THE USE OF DATA ENVELOPMENT ANALYSIS IN THE REGULATION OF UK WATER UTILITIES: WATER DISTRIBUTION$

By

Emmanuel Thanassoulis,

Aston Business School,

University of Aston,

Birmingham B4 7ET, UK.

E-mail: E.THANASSOULIS@ASTON.AC.UK.

Abstract

Regulation is increasingly playing a major role in defence of the public interest in the UK and other economies, in the aftermath of the privatisation of utilities operating in near monopoly environments. This paper gives an account of the use of Data Envelopment Analysis (DEA) by the regulator of water companies in England and Wales in 1994 in the context of setting price limits. DEA is a general purpose linear programming - based method for assessing the productive efficiencies of operating units such as bank branches or schools. The paper details the use of DEA to estimate potential savings in the specific context of water distribution and discusses the use of the results obtained. It also highlights certain generic issues arising in the use of DEA and more generally performance measurement methods in the regulatory context.

Keywords: Data Envelopment Analysis, Performance Measurement, Regulation

---

$ The author is grateful to the regulator of UK water companies (OFWAT) for permission to give this account of analyses undertaken on its behalf. The views expressed in this document are those of the author and no representation is being made that they are necessarily shared by OFWAT. Any shortcomings in the paper are entirely the responsibility of the author.

1. INTRODUCTION

This paper describes an application of Data Envelopment Analysis (DEA) in the context of the regulation of utilities. The application was commissioned by the regulator of water companies in England and Wales during its price review of 1994.

DEA is a linear programming - based method for assessing the productive efficiencies of homogenous units such as bank branches, schools and hospitals which perform a given function. Each unit represents an observed correspondence of input - output levels. DEA uses a linear programming approach to ascertain the potential for input reduction at a unit given its output levels, or alternatively the potential for output augmentation given its input levels. DEA models were first developed by Charnes et al. (1978). This paper assumes only an introductory familiarity with DEA. It avoids in large measure appeal to technical terms, focusing instead on the interpretation of the results, the insights gained and issues arising within the regulatory context of the application. Fuller introductions to DEA can be found in Land (1991) and in Boussofiane et al. (1991).

Regulation of utilities has become increasingly important in the UK economy. Since the early 1980’s there has been a massive programme of privatisations of publicly owned assets in the UK, partly based on the belief that asset ownership has an impact on operating efficiency. This trend has continued world-wide. Megginson et al. (1994) report that in ten years from 1980 nearly 7000 state-owned enterprises were privatised world-wide, valued at over US$185 billion.

Despite the declared aim that privatisations should reinforce market competition there remain significant barriers to such competition at least so far as many utilities are concerned. For example in the case of the water industry in the UK, most houses and businesses can only be supplied through their connection to the distribution mains of their local water company even if they could buy their water from a different supplier. Clearly, if unfettered, the local water company can exert leverage to the detriment of competition through the charges it levies for the use of its mains. Similar arguments hold in the case of the supply of gas, electricity and telecommunications. In order to counter such monopolistic powers the UK government has put in place a regulatory structure. The regulator aims to in effect simulate competition. One of the weapons in the armoury of the regulator is price control which limits the prices for certain services and the annual price rises each regulated company can charge. There are regulators in the UK for the water, electricity, natural gas and telecommunications utilities. In the case of water companies the regulator is the Office of Water Services, known as OFWAT.

Price limits are set by OFWAT once every ten years after so-called "Periodic Reviews" of water companies. However, the regulator may also effect an interim price review after five years from the previous review, if s/he deems it appropriate. Moreover, the regulatory framework also provides for prices to be reviewed even sooner if specific circumstances arise, such as the imposition of new obligations on companies. In 1994, five years after the 1989 privatisation of water utilities in the UK, OFWAT conducted the first periodic review of the water companies in England and Wales. As part of this review price limits were set for each company. This paper describes the use of DEA during the first periodic review of water companies, in 1994.

One of the key questions regulators need to address is whether operating cost savings are in principle feasible. Such savings can then be factored into the permissible annual price increases the regulator announces. DEA is a particularly powerful tool for this aim. The paper presents an adaptation of a generic DEA model for the purpose at hand and outlines the use made of the results obtained. The paper also highlights special issues arising in using DEA as a tool for comparative performance measurement in the regulatory context. The paper should prove of particular interest to those concerned with regulation and more generally with comparative performance measurement.

The paper is laid out as follows. Section 2 describes the framework used by OFWAT for setting price limits for water companies in 1994. Section 3 discusses the choice of input - output variables for use in a DEA assessment of the water distribution function of water companies. Section 4 presents an adaptation of the generic DEA model to estimate potential operating cost savings in water distribution. Section 5 charts the refinement of the DEA model used and section 6 describes the use made of the results obtained. Finally section 7 highlights certain special issues arising in the use of DEA in the regulatory context.

2. THE FRAMEWORK USED BY OFWAT FOR SETTING THE PRICE LIMITS ANNOUNCED IN 1994

In its 1994 Periodic Review OFWAT announced annual price change limits, the limit for the jth company being

$$RPI + K_j \quad (1),$$

where 'RPI' is the Retail Prices Index which measures price inflation in the UK economy and K_j is the permitted company - specific variation from RPI, determined by OFWAT.

So far as potential operating efficiency savings are concerned, OFWAT integrated two components in each company's K - factor: a water - industry - wide component and a company - specific component. (See OFWAT 1998, p 20.) The water - industry-wide component reflects expected savings as a result of economy - wide technological progress. OFWAT estimated the scope for such savings by studying efficiency trends in the UK economy generally. The second component reflects the scope for company - specific savings over and above the water - industry - wide savings, made possible by virtue of improved operating efficiency of the company concerned. **It is with respect to estimating these company - specific savings that the DEA analysis described here was used by OFWAT**.

In order to estimate the potential savings expected through improved operating efficiency, OFWAT decided to estimate potential savings by company function. Figure 1 shows the functions constructed for this purpose. Assessment by functions makes it possible to use simpler estimating models since a model for estimating potential savings at company level would require many variables to reflect the totality of company activities. This is unwieldy given the small number of companies under assessment. The potential savings estimated at function level were later aggregated to estimate potential savings at company level.

Figure 1 about here please

At the time of the 1994 periodic review only ten companies had all of the functions depicted in Figure 1. Such companies are referred to as Water and Sewerage Companies or WASCs. The remaining companies (22 in number) only had clean water functions. They are referred to in this paper as Water only Companies of WoCs. This paper describes the derivation of the DEA - based estimates of potential savings in the function of Water Distribution.

3. POTENTIAL INPUT - OUTPUT VARIABLES IN WATER DISTRIBUTION

A fundamental stage in any assessment by DEA is the identification of a set of input and a set of corresponding output variables. The inputs reflect the resources used in the course of procuring the outputs by the units being assessed. Environmental factors affecting the efficiency of transformation of controllable inputs into outputs should also be accounted for in the assessment. The efficiency measures derived will then reflect the excess, if any, in inputs each unit uses given the output levels it secures and the environmental conditions in which it operates.

In order to identify suitable input - output variables in the water distribution function it is first necessary to delineate the function from the rest of the activities of a water company. Water distribution is the third and final stage in the water supply service depicted in Figure 2.

Figure 2 about here please

The first stage concerns the abstraction of water from reservoirs or bore holes while the second stage involves the treatment of the water abstracted to make it potable or otherwise usable. Water distribution as a function therefore begins with the water input coming from the water treatment plants and ends at the point where water is delivered to clients. Water input into the distribution system is known as ‘Distribution Input’ while water delivered to clients is known as ‘Water Delivered’. The two quantities are not equal because of the loss of substantial amounts of water through leaks in the system. (See OFWAT 1994d Table 3.) Thus the function being modelled is that of receiving water from the water treatment works and delivering it to clients. The water is generally delivered by pumping it to high storage points and then using gravity to take it to clients. This is done by means of the distribution mains which sometimes leak and require periodic maintenance.

Drawing on the nature of the water distribution function outlined above and on discussions with OFWAT the following potential input - output set was constructed.

Table 1: Potential input - output variables in the distribution of water.

Input	**Potential Outputs**
OPEX	PROPERTIES, LENGTH OF MAIN, WDELA, MEASN, REMWDA and BURSTS.

Input

Only one input is listed, that of OPerating EXpenditure or OPEX. OPEX excludes capital investment in renewing and maintaining infrastructure such as the water mains. OPEX as used here encapsulates all variable resource expended in conveying the water from the water treatment works to the customers, except for power costs. Power costs were excluded from OPEX because they could be modelled separately (see OFWAT 1994d Appendix 3).

One of the main reasons for the exclusion of capital cost from the input(s) modelled was that "...(OFWAT) ... *has seen no convincing evidence that relatively high operating expenditure can be explained by relatively low capital expenditure or vice versa*". (OFWAT 1994c p.30.) In the absence of trade offs between capital and operating expenditure the two can be modelled separately. (See also OFWAT 1993 p. 1-2 for further reasons for excluding capital costs from those modelled.)

The operating costs modelled here are generally clearly identifiable and uniformly defined across all companies. Companies face similar staff and materials prices. Thus, once environmental conditions and output levels are taken into account, any remaining cost differences will reflect differing operating efficiencies between companies.

Outputs

The five potential outputs listed in Table 1 were seen as the main factors which can explain OPEX differences between companies. Many of them are highly correlated. The choice of a suitable subset of the potential output variables for our purposes is addressed in the next section. The rationale behind each potential output variable is as follows.

PROPERTIES reflect the number of supply connections served by a company while the LENGTH OF MAIN reflects the dispersion of clients. These two variables capture the scale size of the water distribution network and so we would expect them to influence OPEX.

The amount of water delivered, <u>WDELA</u>, is a measure of the work done by companies in conveying water and so it too should influence the OPEX level.

Water delivered to clients was for the most part not measured in the UK during the period being modelled. Most households paid a water levy commensurate with the 'ratable' value of their property. On the other hand water delivered to businesses was normally measured. Thus the total amount of water delivered by each company is an estimate of the sum of measured and unmeasured water delivered. There were two alternative estimates of water delivered in total by a company, one provided by each company and the other deduced by OFWAT. The latter was used in the analysis.

The two variables of <u>MEASN</u> and <u>REMWDA</u> in Table 1 break down water delivered respectively into that which is measured and the remainder which is estimated. The rationale behind this split is the fact that measured water is delivered almost exclusively to non - households (i.e. businesses). Such water is delivered in large volumes per customer. Thus unit costs of measured and non - measured water should differ which argues for their use as separate variables. Finally, <u>BURSTS</u> is a potential output because it reflects expenditure incurred in repairs to mains bursts.

It should be noted that OFWAT's input into the choice of the input - output variables in Table 1 was undoubtedly influenced by the concurrent but independent econometric analyses (see OFWAT 1993) which identified most of the potential outputs above as explanatory factors of OPEX. (The econometric and DEA analyses reported direct to OFWAT which ran the two independently of each other. We compare the outcomes of the two analyses later in this paper.)

4. AN ADAPTATION OF THE DEA METHOD TO ESTIMATE POTENTIAL SAVINGS IN WATER DISTRIBUTION

Let us assume that cost - efficient water distribution is characterised by constant returns to scale (CRS). (The validity of this assumption will be tested later.) The generic DEA model developed by Charnes et al. (1978) for assessing the DEA - efficiency h_{j_0} of company j_0 under CRS is

$$h_{j_0} = \underset{u_r\, v_i}{Max} \sum_{r=1}^{s} u_r y_{rj_0}$$

$$\text{subject to} \sum_{i=1}^{m} v_i x_{ij_0} = 1 \qquad \text{(M1)}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \qquad j = 1 \ldots j_0 \ldots N$$

$$u_r \geq \varepsilon, \quad v_i \geq \varepsilon, \ \forall \text{ i and r},$$

where *N* is the number of firms, the jth one using input levels x_{ij} i = 1...m to secure output levels y_{rj} r = 1... s. ε is a non - Archimedean infinitesimal while u_r and v_i are respectively the weight to be assigned to output r and input i.

The generic model in (M1) was adapted for assessing potential savings in water distribution in a manner which makes the interpretation of the model more readily transparent. Let us for the time being use PROPERTIES, LENGTH of main and estimated water delivered, WDELA, as the output variables in the assessment. Given that we have a single input, OPEX, if we denote by *v* its weight we will have from the normalisation constraint in (M1) *v* = 1/OPEX j_0. Using this value of *v* and setting $UC_r \equiv u_r \times OPEX\ j_0$ model (M1) becomes

$$h_{j_0}\, OPEX\ j_0 = \underset{UC_{prop} UC_{wde} UC_{len}}{Max} UC_{prop}\ PROPS\ j_0 + UC_{wde}\ WDELA\ j_0 + UC_{len}\ LEN\ j_0$$

subject to:

$$UC_{prop}\ PROPS_j + UC_{wde}\ WDELA_j + UC_{len}\ LEN_j \leq OPEX_j \quad j = 1 \ldots j_0 \ldots 32 \quad \text{(M2).}$$

$$UC_{prop}, UC_{wde}, UC_{len} \geq \varepsilon.$$

$PROPS_j$, $WDELA_j$, LEN_j and $OPEX_j$ are respectively the level of properties, water delivered, length of main and OPEX at company j. There were a total of 32 companies in the assessment. UC_{prop}, UC_{wde}, and UC_{len} can be seen as the unit cost (UC) of properties, water delivered and length of main respectively. The estimated efficient OPEX level of company j_0 is

$$EFF\_OPEX\ j_0 = UC^*_{prop}\ PROPS\ j_0 + UC^*_{wde}\ WDELA\ j_0 + UC^*_{len}\ LEN\ j_0 \qquad (2),$$

where UC^*_{prop}, UC^*_{wde} and UC^*_{len} are respectively the optimal values of UC_{prop}, UC_{wde}, UC_{len} in (M2). **Where we have EFF_OPEX j_0 = OPEX j_0 company j_0 is DEA - efficient**. This is deduced from the fact that model (M2) has identified a set of unit output costs UC^*_{prop}, UC^*_{wde} and UC^*_{len} which when applied to the output levels of the company justify its observed OPEX level in full. Yet, the same unit output costs do not permit any other company to justify via its output levels an OPEX level in excess of that which it actually incurs. **If EFF_OPEX j_0 < OPEX j_0 company j_0 is DEA - inefficient and its estimated potential savings in OPEX are OPEX j_0 -EFF_OPEX j_0.**

5. REFINING THE DEA MODEL USED

Three major issues needed to be addressed in finalising the DEA model used:

- Whether all companies should be used to define acceptable unit output costs;
- The exact output variables to be used and;
- The assumption to be maintained in respect of returns to scale.

5.1 The companies to be used to define acceptable unit output costs

As has already been noted there is a dichotomy of companies between water only (WoCs) and water and sewerage (WASCs) companies. There are only 10 WASCs but they account for some 75% of water delivered in England and Wales. Yet, initial runs with model (M2) indicated that it is mostly WoCs which de facto define the sets of acceptable unit output costs. As WoCs account for a minority of water delivered this it was felt might be unacceptable to WASCs. They could point out to the fact that WoCs are generally small companies which fail to reflect the more complex structure of WASCs. **Therefore, it was decided that only WASCs will be permitted to define acceptable unit output costs.** This is equivalent to removing dynamically from (M2) any constraint which relates to a WoC and turns out to be binding.

This decision means that a conservative view is being taken on potential savings at both WASCs and WoCs. Figure 3 illustrates the point using only two output variables to make a graphical explanation possible. The two output variables are LENGTH and PROPERTIES. Figure 3 shows the observed LENGTH and PROPERTIES per unit of OPEX at a company. The outer boundary enveloping all companies corresponds to the case when WoCs are permitted to define acceptable unit output costs. The inner boundary consisting of the solid thin and thick lines corresponds to the case when only WASCs are permitted to define acceptable unit output costs. The DEA - efficiency rating h_{j_0} which model (M2) would yield in respect of company A is OA/OB when WoCs are not permitted to define acceptable unit output costs. In contrast, when WoCs are permitted to define acceptable unit output costs company A has DEA - efficiency rating h_{j_0} = OA/OC. (For a discussion of how DEA - efficiency is measured by graphical means in the manner illustrated in Figure 3, see Barrow and Wagstaff (1989) p. 84.)

Figure 3 about here please.

Clearly the DEA - efficiencies will always be higher with reference to the inner rather than the outer boundary and so using WASCs to define acceptable unit costs leads to estimates of lower potential OPEX savings by companies than if WoCs were also permitted to define acceptable unit costs. (For a fuller discussion of the 'duality' between the unit costs in (M2) and the 'production space' depicted in Figure 3 see Thanassoulis (1996).)

Although the estimated potential savings are reduced by not permitting WoCs to define acceptable unit output costs, the ranking of companies on efficiency is not affected much. Table 2 shows the relevant rank correlation coefficients for alternative subsets of output variables.

Table 2: Rank correlation coefficients between the case when WoCs are and are not permitted to define acceptable unit output costs.

Output Set	Correlation Coefficient (2-tailed test significance level)
{LEN, PROPS}	0.9615 (0.000)
{LEN, WDELA}	0.9969 (0.000)
{LEN, PROPS, WDELA}	0.9769 (0.000)
{LEN, PROPS, BURSTS}	0.9772 (0.000)
{LEN, PROPS, MEASN, REMWDA}	0.9801 (0.000)

The high positive correlation coefficients in Table 2 suggest that though the absolute levels of estimated unit output costs reduce, their relative levels remain largely unchanged when we move from the WoC to the WASC defined unit output costs under the output sets in Table 2.

5.2 The choice of a subset of outputs

Table 3 shows the correlation coefficients between the output variables in Table 1. OPEX, is highly and positively correlated with all potential outputs and this supports the choice of the output variables. The outputs are also generally highly and positively correlated with each other. The larger correlation coefficients are highlighted and in such cases only one of the two variables concerned need be used in the DEA model.

Table 3: Correlation coefficients

	OPEX	**WDELA**	**MEASN**	**PROPS**	**LENGTH**	**BURSTS**
WDELA	0.946					
MEASN	0.882	**0.981**				
PROPS	0.922	**0.995**	**0.984**			
LENGTH	0.847	0.941	0.972	0.951		
BURSTS	0.680	0.812	0.876	0.835	0.897	
REMWDA	0.961	**0.997**	0.963	**0.989**	0.917	0.777

It was decided to construct different potential self contained sets of output variables and observe the nature of any differences in the assessments of companies they yield. Table 4 shows three initial output sets constructed from those in Table 1.

Table 4: Three output sets for assessing potential savings in water distribution

Set	**Outputs**
1.	PROPERTIES, LENGTH OF MAIN and WDELA
2.	PROPERTIES and LENGTH OF MAIN
3.	LENGTH OF MAIN and WDELA

In all cases the input was OPEX. The first output set uses what are deemed to be the essential variables for explaining OPEX in the distribution of water. That is the properties served, their dispersion and the quantity of water delivered. Sets 2 and 3 replicate set 1 but they drop water delivered and properties respectively, as they are highly correlated. Any companies which buck the trend in correlation between properties served and water delivered should show very different efficiency ratings between these two output sets.

Model (M2) was solved in respect of each one of the 32 companies and output sets in Table 4. The results obtained appear in Appendix 1. Efficiencies and ranks are remarkably stable across the three output sets in Table 4, except for three companies. Two of them, companies C31 and C30, are ranked much better when WDELA rather than properties is used as an output variable. This is because these two companies have very large amounts of water delivered for the properties they serve because of the unusually large component of measured water they deliver. (See Appendix 3.) A third company, C5 has the lowest proportion of measured water delivered and so its ranking on efficiency suffers when water delivered is used as an output variable instead of properties served. Since measured water is delivered to businesses in large volumes per client, and therefore reflects lower expenditure than the same volume of water delivered to households, the output set {LENGTH, PROPERTIES } is thought to give the more accurate reflection of company cost - efficiency between the output sets in Table 4.

Next, the impact of including BURSTS and splitting WDELA into water delivered measured to non - households (MEASN) and the remainder of water delivered, (REMWDA) was assessed. The output sets used were those in Table 5. Appendix 2 shows the related company efficiencies and ranks. The Appendix also shows the largest changes in efficiencies and ranks of each company across the output sets in Table 5.

Table 5: Output sets accounting for measured water and bursts of main

Set	Outputs
1	PROPERTIES, LENGTH of main
2	PROPERTIES, LENGTH of main and BURSTS
3	PROPERTIES, LENGTH of main, MEASN and REMWDA = (WDELA-MEASN)

Only four companies, namely C29, C26, C30 and C31 change efficiency rating by more than 11 percentage points between the three assessments (see column D1 in Appendix 2) and the same four plus C3 change ranking by more than 4 places between the three assessments. Thus for the vast majority of companies it makes little difference to their efficiency rating or their ranking which one of the output sets in Table 5 is used. It is worth exploring the reasons why this is not so for the few companies identified above.

The reason why C31 and C30 attain high efficiency when water delivered is used as an output variable either in aggregate or disaggregate form has already been explained above. Similarly, C29 attains a better ranking and efficiency rating when water delivered is used as an output variable because it has the third largest proportion of measured water after C31 and C30 (see Appendix 3). C26 attains a particularly high efficiency rating when bursts are used as an output variable because it has by far the largest number of bursts per km of main, as can be seen in Appendix 3. Given that the data for bursts is atypical in the case of C26 and it is the only company which is affected substantially in ranking when bursts are added as an output variable, it was decided that bursts would be dropped as a potential output variable. In light of the foregoing observations the decision was made to adopt

{PROPERTIES, LENGTH, WDELA}.

as the output set for the analysis. The variable WDELA was added to LENGTH and PROPERTIES because it does give companies which deliver large volumes of water for their number of properties or length of main the ‘benefit of the doubt’ on cost efficiency.

5.3 Tests for the nature of returns to scale

Companies vary very substantially in size. On each output variable, the largest company is nearly 100 times the size of the smallest company (see Appendix 4). Clearly the issue of returns to scale is important. Given the regulation context in which the assessment is being undertaken, the argument can be advanced that we should assume constant returns to scale irrespective of the nature of returns to scale actually characterising efficient operation. It is for companies to identify the most cost - efficient scale size. This issue is elaborated in the next section. Given that scale size at this stage in the life of the water companies is largely as inherited and outside managerial control in the short run, it was decided to test whether there is evidence that there are economies of scale.

Regressing the natural log of OPEX on the natural logs of PROPERTIES, LENGTH and WDELA gives the model in (3).

LN[$]_OPEX=	0.597	+ 0.314	LN_PR	+ 0.289	LN_LEN	+ 0.402	LN_WD	(3)
St Dev	0.775	0.367		0.189		0.346		
p	0.447	0.4		0.137		0.225		

[$] LN = Log to base *e*.

The model in (3) explains 96.3% of the variation of the natural log of OPEX levels across companies. The sum of the partial regression coefficients is 1.005 and the size of their standard deviations readily leads to the conclusion that the sum is not statistically significantly different from 1. Thus model (3) supports the hypothesis of constant returns to scale in water distribution, though strictly speaking it does not relate to the DEA - efficient boundary.

6. USING THE ASSESSMENT RESULTS

The DEA efficiencies in water distribution used were those obtained under an assumption of constant returns to scale, using {PROPERTIES, LENGTH of main and WDELA } as outputs and not permitting WoCs to define acceptable unit output costs. The data used related to the reporting year 1992/3. The efficiencies and ranks of the companies appear in Table 6. The potential OPEX savings for a company with efficiency under 100% equal the difference between its observed and estimated efficient OPEX level (see section 4). The potential OPEX savings across the full set of companies amounted to £144m on base modelled expenditure of £540.5m in 1992/3 prices, that is 26.67%.

Table 6: Efficiencies and ranks of companies

	Rank	Efficiency (%)			Rank	Efficiency (%)
C32	1	124.11		C22	17	75.92
C16	2	107.57		C29	18	73.15
C25	3	103.82		C24	19	71.07
C19	4	100		C8	20	71.01
C27	4	100		C2	21	69.27
C30	6	99.63		C3	22	67.79
C31	7	99.55		C11	23	63.19
C15	8	96.99		C6	24	62.29
C4	9	92.73		C12	25	61.49
C1	10	91.24		C9	26	60.15
C18	11	89.4		C13	27	58.16
C5	12	84.16		C7	28	57.36
C26	13	83.28		C21	29	54.52
C20	14	83.24		C10	30	52.4
C28	15	81.4		C14	31	52.14
C23	16	80.35		C17	32	44.33

As noted earlier, OFWAT also used econometric estimates of the potential efficiency savings in water distribution. It is interesting to contrast the DEA and econometric results. In the case of Water Distribution the following econometric model was used by (OFWAT 1994e) to estimate potential savings:

$$\text{Distribution Expenditure (£000)} = 17.84\ \text{WDELA}^{0.61}\ \text{LENGTH}^{0.37}\ e^{-1.3\ \text{PMNH}} \qquad (4)$$

where PMNH is the proportion of WDELA (Mega litres per day) delivered measured to non - households. LENGTH of main is in km. The expression in (4) is compatible with the constant returns to scale assumption in the DEA model. PMNH is scale invariant while the exponents of WDELA and LENGTH add to 0.98 which is not statistically significantly different from 1 as can be readily deduced from their standard errors in OFWAT (1994e).

The econometric model in (4) and the DEA model used produce similar results. The ranks of companies on efficiencies have a correlation coefficient of 0.868 which is positive and strong. Both approaches use length of main and water delivered as OPEX drivers. The DEA model uses also PROPERTIES while the econometric model uses PMNH. PMNH captures within the econometric model the impact of PROPERTIES in the DEA model. To see this note that DEA estimates higher OPEX levels for larger numbers of properties served per unit of water delivered. Model (4) does so too for PMNH is low when the number of properties served is high relative to water delivered and this leads to higher estimated OPEX levels by (4). (The more the properties served per unit of water delivered the less the amount of water delivered measured in bulk to businesses which reduces the value of PMNH.)

One key difference between the DEA and econometric estimates of efficient OPEX levels is that DEA allows for varying unit output costs by company, depending on its mix of driver levels. Mix here means the ratios driver levels are to each other. In contrast, the econometric model in (4) is the same irrespective of whether a company operates in a context of say high volume of water delivered but short length of main or the other way round. DEA on the other hand would permit two companies operating in two such different environments to claim that operating conditions differ and so efficient practices differ in the two environments. This in turn leads to different sets of efficient (in the economic sense) unit output costs in the two environments. This is evident if we re-examine (2) where the efficient OPEX level for company j_0 was

$$\text{EFF\_OPEX}\, j_0 = UC^*_{prop}\ \text{PROPS}\, j_0 + UC^*_{wde}\ \text{WDELA}\, j_0 + UC^*_{len}\ \text{LEN}\, j_0 \qquad (2).$$

The unit output costs UC^*_{prop}, UC^*_{wde} and UC^*_{len} yielded by (M2) can be different for each company j_0. The model selects unit output costs which maximise the company's estimated efficient OPEX level (and so they minimise the potential cost savings that can be demanded from it) subject to no company being able to justify under the unit costs higher expenditure than it incurs. The model will generally select different unit costs for companies with different mix of output levels. A further difference between the DEA and econometric approaches is that the DEA estimated OPEX level of a company is efficient in the sense that no company or combination of companies can for the same output levels justify a lower

OPEX level. In contrast, the econometrically estimated OPEX level of a company is only what its output levels would justify on average in the industry.

Precise details of how OFWAT used the DEA results to support the econometric results are not in the public domain but there are some indications. OFWAT (1995 p. 414) notes that the DEA estimates of efficient cost levels in distribution and resources and treatment (the latter not detailed here)

> "*were added to overall average* (clean) *water business activities costs and the result divided by the actual distribution, treatment and business costs to give an overall* (DEA - based) *efficiency ratio* (of clean water operations). *In most cases the results* (on company efficiency on clean water operations) *were similar to those of the regressions. If they were significantly better, the Director* (of OFWAT) *moved the company up one band* (on efficiency in clean water)."

Once the ranks on efficiency were obtained, OFWAT took further factors into account before arriving at the final price determinations. In particular, (see OFWAT 1994c p. 31)

> "*The Director has also taken into account companies' submissions on the effect of their special circumstances; evidence on individual company costs and efficiency improvements incorporated in SBPs* (strategic business plans); *reviews of actual operating expenditure for individual companies in the first period* (that is from privatisation in 1989 to the time of the Price Review in 1994) *compared with actual expenditure up to 1988-89 and the projections then assumed; and an assessment of comparative levels of service*".

(Bracketed text has been added by way of explanation of the background to the quotations given in italics.) Factors of this type influenced jointly with the quantitative estimates of potential efficiency savings the final price determinations announced by OFWAT. OFWAT ultimately set savings targets which would mean inefficient companies would cover 25% - 35% of their distance from the *best performing company or companies* (OFWAT 1998, p. 22). The efficiencies estimated by DEA and econometric models clearly fed into identifying the best performing companies and the distances of the remaining companies.

7. SOME SPECIAL ISSUES IN COMPARATIVE PERFORMANCE MEASUREMENT IN THE REGULATORY CONTEXT

The application reported in this paper brings to the fore certain important issues in using comparative efficiency assessments methods such as DEA within a regulatory framework.

Should the regulator always estimate potential savings assuming constant returns to scale?

Given that the regulator's duty is to safeguard the public interest, an argument can be advanced that assessments should be under constant returns to scale irrespective of the nature of returns to scale actually characterising efficient production. This view treats scale size as controllable by management who should therefore not be permitted to pass on to the public any inefficiencies consequent on operating plant and equipment at uneconomic scale size. There are, however, two counter - arguments to the foregoing statement.

Firstly, if scale size is dependent on such contextual variables as population served, dispersion of population and so on, companies do not have control over their scale size except though mergers and acquisitions, themselves subject to regulatory approval. Secondly, in the case of UK utilities, the scale size of assets (treatment works, pumping stations etc.) each company inherited on privatisation were beyond its control. Assets have long lives and so in the short term management cannot change their scale of operation to exploit returns to scale. Thus, arguably, in the short term at least, companies cannot move to a "most productive scale size" (see Banker (1984)) and so should be assessed given their current scale size.

There is an incentive for companies to move to a more economic scale of operation even if the regulator is using variable returns to scale to estimate potential efficiency savings. Companies will benefit, until the next periodic review, from any savings they make due to scale size changes, not reflected in the savings targets set by the regulator during the preceding review. The question remains open, however, as to whether the regulator should adopt an approach in price reviews which reflects an element of compulsion for companies to move to more economically efficient scale sizes.

How far should the weight of inheritance be permitted to perpetuate any cost inefficiency?

OFWAT sought efficiency savings given the water distribution network used by each company. This approach, obviously cannot identify a cost - inefficient network configuration (e.g. positioning of pumping stations and water towers so as to lead to uneconomic lengths of main). The issue arises in other assessments in regulation too. For example OFWAT sought efficiency savings given the treatment processes for sewage and water used by each company. This approach, again cannot identify any cost inefficiency attributable to using cost - ineffective treatment methodologies. The argument against putting the onus on companies to use cost - efficient networks (or treatment processes) is that assets were inherited at privatisation and so there is little management can do in the short term to change them. This is essentially the same argument as advanced above for not assessing companies always under constant returns to scale. The question therefore remains open as to how, and when, the regulator can factor into price determinations an element of compulsion for companies to move to more cost - efficient distribution network configurations.

8. CONCLUSION

Regulation is inextricably intertwined with the rising tide of privatisations of state - owned utilities in recent decades. So long as there remain physical or other barriers to effective competition regulation of privatised utilities is seen as the main defence of the public interest against potential abuse of monopoly power.

This paper has given an account of the use of DEA during the price review conducted in 1994 by OFWAT, the regulator of water companies in England and Wales. The account concerns the use of DEA to estimate potential savings of operating expenditure in the distribution of water. The paper gives an adaptation of a generic DEA model which makes the DEA - based estimates of potential cost savings transparent. The paper charts the refinement of the model used in terms of the choice of acceptable referent companies, the choice of input - output variables and the assumption on returns to scale to be maintained. Following these refinements the cost efficiency of each company is estimated. The results are contrasted with alternative efficiency estimates obtained by OFWAT using an econometric approach. The two sets of results fed into the price limits OFWAT announced in 1994, to cover the period from 1995 to 2005. (In the event OFWAT announced an interim price review for 1999.) The paper concludes with certain generic issues which need to be addressed on the use of comparative efficiency assessments in the context of regulation.

APPENDIX 1: EFFICIENCIES AND RANKS UNDER THE OUTPUT SETS IN TABLE 4

The larger of the two efficiencies under {LENGTH, PROPS} and {LENGTH WDELA} is the efficiency under {PROPERTIES, LENGTH and WDELA}.

	EFFICIENCIES (%)$			**RANKS**		
	PROPS and LENGTH A	**WDELA and LENGTH B**	**Diff'nce B-A**	**PROPS and LENGTH C**	**WDELA and LENGTH D**	**Diff' nce D-C**
$$C31	**64.77**	**99.55**	**34.78**	**22**	**6**	**-16**
C30	**81.38**	**99.63**	**18.25**	**14**	**5**	**-9**
C29	70.09	73.15	3.06	18	16	-2
C10	48.88	52.40	3.52	31	30	-1
C7	57.36	54.54	-2.82	28	27	-1
C13	57.64	58.16	0.52	27	26	-1
C11	61.91	63.19	1.28	24	23	-1
C2	68.75	69.27	0.52	20	19	-1
C24	69.93	71.07	1.14	19	18	-1
C22	75.92	73.72	-2.20	16	15	-1
C20	83.24	82.75	-0.49	12	11	-1
C19	100.00	100.00	0.0	4.5	3.5	-1
C27	100.00	100.00	0.0	4.5	3.5	-1
C25$	103.82	103.82	0.0	3	2	-1
C17	43.36	44.33	0.97	32	32	0
C21	54.52	53.46	-1.06	29	29	0
C12	61.49	59.36	-2.13	25	25	0
C32	124.11	108.96	-15.15	1	1	0
C14	52.08	52.14	0.06	30	31	1
C6	61.94	62.29	0.35	23	24	1
C3	67.79	64.85	-2.94	21	22	1
C28	81.40	80.67	-0.73	13	14	1
C1	91.24	89.65	-1.59	8	9	1
C4	92.73	91.99	-0.74	7	8	1
C9	60.15	53.54	-6.61	26	28	2
C23	80.35	72.63	-7.72	15	17	2
C26	83.28	80.95	-2.33	11	13	2
C8	71.01	68.62	-2.39	17	20	3
C18	89.40	81.55	-7.85	9	12	3
C15	96.99	88.99	-8.0	6	10	4
C16	107.57	96.14	-11.43	2	7	5
C5	**84.16**	**66.38**	**-17.78**	**10**	**21**	**11**

$ The *efficiency rating* of each company is h_{j_0} x 100 where h_{j_0} is as yielded by the relevant instance of model (M2). For WoCs 'outside' the efficient boundary their 'super - efficiency' is shown, computed using the models developed by Andersen and Petersen (1993).)

$$ For reasons of confidentiality companies are identified only as C1 ... C32.

APPENDIX 2: EFFICIENCIES AND RANKS ACCOUNTING FOR MEASURED WATER AND BURSTS OF MAIN

	EFFICIENCIES (%)				**RANKS**			
	PROP'S and LENGTH A	**PROP'S, LENGTH, BURSTS B**	**PROP'S, LENGTH, MEASN, REMWDA C**	**Difference D1 = Max (A,B,C) - Min (A,B,C) D1**	**PROP'S and LENGTH E**	**PROP'S, LENGTH, BURSTS F**	**PROP'S, LENGTH, MEASN, REMWDA G**	**Difference D2 =Max (E,F,G) - Min (E,F,G) D2**
C17	43.36	43.36	45.26	1.9	32	32	32	0
C6	61.94	64.13	68.36	6.42	23	23	23	0
C28	81.4	81.4	87.09	5.69	13	13	13	0
C21	54.52	54.52	57.88	3.36	29	30	29	1
C9	60.15	60.15	60.15	0	26	27	27	1
C12	61.49	63.42	61.49	1.93	25	24	25	1
C2	68.75	68.75	75.69	6.94	20	21	20	1
C8	71.01	71.01	80.05	9.04	17	18	18	1
C1	91.24	91.24	99.31	8.07	8	9	8	1
C24	69.93	69.93	73.86	3.93	19	20	21	2
C23	80.35	80.35	80.35	0	15	15	17	2
C18	89.4	89.4	89.4	0	9	10	11	2
C19	100	100	100	0	5	6	7	2
C27	100	100	100	0	5	6	7	2
C25	103.82	107.77	104.47	3.95	3	3	5	2
C16	107.57	108.62	107.57	1.05	2	2	4	2
C32	124.11	124.11	124.11	0	1	1	3	2
C10	48.88	48.88	59.84	10.96	31	31	28	3
C14	52.08	59.4	53.8	7.32	30	28	31	3
C13	57.64	57.64	61.26	3.62	27	29	26	3
C11	61.91	61.91	68.63	6.72	24	25	22	3
C22	75.92	78.49	76.37	2.57	16	16	19	3
C4	92.73	99.42	92.73	6.69	7	7	10	3
C15	96.99	96.99	96.99	0	6	8	9	3
C7	57.36	60.5	57.36	3.14	28	26	30	4
C20	83.24	86.23	84.24	2.99	12	11	15	4
C5	84.16	84.16	84.8	0.64	10	12	14	4
C3	**67.79**	**72.19**	**67.79**	**4.4**	**21**	**17**	**24**	**7**
C29	**70.09**	**70.09**	**88.78**	**18.69**	**18**	**19**	**12**	**7**
C26	**83.28**	**100**	**83.28**	**16.72**	**11**	**6**	**16**	**10**
C30	**81.38**	**81.38**	**144.16**	**62.78**	**14**	**14**	**2**	**12**
C31	**64.77**	**64.77**	**161.55**	**96.78**	**22**	**22**	**1**	**21**

APPENDIX 3: INDEX OF MEASURED WATER DELIVERED TO NON - HOUSEHOLDS AND BURSTS PER KM OF MAIN

		Proportion WDELA measured to non households: Largest proportion = 100.		Bursts per Km of main: Largest ratio = 100.
1.	**C5**	**26.96704**	C27	16.9326
2.	C6	36.71576	C28	19.91121
3.	C8	40.83457	C25	27.7417
4.	C9	42.18609	C17	29.97123
5.	C10	43.89907	C24	30.2538
6.	C7	44.06989	C10	30.89259
7.	C1	48.15926	C5	31.44529
8.	C23	49.25481	C29	31.57828
9.	C24	49.26101	C30	31.64834
10.	C2	50.05132	C20	32.95211
11.	C11	51.03481	C9	35.23922
12.	C32	51.24907	C21	35.46706
13.	C21	51.31253	C23	37.52909
14.	C22	53.70771	C1	38.25541
15.	C13	54.96224	C2	38.39765
16.	C14	55.6624	C12	39.59067
17.	C3	55.82996	C32	41.18974
18.	C4	56.248	C4	42.42495
19.	C12	56.33358	C11	42.46931
20.	C15	56.93087	C22	42.56363
21.	C16	57.03691	C15	47.82609
22.	C17	58.5451	C13	48.46348
23.	C18	59.19453	C7	51.39566
24.	C19	61.6235	C3	53.10935
25.	C20	61.67805	C6	55.14901
26.	C25	61.92312	C31	64.60475
27.	C26	62.85737	C14	65.21681
28.	C27	70.18305	C18	65.21728
29.	C28	78.61721	C19	65.9037
30.	***C29***	***79.48386***	C8	69.86327
31.	**C30**	**88.59613**	C16	73.45636
32.	**C31**	**100**	**C26**	**100**

APPENDIX 4: SIZE OF COMPANIES ON EACH ONE OF THE THREE OUTPUTS: LARGEST LEVEL = 100.

Properties	Length of main	Water delivered
1.165571	1.222458	1.09159
1.358685	1.375516	1.344719
1.896401	2.309366	1.671307
1.96338	2.387145	1.828518
2.067293	2.516694	1.842087
2.283568	2.547955	2.051234
3.049765	3.4678	2.585566
3.367762	4.635738	2.686162
3.425352	5.261723	2.95286
3.650391	5.449544	3.241549
4.131455	5.649369	3.995789
5.313615	6.045267	5.471049
5.968701	6.661748	5.960463
6.716745	8.148056	6.293602
8.010016	10.23884	6.460171
8.328326	11.14368	8.114165
8.620344	11.99025	8.248918
13.6025	13.83269	13.03778
14.64977	15.0629	13.2095
15.33646	15.62886	14.04188
15.6651	19.53007	14.46719
17.85352	22.24059	16.46415
17.99562	26.03151	16.47023
20.58529	27.27323	17.37794
28.52833	30.75078	26.7629
29.68482	35.53558	26.81811
35.65258	60.96786	35.13721
48.85759	68.42466	47.04878
57.19624	77.28198	49.93988
87.90829	82.52495	74.19675
90.61033	97.50381	82.011
100	100	100

Each column reflects the sorting of the companies on size on the output concerned and so rows do not necessarily correspond to companies.

REFERENCES

Andersen, P. and Petersen, N.C., A Procedure for Ranking Efficient Units in Data Envelopment Analysis, *Management Science*, Vol. 39/10, (1993) 1261-1264

R. Allen, A. Athanassopoulos, R. G. Dyson and E. Thanassoulis, Weights restrictions and value judgements in data envelopment analysis: evolution, development and future directions, *Annals of Operations Research, Vol. 73* (1997) 13 -34**.**

Banker, R. D."Estimating Most Productive Scale Size Using Data Envelopment Analysis", *European Journal of Operational Research*, 17, (1984) 35-44.

Banker, R.D., Charnes, A. and Cooper, W.W., "Some models for estimating technical and scale inefficiencies in data envelopment analysis.*"*, *Management Science,* 30/9, (1984) 1078-1092

Barrow M. and Wagstaff A., Efficiency measurement in the public sector: An appraisal, *Fiscal Studies*, (1989), 72-97.

Boussofiane, A., Dyson, R.G. and Thanassoulis, E., Applied Data Envelopment Analysis, *European Journal of Operational Research*, 52 (1991), 1-15.

Charnes, A., Cooper, W. and Rhodes, E., "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2, (1978), 429-444.

Land A. "Data envelopment analysis". In S. C. Littlechild and M. F. Shutler (eds) *Operations Research in Management*, Prentice Hall, 1991, 65 - 76.

Megginson, W. L., R. C. Nash and M. V. Randenborgh (1994), "Performance of Newly Privatised Firms: An International Empirical Analysis", *The Journal of Finance*, XLIX/ 2 (June 1994), 403-452.

OFWAT "Research Paper No. 2 — Modelling Water Costs 1992-3", Prepared by M. Stewart for OFWAT, 7 Hill Street, Birmingham B5 4UA, England (1993).

OFWAT "Research Paper No. 3 — Modelling Sewerage Costs 1992-3", Prepared by M. Stewart for OFWAT, 7 Hill Street, Birmingham B5 4UA, England. (1994a).

OFWAT "Research Paper No. 4 — Modelling Sewage Treatment Costs 1992-3", Prepared by M. Stewart for OFWAT, 7 Hill Street, Birmingham B5 4UA, England. (1994b).

OFWAT "Future Charges for Water and Sewerage Services: The outcome of the Periodic Review", OFWAT 7 Hill Street, Birmingham B5 4UA, England, (1994c).

OFWAT "1993-94 Report on the Cost of Water Delivered and Sewage Collected", OFWAT 7 Hill Street, Birmingham B5 4UA, England, (1994d).

OFWAT "Addendum to the 1993-94 Report on the Cost of Water Delivered and Sewage Collected", OFWAT 7 Hill Street, Birmingham B5 4UA, England, (1994e).

OFWAT "South West Water Services Ltd", HMSO, (1995).

OFWAT "1999 PERIODIC REVIEW: Assessing the Scope for Future Improvements in Water Company Efficiency - A Technical paper" HMSO, (1998).

Thanassoulis E., "Assessing Police Forces in England and Wales Using Data Envelopment Analysis", *European Journal of Operational Research*, 87, (1995) 641 - 657.

Thanassoulis, E. "A data envelopment analysis approach to clustering operating units for resource allocation purposes", *Omega, International Journal of Management Science*, 24/ 4, (1996) 463 - 476.