Top down benchmarking of UK Gas Distribution Network Operators

A Report by Europe Economics to OFGEM

Europe Economics
Chancery House
53-64 Chancery Lane
London WC2A 1QU
Tel: (+44) (0) 20 7831 4717
Fax: (+44) (0) 20 7831 4515
www.europe-economics.com

10 April 2007

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Introduction

1 As part of the first Gas Distribution Price Review (GDPCR) Ofgem commissioned Europe Economics to form a view on the relative efficiency of each Gas Distribution Network operator (GDN).

2 The terms of reference require benchmarking of total employment costs and of total operating expenditure (opex) using regression and Data Envelopment Analysis (DEA) together with any other appropriate analysis in order to form a view on the comparative performance of the GDNs. The study should also look at the efficiency of the gas distribution sector as a whole relative to other similar sectors.

3 It was recognised at the start of this work that there were significant data issues which would constrain the scope of this analysis and place limits on the reliability of the conventional approaches to such a benchmarking exercise. We have therefore tested a range of different forms of analysis and models in order to identify the most appropriate top down productivity measures from which conclusions for future performance can be drawn.

Data and opex benchmarking

4 Data on controllable operating costs were taken from the Business Plan Questionnaires (BPQs) completed by each of the GDNs as adjusted by Ofgem.

5 Historic outturn data are available for 2004/05 to 2005/06. Ofgem has separately investigated the reliability of the data in 2004/05 and found that these were significantly affected by the industry changes and were not suitable for this analysis. The benchmarking is therefore based on outturn data for 2005/06 and estimated outturn data for 2006/07.

6 We carried out an initial comparison of GDN employment costs but the data available us on direct and indirect employment costs for the different GDNs were not comparable and we did not pursue the employment benchmarking analysis.

7 GDNs were compared on the basis of unit operating costs taking a number of different output measures. This provides a straightforward initial analysis of comparative efficiency and can be carried out with a limited data set. However this analysis cannot adjust for differences in factors such as customer density which are outside the GDNs' control. The rankings of the GDNs show considerable variation depending on the output measure used. The range of efficiency levels is narrower in the analysis by ownership group but this may suggest that grouping tends to mask underlying productivity differences.

8 Regression analysis was used to explore the relationship between total controllable costs and a number of possible explanatory variables which are considered likely to be the

main cost drivers. We considered volumes of gas distributed, total number of customers, network length, the proportion of non domestic customers and customer density.

9 Our analysis used the Corrected Ordinary Least Squares (COLS) procedure. We estimated four separate regression models. The first two models (COLS1 and COLS2) have the merit that the significance of the output variables is determined within the model rather than being externally imposed through the composite scale variable used in models COLS3 and COLS4. COLS1 and COLS2 are our preferred models

10 Small sample size issue can be a significant problem in the case of regression analysis. However, the results for our preferred models show a high level of 'fit' and are robust to the most serious misspecification problems that could invalidate the results. We have carried out analysis to test whether results were unduly influenced by one or two individual observations or 'outliers' and adjustments were made to allow for such effects.

11 COLS1 and COLS2 control for different variables, leading to some difference in rankings, but statistically there is little to choose between them. The two models show changes in the rankings between the two years. For most companies there is a reduction in efficiency in the second year although Wales and West and Northern show some improvement.

12 We carried out a DEA review. This involves the use of linear programming methods to construct an efficiency frontier against which individual GDNs can be compared. Models were constructed using the same cost drivers as used in the regression analysis.

13 In general, the DEA analysis shows that in the first year a number of companies are on or close to the efficiency frontier but in the second year almost all companies show a lower level of efficiency. The rankings of most GDNs are broadly similar for all models

14 A number of considerations need to be taken into account in drawing conclusions from the DEA work. First it is not possible to verify the accuracy and the quality of the fitted frontier because of the difficulty of applying hypothesis testing to DEA results. Secondly, the very small sample size available for this report limits the reliability of the results. Small sample size is a particular concern with DEA techniques which are less likely to be efficient with a small sample. With a small sample there is a tendency for all firms to be shown to be relatively efficient.

15 Multi-factor productivity indices (MPI) can overcome some of the shortcomings of partial unit cost indices. In order to aggregate the outputs in a total output index we have allocated a weight to each output. Alternative weighting patterns have been drawn from other studies of utility networks.

16 Three MPI models were considered using different weights. These models show reasonably consistent rankings of GDNs in each of the two years but some variation in the rankings between years. As with the regression analysis and DEA most companies show a decline in efficiency in 2006/07.

17 Overall these indices are very sensitive to the assumptions made about the weights to be attached to the individual output variables. In addition these indices do not control for the existence of economies of scale. At best therefore MPI should only be used as a cross-check on the values and rankings identified in the other forms of analysis.

18 We carried out a rank correlation analysis to test the extent to which different approaches produced similar rankings of the GDNs.

19 The results are mixed. All the correlation coefficients are statistically significant suggesting that on this basis alone there is little to choose between the models. There is a higher degree of correlation between the rankings in the four regression models than between the DEA and MPI estimates. The correlation between the regression models and the DEA results is higher than it is with the MPI analysis.

20 Overall the regression analysis appears to be the most robust of the different approaches investigated and in our view should be considered as the most reliable and robust in this case. Our preference is to use the COLS1 and 2 models which make the best use of the available data without the use of predetermined weights between potential explanatory variables. We consider that equal weight should be given to the comparative efficiency estimates from these preferred regression models.

Nature of work comparisons

21 Nature of work comparison is used to estimate the rate of productivity and efficiency growth in a sector as a whole rather than for an individual company. It is important to assess the scope for efficiency improvement of the industry efficiency frontier as well as understanding the scope for companies to catch with the most efficient performers. The time profile of productivity growth in similar sectors provides useful insights into the scope for efficiency improvements for gas distribution.

22 We have constructed a benchmark based on a weighted average of productivity estimates for the comparators in the wider economy which captures the potential for long-term productivity improvements in the gas distribution sector. This reflects the underlying scope for productivity improvement achievable through technological progress and improvements in working practices.

23 The results of this analysis indicate that, on the basis of the nature of work that the gas distribution industry undertakes, we could expect the industry to achieve gains in operating expenditure productivity of between 1.9 per cent and 3.7 per cent per annum over and above the underlying growth in productivity in the economy as a whole. These results can be combined with the analysis of comparative efficiency assessment to calculate the scope for efficiency improvement of individual companies.

Frontier shift and efficiency saving

24 The total scope for efficiency saving can be split into two components, frontier shift and catching up. Frontier shift represents the scope for cost reductions that is due to the

movements in the industry efficiency frontier. Catching up accounts for the remaining part of the total scope for efficiency savings. In order to separate out these two components we have combined our estimates of industry wide efficiency potential with the efficiency rankings derived from our regression analysis.

25 To avoid overstatement of efficiency savings at the company level, the COLS procedure has been adjusted by assuming that all GDNs with an efficiency level equal to the highest quartile are fully efficient and we have re-scaled the efficiency estimates with respect to this target.

26 We have considered two baselines for the comparison. Efficiency levels in general declined in 2006/07. Therefore, restricting the selection of the benchmark company to the upper quartile in 2006/07 only may reduce the average catching up for the sector. Alternatively, there is some merit in using an average of the estimates for both 2005/06 and 2006/07 as the benchmark. This could reduce the possibility that year specific shocks would play too large a role in shaping the industry efficiency frontier. We present results showing the scope for catching up using both of these approaches.

27 We have prepared estimates of the scope for annual efficiency improvements in each GDN using the upper and lower bounds of the nature of work productivity projections and the averaged results from the COLS1 and COLS2 models. These show how much each GDN could, on average, be expected to cut its operating expenditure on top of the economy wide productivity growth each year over the next five year period.

Company specific efficiency improvement: baseline 2006/07[1]

.	Average annual % reductions over five years					
	NWLB			NWUB		
GDN	Frontier shift	Catching up	Total	Frontier shift	Catching up	Total
East of England	1.00%	0.90%	1.90%	2.82%	0.90%	3.70%
London	1.00%	3.12%	4.09%	2.82%	3.12%	5.85%
North West	1.00%	2.18%	3.16%	2.82%	2.18%	4.94%
West Midlands	1.00%	0.33%	1.33%	2.82%	0.33%	3.14%
Northern	1.00%	0.00%	1.00%	2.82%	0.00%	2.82%
Scotia-South	1.00%	0.46%	1.46%	2.82%	0.46%	3.27%
Scotia-Scotland	1.00%	0.01%	1.02%	2.82%	0.01%	2.83%
Wales and West	1.00%	0.00%	1.00%	2.82%	0.00%	2.82%

Source: Europe Economics

1 The totals in these tables represent the compound effect of the frontier shift and catch up productivity estimates.

28 Using 2006/07 as baseline North West and London are the GDNs that could be expected to have the biggest reductions in operating expenditure, while Northern and Wales and West could be broadly expected to match the scope for frontier shift.

29 Taking the average of 2005/06 and 2006/07 as baseline, the results broadly confirm those reported in the previous Table, with London and North West being the distributors with most scope for efficiency improvement in the next five years. One notable difference is in the estimates for Wales and West which appears as one of the most efficient companies based on 2006/07 performance but is at the lower end of the range using the baseline for both years.

Company specific efficiency improvement: baseline average 2005/06-2006/07

.	Average annual % reductions over five years					
	NWLB			NWUB		
GDN	Frontier shift	Catching up	Total	Frontier shift	Catching up	Total
East of England	1.18%	0.24%	1.42%	3.00%	0.24%	3.23%
London	1.18%	1.61%	2.77%	3.00%	1.61%	4.56%
North West	1.18%	1.58%	2.75%	3.00%	1.58%	4.53%
West Midlands	1.18%	0.00%	1.18%	3.00%	0.00%	3.00%
Northern	1.18%	0.00%	1.18%	3.00%	0.00%	3.00%
Scotia-South	1.18%	0.68%	1.85%	3.00%	0.68%	3.66%
Scotia-Scotland	1.18%	0.19%	1.38%	3.00%	0.19%	3.19%
Wales and West	1.18%	1.35%	2.52%	3.00%	1.35%	4.31%

Source: Europe Economics

30 There are arguments in favour of each baseline. At this stage in the analysis we suggest that Ofgem should give consideration to both sets of results which can be used in comparison with other work which is in hand on the analysis of operating expenditure.

1 INTRODUCTION

1.1 As part of its work on the Gas Distribution Price Control Review (GDCPR) Ofgem commissioned Europe Economics to carry out a top down benchmarking exercise of the efficiency of the gas distribution network companies (GDNs) including benchmarking of total employment costs. This analysis should allow conclusions to be drawn about the potential for future improvements in efficiency. The principal focus was on operating expenditure but we have also provided some estimates of total factor productivity. An initial literature review was carried out which is reported in Appendix 2.

1.2 It was recognised at the start of this work that there were significant data issues which would constrain the scope of this analysis and place limits on the reliability of the conventional approaches to such a benchmarking exercise. We therefore set out to test a range of different forms of analysis and models in order to identify the most appropriate top down productivity measures from which conclusions for future performance could be drawn.

1.3 This report sets out the data sets used in the analysis and provides detailed technical reports on the main forms of benchmarking and estimation carried out. These are:

- Opex benchmarking, including regression analysis, data envelopment analysis and multi-productivity indices and a correlation analysis between these different approaches and models used

- Nature of work comparison estimates of potential for growth in total factor productivity

- Breakdown of opex productivity estimates between shifts in the efficiency frontier and catch-up effects

1.4 Conclusions are drawn on the preferred measures of productivity and the potential for future productivity growth.

2 DATA

2.1 A fundamental stage in any assessment of efficiency is the identification and the gathering of a set of relevant input and output variables. This requires the collection of a significant amount of cost data and other related information.

2.2 Data sources available for our benchmarking assessments were the following: Business Plan Questionnaires (BPQs), which have been completed by each of the Gas Distribution Networks (hereafter GDNs); and supplementary questions and responses and information provided by Ofgem.

2.3 In this section we briefly describe the type of sources we have used in the report for each of the data sets necessary for our study.

2.4 Historic outturn data are available for 2004/05 to 2005/06. However, Ofgem has separately investigated the reliability of the data in 2004/05 and found that the data for that year have been significantly affected by the industry changes at the time, and therefore were not suitable for this analysis. As the aim of the study is to estimate the efficiency factors for the coming price period, the 2004/05 data might not be appropriate as all GDNs were still under National Grid ownership.

2.5 The analysis is therefore based on outturn data for 2005/06 and estimated outturn data for 2006/07. The 2006/07 observations are company projections based on actual data for the first part of the year. Ofgem's agreed adjustments to the BPQ data for these years have been taken into account. It will be possible for the analysis to be updated when the full year data for 2006/07 is available later in the year. All financial data are expressed in 2005/06 prices.

Opex benchmarking

2.6 For the Opex benchmarking the input variables are the controllable operating costs at both GDN and ownership group level. These are the costs submitted in the BPQs subject to a number of accounting adjustments by Ofgem to bring them onto a consistent basis.

2.7 Table 2.1 and Table 2.2 show controllable operating costs per GDN and per group as adjusted by Ofgem.

Table 2.1: Total controllable costs (TCC) per GDN after Ofgem accounting adjustments (£ million)

	Revised controllable costs	
	2005-06	2006-07
East of England	104.6	121.4
London	67.2	84.1
North West	82.5	94.7
West Midlands	55.5	65.4
Northern Gas	76.1	73.8
South	110.2	116.4
Scotland	58.7	63.3
Wales and West	84.0	80.2

Source: Ofgem

Table 2.2: Total controllable costs (TCC) per Group (£ million)

	Revised controllable costs	
	2005-06	2006-07
NGGD	309.7	365.6
Northern Gas	76.1	73.8
Scotia	168.9	179.8
Wales and West	84.0	80.2

Source: Europe Economics calculations

2.8 The set of output variables includes the main cost drivers (as suggested by the literature review) which are: length of the network; gas throughput and breakdown of customers. We have also considered other variables: customer density, the number of customers per kilometre of network and the ratio of non domestic customers to the total number of customers.

2.9 Table 2.3 and Table 2.4 show the network length per GDN and per Group. These show the total network length at all pressure levels

Table 2.3: Network length per GDN (Km)

	2005-06	2006-07
East of England	49,346	49,409
London	23,373	23,410
North West	34,521	34,550
West Midlands	24,199	24,220
Northern Gas	36,670	36,664
South	49,959	50,036
Scotland	22,982	23,061
Wales and West	34,634	34,696

Source: Ofgem, BPQs

Table 2.4: Network length per Group (Km)

	2005-06	2006-07
NGGDD	131,438	131,588
Northern Gas	36,670	36,664
Scotia	72,941	73,097
Wales and West	34,634	34,696

Source: Ofgem, BPQs

2.10 Gas throughput has been collected on a Local Distribution Zone (LDZ) basis from the BPQ. Throughput for each GDN has then been calculated by aggregating throughput of the relevant LDZs.

2.11 Table 2.5 and Table 2.6 show the level of throughput of gas distributed per GDN and Group.

Table 2.5: Throughput of gas distributed per GDN (GWh)

	2005-06	2006-07
East of England	130,860	127,278
London	68,357	64,882
North West	87,334	84,438
West Midlands	60,117	57,747
Northern Gas	89,422	88,016
South	123,620	118,868
Scotland	61,226	60,603
Wales and West	81,334	80,578

Source: Ofgem, BPQs

Table 2.6: Throughput of gas distributed per Group (GWh)

	2005-06	2006-07
NGGDD	346,668	334,345
Northern Gas	88,300	88,016
Scotia	184,846	179,471
Wales and West	81,334	80,578

Source: Ofgem, BPQs

2.12 Customer numbers in 2005/06 have been collected from the GDN interruption reports for 2005/06.

2.13 We have used Ofgem estimates for customer numbers for 2006/07. These have been assumed to grow in the same proportion as the growth in the number of services which has been provided in the BPQs. [2]

2.14 Table 2.7 and Table 2.8 show the breakdown of customers per GDN and at group level.

[2] The service used was "connecting pipe per customer" that is assumed to be the best proxy until the actual data becomes available for 2006-07 (update in the summer).

Table 2.7: Total number of customers per GDN

	Total customers		Domestic customers		Non Domestic customers	
	2005-06	2006-07	2005-06	2006-07	2005-06	2006-07
East of England	3,896,365	3,912,923	3,785,511	3,801,598	110,854	111,325
London	2,258,665	2,272,145	2,170,116	2,183,067	88,549	89,077
North West	2,661,270	2,666,170	2,582,472	2,587,227	78,798	78,943
West Midlands	1,925,234	1,930,320	1,869,745	1,874,685	55,489	55,636
Northern Gas	2,480,147	2,489,164	2,408,712	2,416,987	71,435	72,177
South	3,977,807	4,001,567	3,852,875	3,873,919	124,932	127,648
Scotland	1,725,717	1,753,133	1,675,398	1,701,545	50,319	51,587
Wales and West	2,387,595	2,429,009	2,315,129	2,345,320	72,466	83,689

Source: Ofgem, Interruptions report 2005-06 and Ofgem estimations

Table 2.8: Total number of customers per Group

	Total customers		Domestic customers		Non Domestic customers	
	2005-06	2006-07	2005-06	2006-07	2005-06	2006-07
NGGD	10,741,534	10,781,558	10,407,844	10,446,577	333,690	334,981
Northern Gas	2,480,147	2,489,164	2,408,712	2,416,987	71,435	72,177
Scotia	5,703,523	5,754,700	5,528,273	5,575,464	175,251	179,235
Wales and West	2,387,595	2,429,009	2,315,129	2,345,320	72,466	83,689

Source: Ofgem, Interruptions report 2005-06 and Ofgem estimations

Productivity data for the Nature of Work comparison

2.15 The data source used for the nature of work analysis is the National Institute Sectoral Productivity (NISEC02) dataset. The productivity dataset published by the National Institute of Economic and Social Research (NIESR) contains sectoral data for the UK (as well as for Germany, France, Japan and the US). The data set was updated in 2002. It covers the period from 1950 to 1999 and includes up to 30 sub-sectors of the economy in addition to whole economy aggregates.

2.16 The NISEC02 dataset includes data for labour productivity, capital stock and total factor productivity (TFP). The dataset allows for calculation of rates of change, as well as allowing comparison of levels of efficiency across countries. Considerable care is taken to secure a high degree of comparability of data both within country and across countries. However, it should be noted that data are less comparable for the periods prior to 1989 because it was only possible to make substantial and necessary revisions for the later years.

2.17 The labour productivity series are based on sectoral GDP at market and constant prices and total hours worked, with establishment based (as opposed to household surveys) data on hours worked being the source. Productivity levels are then converted into the same currency using 1996 PPP figure extracted from OECD and Eurostat.

2.18 Capital stock is calculated for six asset types. In doing so, the perpetual inventory method (PIM) is employed in accordance with the methodology proposed by the US Bureau of Economic Analysis. Capital stock estimates are adjusted for changes in composition.

2.19 When determining TFP, it is necessary to attribute appropriate weights between labour productivity and capital per hour. The dataset uses the share of labour compensation in GDP for labour (based on total non-wage labour costs including that for self-employed labour and one minus the share of labour in GDP for capital).

2.20 The NISEC02 dataset provides UK TFP data for the sectors listed in Table 4.1 below. We have included some of the manufacturing sub sectors into the analysis as their nature of work might reflect more closely, albeit still at a high-level, some components of the gas distribution industry than “manufacturing” as a whole.

3 OPEX BENCHMARKING

Introduction

3.1 The aim of this chapter is to explore alternative techniques for benchmarking the operating expenditure of the eight GDNs, to present the results of the analysis and assess the most robust approach.

3.2 Relative efficiency can be estimated with different techniques and methodologies, ranging from simple unit cost ratios to index number techniques, data envelopment analysis (DEA) and regression analysis (RA).

3.3 From the academic literature it is clear that different efficiency estimation methodologies can yield significantly different results in terms of both efficiency rankings and levels.[3]

3.4 There is no single methodology that is considered clearly superior to another; however, some techniques are likely to provide more reliable results than others under particular circumstances.

3.5 While it is good practice to undertake efficiency analysis with more than one technique – because if different methodologies yield comparable results the whole reliability of the efficiency evaluation exercise would be enhanced- it is also important to assess carefully whether the conditions necessary for a particular methodology to deliver meaningful results are in place.

3.6 In this chapter we set out:

(a) The main efficiency measurement techniques considered in this Report, focusing on their relative strengths and weaknesses given the available dataset for the eight GDNs "observed" in two separate years, 2005/06 and 2006/07);

(b) The results, in terms of efficiency rankings and levels for each of the methodologies considered;

(c) The results of the correlation analysis between the efficiency rankings produced by the different methodologies; and

(d) Our assessment of the relative efficiency of the eight GDNs.

---

3 See, for example, A.N. Berger (1998).

Methodology

Unit cost ratios

3.7 The simplest category of benchmarking techniques that is available in the literature is simple partial ratios, such as opex per unit of gas distributed, opex per customer, opex per km of network.

3.8 The main advantage of these partial efficiency indicators is that they are easy to compute and to understand and, as such, they are often used to form a preliminary view of company efficiency. Furthermore, when different indicators provide similar answers in terms of efficiency levels and rankings one can place some confidence in the results of the whole benchmarking exercise.

3.9 The main disadvantage of these simple efficiency indicators is that they provide only a partial picture of companies' relative efficiency.

3.10 For instance, analysing the opex efficiency of a group of companies could give different results depending on the scaling variable: for instance, different companies could have different values in their opex to volume ratios not only because of differences in efficiencies, but also because of differences in population densities or in customer composition (which could be relatively more skewed towards, for instance, industrial users than households). Furthermore, if scale economies were important, an analysis based on unit cost ratios (whatever ratio one might consider) would not take differences in scale into account because as it cannot control for any other factor that might impact on a GDN's costs.

3.11 In the following sections, we report unit controllable costs ratios as a preliminary analysis, but we do not expect any strong correlation with the results produced by regression analysis and DEA.

Regression analysis

3.12 Regression analysis has been widely used in regulatory analysis in the UK, including use by Ofgem in the electricity distribution sector, Ofwat in the case of the water and sewerage industry and by Ofcom in the telecommunication sector.[4]

3.13 There are two main regression methodologies that can be used to benchmark the efficiency of a group of companies, namely Corrected Ordinary Least Squares (COLS) and Stochastic Frontier Analysis (SFA).

3.14 These are both regression based methodologies, and therefore share the advantages and limitations of regression analysis.

4 In the latter case, Ofcom has benchmarked BT against a sample of US large telecommunication providers.

3.15 Their main advantage, with respect to DEA, is that they are statistical techniques and, therefore, unlike DEA, they allow the significance of cost drivers to be tested.

3.16 Cost drivers are initially selected on the basis of engineering knowledge of the industry, previous study and data availability. In general, researchers follow the general to specific methodology, which amounts to 'over-fitting' the model, by including all cost drivers that, a priori, could be expected to affect costs and eliminating at each step the most statistically insignificant. At the same time ensuring that the model passes conventional statistical tests such as those designed to investigate the absence of heteroskedasticity, functional form mis-specification and non-normality of the error term. At each step, the previously eliminated cost drivers are re-inserted into the model, and the process stops when only the significant cost drivers remain in the model.

3.17 Secondly, in the case of regression analysis it is possible, by inspecting the residuals with appropriate statistical techniques, to detect the existence of outliers.

3.18 COLS is used to estimate deterministic frontiers: the methodology amounts to regressing operating expenditure on a set of cost drivers using the methodology of Ordinary Least Squares (OLS) and then adjusting the constant term in order to have only one firm (the one with the largest negative residual) with an adjusted residual of zero and all the others (which are deemed as inefficient) with a positive residual.

3.19 The main advantage of COLS is that it is relatively simple to apply and to understand. This is the methodology used by Ofgem in the electricity distribution sector and by Ofwat in the case of the water and sewerage sectors.

3.20 However, the residual is entirely attributed to inefficiency, with no role for random noise, measurement errors, omitted variables, or company specific shocks (e.g. exceptional costs).

3.21 The obvious consequence is that the COLS procedure might produce efficiency estimates and, therefore, targets that are biased.

3.22 However some statistical tests can be used to ensure that the most important misspecification problems, -such as the wrong functional form imposed in the model, the existence of heteroskedasticity and non-normality of the error terms, or the presence of outliers — which could contribute to what is in fact an unachievable industry frontier — are detected and appropriate corrections undertaken.

3.23 Notwithstanding the statistical tests that can be performed to increase the degree of confidence in the efficiency estimates produced by the COLS approach, the assumption that the estimated residuals can be fully interpreted as inefficiency is unsatisfactory and likely not to be met in practice.

3.24 This issue can be tackled in two ways.

- The first approach is to use the SFA methodology, whereby the overall residual is split into two components, one accounting for inefficiency and the other for random noise. However, this is done at the cost of specifying a more complex model which involves distributional assumptions for both the noise and the inefficiency component of the error term. SFA models require large sample sizes to be implemented, even larger than those that researchers would consider acceptable for the use of COLS. Given the small sample available in this study, SFA does not appear to be feasible methodology to adopt

- The second approach to deal with the existence of potential noise and measurement errors not accounted for in the case of COLS is to recognise that the noise does exist and that, therefore, for most companies it might not be feasible to catch up with the industry frontier identified by the COLS method.

- A pragmatic approach is to take the upper quartile firm as defining the industry frontier, considering each firm with an efficiency level higher than the upper quartile as fully efficient and rescaling the efficiency levels of the remaining companies with respect to the upper quartile efficiency level rather than with respect to the frontier identified by the COLS method. This results in a more conservative estimate for the scope of catch up.

- This is the method that Ofgem has undertaken in the case of the electricity distribution industry. Ofwat has followed a similar approach, "discounting" the residuals identified with the COLS approach by 10 per cent in the case of the water sector and 20 per cent in the case of the sewerage sector.

Data Envelopment Analysis

3.25 Data Envelopment Analysis (DEA) employs linear programming techniques to build a frontier which envelops the input-output data of a set of companies which use those inputs to produce those outputs using a common technology.[5]

3.26 In its simplest form, DEA recognises a firm as fully efficient if no firm or linear combination of firms is able to produce more of each output (given the level of input) or to use less of each input (given the level of output).

3.27 A DEA model can be input oriented (i.e. it measures efficiency by considering the feasible proportional reductions in all inputs, given the level of outputs) or output oriented (i.e. it

---

[5] Linear programming techniques have been used in the efficiency literature since the pioneering work by Charnes, Cooper and Rhodes (1978). A linear programming technique is a model which has an objective function (such as an efficiency measure like a ratio of outputs to inputs) which is a linear function of the decision variables of the problem (such as the inputs used in the production process), and that has to be maximised or minimised.

measures efficiency by considering the feasible proportional increase in all outputs, given the level of inputs). If the output level is largely outside the control of the firm, as in the case of the GDNs, an input orientation is to be preferred.[6]

3.28 There are two basic DEA models, constant returns to scale (CRS-DEA) and variable returns to scale (VRS-DEA). The first model would only be appropriate if all GDNs were operating at the efficient scale of operations: if this is not the case, then the efficiency estimates would be affected by differences in scale.

3.29 The VRS-DEA model, unlike CRS-DEA, compares each firm only with firms of similar size, therefore, by construction, it produces efficiency scores that are, on average, no lower than these produced by the CRS-DEA model.

3.30 The main advantages of DEA are:

– It can in principle easily accommodate many inputs and outputs.

– It is less reliant on the availability of a rich sample size than some regression based approaches like stochastic frontier analysis.

– It is a non-parametric methodology and therefore does not require the imposition of any strong a priori restriction on the functional form that links costs to outputs and cost drivers, as is implicit in conventional regression analysis.[7]

– DEA allows for consideration of the impact of environmental variables on efficiency. This can be done by inserting environmental variables directly into the linear program as additional non-controllable inputs and outputs. Alternatively, the impact of non-controllable environmental variables can be analysed with a two-step methodology which consists in regressing DEA efficiency scores on environmental variables.

– It is not subject to the multi-collinearity problems which often affect regression analysis using small samples.[8]

---

6 It should be noted that the efficiency rankings of the company would be preserved by models with different orientation, although the efficiency estimates would not (if a variable returns to scale model is used).

7 Functional form means the mathematical relationship that links inputs to outputs (or outputs and cost drivers to costs). For instance, simple regression models assume that costs and outputs and cost drivers are linearly (or log-linearly) related.

8 In regression analysis, there are multicollinearity problems when two or more regressors are strongly correlated with each other. Especially in small samples it is difficult to find enough variation in the data to allow the identification of the coefficients for those regressors affected by multicollinearity problems. For instance, it might be difficult to estimate the separate effects that customers, volumes of gas and length of main have on operating expenditure given the quite high correlation among the three variables.

3.31 DEA also has limitations:

- It is not a statistical technique. Therefore, any hypothesis testing is precluded: the implication of this is that we cannot test whether or not an output variable included in the analysis is relevant or not to explaining the companies' costs. The choice of the variables to consider in the analysis depends, ultimately, on engineering knowledge of the industry, on previous studies or, as in this report, on the results of the regression analysis.

- It does not allow for measurement errors: each firm's distance from the frontier is interpreted as inefficiency, with no role for random noise or measurement error. Measurement errors and noise can have a strong impact on the shape and the position of the frontier and, therefore, on the efficiency scores computed for each company.[9]

- DEA tends to put each firm in the "best possible light": in practical terms, DEA tends to compare each firm only with a subset of firms that are "similar" to the firm under analysis.[10] When the number of firms in the sample is low, it is possible that a DEA analysis would tend to identify many firms as unique under some dimensions and, therefore, to consider each of them as fully efficient. When the sample size is low, and the number of inputs considered relatively large, most of the firms would tend to appear fully efficient (as the likelihood that each firm will dominate all the others in relation to a particular input or output gets larger with the number of inputs and outputs included in the analysis). This problem is likely to be particularly severe in the case of the VRS model, as each GDN would be benchmarked only against firms of similar size.

- DEA is particularly sensitive to the potential exclusion of some input/output from the analysis.

Multilateral indices of productivity

3.32 Index number techniques can be used to compute multilateral indices of productivity, which are a refinement of unit cost ratios. They can combine multiple inputs and outputs to provide an overall assessment of the relative productivity level of different firms.

3.33 For instance, it is possible to compute an index of the relative productivity of total controllable costs for the GDNs by using total controllable costs as the only input to production with gas volumes, number of connections and network length, appropriately

---

9 For example, at least in the conventional DEA models used in the regulatory practice, it is not possible to conduct statistical analysis aimed at assessing whether or not a particularly high or low efficiency scores should be considered as outliers, as in conventional regression analysis.

10 See Coelli (1998)

weighted, as the outputs.[11] This approach has recently been used in New Zealand by the Commerce Commission to benchmark the New Zealand transmitters and distributors of both gas and electricity.

3.34 In this report, we have constructed multilateral indices of productivity using controllable operating costs. As such, it excludes capital inputs and, does not cover total factor productivity.

3.35 The main advantages of using multilateral indices of productivity are that:

- It is possible to accommodate, at least in part, differences in the companies' operating environment, considering network length, the number of connections and volumes delivered, so that companies that operate in areas with low energy density or low population density are not unfairly penalised.

- The robustness of index number theory does not rest on the availability of a large sample size. It is therefore particularly relevant in this case, where the sample size is made up of only eight firms observed over a period of two years.

3.36 However, index number techniques also have a number of drawbacks:

- They are non-statistical and therefore do not allow one to take into consideration the uncertainty which surrounds the productivity and efficiency results.

- They do not allow one to control for other variables that might impact on costs and, therefore, on efficiency. For example, they do not control for the impact that scale economies might have on costs.

- Finally, index number techniques require weighting the inputs and outputs considered in the analysis: in our case, given the choice of considering one input only, we have to decide on the appropriate output weights. We cannot separately identify the shares of revenue attributable to volumes, customer numbers and network length, respectively. An alternative that is commonly adopted is to use the cost elasticity of volumes, customers and network length derived from econometric estimates: this information should be derived, ideally, from econometric work undertaken for the sample under consideration. If this is not possible, as is the case here, econometric estimates from other countries could be used. However, this may bias the results as the relationship between costs and volumes of gas distributed, customer served and network length can depend on sample specific characteristics.

---

[11] The theoretical underpinning of this methodology is the multilateral total factor productivity index developed originally by Caves, Christensen and Dewert (1982) and used in many studies seeking to compare the relative levels of productivity of different firms or countries.

3.37 To sum up, the index number approach provides an improvement on the simple comparison of unit cost ratios. However, it also shares some of its drawbacks. As a consequence, it may provide a useful sense check on the results derived from more advanced techniques like Data Envelopment Analysis (DEA) and regression analysis, especially considering that the sample size of the GB GDNs is small.

Unit controllable costs: results

3.38 The scope of the unit cost analysis is to investigate at preliminary level whether there are differences in operating unit costs that may contribute to benchmarking the efficiency of the industry over the time.

3.39 Unit controllable costs are calculated as partial ratios of total controllable costs over three different indicators which are the main output variables thought to be appropriate for the gas distribution sector. These are: a) network length, b) volumes of gas distributed; and c) total number of customers.

3.40 Table 3.1 to 3.3 show controllable unit costs for the gas distribution services in 2005/06 and 2006/07, ranked at 2005/06 in ascending order of costs per kilometre of main, costs per throughput of gas distributed in GWh and costs per number of customers served for GDNs owner levels. Rankings in both years are shown. Figures 3.1 to 3.3 represent these graphically.

Table 3.1: GDN's unit controllable costs per network length (£ /Km)

	2005-06	Rank	2006-07	Rank	Rate of change
Northern Gas	2,075	1	2,013	1	-3%
EoE	2,119	2	2,457	4	14%
South	2,206	3	2,327	3	5%
West Midlands	2,292	4	2,700	6	15%
North West	2,389	5	2,741	7	13%
Wales and West	2,427	6	2,311	2	-5%
Scotland	2,556	7	2,746	8	7%
London	2,876	8	3,595	5	20%

Source: Europe Economics

Table 3.2: GDN's unit cost per customers (£)

	2005-06	Rank	2006-07	Rank	Rate of change
EoE	27	1	31	3	13%
South	28	2	29	1	5%
West Midlands	29	3	34	5	15%
London	30	4	37	8	20%
Northern Gas	31	5	30	2	-3%
North West	31	6	35.5	6	13%
Scotland	34	7	36.1	7	6%
Wales and West	35	8	33	4	-7%

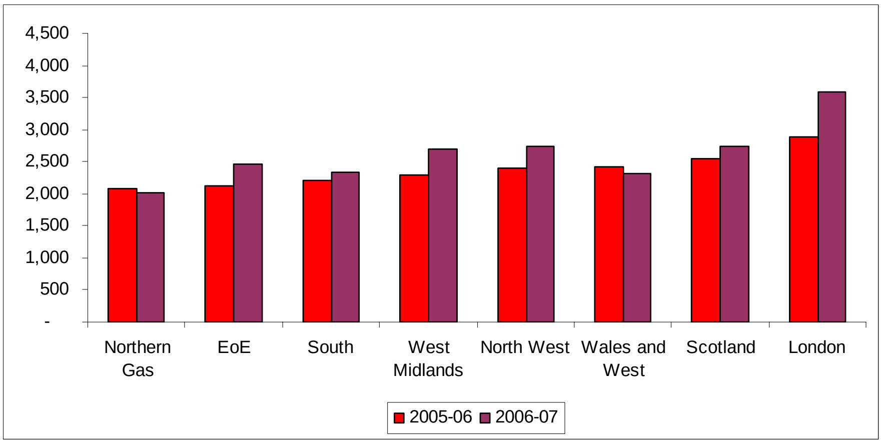
Source: Europe Economics

Table 3.3: GDN's unit cost per throughput (£ / GWh)

	2005-06	Rank	2006-07	Rank	Rate of change
EoE	799	1	954	2	16%
Northern Gas	851	2	838	1	-3%
South	891	3	980	3	9%
West Midlands	923	4	1,132	7	19%
North West	944	5	1,122	6	16%
Scotland	959	6	1,045	5	8%
London	983	7	1,297	8	24%
Wales and West	1,033	8	995	4	-4%

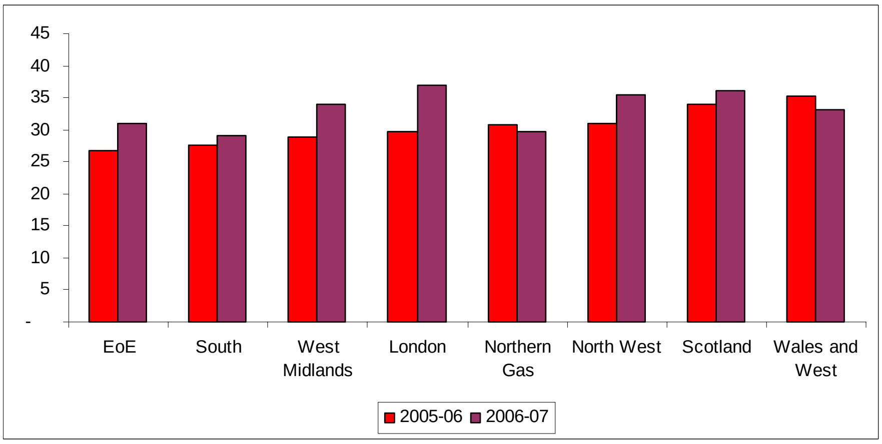
Source: Europe Economics

Figure 3.1: GDN's unit cost per network length (£ / Km)



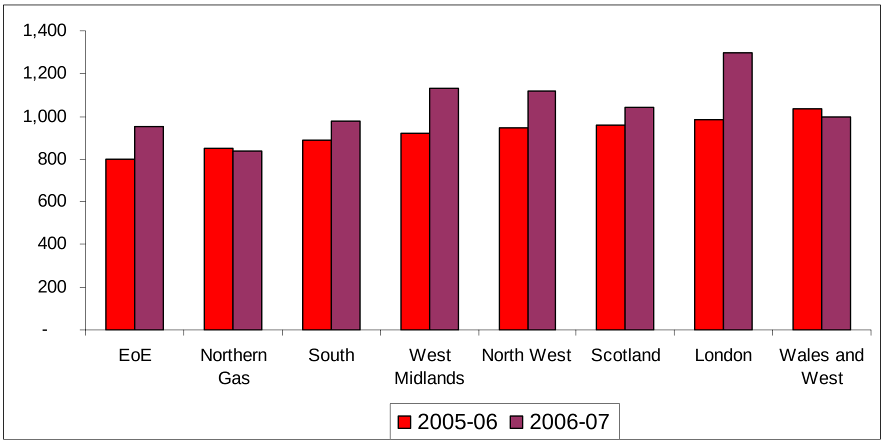
Source: Europe Economics

Figure 3.2: GDN's unit cost per customers (£)



Source: Europe Economics

Figure 3.3: GDN's unit cost per throughput (£ / GWh)



Source: Europe Economics

Level of unit costs per GDNs

3.41 In 2005/06:

- Unit costs per network length were the lowest in Northern Gas (2,075 £/Km) and highest in London (2,876 £/Km);
- The lowest level of unit controllable costs per customer was £27 per customer in East of England. Wales and West had the highest level of unit costs per customer.
- Wales and West was the GDN with the highest level of unit costs per volume of gas distributed while East of England had the lowest rate. All GDNs had relatively high level of unit costs per throughput of gas.

Variation of unit costs per GDN

3.42 At GDN level unit costs generally increased between 2005/06 and 2006/07 (with the exception of Northern Gas and Wales and West).

- In 2006/07 London had the highest level of cost variation per km as well as the highest unit cost per network length while Northern Gas and Wales and West were not only the GDNs to reduce their level of unit costs but also the one with the highest level of efficiency.
- London had the largest increase in controllable costs per customer over the two years and also the highest level of costs in 2006/07 while South was the GDN with the lowest unit costs per customer.

– Unit costs per volume of gas distributed significantly increased in 2006/07 for all of the GDNs apart from Wales and West and Northern Gas: the latter had also the lowest level of unit costs. In 2006/07 London had both the highest level of costs per volume of gas delivered and the highest increase on the previous year.

3.43 Unit cost analysis at Group level is set out below. Tables 3.4 to 3.6 show controllable unit costs for the gas distribution services in 2005/06 and 2006/07, ranked at 2005/06 in ascending order of costs for the same output measures as described above. Figures 3.4, to 3.6 represent these graphically.

Table 3.4: Group unit controllable cost per network length (£ / Km)

	2005-06	Rank	2006-07	Rank	Rate of change
Northern Gas	2,075	1	2,013	1	-3%
Scotia	2,316	2	2,459	3	6%
NGGD	2,356	3	2,779	4	15%
Wales and West	2,427	4	2,311	2	-5%

Source: Europe Economics

Table 3.5: Group unit controllable cost per customers (£)

	2005-06	Rank	2006-07	Rank	Rate of change
NGGD	29	1	34	4	15%
Scotia	30	2	31	2	5%
Northern Gas	31	3	30	1	-3%
Wales and West	35	4	33	3	-7%

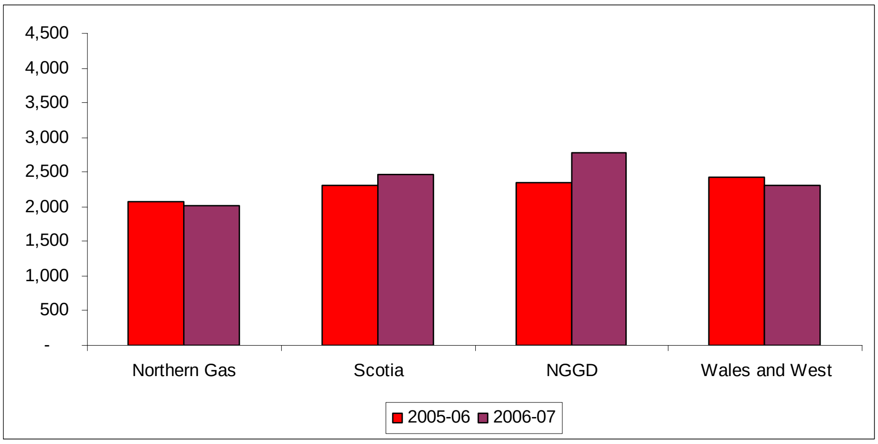
Source: Europe Economics

Table 3.6: Group unit controllable cost per throughput (£ / GWh)

	2005-06	Rank	2006-07	Rank	Rate of change
Northern Gas	851	1	838	1	-3%
NGGD	893	2	1,094	4	18%
Scotia	914	3	1,002	3	9%
Wales and West	1,033	4	995	2	-4%

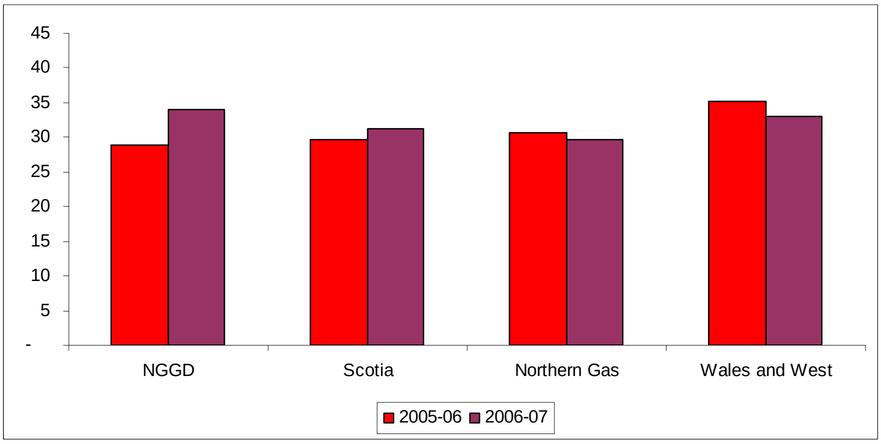
Source: Europe Economics

Figure 3.4: Group unit controllable costs per network length (£ / Km)



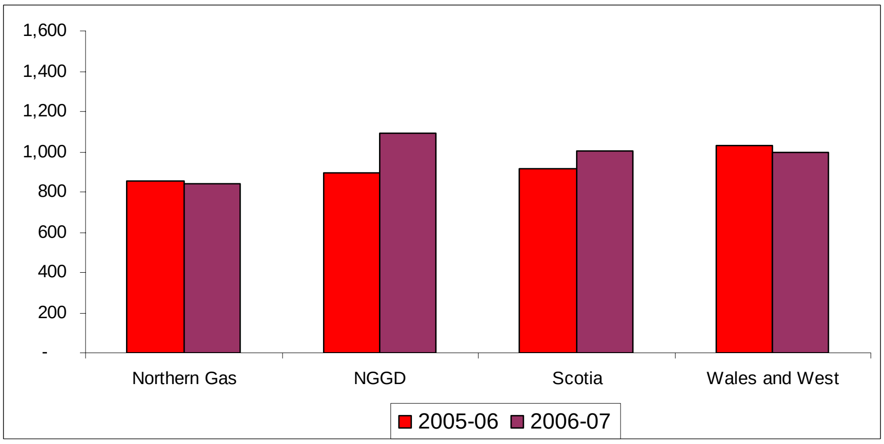
Source: Europe Economics

Figure 3.5: Group unit controllable costs per customers (£)



Source: Europe Economics

Figure 3.6: Group unit controllable cost per throughput (£ / GWh)



Source: Europe Economics

Level of unit costs per Group

3.44 At Group level in 2005/06 Northern Gas has the lowest unit costs on all of the measures considered with exception of unit costs per customers, where NGGD is the most efficient. Wales and West has the highest unit costs in 2005/06. The ranking of Scotia varies depending on the measure used.

Variation of unit costs per Group

3.45 The level of unit costs per ownership group generally increased in 2006/07, with the exception of Wales and West and Northern Gas: for all measures used Northern was the most efficient GDN and NGGD the least in 2006/07.

– Unit controllable costs per network length, customers and volume of gas distributed all followed a similar pattern: for all measures NGGD was the group that increased its controllable cost most over the two years while Wales and West slightly decreased its costs; In 2006/07 Northern Gas was the distributor with the lowest level of unit costs and NGGD the highest.

3.46 As noted above partial productivity measures do not provide a complete view of performance because they do not consider the various trade-offs between inputs and outputs. The rankings of the individual GDNs vary depending on which measure is considered. There have been some significant changes in costs between the two years and as a result there are also changes in the rankings between the two years both for individual GDNs and by Group. At the Group level it appears that there has been some narrowing of the range of unit costs between the highest and lowest. However this is far less apparent at the individual GDN level with an increase in the range on some measures. This suggests that comparison at Group level is likely to hide variations in efficiency.

Regression analysis: results

Introduction

3.47 The regression analysis we have carried out explores the relationship between total controllable costs and a number of possible explanatory variables which are considered likely to be the main cost drivers. Further details on our approach and supporting information on our results are set out in Appendix 1 which provides technical detail.

3.48 The choice of the cost drivers is based on a review of academic papers that have undertaken econometric estimates of gas distribution network cost functions but is constrained by data availability and sample size considerations.

3.49 We have considered volumes of gas distributed (VOL), total number of customers (CUST), network length (LEN), the proportion of non domestic customers (PNDC), (as a proxy for the importance of large users in the customer base of each GDN), and customer density (CD), expressed as the ratio between total customers and network length as potential cost drivers.

3.50 With the exception of the proportion of non domestic customers, all explanatory variables have been entered, as is general practice in econometric analysis of cost functions, in logarithmic form.

3.51 Our analysis uses the Ordinary Least Squares (OLS) technique which we have then adapted using the Corrected OLS (COLS) procedure. The COLS procedures amounts to a correction of the OLS results through a displacement of the constant, so that all companies apart from the one with the largest negative residual have a positive corrected residual.

3.52 The main drawback of this COLS procedure is that each company's residual is ascribed to inefficiency. To limit the possible overestimation of the extent of inefficiency that is present in the sample, one option is to compute relative efficiency with respect to the upper quartile of efficiency, rather than on the basis of the company with the lowest negative residual. We have taken account of the possibility that the COLS procedure may overstate the extent of inefficiency in the sample in drawing the efficiency frontier from the regression results in Section 5.

3.53 A range of tests have been carried out to assess whether the relationships identified meet the underlying conditions for statistical reliability. These tests have not detected any significant violations of these conditions.

Results from the econometric models

3.54 We ran regressions using two sets of regressors. In the first case, we included VOL, CD, and PNDC. The general to specific methodology –once the outliers observations were removed from the analysis[12] - rejected CD and PNCD and therefore we have adopted as model COLS1, a model with VOL as the only regressor. The results are reported in Table 3.7

3.55 The model selected in this case has the volume of gas distributed (expressed in natural logarithms) as the scale variable. The R squared is large (0.87) indicating a good overall fit for the relationship and the statistical tests do not identify major evidence of heteroskedasticity, functional form misspecification or non-normality of the error term which are the key statistical tests.

Table 3.7: Estimation of COLS1

	Ln(TCC)
LnVOL	0.82***
Constant	-4.91***
Adjusted R-sq	0.87
F test	85.79***

Source: Europe Economics

3.56 The volume coefficient is 0.82, implying that there are some economies of scale in this sample with respect to volumes delivered (significant at the 6 per cent level): the existence of scale economies would suggest that even this very simple model, with only

12 The results of the Model with London are reported and discussed in Appendix 1; London's efficiency has been computed with an out of sample prediction.

one regressor, is an improvement over unit cost comparisons, as the latter neglect the impact of scale economies on unit costs.

3.57 The application of the second general to specific methodology, without restrictions on relationship between length and customer numbers, selected the model COLS2 reported in Table 3.8 below. West Midland was found to be a potential outlier in this case and it was therefore dropped from the regression (the impact on GDNs' efficiency is in this case almost negligible, as the correlation coefficient between the efficiency ranking with and without West Midland in the regression is as high as 0.99.

3.58 As we can see, the only regressor selected by the general to specific methodology is CUST, which acts as the scale variable in the model. The R squared for the regression is 0.90. The CUST coefficient is 0.75 and it is significantly lower than 1 (at the 1 per cent level of confidence), suggesting the existence of scale economies in this sample. Statistical tests are generally positive.

Table 3.8: Estimation of COLS2

	Ln(TCC)
LnCUST	0.75***
Constant	-6.69***
Adjusted R-sq	0.90
F test	126.4***

Source: Europe Economics.

3.59 Due to multicollinearity problems we cannot test for the inclusion of customer density in this model.

3.60 Models COLS1 and COLS2 are our preferred econometric models that have been selected applying the general to specific methodology.

3.61 As a further robustness check, we have estimated two additional models, which we have called COLS3 and COLS4.

3.62 These are both based on a similar assumption to that employed by Ofgem in the case of the 2004 electricity distribution price control review, where the only regressor included in the econometric model was a composite scale variable[13], which weighted together volumes of gas distributed, customers and length of mains. Further details on the composition and results for these models are given in the technical annex.

13 The methodology for building the composite scale variable is explained, for instance, in CEPA (2003).

3.63 Referring to the next sections for a rank correlation analysis we note, looking at Table 3.9, that the rankings generated by the different Models tend to be quite similar: for instance, no GDN is ranked in the top two and in the worst two by two different Models.

3.64 There are significant changes in the rankings between the two years, most notably for Wales and West which moves from the bottom to the top of the rankings in the second year and for London –and, to a lesser degree, the other NGG GDNs- whose position deteriorates. We can note that the Model COLS4 produces efficiencies and ranking that are very similar to those generated by Model COLS2: this is a consequence of the fact that in Model COLS4 the composite scale variable gives 86 per cent of weight to customers, which is the only regressor in Model COLS2.

3.65 On average, all models display similar average efficiency levels..

Table 3.9: GDNs' efficiency levels

GDN	Year	COLS1	Rank	COLS2	Rank	COLS3	Rank	COLS4	Rank
East of England	1	0.98	2	0.90	3	0.95	3	0.94	2
London	1	0.89	5	0.93	2	0.84	7	0.93	3
NW	1	0.89	6	0.86	7	0.89	6	0.88	7
West Midland	1	0.97	3	1	1	0.99	1	1	1
Northern	1	0.98	1	0.88	4	0.97	2	0.90	4
Scotia-south	1	0.88	7	0.87	6	0.90	5	0.90	5
Scotia-Scotland	1	0.93	4	0.87	5	0.90	4	0.88	6
Wales and West	1	0.82	8	0.78	8	0.84	8	0.79	8
East of England	2	0.82	4	0.78	6	0.82	6	0.81	6
London	2	0.68	8	0.75	8	0.66	8	0.74	8
NW	2	0.75	7	0.75	7	0.77	7	0.76	7
West Midland	2	0.80	6	0.85	2	0.84	5	0.85	3
Northern	2	1	1	0.91	1	1	1	0.94	1
Scotia-south	2	0.81	5	0.83	3	0.85	3	0.85	2
Scotia-Scotland	2	0.86	2	0.82	5	0.84	4	0.82	5
Wales and West	2	0.86	3	0.82	4	0.88	2	0.84	4
Average	1	0.92		0.89		0.91		0.90	
Average	2	0.82		0.81		0.83		0.82	

Source: Europe Economics

3.66 Table 3.10 below reports the Results at the level of Ownership groups. In order to calculate efficiency at the ownership group level, we have simply weighted the individual GDN's efficiency scores by the volumes of gas distributed.

Table 3.9: Group efficiency levels

Owner	Year	COLS1	Rank	COLS2	Rank	COLS3	Rank	COLS4	Rank
NGGD	1	0.94	2	0.91	1	0.92	2	0.93	1
Northern	1	0.98	1	0.88	2	0.97	1	0.90	2
Scotia	1	0.90	3	0.87	3	0.90	3	0.89	3
Wales and West	1	0.82	4	0.78	4	0.84	4	0.79	4
NGGD	2	0.77	4	0.78	4	0.78	4	0.79	4
Northern	2	1	1	0.91	1	1	1	0.94	1
Scotia	2	0.83	3	0.83	2	0.85	3	0.83	3
Wales and West	2	0.86	2	0.82	3	0.88	2	0.84	2

Source: Europe Economics

3.67 In the first year, all models tend to rank Wales and West low relative to others, while its performance improves in the second year. . The performance of NGGD deteriorates in the second year. The companies overall tend to perform less well in the second year than the first with the exception of Northern and Wales and West. In 2005/06 NGGD and Northern perform best overall.

Data Envelopment Analysis

3.68 Data Envelopment Analysis (DEA) involves the use of linear programming methods to construct a non-parametric frontier over the data in order to calculate efficiencies relative to this frontier.

3.69 Theoretically only efficient companies are on the frontier. Efficiency consists of two components: technical efficiency, which is the ability of a firm to produce the maximum output from a given set of inputs, and allocative efficiency which is the ability of a firm to use the inputs in optimal solutions given their respective prices in the market and the production technology. Economic efficiency is the measure of these two efficiencies combined together.[14]

3.70 As discussed in paragraphs 3.25 -31 there are two types of DEA techniques that can be used both of which are calculated in this study:

– Constant Returns to Scale (CRS): where firms of any size are benchmarked against each other. This measure of efficiency contains both technical and scale efficiency

– Variable Returns to Scale (VRS): where firms of small size are benchmarked against small firms and big firms against big firms.

3.71 In our view the efficiency scores obtained from the VRS model are more relevant for an analysis of the gas distribution sector. This is because it is difficult to change scale of operation in the short run and the frontiers calculated with VRS do not penalise firms that are not operating at the efficient scale level.

3.72 In this report DEA is based on the concept of radial efficiency where the technical efficiency is measured along a ray from the origin to the observed production point: relative proportions of inputs or outputs are held constant. This has the advantage of considering efficiency measures as unit invariant which means that the units of measurement do not change the level of efficiency.[15] We identify the technical inefficiency as a proportional reduction in input usage using a DEA input oriented model.

---

14 This can be represented as a general ratio:

$$\text{efficiency ratio} = \frac{weightedsumofoutputs}{weightedsumofinputs} = \frac{\sum_i q_i y_i}{\sum_j p_j x_j}$$

Where y's and x's are outputs and inputs while q's and p's are weights to be calculated by the DEA technique. The linear problem is to chose weights which maximise the efficiency scores subject to the constraint that the score can not be higher then one. Each company is given a single efficiency score from zero to one and no company gets a score higher then one. The closer the score is one the more efficient the firm.

15 There are also non-radial efficiencies measures, such as additive and maximum average or minimum average which however not being units invariant are better to be avoided.

3.73 A fundamental stage in any assessment of efficiency by DEA is the identification of a set of input and output variables as well as environmental variables which might affect the efficiency.

3.74 In the gas distribution sector there are several potential subsets of variables. For inputs the choice is constrained to the OPEX (total controllable costs) but we have considered three potential outputs, network length, volume of gas delivered, and total number of customers; and two environmental variables, customers density and proportion of residential customers to total customers.

3.75 In contrast to the regression analysis there is no statistical test (decision rule) to determine the significant cost drivers and in this analysis we have chosen them to align with our other analysis. This is in line with other empirical studies in the literature where DEA has been used as a cross-check on results from the regression analysis.

3.76 Therefore, on the basis of the results from the regression analysis we have developed two main models. These are:

- DEA 1: where the main cost driver is the volumes of gas distributed (VOL) and total controllable cost (TCC) is the input measure included. This would correspond to the COLS 1 of the regression analysis.

- DEA 2: where the main cost driver is the number of customers (CUST) and total controllable cost (TCC) is the input measure included. This would correspond to the COLS 2 of the regression analysis.

3.77 We have also calculated efficiency scores for two additional models DEA 3 and DEA 4 which are based on composite variables: adjusted customer one (ADJSTCUST1) and adjusted customers two (ADJUSTCUST2). These are similar to the models used by Ofgem in electricity distribution models. These correspond to COLS3 and COLS4 of the regression analysis:

Table 3.10: DEA main models

	DEA 1	DEA 2	DEA 3	DEA 4
Input	TCC	TCC	TCC	TCC
Output	VOL	CUST	ADJSTCUST1	ADJUSTCUST2

Source: Europe Economics

3.78 Efficiency scores at individual GDN level based on the VRS models are reported in Table 3.11. [16]

Table 3.11: DEA efficiency scores per GDN

		DEA1		DEA2		DEA3		DEA4	
company	Year	VRS	Rank	VRS	Rank	VRS	Rank	VRS	Rank
East of England	1	1	1	1	1	1	1	1	1
London	1	0.91	5	0.95	4	0.84	7	0.94	5
North West	1	0.89	7	0.90	7	0.89	6	0.90	7
West Midland	1	1	1	1	1	1	1	1	1
Northern	1	0.98	3	0.91	6	0.97	3	0.92	6
South	1	0.90	6	1	1	0.95	4	1	1
Scotland	1	0.96	4	0.94	5	0.94	5	0.94	4
Wales and West	1	0.83	8	0.80	8	0.84	8	0.80	8
East of England	2	0.84	5	0.87	4	0.86	5	0.86	4
London	2	0.70	8	0.76	8	0.67	8	0.75	8
North West	2	0.75	7	0.78	7	0.77	7	0.78	7
West Midland	2	0.85	4	0.85	5	0.85	6	0.85	6
Northern	2	1	1	0.94	2	1	1	0.95	1
South	2	0.82	6	1	1	0.89	2	0.94	2
Scotland	2	0.88	2	0.88	3	0.88	4	0.88	3
Wales and West	2	0.86	3	0.85	6	0.88	3	0.85	5
Average	1	0.93		0.94		0.93		0.94	
Average	2	0.84		0.87		0.85		0.86	

Source: Europe Economics

16 DEA calculations can be conducted using a number of different computer programmes. We use Efficiency Measurement System (EMS). produced by H. Scheel University of Dortmund, Germany.

3.79 Table 3.12 reports the results at the ownership group level rather than at the individual GDN level. These are the sum of DEA variable returns to scale calculated at firm level weighted for each group on the basis of volumes of gas distributed by the GDNs.

Table 3.12: DEA efficiency scores per ownership group

Group	Year	DEA1		DEA2		DEA3		DEA4	
		VRS	Rank	VRS	Rank	VRS	Rank	VRS	Rank
NGGD	1	0.93	2	0.95	2	0.92	3	0.94	2
Northern	1	0.98	1	0.91	3	0.97	1	0.92	3
Scotia	1	0.92	3	0.98	1	0.95	2	0.98	1
Wales and West	1	0.83	4	0.80	4	0.84	4	0.80	4
NGGD	2	0.79	4	0.82	4	0.80	4	0.82	4
Northern	2	1	1	0.94	2	1	1	0.95	1
Scotia	2	0.84	3	0.96	1	0.88	2	0.92	2
Wales and West	2	0.86	2	0.85	3	0.88	2	0.85	3
Average	1	0.92		0.91		0.92		0.91	
Average	2	0.87		0.89		0.89		0.89	

Source: Europe Economics

3.80 The relative efficiency measures calculated reflect the excess (if any) in inputs usage given the output levels and the environmental conditions in which each company operates. Scores of 1 indicate companies on the frontier operating at the most efficient level, scores below 1 indicate the potential for cost savings from moving to the efficiency frontier.

Summary of DEA results

3.81 In general terms, the scores in Tables 3.12 and 3.13 tend to show that while in the first year a number of companies are on or close to the frontier, in the second year almost all companies decrease their relative level of efficiency within the industry allowing scope for higher cost saving in 2006/07 than in 2005/06

3.82 In 2005/06 the rankings of most GDNs are broadly similar for all models and the average efficiencies, as we noted, are higher than in 2006/07.

3.83 With DEA 1 there are 3 companies operating on the frontier, two in the first year and one in the second. However none of the fully efficient GDNs in 2005/06 remains on the frontier in the second year: Northern becomes the most efficient and London is not only the distributor with the largest decrease in its level of efficiency over the two years, but also the least efficient in 2006/07. East of England and West Midland also move away from the frontier performing less well in the second year while Wales and West's performance shows a slight improvement.

3.84 With DEA 2 the number of fully efficient firms increases to four: these are East of England, West Midland and South in 2005/06 and only South in 2006/07.

3.85 In the second year East of England and West Midland performance declined with scores of 0.87 and 0.85 respectively. In both years Wales and West is not particularly efficient but it improves its performance in the second year. In DEA2, as in model 1, a comparison of performance with the other companies shows that London has the lowest efficiency score in 2006/07.

3.86 In DEA 3 and DEA 4 the efficiency scores are generally similar to each other and to DEA1 and DEA2. Efficiency scores, however, are slightly lower in the second year.

3.87 All the distributors that operate within NGGD group in DEA3 decrease their level of efficiency in 2006/07: London has the largest decrease in efficiency and it is also the worst performer over the two years. North West is relatively inefficient in both years.

3.88 In both DEA 3 and DEA 4 Wales and West increased its efficiency in the second year as in the other models. In DEA 4, East of England, West Midland and South are fully efficient in 2005/06 while London appears as the worst performer. In the second year both DEA 3 and DEA 4 show Northern Gas on the frontier.

3.89 For all models in 2005/06 East of England and West Midland are always on the frontier and Wales and West is always the least efficient. Average levels of efficiency are similar in all models and show a fall in 2006/07. In 2006/07 London is the least and Northern Gas the most efficient GDN (with exception of DEA 2 where it is South that lies on the frontier). .

3.90 At Group level in 2005/06 average efficiencies remain quite high for Northern Gas in DEA1 and DEA 3 and at a low level for Wales and West in all models. Variations in the second year are evident especially in terms of the worst performers: NGGD appears as the least efficient, given the decreasing levels of efficiency among the GDNs.

3.91 A number of considerations need to be taken into account in drawing conclusions from this work.

3.92 First of all it is important to note that, given the asymmetric information on the production function, the true efficiency of a company is unobservable and it is not possible to verify the accuracy and the quality of the fitted frontier because of the difficulty of applying hypothesis testing to DEA results.

3.93 Secondly, the very small sample size available for this report limits the reliability of our results. Small sample size is problematic for both RA and DEA but is a particular concern with DEA which, as a non-parametric method, is less likely to be efficient with a small sample.

3.94 DEA as a general characteristic tends to put the companies in the best of possible lights. This feature is enhanced in the presence of outliers DEA is very sensitive to the presence of outliers in any one or more cost drivers used as dataset. This is because there is the possibility that it labels as fully efficient a company if it does not find any other with similar feature. With a very small sample size there is a tendency for all firms to be shown to be relatively efficient. This is most obvious in the results shown above for DEA2

Multilateral Indices of Productivity

3.95 A different approach to assessing the relative levels of productivity across firms is the multilateral productivity index developed by Caves, Christensen and Diewert (CCD) to compare the levels of total factor productivity of different cross sectional observations. [17]

3.96 The CCD model has been developed and used to assess the relative levels of total factor productivity. However, as the benchmarking exercise is limited to opex, and in order to ensure comparability with the DEA and regression results, we have adapted this model to assess the opex productivity levels of the GDNs.

3.97 This is achieved by assuming that the GDNs use a single input, total controllable costs, to produce up to three outputs, volumes of gas distributed, number of customers and network length.

3.98 To aggregate the outputs in a total output index using indexing procedures we have to allocate a weight to each output. One way of doing this is using econometrics to estimate the relative shares of cost elasticities derived from an econometric cost function. In this case, however, given the limitation of data, we have used cost elasticity shares derived from previous studies.

---

[17] See Caves, Christensen and Diewert (1982). The main advantage of the CCD index is that it has the important property of transitivity, i.e. the relative levels of productivity of firms a and b do not depend on whether they are compared directly or via a third firm, c. The CCD index can be represented, in this case, as:

$$\ln TFP_m / TFP_n = \sum_i (R_{im} + R_i^*)(\ln y_{im} - \ln y_i^*)/2 -$$
$$\sum_i (R_{in} + R_i^*)(\ln y_{in} - \ln y_i^*)/2 - \sum_i (\ln x_m - \ln x^*) + \sum_i (\ln x_n - \ln x^*)$$

where y_{im} stands output *i* for GDN *m*, y_{in} stands for output *i* for GDN *n*, y_i^* is the geometric average of output *i* over all GDNs and time periods, R_i^* is the geometric average of the revenue share for output *i* over all GDNs and time periods, R_{im} is the revenue share of output *i* for GDN m which is, in our model, equal to that of GDN *n*, x^* is geometric average of controllable costs over all GDNs and time periods, while x_m and x_n are controllable costs for GDN *m* and *n*, respectively.

This formula gives the relative change in productivity levels between two adjacent units.

3.99 We have built three models. The first one (Model MPI1) is based on the assumption that the GDN produce two outputs: customers and volumes of gas distributed; the weights of which are the relative shares of cost elasticities derived from the Pacific Economics Group's (2001) large sample of US gas distributors over several years. These were respectively 86 per cent for customers and 14 per cent for volumes of gas distributed.[18]

3.100 Model MPI2 has been built by assuming that the GDNs produce three outputs: network length, to which we assigned a weight of 50 per cent, customers and volumes of gas distributed which received a weight of 25 per cent each based on the weights used by Ofgem in devising a composite scale variable for analysing electricity distribution.[19]

3.101 Finally, Model MPI3 is a modification of MPI1 whereby we have adjusted customers and volume figures for each GDN such that all GDNs have a customer density (expressed as the ratio of customer to network length) and an output density (expressed as the ratio of volumes distributed to network length) equal to that of the GDN with values of customer and output density most similar to the sample average, in order to try to put all GDNs on an equal ground in terms of output and customer density. The ad hoc nature of the adjustment could, however, severely affect the reliability of the results from this model.

3.102 The results for Models MPI1, MPI2 and MPI3 are reported in Table 3.13.

18 See Meyrick (2004)

19 See Ofgem (2004) Electricity Distribution Price Control Review.

Table 3.13: GDN's MPI efficiency scores

Owner	Year	Model1	Rank	Model2	Rank	Model3	Rank
East of England	1	1	1	1	1	0.95	2
London	1	0.89	4	0.79	8	0.70	8
North West	1	0.86	6	0.87	5	0.84	5
West Midlands	1	0.92	3	0.91	4	0.88	4
Northern	1	0.88	5	0.96	2	0.97	1
Scotia-South	1	0.96	2	0.95	3	0.91	3
Scotia-Scotland	1	0.79	7	0.820	6	0.79	7
Wales and West	1	0.76	8	0.819	7	0.83	6
East of England	2	0.86	3	0.857	4	0.82	4
London	2	0.71	8	0.63	8	0.56	8
North West	2	0.749	6	0.75	7	0.734	7
West Midlands	2	0.78	5	0.77	5	0.75	5
Northern	2	0.912	1	0.99	1	1	1
Scotia-South	2	0.907	2	0.89	2	0.86	3
Scotia-Scotland	2	0.746	7	0.76	6	0.733	6
Wales and West	2	0.81	4	0.861	3	0.87	2
Average	1	0.88		0.89		0.86	
Average	2	0.81		0.81		0.79	

Source: Europe Economics

3.103 The results of the first two models show a similar level of average productivity across the models in each year.

3.104 At GDNs level the first result to note is that MPI models tend to yield an average level of productivity that is slightly lower than the average efficiency identified by regression-based Models and DEA based models.

3.105 Models MPI2 and MPI3, that take into account the length of the network, either directly as an output, or indirectly by adjusting the customers and volumes figures to give an average customer and volume density to each GDN, result in higher productivity levels for those distributors that have a sparse network, like Wales and West.

3.106 East of England and Northern appear to be among the most efficient GDNs and Wales and West and London the less efficient. There is a general decline in productivity in the second year for most of the GDNs apart from Wales and West and Northern.

3.107 Table 3.14 reports the results by ownership group rather than at the GDN level.

Table 3.14: Group level MPI efficiency scores

Owner	Year	Model1	Rank	Model2	Rank	Model3	Rank
NGGD	1	0.92	2	0.85	3	1	1
Northern	1	0.97	1	0.97	1	0.95	3
Scotia	1	0.91	3	0.87	2	0.97	2
Wales and West	1	0.83	4	0.83	4	0.83	4
NGGD	2	0.77	4	0.72	4	0.85	4
Northern	2	1	1	1	1	0.99	1
Scotia	2	0.85	3	0.82	3	0.92	2
Wales and West	2	0.87	2	0.87	2	0.88	3
Average	1	0.91		0.88		0.95	
Average	2	0.87		0.85		0.92	

Source: Europe Economics

3.108 The figures suggest a small variability of rankings across models: Northern appears as the most efficient group with the exception of Model 3 where NGGD performs particularly well. In 2006/07, however, NGGD's performance would appear to deteriorate significantly.

3.109 Wales and West is the lowest performer in year 1 in all models although its performance does improve year on year in Model 1 and 2.

3.110 Overall these measures of performance are very sensitive to the assumptions made about the weights to be attached to the individual output variables. In addition these indices do not control for the existence of economies of scale. At best therefore they should only be used as a cross-check on the values and rankings identified in the other forms of analysis.

Rank correlation coefficients analysis

3.111 In general, the conclusions that can be drawn on opex efficiency will be stronger if different estimation methodologies and models yield similar results in terms of efficiency levels and rankings. In particular, if different models and methodologies tend to identify the same GDNs as the best and worst performers this could be considered as evidence that the rankings do not depend significantly on the exact methodology or model chosen for the efficiency evaluation exercise. However these results need to be treated with caution given other interdependencies between the models.

3.112 **Error! Reference source not found.** reports the rank correlation coefficients for the results from the different Models we have estimated in the sections above.[20]

Table 3.16: Spearman rank correlation coefficients

	COLS1	COLS2	COLS3	COLS4	MPI1	MPI2	MPI3	DEA1	DEA2	DEA3	DEA4
COLS1	1.00										
COLS2	0.85***	1.00									
COLS3	0.89***	0.81***	1.00								
COLS4	0.87***	0.98***	0.86***	1.00							
MPI1	0.68***	0.78***	0.75***	0.86***	1.00						
MPI2	0.77***	0.67***	0.87***	0.76***	0.89***	1.00					
MPI3	0.69***	0.53**	0.86***	0.64***	0.76***	0.95***	1.00				
DEA1	0.96***	0.92***	0.89***	0.92***	0.70***	0.72***	0.62***	1.00			
DEA2	0.64***	0.80***	0.70**	0.84***	0.87***	0.71***	0.52**	0.69***	1.00		
DEA3	0.85***	0.76***	0.96***	0.84***	0.79***	0.91***	0.85***	0.86***	0.76***	1.00	
DEA4	0.80***	0.87***	0.81***	0.91***	0.88***	0.77***	0.62**	0.84***	0.95***	0.86***	1.00

Source: Europe Economics

3.113 Statistically significantly positive and, possibly, near to one, correlation coefficients can be interpreted as evidence that different estimation methodologies and models tend to produce similar efficiency rankings.

3.114 The results presented in **Error! Reference source not found.** are mixed. All the correlation coefficients are statistically significant suggesting that on this basis alone there is little to choose between the models. There is a higher degree of correlation between the rankings in the four regression models than between the DEA and MPI estimates. The correlation between the regression models and the DEA results is higher than it is with the MPI analysis.

Conclusion on alternative models

3.115 Each of the approaches to measuring and comparing the efficiency of the GDNs has strengths and weaknesses which have been discussed above.

3.116 Comparison of simple measures of unit cost provides a straightforward starting point but it is clear that rankings vary depending on the output measure considered. The more sophisticated use of multi-productivity indices brings together a wider range of influences on efficiency in a single measure but the rankings are very dependent on the weighting assumptions which, in this analysis are derived from external sources and are not specific to the GDNs. In addition they do not make any allowance for economies of scale.

[20] We have used the Spearman rank correlation coefficient which is a non-parametric measure of association between two variables which can take values between -1 and 1, indicating perfect negative and positive associations, respectively. In Table 3.18 ***, ** and * stands for statistical significance at 1 per cent, 5 per cent and 10 per cent levels, respectively.

3.117 The index based comparisons of productivity which have been summarised above are therefore, at best, useful as a starting guide to relative performance but do not provide a basis for assessing the scope for efficiency improvements.

3.118 DEA models are, in principle, better suited to this task. They are used to estimate an efficiency frontier against which the performance of individual companies can be judged. However this approach is not well suited to analysis of small samples, such as we have in this case. With small samples DEA has a tendency to group firms with similar characteristics at or close to the efficiency frontier. In addition it is not possible to use statistical tests to assess the reliability of findings. In these circumstances it is best to use the DEA analysis alongside the regression analysis.

3.119 The small sample size issue can also be a significant problem in the case of regression analysis. However, the regression results set out above show a high level of 'fit' and are robust to the most serious misspecification problems that could invalidate the results. We have carried out analysis to test whether test whether results were unduly influenced by one or two individual observations or 'outliers'. Adjustments have been made to adjust the models for outlier effects.

3.120 Therefore, while all the methodologies considered in this report have pros and cons, we believe that, all in all, regression analysis is the methodology that should be considered as the most reliable and robust in this case.

3.121 We have estimated four separate regression models. COLS1 and COLS2 have the merit that the significance of the output variables is determined within the model rather than being externally imposed in the composite scale variables used in COLS3 and COLS4.

3.122 There is a high degree of correlation between the rankings produced by these models but our preference is to use the COLS1 and 2 models which make the best use of the available data without the use of predetermined weights between potential explanatory variables.

3.123 COLS1 and COLS2 control for different variables, leading to some difference in rankings, but statistically there is little to choose between them. In our view Ofgem should assign equal weight to the comparative efficiency estimates from these two models.

3.124 In Section 5 we report the company specific efficiency factors derived from COLS1 and COLS2 individually and their average in order to draw conclusions on the scope for individual companies to 'catch-up' with the most efficient over the next five years and on the scope for growth in efficiency for the sector as a whole.

4 NATURE OF WORK PRODUCTIVITY TREND

4.1 This section discusses the use of high-level data to investigate long-term productivity trends possible in gas distribution.

Introduction

4.2 Nature of work comparison is used to estimate the rate of productivity and efficiency growth in a sector, rather than for an individual company. In regulatory price reviews, it is important to understand the scope for the most inefficient companies catching up to the most efficient ones, but also it is crucial to assess the scope for efficiency improvement of the industry efficiency frontier. The time profile of productivity growth in similar sectors can therefore provide useful insights for the scope for efficiency improvements for a given sector.

4.3 This section describes the construction of a comparator benchmark intended to capture the determinant of long-term productivity improvements for the gas distribution industry. This factor is intended to reflect the underlying scope for productivity improvement achievable through technological progress and improvements in working practices. The "nature of work" of gas distribution is an observable variable that should be correlated with these underlying factors.

4.4 We follow the approach developed in Europe Economics (1998 and 2003) of constructing a "nature of work" benchmark that combines productivity estimates in sectors of the economy in proportions that reflect the nature of work of the gas distribution industry. The analysis should be understood as a framework for combining data about potential comparator industries by assuming that the nature of work in each comparator is the main determinant of productivity improvements.

4.5 The output of the analysis is a long-term trend of productivity improvement in the gas distribution industry. This can be combined with the results of the above comparative efficiency analysis to produce an estimate of the frontier shift and company specific catch-up factors, as discussed in Section 5.

4.6 The rest of this section is structured as follows. We first discuss the data sources and possible adjustments to the raw data. Next we investigate how the gas distribution industry could be described in terms of the nature of work undertaken at different stages. The next step is to match the components of the gas distribution to the available comparator sectors, so that the benchmark can be constructed. In doing this, we account for the effect of privatisation on the industries in the data set to get at a trend productivity growth without effects of privatisation. Finally we consider adjustments for capital substitution, input price movements, economies of scale and add a projected ongoing effect of privatisation and regulatory change back into the estimate.

Data Sources

4.7 An important aspect of the nature of work analysis is to use reliable data effectively, rather than risking spurious accuracy in the amalgamation of a complex selection of comparators that must rely on less robust data. The exercise is inherently a very high level one. Nevertheless, the data available data conditions the analysis in terms of industry breakdown and comparator identification.

4.8 The comparator data source used for the analysis is the same as used for Europe Economics (2003): the National Institute Sectoral Productivity (NISEC02) dataset. This is briefly described below.

Overview of the NISEC02 dataset

4.9 The productivity dataset published by NIESR contains sectoral data for the UK (as well as for Germany, France, Japan and the US). The data set which was updated in 2002, spans from 1950 to 1999 and includes up to 30 sub-sectors of the economy in addition to whole economy aggregates. The fact that the dataset only runs up to 1999 is not of great concern due to the long-term nature of the analysis.

4.10 Reliable and broad productivity datasets are hard to come by. The only alternative source of UK productivity data is the University of Groningen productivity database. However, that dataset too runs only up to 2002, and does not include such a long run of historical data as the NISEC02 dataset. Further, the Groningen database does not include total factor productivity (TFP) estimates for different sectors of the UK economy. We have therefore relied on the NIESR data.

4.11 The analysis focuses on TFP because, provided that the data used are reasonable for the intended purpose, this should be the best measure of productivity improvement for the nature of work analysis (even if we then wish to make specific assumptions to move to an operating expenditure figure). We have therefore used the NISEC02 data in the analysis.

4.12 The NISEC02 dataset includes data for labour productivity, capital stock and TFP. The data allow for calculation of rates of change, as well as allowing comparison of levels of relative efficiency across countries. Considerable care is taken to secure a high degree of comparability of data both within country and across countries. However, it should be noted that data are less comparable for the periods prior to 1989 because substantial, and necessary, revisions were only possible for this latter period.

4.13 Labour productivity series are based on sectoral GDP at market and constant prices and total hours worked, with establishment based (as opposed to household surveys) data on hours worked being the source. Productivity in levels is then converted into the same currency using 1996 PPP figure extracted from OECD and Eurostat.

4.14 Capital stock is calculated for six asset types. In doing so, the perpetual inventory method (PIM) is employed in accordance with the methodology proposed by the US Bureau of Economic Analysis. Capital stock estimates are adjusted for changes in composition.

4.15 When determining TFP, it is necessary to attribute appropriate weights between labour productivity and capital per hour. The dataset uses the share of labour compensation in GDP for labour (based on total non-wage labour costs including that for self-employed and one minus the share of labour in GDP for capital).

4.16 The NISEC02 dataset provides UK TFP data for the sectors listed in Table 4.1 below. We have included some of the manufacturing sub sectors into the analysis to more closely reflect at a high-level some components of the gas distribution industry than "manufacturing" as a whole.

Table 4.1: Sectors with TFP estimates available in the NISEC02 dataset

Broad sectors with TFP estimates	Manufacturing sub-sectors with TFP estimates
Agriculture, forestry & fishing	Coal & petroleum products
Mining & extraction	Chemicals & Allied products
Manufacturing	Chemicals
Construction	Rubber & Plastics
Transport & communications	Basic Metals & Fabricated Metal products
Transport	Total machinery & equipment
Communications	Machinery
Distributive trades	Office equipment
Whole sale trade	Electrical & Electronic equipment
Retail trade	Motor vehicles
Hotels & catering	Other transport equipment
Financial and business services	Optical equipment & instruments
Financial intermediation	Textiles, Clothing & Leather
Business services	Food, Drink & Tobacco
Miscellaneous personal services	Other Manufacturing
Non-market services	Non-Metallic mineral products
Total Economy	Wood Products
Total Market Sectors	Paper & Printing
	Furniture & Miscellaneous Manufacturing

4.17 In principle, it might be appropriate to adjust these data for economies of scale before use in the analysis. However, these would need to be estimated at the production unit level, and sector-level input or output growth figures are not relevant for that purpose as they could reflect an increase in the number of production units as well as an increase in scale of production. It is reasonable to think that in a competitive environment and in absence of relevant fix costs and barriers to entry, industries are generally around the efficient scale of production.

4.18 However, for network infrastructure companies, for which fixed costs and barriers to entry may be significant, economies of scale are likely to be relevant. This would tend to suggest that an adjustment to the TFP figures is needed for sectors such as transport and communications. However, in these sectors there has been significant liberalisation and entry, leading to a countervailing trend towards smaller units. The net effect is unclear, and we have chosen to make no adjustment for economies of scale.

Components of Gas Distribution

4.19 The first stage of analysis in construction of a composite benchmark (one that draws together productivity evidence from a range of comparator industries) is to consider how to divide the activities of gas distribution industry to allow for identification of comparator sectors and services.

4.20 By breaking down the activities of the gas distribution industry into different components, based on markedly different business activities, it is possible to identify comparator sectors within the dataset more readily than if the industry is considered as a whole.

4.21 There are two requirements on the activities that the gas distribution is divided into. First, they have to be meaningfully separate activities with possible comparators in the data set. Second, we have to be able to reliably measure the proportion of the different activities within gas distribution.

4.22 Following discussion with Ofgem we concluded that the best available division to satisfy the two criteria is provided in the Business Planning Questionnaires (BPQ) filled in by the companies. The BPQs break down the gas distribution into:

(a) Capital and replacement expenditure;

(b) Work Management;

(c) Emergency and Repairs;

(d) Support services and Indirect operating expenditure; and

(e) Maintenance and other works.

4.23 This breakdown is sufficient to capture the key differences in activities that are likely to lead to differences in productive improvement within different operations, while still being general enough to be compatible with the available breakdown of the comparators in the data source. Further, this BPQ based categorisation is valuable because it allows the relative weights of the industry components to be based on relative cost allocations, combined from the BPQs submitted for the industry as a whole.

4.24 Ofgem provided us with the relative weights of the above components in the BPQ, for 2005/06 submissions, given in Table 4.2 below.

Table 4.2: Weights for industry components based on 2005/06 BPQ:

Industry component	Component weight
Capital and replacement expenditure	55.9%
Work management	12.5%
Emergency & Repairs	11.6%
Support services and Indirect opex	13.0%
Maintenance and other	7.0%

4.25 Of course there will be variation among the companies from these industry average weights. We undertake sensitivity analysis below to see how the results change if different weights are used for the components.

Matching Components to Comparator Services

4.26 The industry components identified in Table 4.2 can be mapped to services by a high-level comparison of activities between industry components and potential comparator sectors in the dataset. We now discuss the nature of work of the identified industry components and set out the comparator sectors for each. The comparator sectors have been identified in collaboration with Ofgem, drawing on their in-house sector expertise. As mentioned above, the gas distribution was broken down into activities according to the BPQ, which defines the categories as follows.

Work management

4.27 Work management includes staff and other non-operational costs associated with:

(a) Asset management including network integrity, planning and design;

(b) Operations Management including supervisory costs;

(c) Contract Management, managing the relationship with engineering contractors and other bought in services;

(d) Customer Management, managing the processes that interface with consumers and shippers;

(e) Network Support, costs associated with engineering back office e.g. records management and work scheduling processes;

(f) Health, safety and Environment;

(g) Network policy;

(h) Safety & Engineering; and

(i) Call centres.

4.28 This category therefore relates to the softer management expenses of the gas distribution network, but also includes some engineering and design components. The comparator sectors for this activity are **Business services, Engineering, and Communications**.

Emergency and Repairs

4.29 The emergency and repair category corresponds to the cost of service to responding and making safe all gas escapes, and then carrying out repairs to external gas mains and services.

4.30 This category therefore is akin to activities of networked utilities and the construction industry in terms of repairs and emergency repairs. The comparator sectors used are **Utilities and Construction**.

Support services and Indirect operating expenditure

4.31 These cover a wide range of internal support and indirect services, ranging from day to day information provision and data centres to legal services, audit costs and property management.

4.32 The comparator used for this category is **Business services**.

Maintenance and other

4.33 This category captures the low pressure storage and maintenance and DN pipeline and AGI/PRS maintenance. It also captures the costs associated with other leakage control activities, including mains surveys, gas conditioning and pressure profiling, mains service repairs and maintenance, and instrumentation repairs and maintenance are also covered.

4.34 The activities here are similar to Emergency and Repairs activities, with the addition of instrumentation repairs and more routine maintenance monitoring activities. The comparator sectors used are **Utilities, Engineering, and Construction**.

Capital expenditure and replacement expenditure

4.35 Capital expenditure (capex) is defined as

> […] investment in assets whose benefits can be expected to last for some years, such as high-pressure pipelines and lower pressure mains. This includes expenditure on extending or reinforcing the pipe networks or adding new connections. For pipelines and mains a further distinction can be made between capital and replacement expenditure.

4.36 Replacement expenditure (repex) is:

> […] expenditure on replacing component mains and services, where the replacement may lead to an increase in the capacity or extend the life of the network but is principally driven by regulatory requirements, particularly heath and safety.

4.37 Both are therefore the expenditure associated with construction or replacement of pipelines and mains, or components thereof. The comparator sectors used for this category are **Construction, Engineering and Utilities**.

Adjustment for business services

4.38 Some care has to be taken in the transition from the available comparator sectors to comparator services. Suppose, for example, that nature of work analysis found that a benchmark for a particular industry could be constructed on the basis that it is 50 per cent communications services and 50 per cent business administration services. An estimate could be calculated using equal weights for both communications and business administration services. However, such a calculation would not take into account the fact that the communications sector itself is likely to contain a non-negligible proportion of business support services. A more accurate estimate would be provided by reducing the weight attached to the business services sector figure in the benchmark to reflect the fact that the communications sector figure includes some of the “business services” activity.

4.39 By taking a view on what fraction of the communications sector is “pure” communications service activity, and what fraction business support services, we can adjust for the potential double counting.

4.40 We therefore distinguish between, for example, communications services, representing the pure communications activity, from the communications sector. Communications services make up the majority of the sector for which there is comparator data, but this sector also contains some generic financial or business services. Thus the communications sector productivity growth data include X per cent construction services productivity growth and Y per cent business services productivity growth.

4.41 Within the dataset shown in Table 4.1 the only service that features in several comparator sectors is financial and business services. Thus we consider that each comparator sector in the dataset comprises a specific main service and also a financial and business services component.

4.42 Suppose that X represent the productivity growth between two periods that we wish to estimate for a comparator service; Y represents the observed productivity growth in the financial and business services sector; Z represents the observed productivity estimate in the associated comparator sector between those periods; and that a percentage “p” of the productivity growth in the comparator sector is considered to reflect the productivity growth in financial and business services. The observed productivity growth Z can then be expressed as:

Z = (1-p)X + pY

4.43 Rearranging the expression to give X from observed variables Z and Y (for assumed p):

X = (Z - pY)/(1 - p)

4.44 However, it is difficult to obtain data on “p” for each sector. In the absence of data, we adopt the same assumption as in Europe Economics (2003) that 5 per cent of comparator sectors reflect productivity improvements in financial and business services, with the exception of communications sector, of which 15 per cent is assumed to be financial and business services.

4.45 The calculations undertaken for the nature of work benchmark are therefore based on the raw TFP data provided in the NISEC02 dataset, adjusted according to this method. Because of the uncertainty surrounding these assumptions, we later test the effect of this adjustment on the results.

Construction of Composite Benchmarks

4.46 The benchmark for the nature of work analysis is a weighted average of productivity estimates for the comparators identified above, with relative weights determined for each comparator service by the weight attributed to the industry component.

4.47 Where two or more comparators are identified for an industry component, the productivity growth attributed to that component is taken equally from the relevant comparators (a simple average). We appreciate that this division within the industry sub-section is somewhat arbitrary, but found no clear method that could be adopted while maintaining the high level top-down nature of the analysis. Also, the category “Engineering” does not exist as such in the NISEC02 dataset, instead we use category “Total machinery and equipment” as the source of engineering type data. In any event, we address this issue in the sensitivity analysis by testing the inclusion of different comparator services particularly Engineering.

4.48 Table 4.3 shows the structure of the benchmark for the gas distribution industry, summarising the above discussion. Overall, construction, engineering and utilities comparators will carry the largest weight, with communications and business services making up the rest of the gas distribution activities.

Table 4.3: Structure of the nature of work benchmark

Industry component	Weight		Comparators	
Capex and repex	55.9%	Construction	Engineering	Utilities
Work management	12.5%	Business services	Engineering	Communications
Emergency & Repairs	11.6%	Construction	Utilities	
Support services and Indirect	13.0%	Business services		
Maintenance and other	7.0%	Construction	Engineering	Utilities

Calculating productivity estimates

4.49 Using the above weights and components, we calculate an index of overall TFP growth and growth out performance of the economy, defined such that the index changes from one period to the next according to the weighted average growth between those periods in each comparator. The intuition behind these indices is that if the productivity growth in the gas distribution industry had followed the productivity growth in the comparator services in proportion to the comparator weights defined for the benchmark, productivity for the gas distribution industry would have followed the path indicated by the index.

4.50 The dataset provides TFP data from 1950 to 1999. The table below includes the composite TFP growth, and the TFP out performance over the economy for the whole period, and for two more recent sub periods, calculated as the average annual growth.

Table 4.4: Overall TFP growth and out performance

Time period	TFP growth (%)	TFP out performance
1950-1999	1.9%	0.7%
1973-1999	2.1%	0.7%
1989-1999	1.9%	0.8%

4.51 The above table shows that both the composite TFP growth and out performance have changed slightly during the period. This could be for various reasons, including the changing nature of the whole economy through such a long time, and, particularly for the more recent period, changes in the regulatory environment of some of the comparator industries. In choosing the time interval to use we have to balance the need for a long time series of data to get at an underlying productivity improvement against possible fundamental changes in the trend. It could be argued that data for the earlier part (in the 1950s and 1960s) is no longer relevant, for example due to the effect that oil price shocks might have had on industries or other changes placing more focus on efficiency. On the other hand, the very latest period could simply be too short to be taken as indication of long-term productivity trends due to short-term fluctuations. Our preferred time period therefore is from 1973 to 1999.

4.52 In addition to underlying gradual changes in the data, the privatisation of network industries in the late 1980s and early 1990s could have had a significant effect on the TFP growth trends, particularly as Utilities and Communications industries form a part of the comparator set. The so called "privatisation effect" does indeed appear to have a significant and sustained effect, as discussed and measured in for example Europe Economics (2003). In so far as this effect has transitory components, it is important to remove it from the TFP benchmark. Controlling for this (and potentially other) effects can be achieved by the use of regression analysis, discussed below.

The nature of work TFP benchmark

4.53 In order to assess the nature of work productivity trend free of transitory effects of privatisation of the network industries, we have used a regression methodology to control for potential effects of privatisation.

4.54 In order to accomplish this, we constructed a privatisation index to capture the degree to which companies in a given sector have been privatised and may therefore display the privatisation effect. This index was constructed to have higher values near privatisation for each industry, with decaying effect such that after 15 years after privatisation of the industry the index would be zero. The comparators affected were the Utilities and Communications industries.[21]

4.55 The data are in a cross sectional time series panel, where we are interested in the individuals that make up the panel. The estimation method used is panel data corrected pooled OLS (or time series cross section OLS), with dummy variables estimated as the coefficients for the comparator industries. The coefficients γ from the equation can therefore be used to calculate the equivalent to an average TFP out performance for the comparator services, with the effect of privatisation removed. These can then be combined according to the weights in Table 4.3 to produce the benchmark TFP trend for the gas distribution industry.

4.56 The table below gives the estimated nature of work TFP growth and out performance benchmarks. As expected, this figure is lower than the estimate for the same time period in Table 4.4.

21 The regression estimated was:

$$OATFP_{it} = \text{constant} + \beta PI_{it} + \gamma IND_i + \varepsilon_{it}$$

OATFP represents the TFP growth out performance figure for each comparator service (i.e. sector data adjusted for business services);
PI is the privatisation index, taking values of zero through out for industries other than Utilities and Communications; and
IND is a dummy variable for the comparator industries, except for Business services that is captured by the constant.

Table 4.5: The nature of work TFP benchmark

Time period	TFP growth	TFP out performance
1973-1999	2.0	0.6

4.57 It should be emphasised that this methodology is used only to gain benchmark TFP growth estimate absent distorting effects of privatisation. However, the estimated values of the PI privatisation coefficient are not used in the analysis. The estimation here has been done on TFP, whereas for the present purpose we would be interested in the effect of privatisation on operating expenditure efficiency. Our view is that the most accurate way of estimating the privatisation effect is to use more direct information on the privatised sectors, consistently with Europe Economics (2003), discussed below.

Sensitivity analysis

4.58 We conducted sensitivity analysis of the results for the different assumptions made in building the nature of work TFP benchmark. In particular, we investigated the effect of:

(a) Change in the relative weights of the identified gas distribution industry components, as they might be expected to change over time.

(b) Choice of comparator sectors — in particular we estimated the impact that including Engineering has on the results

(c) The business services adjustment, and what impact omitting it has (as the proportion of business services in the other sectors has is difficult to verify).

4.59 Changing the weighting such that capital and replacement expenditure falls to 45 per cent with the weight of the other components adjusting proportionally only has a small effect on the estimated out performance: it moves from 0.6 to 0.54 per cents per annum. Further reducing the weight to 30 per cent moves the result only to 0.50 per cents per annum. The result seems to be quite robust to the precise weighting of the industry components, at least within a reasonable range that the weighting might move from year to year.

4.60 Removing Engineering from the comparator set has a similar effect, reducing the estimated out performance to 0.47 per cent per annum. The rationale for this sensitivity is the perhaps imperfect correspondence of the data to engineering activity.

4.61 Omitting the adjustment for business services also tends to reduce the estimated TFP out performance, but only very slightly. The category Financial and business services is one of the few sectors underperforming the total economy during the period. Keeping in mind that only 5 per cent of the comparator sectors were assumed to be business services (where as the weighting above would give more than 10 per cent for gas distribution), the nature of work TFP benchmark with the business services adjustment could be conservative.

Adjustments for OPEX productivity

4.62 The nature of work analysis undertaken in this section has focused on TFP because, provided that the data used are reasonable for the intended purpose, this should be the best measure of productivity improvement for the nature of work analysis (even if we then wish to make specific assumptions to move to an operating expenditure figure).

4.63 The aim of the current exercise, however, is to estimate operating expenditure productivity growth. This requires adjustments for capital substitution, and possibly for differences in input prices and economies of scale, discussed below.

Capital substitution

4.64 The effect of capital substitution has to be taken into account to move from the TFP figures towards forward-looking operating expenditure efficiency. Since measures of operating expenditure generally include a higher proportion of labour inputs to capital inputs compared to the total cost measures that should underpin TFP estimates, it is important to analyse the effects of capital substitution. We can use the labour productivity data in the NISEC02 dataset for this.

4.65 Capital substitution is at the source of the differences between TFP and labour productivity (LP) measures. In a two-factor model this can be set out as:

$$g_{LP} = g_{TFP} + \alpha_K(g(x_K) - g(x_L))$$

4.66 The above states that the growth in labour productivity (g_{LP}) is the sum of growth in total factor productivity (g_{TFP}) and the difference in growth of capital and labour inputs, the capital substitution ($g(x_K) - g(x_L)$), adjusted by the share of capital in production inputs (α_K)[22]. For example, if the growth in capital inputs was 2 per cent per year and the growth in labour inputs 1 per cent per year, with zero TFP growth, and the capital share of inputs was 50 per cent, the growth in labour productivity would be 0.5 per cent per year.

4.67 We can calculate the growth in labour productivity from the NISEC02 data as well as the growth in total factor productivity as above. With an assumption about the capital share in production inputs, the capital substitution can be inferred as (rearranging the above):

$$(g(x_K) - g(x_L)) = (g_{LP} - g_{TFP})/\alpha_K$$

[22] Strictly this is the rate of marginal rate of substitution between capital and all inputs, which is the same as the capital share of value under assumptions of no constraints preventing adjustments to inputs used.

4.68 We investigate two alternatives. First, an underlying economy wide rate of capital substitution assuming capital share of production of 40 per cent.[23] Second, rate of capital substitution implied by the nature of work comparators, assuming a capital share of 50 per cent in the comparator set (the industries in the comparator set are likely to be more capital intensive than the economy as a whole).

4.69 The NISEC02 data show that the total UK economy labour productivity grew on average by 2.2 per cent per annum from 1973 to 1999, and TFP grew by 1.4 per cent per annum. With the above 40 per cent capital share, this implies a capital substitution factor of 2.1 per cent. For the nature of work comparators, controlling for the privatisation effect, the implied capital substitution factor is 1.7 per cent ((2.81-1.98)/0.5). These estimates of capital substitution can then be used, with an appropriate capital share assumption for gas distribution, to infer the growth in operating expenditure productivity for gas distribution.

4.70 Assuming 70 per cent share of capital in the gas distribution sector, the nature of work benchmark TFP growth should therefore be increased by 1.5 percentage points assuming similar capital substitution for the gas industry as for the economy as a whole, or by 1.2 percentage points assuming a similar capital substitution as in the nature of work comparators. These adjustments, though already conservative by the way of construction, could be further scaled downwards due to the fact that labour is not the only factor in the definition of operating expenditure for this study. On the other hand, the rate of capital substitution in a network industry could be expected to be much higher than in the economy as a whole. We use the range of 1.2 to 1.5 per cent as the effect of the capital substitution on labour productivity.

Relative input price movements

4.71 The relevant input price basket for operating expenditure will have a disproportionate contribution from labour compared to a total cost input price basked (which includes a significant capital element). Given that wages are normally one of the fastest rising input prices in the economy, this will mean that the scope of (real) operating expenditure reduction is somewhat less than it would be for a measure of total costs, to which TFP is most directly relevant.

4.72 Europe Economics (2003) calculates that, economy wide, labour input prices rise by 1.9 per cent per year, and the economy wide input price basket rises by 1.4 per cent a year in line with whole-economy TFP growth. With labour constituting two thirds of an economy-wide input price basked, this implies input prices other than labour increased at zero to 0.5 per cent per year. If the relevant input price basket for operating maintenance

[23] Bank of England (2000), *Inflation dynamics and labour share in the UK*, shows that the labour share of value in the UK has been varied roughly between 60 and over70 per cent over the past 30 years or so. In a two-factor model this implies capital share of value of 40 to 30 per cent. Using a value of 40 per cent leads to a conservative (low) estimate of opex efficiency. http://www.bankofengland.co.uk/publications/externalmpcpapers/extmpcpaper0002.pdf

expenditure were exclusively composed of inputs with the same trend growth rate as labour, a downward adjustment of 0.4 to 0.6 to the scope for efficiency improvement would be required. In practise there are some capital and materials input involved in operating expenditure definition used here (raw materials used for asset maintenance, use of machinery etc) and the adjustment would therefore be less, and we therefore use the lower end of the range, 0.4 per cent.

Economies of scale

4.73 In principle, it might be appropriate to adjust these data for economies of scale before use in the analysis. However, as discussed in paragraph 4.17, we decided not to undertake an adjustment.

4.74 Even though the comparator data does not have economies scale adjustment applied, the benchmark trend could in principle be adjusted, as the econometric models (COLS1 and COLS2) above suggest that the gas distribution industry might display economies of scale in output. If there is demand growth from existing customers or from new customers, then, for a given network size, costs in the presence of scale economies costs are likely to grow less rapidly than output. This would require a further reduction in costs in addition to the estimated productivity growth.

4.75 The adjustment would only be relevant if the scale of production was seen to grow significantly over the period (by having a scale variable, comparative scale is already accounted for in the econometric estimates of comparative efficiency). This would require an estimate of the demand growth as well as an estimate of the level of economies of scale in output. Europe Economics (2003) found that a typical regulatory assumption in electricity networks for example is that a 1 per cent growth in volumes leads to 0.5 per cent increase in costs. Ofgem could use this 0.5 per cent estimate as an upper bound adjustment on top of the productivity growth estimated here. Omitting the adjustment, expected to be small in any case, would not lead to unduly stringent efficiency targets. We have therefore not employed any economies of scale adjustments to the nature of works results.

Privatisation effect

4.76 The above regression analysis removed the effect privatisation on some of the comparators from the estimated TFP improvement. The integrated gas industry (transmission, distribution and supply) was initially privatised in 1986. Transmission and distribution were further separated from supply ownership in 1997. Though the original privatisation occurred already over 20 years ago, the re-organisation of the industry in 1997 could be argued to have incentive effects similar to privatisation, or at least strengthen them, such that the industry could well be affected with an ongoing privatisation effect. The continued success of incentive based price regulation in driving down costs is likely to be the principal mechanism through which the initial stimulus to productivity from privatisation is extended over a much longer period.

4.77 Europe Economics (2003) estimated a privatisation effect combining expenditure data of privatised companies with the results of the nature of work TFP analysis undertaken for the water, sewerage and electricity networks. The conclusion from that analysis was that the privatisation effect is ongoing, and that a suitable range for its effect over the ten years from 2003 constituted a 0.5 to 2.5 per cent per annum reduction in real unit operating expenditure for the water and sewerage industries, and 0.5 to 4.1 per cent for the electricity networks. The ranges drawn were conservative relative to the estimates of the effect produced.

4.78 We therefore have good reason to expect there to be some continuing effect on productivity from privatisation and continuing regulation of network activities. The recent changes, including the division of NGGD into regional distribution companies and the sale of a number of the GDNs in 2005, could extend the longevity of this effect further but such restructuring was not a feature of the earlier study referred to above and is therefore separate from this analysis.

4.79 Using the 2003 estimates as the basis of the privatisation effect for gas distribution, and allowing for a further exhaustion of the effect in the future, we use a range of 0.5 to 2 per cent in our projected improvement in total controllable operating expenditure.

Summary

4.80 This section has undertaken an analysis of the scope of long term productivity improvement in the UK gas distribution by using a nature of work analysis of past productivity trends in comparator sectors.

Table 4.6: The nature of work benchmark for operating expenditure

	Lower	**Higher**
Bench mark TFP out performance 1973-1999	0.6%	0.6%
Privatisation effect	0.5%	2.0%
Capital substitution effect	1.2%	1.5%
Input price effect (possible downward adjustment)	-0.4%	-0.4%
Economies of scale adjustment	0	0
Total	**1.9%**	**3.7%**

4.81 The results in Table 4.6 suggest that, on the basis of the nature of work that gas distribution industry undertakes, we could expect the industry to achieve gains in operating expenditure productivity of between 1.9 and 3.7 per cent per annum over the economy as a whole. These results can be combined with the analysis of comparative efficiency assessment to calculate the scope for efficiency improvement of individual companies.

5 FRONTIER SHIFT AND EFFICIENCY SAVING: DECOMPOSITION FOR EACH GDN

Introduction

5.1 In section 4 we have derived a range of estimates for the scope for total efficiency savings in operating expenditure, on top of the economy's growth in TFP, for the gas distribution sector as whole.

5.2 The total scope for efficiency saving should ideally be split into two components, frontier shift and catching up.

5.3 Frontier shift can be understood as the scope for cost reductions that is due to the movements in the industry efficiency frontier. In other words, there is an element in the sector's scope for efficiency savings that is achievable even by the companies that are operating at or near the industry efficiency frontier.

5.4 Technological progress and imitation of best practices in other sectors are just some of the reasons usually invoked to explain why even companies that are operating with no slack could achieve cost reductions, even on top of the economy-wide productivity growth.

5.5 In turn, catching up accounts for the remaining part of the total scope for efficiency savings: it is made up of the (weighted) average of each GDN's scope for catching up with the industry frontier: GDNs that are relatively inefficient should be expected to cut costs more than their more efficient peers.

5.6 Separating out frontier shift and catching up from the total scope for efficiency savings, while not theoretically difficult, is in practice a challenging task to achieve, both for the sector as a whole and for the individual GDNs.

5.7 Academic studies have used stochastic frontier analysis and DEA Malmqvist index numbers to separately identify technical efficiency (and therefore the scope for catching up) and frontier shift.

5.8 However, both methodologies have drawbacks. First of all, they are data intensive techniques which require considerably larger samples than these that are available in the case of the UK GDNs. Secondly, they necessarily rely on the past performance of the firms in the sample: while this does not raise any issue as far as the estimation of technical inefficiency (and therefore the scope for catching up for each GDN in the sample), the scope for frontier shift could be underestimated, because it does not take into account that a major change in regulation (such as the use of yardstick competition) might provide further incentives to cut costs, even for the GDNs that are relatively efficient.

5.9 For this reasons, we have decided to follow a rather different methodology.

5.10 From our Nature of Work analysis we have derived a range of estimates for the total scope of efficiency savings for operating expenditure on top of the economy productivity growth.

5.11 The benchmarking techniques considered in section 3 have provided us with an estimate for the extent of inefficiency for each GDN relative to a frontier. By taking a weighted average of the scope for catching up with the industry frontier for each GDN we can then derive an estimate for the overall catching up, i.e. the reduction in operating expenditure that is achievable at the level of the sector.

5.12 The difference between the overall scope for efficiency savings and the overall catching up would give the expected rate of reduction in operating expenditure that is achievable even by the frontier companies, i.e. frontier shift.

5.13 For the relatively inefficient GDNs, their expected reduction in operating expenditure could then be computed by adding their individual catching up factor on top of the frontier shift factor.

5.14 In the remainder of this chapter we will:

(a) Discuss how the methodologies developed in section 3 above could be used to derive an estimate for the catching up factor for each GDNs and, therefore for the overall gas distribution sector; and

(b) Use that figure to compute the yearly reductions in opex achievable by the efficient GDNs (frontier shift factor) and inefficient GDNs (frontier shift factor plus GDN specific catching up factor).

The scope for catching up and frontier shift

5.15 Each of the methodologies for assessing efficiency could potentially be used to compute the individual catching up factors and, therefore, the sector-level scope for efficiency savings due to catching up.

5.16 However, as noted in section 3, we have concluded that the regression analysis provides the most robust estimates for the current study and that of the four models considered COLS 1 and COLS 2 provide our preferred format for estimating the scope for catching up for each individual GDN.

5.17 The COLS procedure we have used to compute relative efficiency in section 3 above can be criticised because it amounts to interpreting in the regression' residuals as inefficiency, ignoring other possible reasons that might explain why GDNs costs might diverge from each other: as a result, COLS-like procedures are likely to overstate the scope for catching up that is really achievable by most GDNs.

5.18 Stochastic frontier analysis would be the correct theoretical framework to deal with the existence of both noise and inefficiency in the each GDN's residual. However such models cannot be estimated using such a small sample. Instead we have taken a more pragmatic approach, which consists in recognising that, in practice, not all the efficiency savings identified by the COLS methodology are in fact achievable.

5.19 This can be done by assuming that all GDNs with an efficiency level equal to highest quartile are, in fact, fully efficient (i.e. with an efficiency level of 1) and re-scaling the efficiency estimates with respect to this less challenging target.

5.20 Further refinements are possible. For example, we have seen in Table 3.9 above that efficiency levels have in general declined in 2006/07. Therefore, restricting the selection of the benchmark company to the upper quartile in 2006/07 only may reduce the average catching up for the sector.

5.21 By contrast, there could be some merit in computing the benchmark by taking an average of the two years OLS residuals and then applying the COLS procedure: this approach could reduce the possibility that year specific shocks would play too large a role in shaping the industry efficiency frontier: as a result, the performance of some GDNs could change considerably, given that for some of them the results presented in section 3 clearly indicated rather significant deterioration in performance in 2006/07.[24]

5.22 For comparison we present results showing the scope for catching up using both of these approaches.

5.23 In the first, displayed in Table 5.1, efficiency was computed by identifying the third upper quartile of efficiency in 2006/07 (which were computed according to the COLS formula displayed in section 3 above) and rescaling efficiency levels, such that each GDN with an efficiency lower than that of the third upper quartile, is compared against it, rather than against the pure frontier identified by the COLS procedure, as reported in Table 3.9.

5.24 As, in general, the different models tend to select GDNs in 2005/06 as the frontier, this procedure tends to generate higher levels of efficiency, on average, with respect to those identified in Table 3.9.

---

[24] Of course, given that the total scope for efficiency saving in the sector is fixed –as given by the Nature of Works analysis developed in section 4 above – changing the way we define the frontier and, therefore, the scope for catching up would result in different estimates for the frontier shift component of the total scope for efficiency savings. In particular, using the third upper quartile as the benchmark company would tend to reduce the scope for catching up and, as a result, increase the frontier shift component of the total scope for efficiency savings.

Table 5.1: GDN's efficiency scores for catch-up based on 2006/07(1)

	Year	COLS1	Rank	COLS2	Rank
East of England	2	0.96	4	0.94	6
London	2	0.80	8	0.90	7
NW	2	0.88	7	0.90	7=
West Midlands	2	0.93	6	1	1=
Northern	2	1	1=	1	1=
Scotia-South	2	0.95	5	0.95	5
Scotia-Scotland	2	1	1=	1	1=
Wales and West	2	1	1=	1	1=
Average	2	0.94		0.96	

Source: Europe Economics

5.25 The second set of results was calculated by averaging the residuals for each GDN over 2005/06 and 2006/07 and then applying the COLS formula. We then rescaled the efficiency levels by allowing only the upper quartile efficiency level to set the benchmark for the less efficient GDNs, as explained above.

5.26 This second set of results shown in Table 5.2 should be less sensitive to the possibility that in some years specific shocks might have impacted too much on the "true" efficiency levels of some GDN: the drawback of this procedure is that genuine reductions in efficiency levels occurring in 2006/07 will not be fully captured in this second set of efficiency estimates.

Table 5.2: GDN's efficiency scores for catch-up based on average 2005/06 & 2006/07(2)

	COLS1	Rank	COLS2	Rank
East of England	1	1=	0.97	5=
London	0.87	8	0.97	5=
NW	0.91	7	0.93	7
West Midlands	0.98	4	1	1=
Northern	1	1=	1	1=
Scotia-South	0.95	5	0.99	3
Scotia-Scotland	1	1=	0.98	4
Wales and West	0.94	6	0.93	8
Average	0.96		0.97	

Source: Europe Economics

5.27 The results show that the average level of efficiency[25] is slightly higher than that identified in Table 5.1, perhaps because of the possibility that year specific shocks might have tended to overestimate (underestimate) the efficiency levels of some GDNs in 2005/06 (2006/07).

5.28 The averages reported in Table 5.1 and Table 5.2 are the average levels of efficiency in the sector. On the basis of these figures, Table 5.3 reports, for both sets of results and for each Model, the average scope for catching up (which is defined as the complement to one of the average efficiency figures reported in Table 5.1 and Table 5.2, expressed as percentages of total controllable costs.

Table 5.3: The average catching up factor across GDNs

Owner	Third quartile, 2006/07 only		Third quartile, average of 2005/06 and 2006/07	
	COLS1	COLS2	COLS1	COLS2
% Catching up	6.02	4.0	4.45	2.91

Source: Europe Economics

5.29 The figures reported in Table 5.3 show that the catching up factors for the sector as a whole range from about 2.9 per cent to 6 per cent, depending on the model and the assumption to build the benchmark GDN. These figures should be interpreted as the percentage by which companies should cut their costs, on average, to eliminate any slack or inefficiency so that all GDN would be on the industry efficiency frontier. The overall average figure for this, based on the average inefficiencies derived from the COLS1 and COLS2 models for both approaches, is approximately 4.3 per cent.

5.30 To derive an estimate for frontier shift, we have to combine the results of the Nature of Work analysis with the figures we have just derived for the scope for catching up.

5.31 Before doing that, we should note that the catching up figures should be interpreted as the extent by which, on average, GDNs should cut their cost to get to the industry frontier, over a specified period of time. If we take that specified period of time to be a five years period, we can interpret the catching up figures by saying that, on average, the GDNs should cut their costs by 4.3 per cent to eliminate any inefficiency, in the next five years period.

5.32 From the Nature of Work analysis we have seen that the likely scope for efficiency savings is in the range of 1.9 to 3.7 per cent per year. Using the formula for compound

[25] The sector level efficiency is computed as a weighted average of each GDN efficiency level. The weight that was used was the percentage of total controllable costs of each GDN in the total sector controllable costs.

negative growth, this could be translated, over a five years horizon period, into a reduction in opex (on top of the economy productivity growth), in the range of 9.2 – 17.2 per cent.

5.33 If we call CATCH the average scope for catching up over the next five years period, TES the total scope for efficiency savings over the next five years period and FRONTIER the percentage reduction in costs achievable by frontier firms in the next five years period, we can write TES = FRONTIER*CATCH, which gives: $FRONTIER = TES / CATCH$.

5.34 On the basis of the above formula, Table 5.4 reports, for both sets of results and for each Model, the estimate for frontier shift as the percentage reductions of costs achievable by the frontier companies both in the next five years period and as a yearly rate.

Table 5.4: The frontier shift as percentage reductions of costs achievable

	Upper quartile, 2006/07 only				Upper quartile, average of 2005/06- 2006/07			
	COLS1		COLS2		COLS1		COLS2	
	NWLB	NWUB	NWLB	NWUB	NWLB	NWUB	NWLB	NWUB
Five years	3.08%	11.65%	5.34%	13.72%	4.91%	13.32%	6.42%	14.70%
Annual	0.62%	2.45%	1.09%	2.91%	1.00%	2.82%	1.32%	3.13%

Source: Europe Economics

5.35 In the analysis below we present results expressed in terms of the average annual rate of efficiency improvement over a five years period. We have a range of results because the Nature of Work comparison identified a lower bound (NWLB) for the scope of reduction in real operating expenditure of about 1.9 per cent each year and an upper bound (NWUB) of 3.7 per cent a year, on top of the productivity improvements at the level of the economy as a whole.

5.36 The corresponding upper and lower confidence bounds for frontier shift, in the case of Model COLS1 — and the scope for catching up computed by using only the 2006/07 estimates to define the frontier — are 0.62 per cent and 2.45 per cent.

5.37 In the case of Model COLS2, these two figures are 1.09 per cent and 2.91 per cent, given the slightly lower inefficiency identified by COLS2 with respect to COLS1, and therefore the somewhat smaller scope for catching up.

5.38 If we compute catching up with the second methodology, we have somewhat higher figures for frontier shift, as the average level of inefficiency, and therefore the scope for catching up, is somewhat smaller.

5.39 Combining the figures in Table 5.3 and those in Table 5.4 we can derive, for each GDN, the average annual scope for efficiency saving in the next five years period, on top of the economy TFP growth, as split into a frontier shift and a catching up component. As we do not have a strict preference for either of the two models, and the models do differ somewhat in their ranking and levels ascribed to the companies, we report only the average calculated using them both.

Table 5.5: Company specific efficiency improvement based on upper quartile average, 2006/07 only, average of COLS1 and COLS2

.	Average annual % reductions over five years					
	NWLB			NWUB		
GDN	Frontier shift	Catching up	Total	Frontier shift	Catching up	Total
East of England	1.00%	0.90%	1.90%	2.82%	0.90%	3.70%
London	1.00%	3.12%	4.09%	2.82%	3.12%	5.85%
NW	1.00%	2.18%	3.16%	2.82%	2.18%	4.94%
West Midlands	1.00%	0.33%	1.33%	2.82%	0.33%	3.14%
Northern	1.00%	0.00%	1.00%	2.82%	0.00%	2.82%
Scotia-South	1.00%	0.46%	1.46%	2.82%	0.46%	3.27%
Scotia-Scotland	1.00%	0.01%	1.02%	2.82%	0.01%	2.83%
Wales and West	1.00%	0.00%	1.00%	2.82%	0.00%	2.82%

Source: Europe Economics

5.40 The figures reported in Table 5.5 show how much each GDN could on average be expected to cut its operating expenditure on top of the economy productivity growth each year over the next five year period.[26]

5.41 NW and London are the GDNs that could be expected to have the biggest reductions in operating expenditure, while Northern and Wales and West could be broadly expected to match the scope for frontier shift.

5.42 In Table 5.6 we have repeated the same exercise, but this time using the scope for catching up computed from the average of residuals in 2005/06 and 2006/07 (see Table 5.2).

---

[26] The totals in tables 5.5 and 5.6 represent the compound effect of the frontier shift and catch up productivity estimates.

Table 5.6: Company specific efficiency improvement based on upper quartile average 2005/06-2006/07, average of COLS1 and COLS2

.	Average annual % reductions over five years					
	NWLB			NWUB		
GDN	Frontier shift	Catching up	Total	Frontier shift	Catching up	Total
East of England	1.18%	0.24%	1.42%	3.00%	0.24%	3.23%
London	1.18%	1.61%	2.77%	3.00%	1.61%	4.56%
NW	1.18%	1.58%	2.75%	3.00%	1.58%	4.53%
West Midlands	1.18%	0.00%	1.18%	3.00%	0.00%	3.00%
Northern	1.18%	0.00%	1.18%	3.00%	0.00%	3.00%
Scotia-South	1.18%	0.68%	1.85%	3.00%	0.68%	3.66%
Scotia-Scotland	1.18%	0.19%	1.38%	3.00%	0.19%	3.19%
Wales and West	1.18%	1.35%	2.52%	3.00%	1.35%	4.31%

Source: Europe Economics

5.43 The results broadly confirm those reported in the previous Table, with London and NW being the distributors with most scope for efficiency improvement in the next five years. One notable difference is in the estimates for Wales and West which appears as one of the most efficient companies based on 2006/07 performance but is at the lower end of the range using the baseline for both years.

5.44 There are arguments in favour of each approach. The most recent performance as shown in the 2006/07 results may be a better indicator of the baseline for future efficiency savings. However, the 2006/07 results are based on data which are, in part, estimated and would need to reviewed when full year results are available. We have some concern that the 2005/06 data may have been subject to once off special factors following the separation of the companies. Taking an average of two years may smooth any such effect and has the advantage that it makes greater use of a limited data set.

5.45 At this stage in the analysis we suggest that Ofgem should give consideration to both sets of results which can be used in comparison with other work which is in hand on the analysis of operating expenditure.

APPENDIX 1: REGRESSION ANALYSIS – TECHNICAL BACKGROUND

A1.1 The basic model that we have estimated can be represented as in equation 1:

$$\ln TC_{it} = \alpha + \beta X_{it} + u_{it} \quad (1)$$

TC_{it} is total controllable costs for GDN i at time t, X_{it} is a vector of cost drivers, u_{it} is the error term, α is the constant term, β is a vector of parameters to be estimated and $\ln$ stands for natural logarithm.

A1.2 We have considered volumes of gas distributed (VOL), total number of customers (CUST), network length (LEN), the proportion of non domestic customers (PNDC), (as a proxy for the importance of large users in the customer base of each GDN), and customer density (CD), expressed as the ratio between total customers and network length as potential cost drivers.

A1.3 With the exception of the proportion of non domestic customers, all explanatory variables have been entered, as is general practice in econometric analysis of cost functions, in logarithmic form.[27]

A1.4 Given the small number of data points in our sample, as well as the high correlation which exists between the variables related to the scale of it is unlikely that all of them would enter significantly in a regression of this sort.

A1.5 We have therefore run two general to specific methodologies. In the first we have begun by regression on volume, PDNC, and CD and looked to eliminate the insignificant variables. This imposes a restriction on the relationship between customer numbers and network length.[28] In the other we have begun by regressing on VOL, LEN, CUST and PNDC and then eliminated variables.

A1.6 The model in equation 1 above neglects the panel nature of the dataset and it treats all the observations as if they are independent. In the case of panel datasets, the common

---

27 PNDC is, by construction, expressed in percentage form. The common practice is therefore to not transforming such a variable into natural logs.

28 It is not possible to include network length as an explanatory variable in the first regression due to the inclusion of customer density. The reason is that the log of CD is a linear combination of the log of CUST and the log of LEN (in fact it is just the difference between the two): the inclusion of the log of CD rather than the log of CUST and the log of LEN is based on the assumption that the coefficients of LEN and CUST, which we can call a and b bear the following relation: $a = -b$. Of course, this restriction should be tested rather than imposed a priori. However, given the small sample size, the statistical tests and the degree of multicollinearity in the variables are unlikely to provide us with sufficiently strong confidence in the results. Therefore we consider that a sensible approach would be to impose the restriction and apply the general to specific methodology to this set of regressors and then compare the results with the model produced by the application of the general to specific methodology to the original set of regressors without the imposed restriction.

practice is to assume that the error term has an individual specific component, v_i[29], representing time invariant-individual specific effects, unobserved by the econometrician, which are assumed either as fixed (which lead to the estimation of the fixed effects model) or random (which leads to the estimation of the random effects model), as shown in equation 2 below.

$$\ln TC_{it} = \alpha + \beta X_{it} + v_i + u_{it} \tag{2}$$

A1.7 In principle, the existence of the fixed effects or random effects could be tested. However, in this case the estimation of such models can be difficult to carry out, and the results may be unreliable, because of the limited sample size.[30] Indeed, for the models estimated here, pooled OLS estimation was not rejected in favour of the panel data models.[31] This also ruled out the use of Stochastic Frontier Analysis.

A1.8 We have estimated equation 1 with pooled OLS and then used the COLS procedure to measure the GDN's relative efficiencies. We have taken account of the possibility that the COLS procedure may overstate the extent of inefficiency in the sample in drawing the efficiency frontier from the regression results in Section 3.

A1.9 The COLS residuals have been computed as $\hat{u}_{it}^{\ Cols} = \hat{u}_{it} - \min \hat{u}_{it}$, where $\hat{u}_{it}$ is the residual from the pooled OLS regression in equation 1.

A1.10 Efficiency can be computed as $EFF_{it} = 0 \leq \exp(-\hat{u}_{it}^{\ \mathrm{Cols}}) \leq 1$. $1 - EFF_{it}$ can be understood as the percentage of a GDN's total controllable operating costs that are inefficient.

A1.11 Before discussing our findings, we note that, from a statistical point of view, the general to specific methodology requires that, at each step, the models passes the conventional statistical tests that make the use of OLS, and the standard errors, valid. In particular, at each step we have run the RESET test for functional form misspecification[32], heteroskedasticity[33] and normality tests[34] but we have not detected significant evidence of violations of the assumptions of homoskedasticity, normality and indications of functional form misspecification.

---

29 By individual we mean a cross sectional unit, in this case the GDNs.

30 In particular, due to the small sample size, the indicated random and fixed company specific effects are more likely to be zero. The interpretation of zero company specific errors in normal circumstances would be that all companies are fully efficient, where as really the result is produced by the small sample size.

31 In some models we could not reject the null hypothesis the fixed effects were not jointly equal to zero. However, given the very small within variability of some variables, the panel data fixed effects model is unlikely to identify with any precision the egression coefficients, which was what actually happened.

32 We have run two versions of the test, one using powers of the fitted values and another using powers of the independent variables (see, for instance, Greene, 2003).

33 We have run the Breush-Pagan test and the Cameron and Trivedi tests for heterosckedasticity.

34 We have run the Jarque Bera Normality tests.

A1.12 We ran regressions using two sets of regressors. In the first case, we included VOL, CD, and PNDC. The methodology rejected only PNDC and therefore we have adopted as model COLS1 the model with VOL and CD as regressors, whose results are reported in Table A1.1.

A1.13 The model selected in this case has the volume of gas distributed (expressed in natural logarithms) as the main scale variable. The R squared is large (0.83) and the statistical tests do not identify major evidences of heteroskedasticity, functional form misspecification or non-normality of the error term.

A1.14 We then computed a series of test statistics to explore whether some observations unduly influenced the results.

A1.15 In this case the DFFITS and the Welsh test statistics identified London in year 2 as a possible outlier; while according to the Cook statistics and, at the margin, the Student t residuals, the observations related to the London (both year 1 and 2) GDNs should be considered as influential. Considering the fact that the Model Cols1 would identify London in year 1 as the most efficient, we decided to re-estimate the Model COLS1 without the two observations for London in order not to let London to exert too much influence on the regression parameters, and to compute London's residuals as an out of sample prediction, i.e. using the regression parameters together with London's values for VOL and CD to compute the residuals and, therefore, the efficiency.

Table A1.1: Estimation of COLS1

	Ln(TCC)
LnVOL	0.80***
lnCD	0.46*
Constant	-6.64***
Adjusted R-sq	0.83
F test	38.56***
Reset 1	0.33
Reset 2	0.67
Het test 1	0.1
Het test 2	5.64
JB test	1.03

Source: Europe Economics

A1.16 The main result is that the coefficient for CD is no longer significant (although the coefficient was pretty similar) and, therefore, in a general to specific methodology, it should be omitted and the model re-estimated without the log of CD as a regressor. We note that if we performed the sensitivity analysis at the very beginning of the General to Specific Methodology, we would end up with the same models, i.e. with a model with the log of volume as the only significant regressor. The final results for Model COLS1 are reported in Table A1.2

Table A1.2 COLS1 without London

	Ln(TCC)
LnVOL	0.82***
Constant	-4.91***
Adjusted R-sq	0.87
F test	85.79***
Reset 1	0.09
Reset 2	0.09
Het test 1	0.00
Het test 2	0.77
JB test	0.76

Source: Europe Economics

A1.17 As we can see, there is a small increase in the coefficient for volume, and the adjusted R-squared increased as well. In terms of inefficiency, London in year 1 is still the most efficient. The correlation coefficient between the efficiency levels generated by the Models with and without London is as high as 0.87: as expected, the efficiency of London -computed as an out of sample prediction- drops quite significantly.

A1.18 The second regression analysis selected the model (COLS2) reported in Table A1.3 below.

A1.19 As we can see, the only regressor selected by the general to specific methodology is CUST, which acts as the scale variable in the model. The CUST coefficient is 0.79 and it is significantly lower than 1, suggesting the existence of scale economies in this sample. [35] [36] The value is similar to the coefficient estimated for the volumes scale variable in the previous model.

35 An F test rejects the null hypothesis that the coefficient of lncust is equal to 1 at 2 per cent level

36 We should remember that this very simple cost function is not the correct theoretical framework to draw any robust conclusions on technological relationship such as scale economies.

Table A1.3: Estimation of COLS2

	Ln(TCC)
LnCUST	0.79***
Constant	-7.31***
Adjusted R-sq	0.88
F test	116.10***
Reset 1	0.46
Reset 2	0.46
Het test 1	0.40
Het test 2	1.33
JB test	0.32

Source: Europe Economics.

A1.20 The R squared for the regression is large, although that is mainly a consequence of having two scale related variables in both the left and the right hand side of the regression and as such not surprising.

A1.21 The Reset 1 and Reset 2 tests are the Reset tests for functional form misspecification with powers of the fitted values and powers of the independent variables, respectively. As we can see, the tests do not detect, as in the case of Model COLS1, any evidence of functional form misspecification.

A1.22 Het1 and Het2 are the Breusch-Pagan and Cameron and Trivedi tests for heteroskedasticity, and they both do not detect any evidence of violations of the homoskedasticity assumption.

A1.23 Finally, JB is the Jarque-Bera test for normality of the error term, which is not significant at the conventional levels of confidence.

A1.24 To further test the statistical soundness of the model results (bearing in mind the small sample size), we tested for the existence of outliers and other influential observations, carrying out the DFFITS, Welsch and Cook statistics for influential observations and computing the Students T statistic.[37]

A1.25 The DFFITS and Welsh statistics did not identify any infuential observation. The Cook statistics and the Students T residuals identified West Midlands in year 1 as a possible outlier. We therefore adopted the same procedure we followed for Model COLS1 above and re-ran the Model COLS2 without West Midlands. The results are reported in Table A1.4.

[37] See Maddala (2001), for an exhaustive treatment of statistical analysis of outliers and influential observations.

Table A1.4: Estimation of COLS2 without West Midlands

	Ln(TCC)
LnCUST	0.75***
Constant	-6.69***
Adjusted R-sq	0.90
F test	126.4***
Reset 1	0.05
Reset 2	0.05
Het test 1	0.00
Het test 2	3.30
JB test	0.22

Source: Europe Economics.

A1.26 As we can see, the results were statistically very close to each other. Also in terms of efficiency levels, the correlation coefficient between the two Models (i.e. those with and without West Midlands) is as high as 0.99 which suggests that, all in all, dropping the outlier would not seem to have a great effect on the efficiency rankings. On balance we recommend the versions of both COLS 1 and COLS2 purged of the outlier observations.

A1.27 Another issue is which of the two models (COLS1 and COLS2) should be preferred on statistical grounds. They are non-nested models and the non-nested tests that are available in the literature are likely to be inconclusive, given the small sample size we have been working with.

A1.28 The first model is the product of a restriction that we have not tested for (see above), which makes the results of the second model, where we have not imposed such a restriction, a little bit more general. However, other considerations, such as the correlation with the efficiency rankings generated by other methodologies, might provide us with another criterion to inform our view on the model to be preferred (if any).

A1.29 As a further robustness check, we have estimated two additional models, which we have called COLS3 and COLS4.

A1.30 They are both based on a similar assumption to that employed by Ofgem in the case of the recent electricity distribution price control review, whereby the only regressor included into the econometric model was a composite scale variable[38], which weights together volumes of gas distributed, customers and length of mains.

[38] The methodology for building the composite scale variable is explained, for instance, in CEPA (2003).

A1.31 Models COLS3 and COLS4 differ on the exact weights chosen for the three outputs: in COLS3 we have used the same weights chosen by Ofgem at the last electricity distribution price review, namely 50 per cent to network, and 25 per cent each to volumes and customers (and the resulting composite scale variable is called Composite1); in COLS4 we have used econometric evidence for the US gas distribution sector, cited in Meryck (2004), and we have used 86 per cent for customers and 14 per cent to volumes (and the resulting composite scale variable is called Composite2).

A1.32 The rationale underlying the composite scale variable is that, given the high correlation between customers, network length and volumes of gas distributed, a multicollinearity problem would arise. Rather than using the general to specific methodology, which basically makes the three collinear variables "to fight" against each other in the regression, the composite scale variable is based on a more pragmatic approach, by constructing a composite scale variable that takes into consideration the differences in each GDN network. The results of the regression are dependant on the weights chosen.

A1.33 Table A1.5 reports the regression results for Model COLS3 and COLS4.

A1.34 The most striking difference between the two models is the lower coefficient for the composite scale variable in COLS3 with respect to COLS4, though they each lie well within the conventional confidence interval of each other. No major evidence of functional form misspecification, heterosckedasticity and non-normality are found for both Models Cols3 and Cols4.

A1.35 Also in the case of Models COLS3 and COLS4 we ran some tests to explore the issue of outliers and influential observations. In Model COLS3 each statistics computed (DFFITS, Welsh, Cook and Students t consistently identify London in year 2 as a possible influential observation. By way of contrast, in Model COLS4 West Midland in year 1 was detected as possible outlier by the Cook statistics and the students t. We therefore re-ran Model COLS3 without London in year 2 and Model COLS4 without West Midlands in year 1. The results are reported in Table A1.5. As we can see, the magnitude of the scale variable changes, although the adjusted R-squared is quite stable.

Table A1.5: Estimation of COLS3 and COLS4

	COLS3	COLS3 no outliers	COLS4	COLS4 no outliers
	Ln(TCC)	Ln(TCC)	Ln(TCC)	Ln(TCC)
LnComposite1	0.75***	0.81***		
LnComposite2			0.796***	0.76***
Constant	-6.62***	-7.57***	-7.34***	-6.73***
Adjusted R-sq	0.81	0.90	0.89	0.89
F test	66.69***	127.21***	111.19***	116.40***
Reset 1	0.48	0.09	0.23	0.09
Reset 2	0.48	0.09	0.23	0.09
Het test 1	1.42	0.00	0.38	0.01
Het test 2	1.09	01.61	1.33	2.88
JB test	1.43	2.73	0.35	0.23

Source: Europe Economics

A1.36 In Table 3.9 in Section 3 we report the efficiency levels, for each GDN in each year, computed with the COLS procedure described above. We should note that for all Models, the efficiency estimates refer to the models purged of the outliers.

A1.37 Referring to sections 3 for a correlation analysis of the efficiency rankings, we note that COLS1 and COLS3 and COLS4 produce very similar average efficiency levels in both years, while COLS2 yields somehow lower average.

A1.38 The most apparent results from Table 3.9 are the general deterioration in performance between 2005/06 and 2006/07 and, in particular, that of the NGG, while we note an improvement in performance by Wales and West and Northern..

A1.39 Whether these quite large swings in performance are genuine efficiency changes or forecasting issues for total controllable costs is not yet clear, particularly as the 2006/07 data are based on a part year projection.

A1.40 As a further robustness check of this result, we have estimated the following regression by OLS:

$$\ln TC_{it} = \alpha + \beta CUST_{it} + \gamma PNDC_{it} + \delta_1 NGG + \delta_2 NORTH + \delta_3 SCOTIA + u_{it} \quad (3)$$

A1.41 In equation 3 NGGD is a dummy variable equal to one for the GDNs belonging to NGGD and zero otherwise, WW is a dummy variable equal to one for these GDNs belonging to Northern and zero otherwise and SCOTIA is a dummy variable equal to one for these GDNs belonging to Scotia and zero otherwise.

A1.42 The rationale of estimating equation 3 is that the coefficients of the group dummy variables should tell us the percentage difference in total controllable costs for the different ownership groups with respect to Northern, which is taken as the reference group.[39]

A1.43 Table A1.6 reports the econometric estimates of equation 3

Table A1.6: Dependent variable

	Ln(TCC)	Ln(TCC)
LnCUST	0.80***	
LnVOL		0.84***
NGGD	0.05	0.13*
WW	0.12	0.13*
SCOTIA	0.06	0.17*
Constant	-7.41***	-5.21***
Adjusted R-sq	0.87	0.89
F test	27.02***	29.41
Reset 1	0.72	1.37
Het test 1	0.38	0.06
JB test		2.31

Source: Europe Economics

A1.44 As we can see, all the ownership dummies carry a positive sign, which means that, everything else being equal, they tend to have higher costs than Northern However, no dummy is statistically significant, and therefore we can not reject the null hypothesis that each group has costs that are no higher or lower than Northern[40].

A1.45 Looking at the point estimates of the dummy variables, the coefficient for the dummy for NGGD tells us that its costs are about 5 per cent higher than those of Northern over the two years sample period[41].

A1.46 However, if we use VOL as the scale variable, and we discard London year 2 because it is a possible outlier, there is some evidence that NGG, Scotia and Wales and West tend to have higher costs than Northern: in particular, NGG and Wales and West would have

---

[39] The regression 3 above neglects the panel nature of the dataset and does not take into account that there might be some within correlation between the residuals of the GDN that belong to the same ownership group. In large samples, the existence of within group correlation would be accommodated using cluster robust standard errors. However, the use of cluster standard errors is strictly valid only in large samples, and we have therefore decided to stick to the conventional OLS standard errors.

[40] We also computed a test that the three dummies were jointly equal to zero and we could not reject the null hypothesis at the conventional levels of confidence.

[41] Given the discrete nature of a dummy variable we cannot interpret its coefficient as that of a continuous variable. See Gujarat (2004) i for the correct formula.

costs about 13 per cent higher than Northern and Scotia would have costs about 17 per cent higher than Northern.

APPENDIX 2: LITERATURE REVIEW TABLE

A2.1 In the gas distribution sector output can have a number of different dimensions: the ones that have generally been considered in the applied efficiency measurement literature are the volumes of gas distributed, the number of customers and the length of mains.

A2.2 Therefore, before any description of the efficiency techniques used for benchmarking exercises is provided, it is useful to undertake a literature review, which investigates the cost drivers commonly used in the regulatory and academic literature on benchmarking in the gas distribution sector. The choice of the cost drivers however will also ultimately depend on data availability and the boundaries of the techniques applied.

A2.3 The scope of the literature review is therefore to provide a short summary of main empirical studies that have sought to undertake a benchmarking analysis of a set of gas distributor companies in different jurisdictions, as background for our study.

A2.4 In particular, we are interested in the variables that these different studies have identified as output measures and environmental variables that are relevant for the gas distribution sector.

A2.5 It is not the purpose of this section to provide a full description of the methodologies that have been used in the literature. We note, however, however, that a production frontier is a function that describes the maximum output a firm can produce using any particular set of inputs. This represents the best practice performance of the industry.

A2.6 Production functions are usually estimated using sample of outputs and inputs used by a number of firms. Alternatively, efficiency analysis is conducted using cost functions, whereby firms are benchmarked to assess, given the output produced and other cost drivers outside the company control, which company has the lowest costs.

A2.7 In the applied efficiency measurement literature there is wide range of methodologies to determine the efficiency frontier. The four approaches which have been most often used in comparison or on their own are:

(a) Least squares econometric production models (COLS);

(b) Stochastic Frontier Analysis (SFA);

(c) Total Factor Productivity Indices (TFP);

(d) Data Envelopment Analysis (DEA).

A2.8 These include statistical techniques (A and B) and non-statistical methods (C and D)[42] where the method utilised might depend among other factors on the dataset available, e.g. a limited number of observations available may restrict the usable techniques for benchmarking efficiency to indexing methodologies.

A2.9 In general, however, one of the main conceptual problems is to identify, within a set of variables, which one or ones to take as the measure of output. The literature indicates that the estimates of scale economies and density can be sensitive to the assumptions regarding the variation of output with output characteristics.

A2.10 In the gas sector the output dimensions that have generally been considered in the applied efficiency measurement literature are the volumes of gas distributed, the number of customers and the network length.

A2.11 Relevant cost drivers however might also be environmental variables as:

(e) customer density; and

(f) The proportion of residential customer's number over the total.

A2.12 In the table A2.1 overleaf below shows a summary of the main empirical results of the literature review that we have conducted. In particular we summarise the output variables and the inputs characteristics chosen by the authors, as well as the methodologies, functional forms and data set used.

42 DEA is a technique based on linear programming methods.

Table A2.1: Overview of relevant previous studies

1. Cost efficiency in the Swiss Gas distribution sector	**Author(s)**	Mehdi Farsi, Massimo Filippini, Micheal Kuenzle
	Year	2004
	Sector	Gas distribution
	Relevance to research	High
	Methodology	Stochastic Frontier Analysis - Total cost approach
	Database used	Panel of 26 distribution gas companies operating between 1996 and 2000.
	Functional form	Cobb-Douglas
	Measures of output and other environmental variables	Volume of gas delivered/number of terminal blocks, customer density, area size, number of customers per km network length, network length in km
	Measures of inputs	Firm average annual price per employee, firm average annual price per meter network length, firm average annual energy price per KWh
	Main findings	The results suggest that the utilities could slightly reduce their operating costs by improving efficiency. There is no evidence of significant unexploited scale economies. However the analysis indicates that the estimates of scale economies could be sensitive to the assumptions regarding the variation of output with output characteristics.
2. Comparative benchmarking of Gas Networks in Australia and New Zealand	**Author(s)**	Meyrick and Associates
	Year	2004
	Sector	Gas distribution and transmission
	Relevance to research	High
	Methodology	Total Factor Productivity Indexes (TFP) Partial Productivity Indexes
	Database used	Cross section of 4 New Zealand and 10 Australian gas pipeline distributors 2003 Time-series 1997-2003 for 8 transmission operators
	Functional form	Multi-lateral TFP and partial Indexes
	Measures of output and other environmental variables	Throughput, local access roads (number of connections), system line capacity.
	Measures of inputs	Directly pipeline length (or indirectly dollar measure of the value of assets); directly estimated depreciation rate ODV or depreciated replacement costs of assets DORC (or indirectly residual of revenue loss operating costs).
	Main findings	The limited number of observations available for gas distributors in Australia and New Zealand restrict the usable techniques to benchmarking efficiency to indexing methodologies
3. International Benchmarking for Monopoly Price Regulation: The Case of Australian Gas Distribution	**Author(s)**	Roger Carrington, Tim Coelli, Eric Groom
	Year	2002
	Sector	Gas distribution
	Relevance to research	High

	Methodology	Data Envelope Analysis (DEA) Stochastic Frontier Analysis (SFA)
	Database used	59 Australian and US distributors – 1997 data for U.S. distributors and a combination of financial and calendar year information for the years 1997-98, 1998-99 and 1998 for the Australian distributors.
	Functional form	DEA: Constant Returns to Scale (CRS) and Variable Returns to Scale (VRS)
	Measures of output and other environmental variables	Total deliveries, residential customers number, other customers number
	Measures of inputs	Length (kilometers) of distribution (as capital proxy), Employees.
	Main findings	Benchmarking can help reduce the problem of asymmetric information on efficient costs. Results suggest that there is scope to improve efficiency and the choice of benchmarking technique is not unduly influence the results of the exercise.
4. Technical change and efficiency measures: the post privatisation in the gas distribution sector in Argentina	**Author**	Martin, A. Rossi
	Year	2001
	Sector	Gas distribution
	Relevance to research	High
	Methodology	SFA Correct Ordinary Least Squares (COLS)
	Database used	Panel of 8 Argentinean gas distribution companies 1993-1997
	Functional form	Cobb-Douglas production functions / COLS and Battese and Coelli (1992) model
	Measures of output and other environmental variables	Volume of gas delivered, number of customers/service area, share of residential to total sales, maximum demand
	Measures of inputs	Network length (as a proxy for capital input), number of employee (as labor input), environmental variables such as concession area, ratio residential sales to total sales, maximum demand
	Main findings	In the post privatization in the gas distribution sector in Argentina there is both a catching up effect and a shift in the frontier. This phenomenon holds for the average and also for every firm in the sample.
5. Costs, technology and ownership of gas distribution in Italy	**Author**	Paola Fabbri, Giovanni Fraquelli, Roberto Giandrone
	Year	2000
	Sector	Gas distribution
	Relevance to research	High
	Methodology	Stochastic Frontier Analysis (SFA)
	Database used	Panel of 31 Italian distribution companies 1991-1992
	Functional form	Translog cost function/SURE

	Measures of output and other environmental variables	Volume of gas delivered, network length/ inverse of customers density, average altitude of the service area, population density
	Measures of inputs	Yearly average cost per employee (as labor price), the book-value of equipment divided by the length of the distribution network (as capital price) and the price of material and services is calculated as the residual expenses divided by network length.
	Main findings	The analysis highlights the fact that the number of customers is more important than the amount of gas delivered in explaining the variability of distribution costs. Very low economies of scale, high economies of density, and the significant role of the morphologic and demographic variables characterize the nature of the technology.
6. Parametric and Nonparametric Approaches to Benchmarking the Regulated Firm	**Author**	Gerald Granderson, Carl Linvill
	Year	1999
	Sector	Natural gas transmission companies
	Relevance to research	High
	Methodology	Stochastic frontier Analysis (SFA) Data Envelope Analysis (DEA)
	Database used	11 year panel of 20 US interstate natural gas transmission companies 1977-1987
	Functional form	Translog cost function/ random effect model
	Measures of output and other environmental variables	Volume of gas delivered
	Measures of inputs	n.a.
	Main findings	Results show that using a non parametric approach the inefficiency estimates are lower but the inefficiency ranking stays more or less the same.
7. The marginal cost	**Author**	Jean Thomas Bernard, Dennis Bolduc, Annie Hardy

	Year	1998
	Sector	Gas distribution
	Relevance to research	High
	Methodology	Stochastic Frontier Analysis (SFA)
	Database used	Cross section of 131 Canadian gas extension projects.
	Functional form	Linear Box-Cox cost function
	Measures of output and other environmental variables	Max daily demand, network pipe length
	Measures of inputs	n.a.
	Main findings	Incorporating distribution capital cost aspects in the variable part of the tariff structure makes prices too high in terms of economic efficiency. Capital cost is in fact not separable into a fixed and a variable component. Also elasticity with respect to maximum daily demand is not significant, and elasticity with respect to pipe length is slightly less than one. Maximum daily demand by each consumer class and consumer density per kilometer plays no statistically significant role.

Source: Europe Economics

APPENDIX 3: REFERENCES

Bernard, J.T., Bolduc, D., Hardy, A., “The marginal cost of natural gas distribution pipelines: the case of Societe en Commandite Gaz Metropolitan, Quebec” Energy Journal, 24(5), 425-438, 2002

Carrington. R., Coelli T., Groom, E., “International Benchmarking for Monopoly Price Regulation: The Case of Australian Gas Distribution”, Journal of Regulatory Economics; 21:2 191-216, 2002.

Caves, D.W., Christensen L.R., Diewert, W.E., “Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers”, Economic Journal, Vol. 92, March 1982.

CEPA “Background to work on assessing efficiency for 2005 distribution price control review 2005” September, 2003.

Charnes, A., Cooper, W., Rhodes, E., "Measuring the efficiency of decision-making units," European Journal of Operational Research, 1978.

Coelli, T., Prasada, Rao D.S., Battese, G.E., “An Introduction to efficiency and productivity analysis”, Kluwer Academic Publishers, 1998.

Fabbri, P., Fraquelli, G., Giandrone, R., “Costs, technology and ownership of gas distribution in Italy” Managerial and decision Economics, 2000

Farsi, M., Filippini, M., Kuenzle, K., “Cost efficiency in the Swiss Gas distribution sector”, Centre of Energy Policy and Economics working paper No. 36, 2004.

Granderson, G., Linvill, C., “Parametric and Nonparametric Approaches to Benchmarking the Regulated Firm”, Journal of Productivity Analysis, Volume 12, No.3, November 1999.

Greene, W.H. “Econometrics Analysis” New York University press, 2003

Gujarat, D. N., “Basic Econometrics” McGraw-Hill International edition, 2004

Khumbakar, S. and K., Lovell Stochastic Frontier Analysis, Cambridge University Press, 2000.

Maddala, G.S. “Introduction to Econometyrics”, Wiley edition, 2001

Meyrick and Associates “Comparative benchmarking of Gas Networks in Australia and New Zealand”, Report prepared for Commerce Commission, Wellington, 2004.

Padraja-Chaparro, F., Salinas, P., “On the role of weights restrictions in data envelope analysis” 8, 215-230, 1997.

Rossi, A. M., “Technical change and efficiency measures: the post privatisation in the gas distribution sector in Argentina”, Energy Economics 23 (2001) 295-304.