a given year.

We subsequently employ a user-cost of capital approach, to calculate total capital costs as the sum of the opportunity cost of invested capital and capital depreciation relative to the MEA asset values, and construct the price of physical capital as the user cost of capital divided by the above MEA based measure of physical capital stocks. The opportunity cost of capital is defined as the product of the weighted average cost of capital (WACC) before tax and the companies' average Regulatory Capital Value (RCV). The RCV is the financial measure of capital stock accepted by Ofwat for regulatory purposes. The WACC calculation is broadly consistent with Ofwat's regulatory assumptions and is estimated with the risk free return assumed to be the average annual yields of medium-term UK inflation indexed gilts. The risk premium for company equity and corporate debt was assumed to be 2% following Ofwat's approach at past price reviews. We also allowed for differences in company gearing ratios and effective corporate tax rates, which were calculated as the sum of aggregate current and deferred tax divided by the aggregate current cost profit before taxation. Finally, following the approach in Ofwat's regulatory current cost accounts, capital depreciation was the sum of current cost depreciation and infrastructure renewals charge.

The average number of full-time equivalent (FTE) employees is available from the companies' statutory accounts. Firm specific labour prices were calculated as the ratio of total labour costs to the average number of full-time equivalent employees. Other costs in nominal terms were defined as the difference between

operating costs and total labour costs.² Given the absence of data allowing a more refined break out of other costs, we employ the UK price index for materials and fuel purchased in purification and distribution of water, as the price index for other costs, and simply deflate nominal other costs by this measure to obtain a proxy for real usage of other inputs. Given these input quantity and price measures, we are able to calculate the spatially consistent indices of relative input usage discussed above. As total nominal economic costs are obtained as the sum of total capital costs, labour costs and other costs in nominal terms, division of this sum by the spatially consistent input index, allows construction of spatially consistent input price indices. Finally, economic profits are calculated as the difference between turnover and calculated economic costs.

We now have the necessary set of output and input quantity and price measures, as well as the necessary profit, cost, and turnover measures to proceed with our model. However, as differences in operating characteristics may result in legitimate differences in required inputs to produce a given output, variation in measured spatial productivity may result partially from these differences. We would argue that while most such characteristics, such as density of population supplied, or differences in sources of water supply, etc., have an impact on relative performance, these differences are largely stable over time, and will also have only a small impact on explaining differences between the required inputs of WaSCs. In other words, if we are primarily focussed on measuring changes in relative performance over time, the stability of these differences in characteristics as well as their relatively small impact on input requirements, will not significantly influence trends in relative productivity performance. Nevertheless, as we wish to test the impact of operating characteristics on our model results, and because past research has demonstrated that quality improvements do significantly impact temporal productivity estimates, we also adapt our model to allow for the cross sectional and intertemporal variation in the sewage and drinking water quality.

As is well documented in past studies, the water and sewerage companies have been obliged to carry substantial capital investment projects in order to improve water and sewerage quality and environmental standards. Thus, it is important to measure

² While it would be particularly desirable to disaggregate other input usage data further, and in particular to allow for separate energy and chemical usage inputs, the data available at company level from Ofwat's regulatory return does not allow a further meaningful decomposition of other input usage.

the impact of quality in our profitability, productivity and price performance measures. We therefore calculated quality-adjusted measures of output for water and sewerage services, as the product of water output and a drinking water quality index and sewerage output and sewage treatment quality index respectively.

Following Saal and Parker (2001) the drinking water quality index is calculated as the ratio of the average percentage of each WaSC's water supply zones that are fully compliant with key water quality parameters, relative to the average compliance percentage for England and Wales in 1991. Water supply zones are areas designated by the water companies by reference to a source of supply in which not more than 50,000 people reside. The data were drawn from the DWI's annual reports for drinking water quality for the years ending 1991-2007. Due to changes in some of the drinking water quality standards and the new regulations, we employed six water quality parameters³ that are also employed by Ofwat to reflect how well treatment works and distribution systems are operated and maintained (Ofwat, 2006).

The sewage treatment quality index is defined as a weighted index of the percentage of connected population for which sewage receives primary treatment and the percentage of population for which sewage receives at least secondary treatment. It also implicitly includes the percentage of connected population for which sewage is not treated with a zero weight. This data choice reflects both the availability of consistent data capturing quality trends for the entire 1991-2008 period, and does clearly capture substantial increases in sewage treatment levels, particularly in the earlier part of the sample period. The sewage treatment data were taken from *Waterfacts* for the period 1990-91 to 1995-96 and the companies' regulatory returns for the years 1996-97 to 2007-08.

It is clearly necessary to employ a weighted index of these measures as both the quality and costs of higher treatment levels exceed those associated with non treatment or primary treatment alone. We therefore endeavoured to construct a cost based weighting system, although the necessary data to accomplish this was relatively limited. However, we were able to calculate relative cost measures based on the ratio of sewerage treatment costs to volumes of sewerage treatment, using two alternative cost estimates available from company regulatory returns. One of these alternative

³ The six water quality parameters, which form the Operational Performance Index (OPI) are iron, manganese, aluminium, turbidity, faecal coliforms and trihalomethanes. The resulting drinking water quality index suggests an increase in quality of 10.3 percent between 1991 and 2008 after aggregating the data for all WaSCs.

estimates was based on total sewerage treatment functional expenditure and direct costs for all treatment works, while the other was based on total sewage treatment costs for large treatment works only. These estimates suggest that higher levels of treatment are 1.68 to 2.40 times more costly than primary treatment only. Given this estimate range, we chose to weight the percentage of population receiving secondary treatment of sewage or more twice as much as the percentage receiving primary treatment only. While admittedly, somewhat ad hoc, we emphasize there is some empirical evidence to support these weights. Moreover, we note that it is straightforward to demonstrate that the resulting weighted quality index is nested between an index based solely on the percentage of population receiving at least primary sewage treatment, which would underestimate gains in sewage treatment quality, and one based solely on the percentage of population receiving at least secondary sewage treatment, which would overestimate gains in sewage treatment quality.⁴

Once the quality adjusted water and sewerage outputs are constructed, quality adjusted indices are straightforward to produce, by simply repeating the procedures identified above to first produce spatially consistent quality adjusted output indices ($Y_{i,t}^{S,Q}$). A spatial aggregate quality-adjusted aggregated output price index is then constructed as $P_{i,t}^{S,Q} = R_{i,t}^S / Y_{i,t}^{S,Q}$. We can also derive a spatial implicit quality index ($Q_{i,t}^S$) which measures the implied difference in quality relative to the base firm as $Q_{i,t}^S = Y_{i,t}^{S,Q} / Y_{i,t}^S$. Therefore, quality adjusted spatial outputs and prices can also be respectively expressed as $Y_{i,t}^{S,Q} = Q_{i,t}^S Y_{i,t}^S$ and $P_{i,t}^{S,Q} = P_{i,t}^S / Q_{i,t}^S$, which illustrate that the impact on spatial output quantities will be perfectly balanced by changes in spatial output prices. This also implies that measured economic profitability ($\Pi_{i,t}$) is not influenced by quality adjustment. In contrast, the impact of quality adjustment implies that quality adjusted spatial TFP can be expressed as $TFP_{i,t}^{S,Q} = Q_{i,t}^S TFP_{i,t}^S$, or

⁴ To highlight this, we note that while our weighted index implies an increase in sewage treatment quality of 19.3% for all England and Wales between 1991 and 2008, an index based only on population receiving at least primary treatment would indicate a quality improvement of 13.7% while one based only on the percentage of population receiving at least secondary treatment of sewage would indicate a 25.4% quality improvement. However, our approach not only provides a mid range estimate between these two more extreme measures, but also better reflects the process of improving sewage treatment quality that occurred through both treating previously untreated sewage, and increasing the level of sewage treatment.

equivalently that the quality adjusted excess cost index can be expressed as $E_{i,t}^{R,Q} = E_{i,t}^R / Q_{i,t}^S$. Similarly, quality adjusted regulatory price performance can be expressed as $TPP_{i,t}^{R,Q} = TPP_{i,t}^R / Q_{i,t}^S$ and economic profitability can be decomposed as:

$$\Pi_{i,t} = TPP_{i,t}^{R,Q} TFP_{i,t}^{S,Q} \quad (18)$$

Thus, for example, if we assume that $Q_{i,t}^S < 1$, which implies that firm i has lower measured quality than the base firm, the quality adjusted model will result in $TFP_{i,t}^{S,Q} < TFP_{i,t}^S$, but $TPP_{i,t}^{R,Q} > TPP_{i,t}^R$. This demonstrates that without quality adjustment, $TFP_{i,t}^S$ does not reveal the full extent of the firm's excess costs due to low productivity, while $TPP_{i,t}^R$ also results in a perfectly proportional understatement of the excess of allowed revenues to benchmark costs. However, these relationships also suggest that only if $Q_{i,t}^S$ deviates significantly from 1, will there be significant differences between the results and policy implications of the quality-adjusted and quality-unadjusted models.

5. Empirical Results

Before turning to our model results, we first consider trends in aggregate WaSC turnover, costs and profits since privatization, as reported in Figure 1. As we should expect, these trends closely follow the regulatory cycle. Thus, a substantial economic loss in 1991 was rapidly eliminated and the industry became increasingly profitable until 1994, when Ofwat not only declared that it would exercise its right to review the relatively lax 10 year price caps set at privatisation after five years, but that it would also effectively rescind the price increases of firms which used their full price cap allowance in 1995 in the new five year price review that would come into effect in 1996. Even though the first price review in 1994/95 tightened regulatory price increases, economic profits remained positive despite falling from 565 to 70 million pounds between 1994 and 1998. Moreover, perhaps reflecting increased incentives to contain costs, there is a noticeable shift in cost trend between 1998 and 2000, and this led the aggregate WaSCS to achieve their highest nominal economic profitability in 2000 with profits of 680 million pounds.

The 1999 price review, which set prices for 2001-2005, marked a shift to considerably tighter regulation by Ofwat. Thus, the 10 year trend of above inflation

price increases that had been justified as necessary to fund the industry's capital investment needs, was followed by a substantial reduction in regulatory price caps in 2001, which Ofwat justified as necessary in order to pass cost savings to consumers. This shift in regulation is evidenced by the fall in aggregate turnover from 6,279 to 5,815 million pounds between 2000 and 2001, which is the only example of a nominal decline in aggregate WaSC revenue during the entire 1991-2008 period. When coupled, with the substantial increase in aggregate economic costs in 2001, which can be substantially explained by RPI inflation effects which particularly effect the estimated normal rate of return on invested capital between 1999 and 2002, this resulted in aggregate WaSC losses to 502 million pounds in 2001.

These economics losses do not only reflect a momentary blip in turnover or estimated economic costs in 2001, but rather suggest the institution of a consistently tighter regulatory regime. Thus, despite nominal turnover increasing in every year after 2001, aggregate economic cost increases outstripped allowed revenue increases until 2005 when aggregate economic losses had fallen to 544 million pounds. Moreover, although the implementation of the 2004 price review in 2006 appears to have allowed for a momentarily closer link between regulated revenues and costs, thereby reducing aggregate economic losses to 43 million pounds, subsequent revenue increases have been by far outstripped by increases in economic costs, and by 2008 economic losses again increased to 568 million pounds. Thus, even a straightforward analysis of aggregate WaSC economic profits suggests a shift from a regulatory policy that tolerated above normal returns for the entire period before 2000, to one which set prices resulting in below normal returns after 2000.

Given these general trends, we begin the presentation of our model results with Figure 2, which depicts the geometric average, as well as the range of WaSC profitability over the sample period. We also remind the reader that reported economic profitability reflects actual firm profitability based on firm specific economic costs and revenues, and is in no way influenced by spatial comparisons. These firm specific indices largely confirm the conclusions drawn from Figure 1. Thus, profitability peaked in 1994 when the average company made a profit of 12.9%, the most profitable firm made a profit of 30.2%, and the least profitable firm made an economic profit of 0.04%. Regulatory tightening does appear to have shifted this range downward by 1998, but without substantially tightening the range of observed profitability given that the maximum, average, and minimum profitability respectively

fell to 1.152, 1.016, and 0.913. However, by 2000 economic profitability reached its highest observed levels as the average company made a profit of 13.7% above the normal rate of return, the highest observed economic profit was 30.7%, and the lowest economic profit was 1.5%. While these high profits in 2000 at least partially reflect the observed reduction in total economic costs between 1998 and 2000, we would emphasize the continued wide range in observed profitability before 2000, as well as the lack of a significant number of firms that made economic losses after 1991. Moreover, this high profitability is indicative of what in hindsight were inappropriate regulated prices, as prices appear to have had little relationship to the actual economic costs of firms, let alone benchmark economic costs.

Figure 2 highlights the dramatic shift in regulatory practice implemented in the 1999 price review. Thus, in 2001, not only did average economic profitability fall to a loss of 7.4%, but the range of observed economic profitability tightened substantially as the highest observed profit was 3.6%, and the most extreme loss was 11.3%. This substantially reduced range in estimated economic profitability, which is sustained in every year after 2000, suggests that Ofwat more closely aligned regulated revenues with actual firm costs after the 1999 price review, and particularly in 2006 which was the first year of the current price review period. Moreover, the consistent economic losses realized by many firms during this period also suggest that Ofwat had begun to deliberately set revenues below actual economic costs, so as to better incentivize firms to reduce their excessive regulatory costs.

Before considering our spatial estimates of quality unadjusted ($E_{i,t}^R$) and adjusted ($E_{i,t}^{R,Q}$) regulatory excess costs⁵, which are respectively reported in Figure 4 and Figure 5, we first consider the average and range of the spatial implicit quality index ($Q_{i,t}^S$) over the sample period so that we can highlight the significant role of

⁵ We have not identified firms for confidentially reasons. The same firm is consistently found to have the highest spatial productivity estimates for both quality unadjusted and quality adjusted models in all years, and is therefore modelled as the benchmark most productive firm in each year of our study. Moreover, we note that this same firm was found to have the highest spatial productivity estimates in each year of the study regardless of whether we applied the spatially consistent Fisher indices provided in the main text, similar spatially consistent Tornqvist indices, or the multilateral translog index for WaSCs based on the Tornqvist index developed by Caves et al (1982a). Furthermore, there is little substantive difference between the results regardless of which method is employed.

quality in our results.⁶ Figure 3 demonstrates that over the entire sample period, the once wide divergence in the spatial implicit quality index was eliminated, as the companies improved their drinking and sewerage treatment quality conditions relative to the most productive company. Thus, in 1991, the average and worst performing company's implicit quality index were respectively only 83.7% and 71.5% of the base firm's measured quality, while the highest observed quality exceeded the most productive firm's measured quality by 3.9%. In contrast, in 2008, the average and worst company's quality index were respectively 97.1% and 92.6% relative to the most productive company. Moreover, it is worthwhile to note that despite significant investment in drinking water and sewage quality improvement throughout the sample period, little to no convergence in the average and minimum relative quality index occurred before 1998, while most of this convergence occurred between 1998 and 2003. This is likely to reflect what were in fact considerable lags between the provision of revenues necessary to fund quality improving capital investments, and the actual date when the resulting quality improving assets became operational.

As, $E_{i,t}^{R,Q} = E_{i,t}^R / Q_{i,t}^S$, and since $Q_{i,t}^S < 1$ for almost all the observations in our sample, the geometric average of quality unadjusted excess costs ($E_{i,t}^R$) reported in Figure 4 is always lower than the geometric average of quality adjusted excess costs ($E_{i,t}^{R,Q}$) reported in Figure 5, and the proportional difference is equal to the geometric average of $Q_{i,t}^S$ reported in Figure 3. Moreover, the strong convergence in $Q_{i,t}^S$ documented in Figure 3, explains the considerably lower convergence of unadjusted excess costs over the sample period when compared to the convergence of quality adjusted excess costs. Thus, between 1991 and 2008 average quality unadjusted excess costs only declined from 1.27 to 1.205, thereby suggesting that on average productivity catch up by laggard firms contributed only a 5.39% reduction in WaSC costs. In contrast, average quality adjusted excess costs decreased from 1.518 to 1.242, thereby suggesting a much more considerable 18.18% reduction in average costs attributable to productivity catch up by laggard firms. The latter estimate, which is broadly consistent with estimates of cost savings attributable to eliminating

⁶ The maximum of the spatial implicit quality index is marginally above 1 in almost yeas of the sample. This reflects the fact that while the base firm is chosen based on its superior quality unadjusted and quality adjusted spatial productivity estimates, its spatial implicit quality index is marginally inferior to at least one firm in the sample.

efficiency made by Ofwat, demonstrates the empirical necessity of controlling for quality over the 1991-2008 period. However, the lack of significant quality differences after 2003, when the average of $Q_{i,t}^S$ always exceeds 0.97, suggest that results in this latter period will not be significantly affected by the quality adjustment method employed in this study. Stated differently, this implies that quality adjustment is necessary if we wish to consider long term trends in the industry, but has very little influence on estimates of excess costs in recent years, which is important if we consider it is precisely these latter estimates that are most relevant for the forthcoming 2009 price review.

Given this general discussion of the excess cost estimates, we now discuss their implication when set in context of the regulatory history of the English and Welsh water industry. Despite the wide divergence between estimated quality unadjusted and adjusted excess costs in the early years of the sample, there is nonetheless a remarkable convergence with regard to evidence demonstrating the failure of the regulatory system to eliminate of excess costs during this period. Thus, average quality unadjusted excess costs increased from 27.0% to 27.7% between 1991 and 1995 while average quality adjusted excess costs increased from 51.8% to 53.6%. Even more strikingly, the worst laggard firms saw their quality unadjusted excess costs increase from 47.3% to 59.2% over the same period, while their estimated quality adjusted excess costs increased from 102.1% to 123.3%. Thus, there is clear evidence that during the WaSCs first five years under price cap regulation, little to no improvement in relative productivity/cost performance occurred.

As there is a general consensus that during the 1990s Ofwat's price capping policies was strongest between 1994 and 1997 it is interesting to note that both the quality unadjusted and adjusted excess cost indices fall to a temporary low in 1997, when average quality unadjusted excess costs fell to 20.9% and average quality adjusted excess costs fell to 45.0% of benchmark costs. However, this reduction in excess costs was not sustained in the quality unadjusted model, and its decline was temporarily halted in the quality adjusted model. As a result, if we focus on changes over the formal five year price cap period covering 1996 to 2000, there is a considerable difference in the implications of the unadjusted and quality adjusted results. Thus, the quality unadjusted results reported in Figure 4, provide limited evidence of sustained convergence in average excess costs, which fall from 27.7% to

25.0%, while the quality adjusted results suggest a substantial fall in average excess costs from 53.6% to 41.7%. Moreover, while quality unadjusted results suggest no sustained improvement in the excess costs of the worst laggard firm, which actually saw its excess costs increase from 59.2% to 61.1%, the quality adjusted results suggest a considerable improvement in laggard firm performance, given that the maximum quality adjusted excess cost estimate fell from 123.3% to 69.2%. We would suggest that this difference is indicative of the need to control for substantial differences in quality, as well as convergence in quality, if one wishes to properly measure spatial differences in productivity/excess costs. Nevertheless, given that even the quality adjusted excess cost index does not show sustained improvement before 1998, these results may also suggest that the tightening of regulation also acted to reduce the lag between provision of revenues for quality enhancement programmes and their delivery.

Our results do suggest consistent trends with regard to regulatory excess costs for the five year period covered by the 1999 price review, even if the quality unadjusted results, show a much more dampened reduction in excess costs. Thus, between 2000 and 2005, on average the quality unadjusted excess cost index fell from 1.250 to 1.216, while the quality adjusted index fell from 1.417 to 1.243. Both measures also show relatively large average and laggard firm excess cost reductions in 2001, although it must be noted that the magnitude of the quality adjusted excess cost reduction is influenced by the largest observed annual quality increase for both the average and the lowest observed spatial implicit quality index. It is also notable that both indexes also suggest a considerable improvement in regulatory excess costs in 2005. The excess cost results therefore suggest that firms' cost reducing efforts were concentrated in 2001 in a clear response to the large reduction in maximum allowed prices in the first year of the price period, and in the last year of the period, which may suggest they were working to reduce costs to improve their position for the 2006-10 price determination. Moreover, particularly if we focus on the worst laggard firms as represented by the maximum observed excess cost estimates, which declined from 1.611 to 1.469 in the unadjusted model and from 1.692 to 1.490 in the quality adjusted model, there is fairly clear evidence that the tightening of price caps in the 1999 review led to sustained improvements in the relative productivity performance of laggard firms.

The immediate impact of the 2004 price review, which allowed for an initial increase in operating costs in 2006, has already been observed in Figure 1 and Figure 2, where there is a clear shift from substantial economic losses in 2005 to a near perfect alignment between revenues and estimated economic costs in 2006. This may or may not be justified on the grounds that Ofwat has a duty to maintain the financial viability of firms in addition to its duty to promote efficiency. Nevertheless, our excess cost results suggest that the realignment of regulated revenues with actual costs led to an immediate increase in excess costs relative to benchmark costs in 2006, an increase which is consistent with the reduction in regulatory incentives to reduce costs in 2006. However, as price caps in years subsequent to 2006 reverted to allowing below inflation price increases, it would appear that the industry not only reverted to economic losses, but also began to improve excess costs relative to the base firm. Nevertheless, while geometric average quality adjusted excess costs fell from 27.8% to 24.2% between 2006 and 2008, as firms again worked to improve productivity relative to benchmark levels, the negative impact of the momentary loosening of price caps in 2006 is demonstrated by the fact that average quality adjusted excess costs in 2008 were only 0.1% lower than they were in 2005.

Given this discussion of regulatory excess costs, which measure the excess of actual costs relative to benchmark costs, Figures 6 and 7 now focus our attention on regulatory TPP, which measures the excess of regulated revenues to benchmark costs, thereby allowing a direct estimate of the tightness of regulatory price caps. From 1991 to 1994, both the unadjusted and quality adjusted results quantify what amounted to a significant loosening in regulatory price caps, as the average excess of regulated revenues to benchmark costs respectively increased from 16% to 44.8% and from 38.7% to 74.2%. In contrast, after 1995 the average values of $TPP_{i,t}^R$ and $TPP_{i,t}^{R,Q}$ both suggest a considerable tightening of price caps that persisted until 1998 when they respectively indicate that regulated revenues exceeded benchmark costs by 26.6% and 50.6%. However, during the last two years of the 1996-2000 price cap period, average regulatory TPP again increased, thereby suggesting that price caps had effectively become looser again.

The increased and sustained regulatory demands of the 1999 price review are clearly illustrated in the dramatic fall in estimated regulatory TPP between 2000 and 2001. Thus, in a single year, the average excess of regulated revenues over benchmark

costs respectively decreased from 42.2% to 13.5% and from 61.2% to 20.4% for the unadjusted and quality adjusted models. Moreover, the wide dispersion in regulatory TPP, which suggests a more accommodative policy for laggard firms up to 2000 also came to a sudden end, as the range of allowed excess revenues relative to benchmark costs tightened, and in particular, regulatory TPP for the worst performing firms was reduced more than for other firms. If we focus on $TPP_{i,t}^{R,Q}$, as illustrated in Figure 7 this tightening in the range of regulatory rigour is illustrated by the decline of 51.3% in the maximum value of $TPP_{i,t}^{R,Q}$ while the minimum value only declined by 23.0%.

While our regulatory TPP estimates largely suggest continuity in regulatory policy for the post 2000 period, they do provide evidence for a small loosening of price caps in 2006 followed by a return to tougher price caps. Thus between 2005 and 2006 average $TPP_{i,t}^R$ and $TPP_{i,t}^{R,Q}$ respectively increased from 1.123 to 1.245 and from 1.149 to 1.269, thereby suggesting an average increase of 12% in the excess of allowed revenues to benchmark costs. However, by 2008 average $TPP_{i,t}^R$ and $TPP_{i,t}^{R,Q}$ had respectively fallen to 1.131 and 1.165, thereby demonstrating the return to price caps that were of broadly equivalent tightness to those that had been in place in 2005.

In order to clearly illustrate our underlying model of regulatory price caps, Figures 8 and 9 respectively report the decomposition of average economic profitability into regulatory TPP and regulatory excess costs indices for the unadjusted and quality adjusted models. As both models suggest the same conclusions with regard to changes in regulatory policy over the sample period, we have chosen to focus on the quality adjusted models in the interest of brevity.

As average $TPP_{i,t}^{R,Q}$ substantially exceeds one in all sample years there is no evidence that the WaSCs have ever been subject to a “powerful” price cap regime requiring immediate full catch up to benchmark costs to regain economic profitability. Moreover, the trend in average $TPP_{i,t}^{R,Q}$ suggests that price caps became progressively looser until 1994 when on average regulated revenues exceeded benchmark costs by 74.2 percent. As the quality adjusted excess cost index ($E_{i,t}^{R,Q}$) suggests that, on average, actual costs only exceeded benchmark costs by 54.2 percent in 1994, the lack of progress in reducing $E_{i,t}^{R,Q}$ before 1995, is fully consistent with the weak incentives created by regulatory price caps that allowed for *increased* economic profitability

even in the absence of any effort to improve productivity. Subsequent declines in the average value of $TPP_{i,t}^{R,Q}$ to 1.506 in 1998 demonstrate a substantial reduction in allowed revenues that is indicative of tighter regulation, as firms would at least be required to improve productivity in order to maintain their existing level of economic profitability in the future. However, we emphasize again that even in 1998 average $TPP_{i,t}^{R,Q}$ still exceeded average $E_{i,t}^{R,Q}$, which was 1.483. Moreover, by 2000 average $TPP_{i,t}^{R,Q}$ had been allowed to increase to 1.612 while at the same time $E_{i,t}^{R,Q}$ declined to 1.417. Thus, despite some improvements in incentives after 1994, 1991-2000 can still be characterized as a period of “weak” regulation because $TPP_{i,t}^{R,Q}$ was generally allowed to exceed $E_{i,t}^{R,Q}$, thereby allowing the retention of above normal returns even in the absence of any effort to achieve benchmark productivity levels.

As our $TPP_{i,t}^{R,Q}$ measure clearly indicates that existing price caps in 2000 were highly inappropriate and allowed firms to achieve average economic profitability of 13.7%, the dramatic tightening of price caps that took place in 2001 was more than justified. Nevertheless, as the average and almost all firm specific estimates of $TPP_{i,t}^{R,Q}$ remained in excess of one even after 2001, there is no evidence that Ofwat moved to a “powerful” price cap regime as defined in Section 2. Instead, as clearly illustrated in Figure 9, after 2001, Ofwat effectively moved from a policy of setting price caps which allowed regulated revenues to exceed regulatory excess costs ($TPP_{i,t}^{R,Q} > E_{i,t}^{R,Q}$) to a policy of setting regulated revenues below regulatory excess costs ($TPP_{i,t}^{R,Q} < E_{i,t}^{R,Q}$). Therefore, Ofwat’s dramatic regulatory tightening in 2001 amounted to a move from “weak” price caps that allowed economic profits even without productivity catch up to “catch up promoting” price caps that required the elimination of at least some excess costs in order to regain economic profitability.

While we would emphasize the overall persistence of this new policy for the entire post 2000 period, the near perfect alignment on average between $TPP_{i,t}^{R,Q}$ (1.269) and $E_{i,t}^{R,Q}$ (1.278) in 2006, followed by resumption of a price caps resulting in $TPP_{i,t}^{R,Q} < E_{i,t}^{R,Q}$ in subsequent years suggests that Ofwat refined its price capping model in the 2004 price review. Specifically, our results suggest that Ofwat’s price caps effectively allowed companies their actual economic costs in 2006, but required catch

up productivity improvements in subsequent years. This suggests the effective use of a hybrid regulatory model, where companies' costs in the base year were accepted (rate of return?), but productivity improvements were expected in later years (benchmarking?). However, as the average of $E_{i,t}^{R,Q}$ increased markedly from 1.243 to 1.275 between 2005 and 2006, we would suggest that this hybrid system was inappropriate because the temporary dampening of productivity enhancing incentives led to firms falling further behind the benchmark firm in 2006. Thus, we would argue that a sustained policy of strictly "catch up promoting" price caps may have resulted in regulatory excess costs falling significantly below their 2005 levels by 2008. Instead, Ofwat's apparent temporary dampening of incentives in 2006, would appear to have effectively eliminated 3 years of continued reduction in regulatory excess costs, as illustrated by the average $E_{i,t}^{R,Q}$ of 1.242 in 2008, which was virtually unchanged from its level in 2005.

6. Summary and Conclusions

This study analyzed the impact of regulation on the financial performance of water and sewerage companies in England and Wales over the period 1991-2008. We developed a cross sectional index number technique to decompose actual economic profitability into a spatial productivity and a regulatory price performance index and also measure the spatial implicit impact of quality. The inverse of the spatial productivity index is equivalent to a regulatory excess costs index, which denotes the excess of a firm's actual costs relative to benchmark costs, whereas the regulatory price performance index measures the excess of revenues above benchmark costs. We then demonstrated that increases (decreases) in regulatory price performance are indicative of the loosening (tightening) of price cap regulation. Moreover, we also showed that the relationship between actual economic profitability, regulatory excess costs and regulatory price performance indices can be used to categorize regulatory price caps as “weak”, “powerful” or “catch-up promoting”

The results indicate that throughout the entire 1991-2008 period price caps were never “powerful”, in the sense that they required less productive firms to immediately and fully catch-up to the most productive firm to regain economic profitability. As regulatory TPP increased markedly until 1994, we are able to quantify the extent to which price caps became laxer in the early post privatisation period, and how this offered firms the potential to increase their economic profitability without making any effort to reduce their regulatory excess costs. In contrast, between 1994 and 1998, a substantial reduction in regulatory TPP occurred, thereby quantifying the extent of regulatory tightening after the 1994 price review, as falling regulatory TPP implies that laggard firms must reduce their regulatory excess costs, or would otherwise face a reduction in economic profitability. However, our results suggest a renewed increase in regulatory price performance between 1998 and 2000, suggesting that regulatory incentives once again weakened during this period, and economic profitability reached its peak in 2000. In sum, while our results do suggest substantial regulatory tightening after 1994, we would emphasize that the period 1991-2000 can be characterised as a period of “weak” regulation since allowed regulatory revenues almost always exceeded regulatory excess costs, thereby demonstrating that price caps during this period allowed firms to maintain economic

profitability regardless of whether they made any progress in catching up to benchmark productivity levels.

Our methodology performs particularly well in demonstrating and quantifying the dramatic tightening of Ofwat's regulatory policies in the 1999 price review. Thus, a sharp tightening in regulation in 2001 is quantified as a substantial fall in the ratio of allowed regulatory revenues relative to benchmark costs, as measured by regulatory TPP. Moreover, we also clearly demonstrate that Ofwat's dramatic regulatory tightening in 2001 amounted to a move from "weak" price caps that allowed economic profits even without productivity catch up to "catch up promoting" price caps that required the elimination of at least some excess costs in order to regain economic profitability. Furthermore, while our regulatory TPP index clearly demonstrates a momentary but substantial reduction in regulatory incentives in 2006, which was the first year of the current price review, it also demonstrates a return to tighter regulation in subsequent years. Thus, our results suggest that since 2001 Ofwat has implemented "catch up promoting" price caps since average regulated revenues were always below average regulatory excess costs indicating that the firms were required to eliminate at least some excess costs in order to regain economic profitability. We would also emphasize that as our results also clearly demonstrate a much closer alignment between allowed revenues and benchmark costs after 2001, Ofwat's approach during this period was not only appropriate, but should also be continued in the 2009 price review. .

We finally emphasize that our methodological approach is generally applicable. This is because it allows regulators to assess relative performance in cases where the number of observations is extremely limited, thereby directly providing firm specific evidence of potential productivity catch up as measured by deviation from benchmark productivity levels, as well as evidence of the deviation of regulated revenues from those that would be consistent with benchmark costs. Moreover it also facilitates a backward-looking approach that allows conclusions to be drawn with regard to the effectiveness of price cap regulation. More specifically, using our methodology, regulators and policy makers can determine if past regulatory decisions have not only promoted productive efficiency by providing appropriate efficiency incentives to firms, but also whether they have led to increased allocative efficiency by aligning consumer prices more closely with efficient costs.

Despite the considerable benefits of the methodology employed in this paper, we are nonetheless aware that it only allows for the cross-sectional (spatial) measurement of productivity, regulatory price performance, and profitability; and the subsequent comparison of how these cross sectional measures have changed over time. We therefore plan to extend our approach, by following Hill (2004), which suggests a framework for consistently integrating cross sectional and temporal index numbers. As this will allow the simultaneous measurement of firm specific productivity growth, as well as the spatial relative productivity measures employed in this paper, this extension of the existing methodology would facilitate a forward looking approach that can provide evidence not only with regard to the potential productivity catch up of laggard firms, but also the potential for further improvements in benchmark productivity levels. We therefore emphasize that such an approach would further aid regulators wishing to determine appropriate X-factors for regulated firms, as it would not only provide evidence for potential productivity catch-up, as in the current approach, but would also provide evidence for further potential productivity improvements by benchmark firms.

Figure 1 - Aggregate WaSC Turnover, Costs and Profits: Millions of Pounds

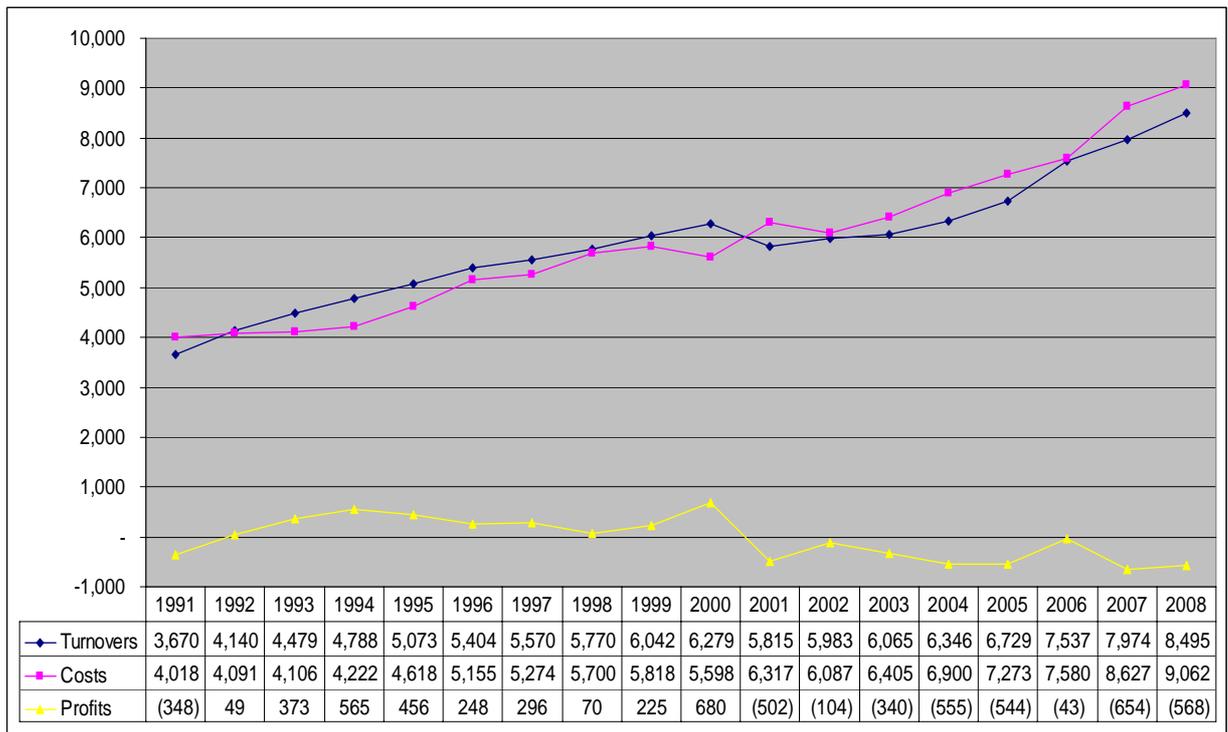


Figure 2 - Economic Profitability: Firm Specific Estimates

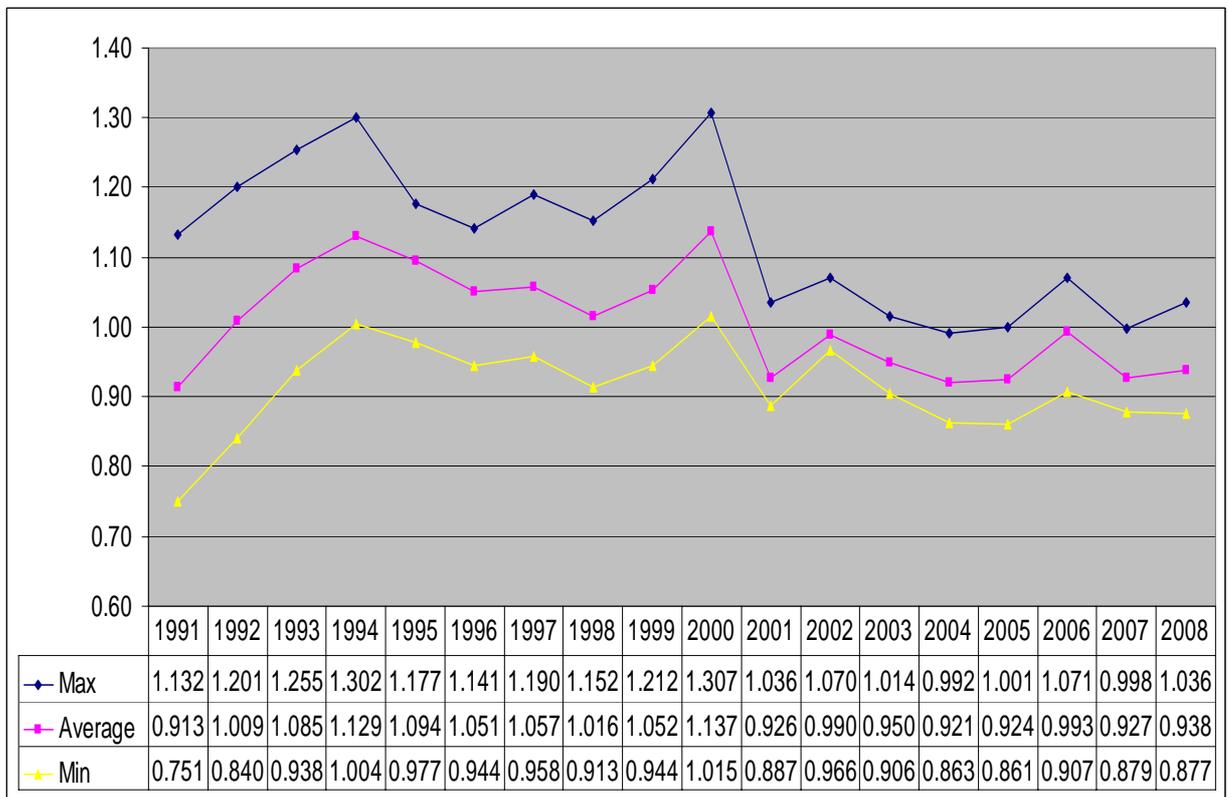


Figure 3 - Spatial Implicit Quality Index: Firm Specific Estimates

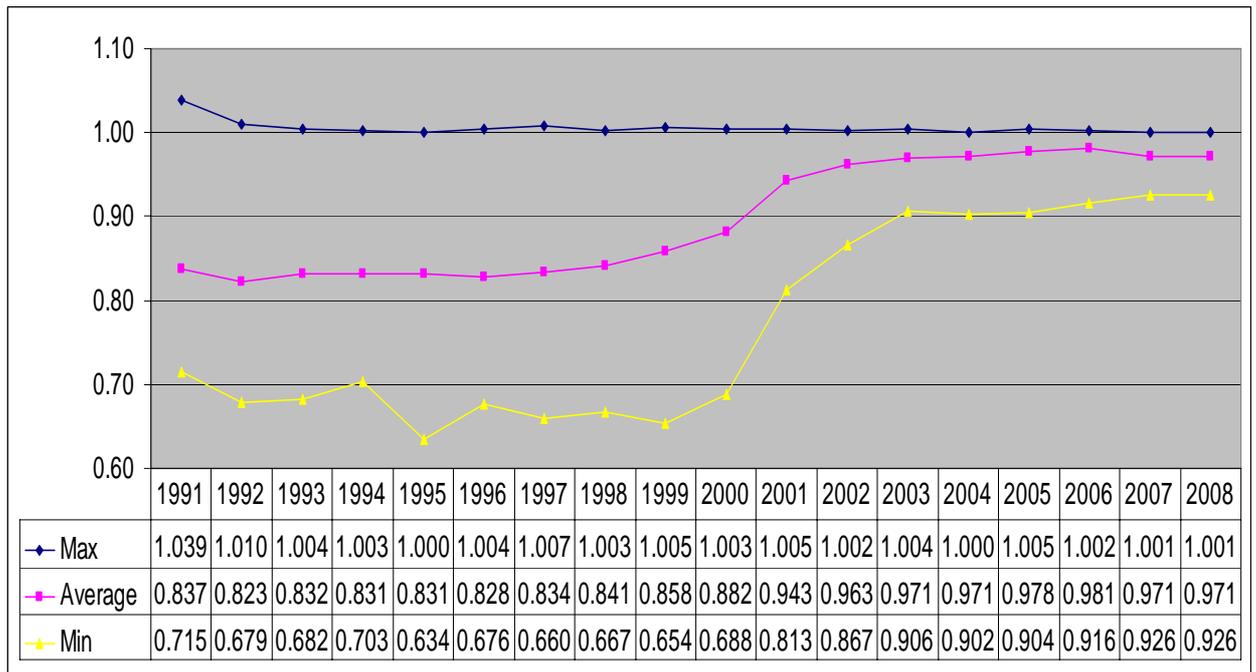


Figure 4 – Regulatory Excess Costs Quality Unadjusted: Firm Specific Estimates

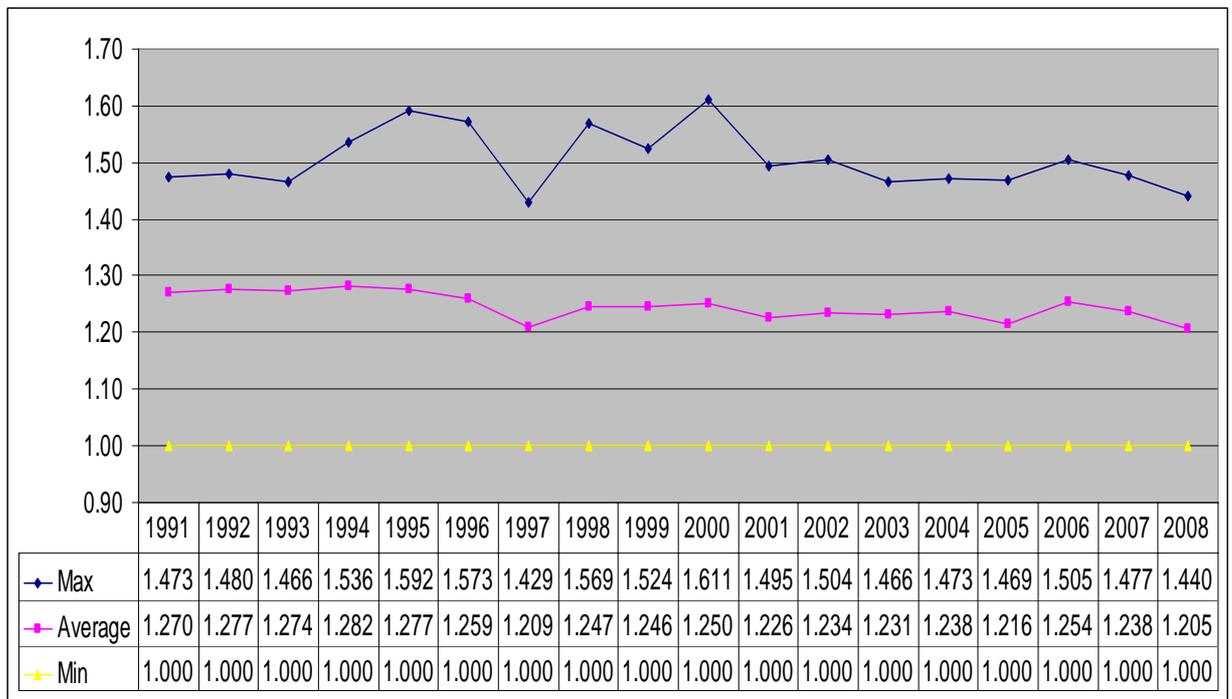


Figure 5 – Regulatory Excess Costs Quality Adjusted: Firm Specific Estimates

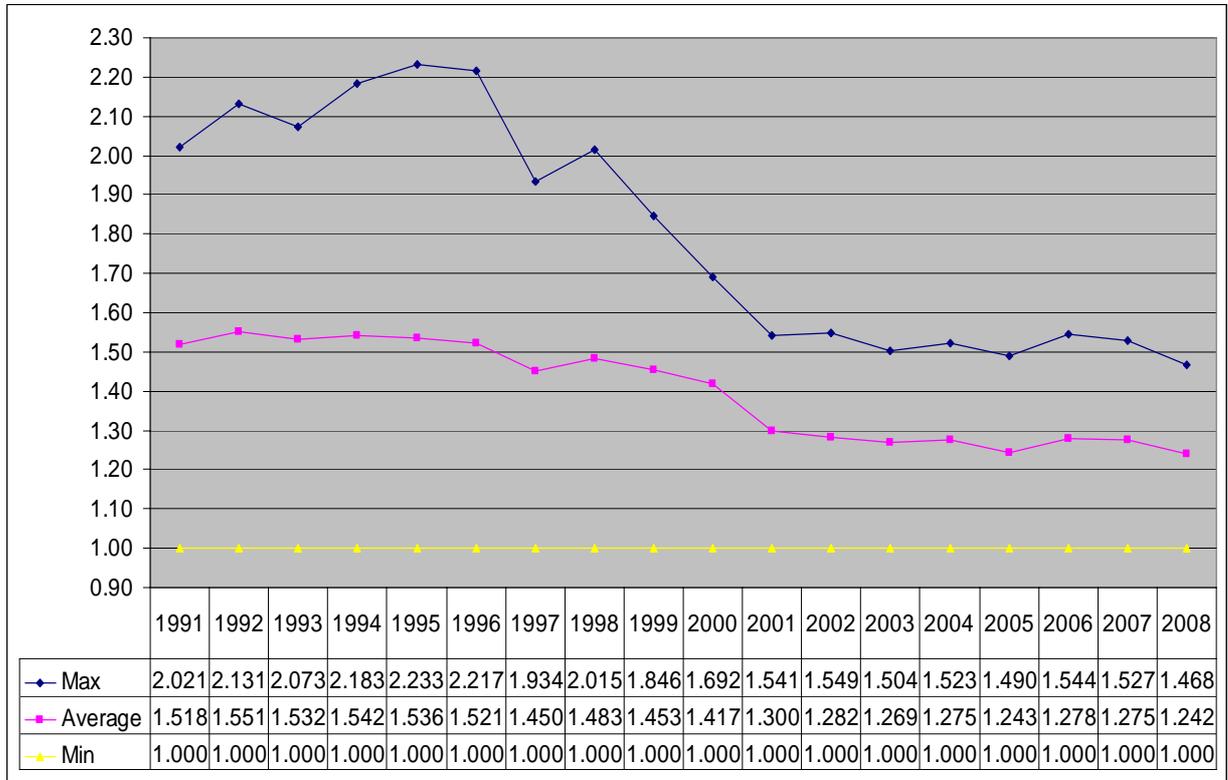


Figure 6 - Regulatory TPP Quality Unadjusted: Firm Specific Estimates

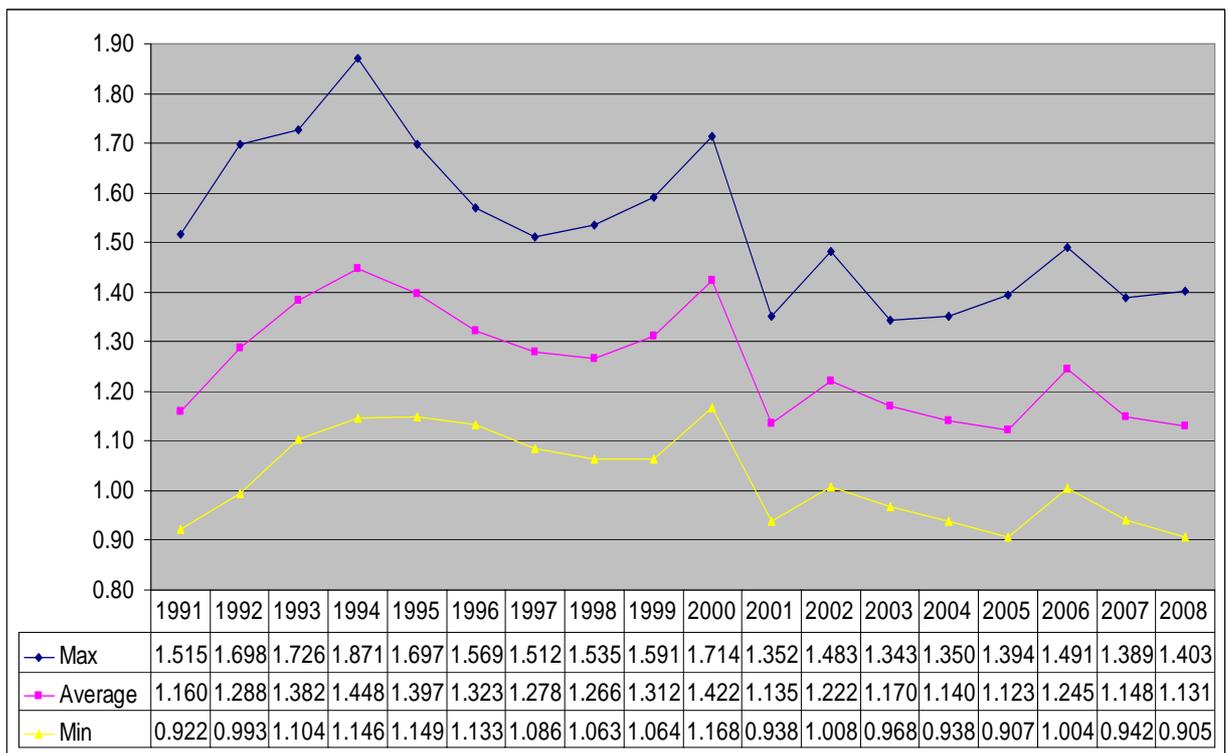


Figure 7 - Regulatory TPP Quality Adjusted: Firm Specific Estimates

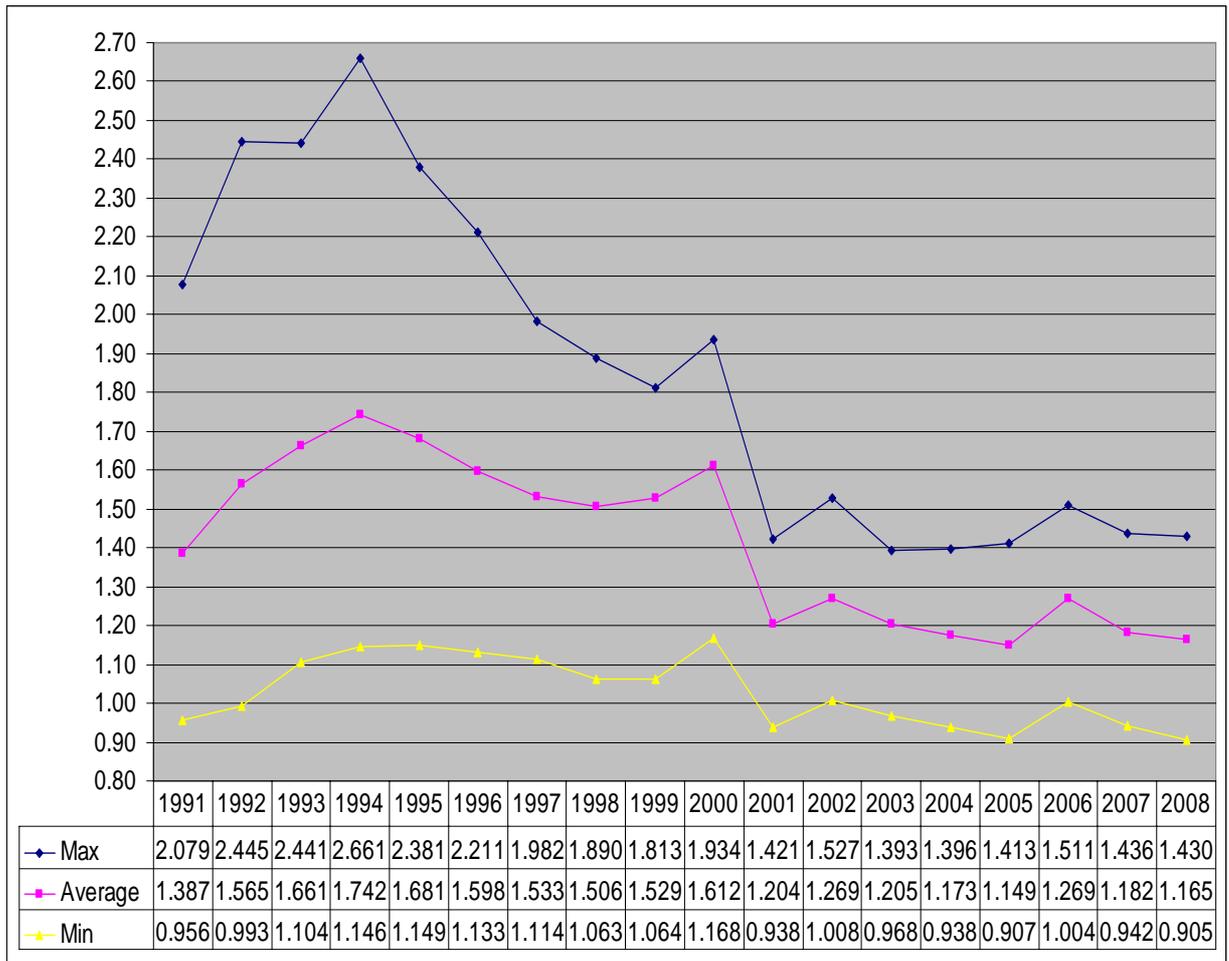


Figure 8 - Quality Unadjusted Economic Profitability, Regulatory TPP and Excess Costs: Geometric Average of Firm Specific Estimates

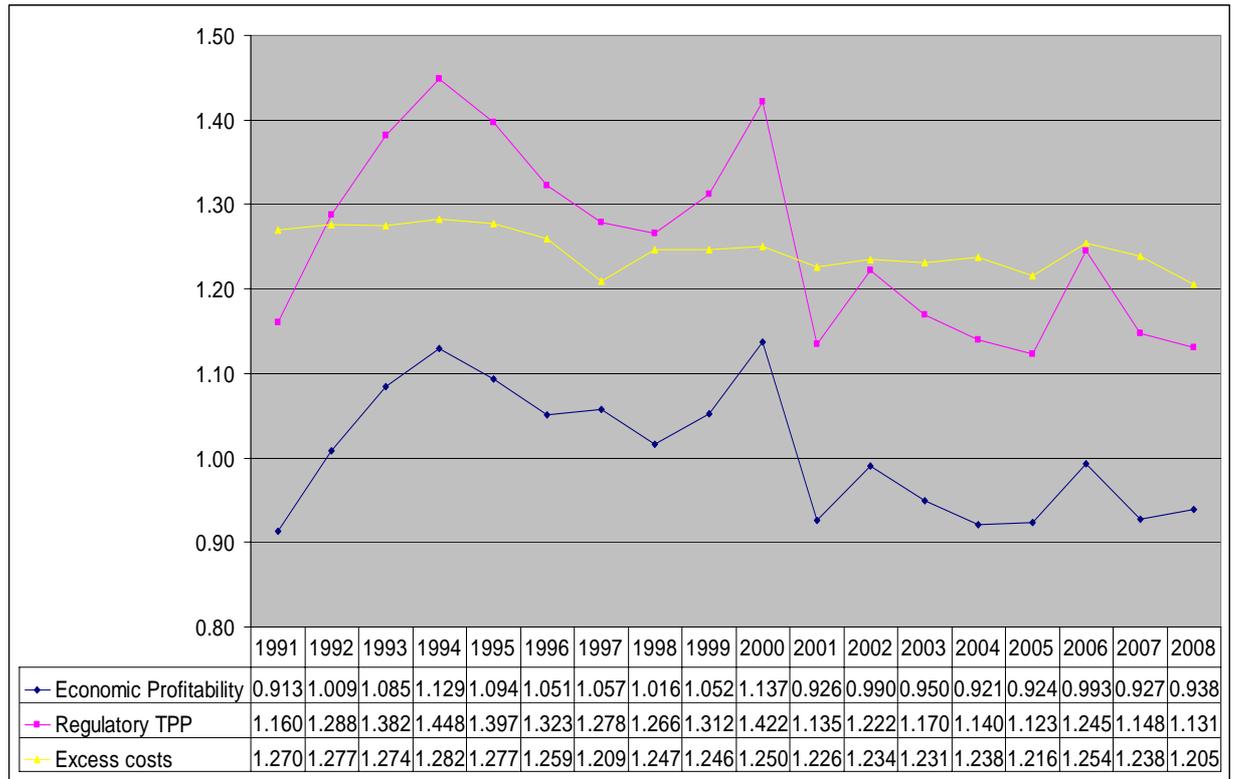
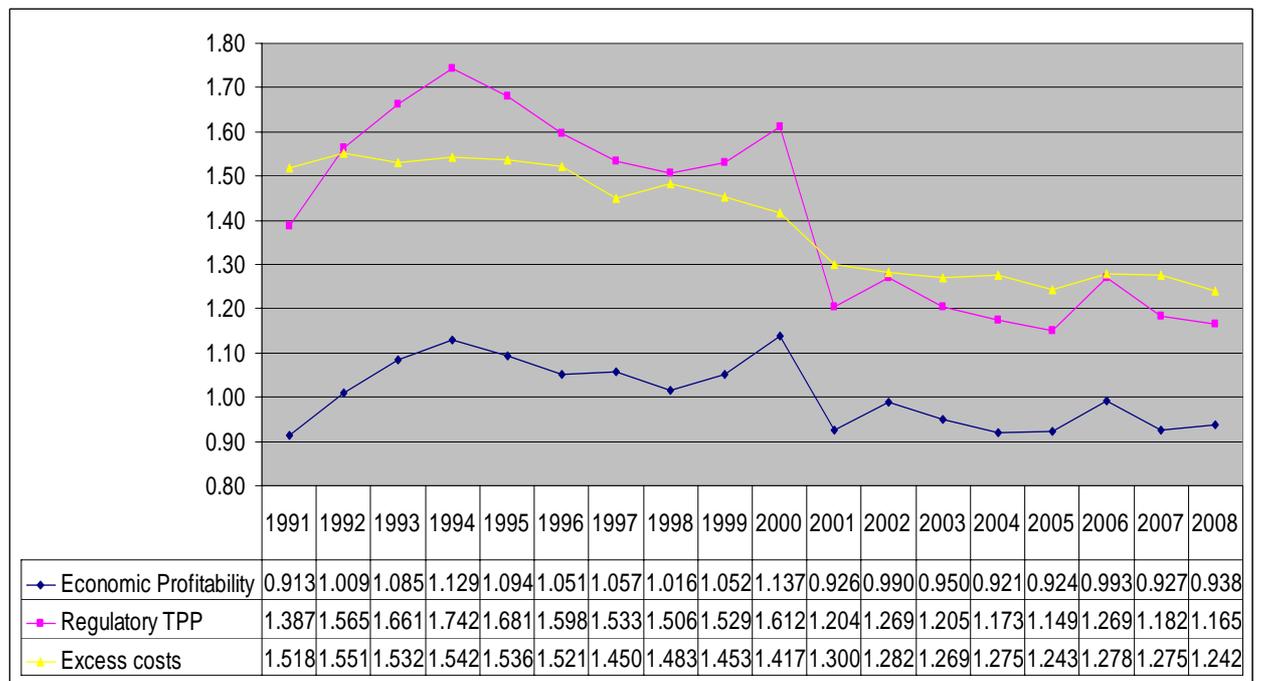


Figure 9 - Quality Adjusted Economic Profitability, Regulatory TPP and Excess Costs: Geometric Average of Firm Specific Estimates



References

- Averch H., Johnson L.L., “Behavior of the firm under regulatory constraint”, *The American Economic Review*, 52 (5), 1052-1069 (1962)
- Ball V.E., Butault J-P and Nehring R., “U.S. agriculture, 1960-96, a multilateral comparison of total factor productivity”, Electric Report from the Economic Research Service (2001)
- Caves D., Christensen L.R. and Diewet E.W., “Multilateral comparisons of output, input and productivity using superlative index numbers”, *The Economic Journal*, 92 (365), 73-86 (1982a)
- Coelli T., Rao P.D.S. O’Donnell C.J. and Battese G.E., *An introduction to efficiency and productivity analysis*, Springer (2005)
- Diewert E.W. and Lawrence D., “Regulating electricity networks: The ABC of setting X in New Zealand”, In *Performance measurement and regulation of network utilities* (ed. T. Ceolli and D. Lawrence), Edward Elgar (2006)
- Elteto O. and Koves P., “On a problem of index number computation relating to international comparisons”, *Statisztikai Szemle*, 42, 507-518 (1964)
- Griffell-Tatje E. and Lovell C.A.K., “Profits & Productivity”, *Management Science*, 45 (9), 1177-1193 (1999)
- Han, S-H and Hughes A.D., “Profit Composition Analysis: A technique for linking productivity measurement & financial performance”, New South Wales Treasury Working Paper, TRP 99-5 (1999)
- Hill R.J., “Constructing price indexes across space and time: The case of the European Union”, *American Economic Review*, 94 (5), 1379-1410 (2004)
- Ofwat, “Future charges for water and sewerage services” Birmingham: Office of Water Services (1994)
- Ofwat, “Future water and sewerage charges 2000-2005; Final determinations” Birmingham: Office of Water Services (1999)
- Ofwat, “Future water and sewerage charges 2005-2010; Final determinations” Birmingham: Office of Water Services (2004)
- Ofwat, *July Returns for the Water Industry in England and Wales*. CD-ROM. Birmingham: Office of Water Services (2006)

Saal D. and Parker D., "Productivity and Price Performance in the Privatized Water and Sewerage Companies in England and Wales", *Journal of Regulatory Economics* 20:1, 61-90 (2001)

Salerian J., "Analysing the performance of firms using a decomposable ideal index number to link profit, prices and productivity", *The Australian Economic Review*, vol 36, no.2, pp.143-55 (2003)

Szulc (Schultz) B.J., "Indices for multiregional comparisons", *Przegląd Statystyczny (Statistical Review)*, 3, 239-254 (1964)

Waters II, W.G. and Tretheway M.W., "Comparing Total Factor Productivity and Price Performance: Concepts and Application to the Canadian Railways", *Journal of Transport Economics and Policy*, 33 (2), 209-220 (1999)