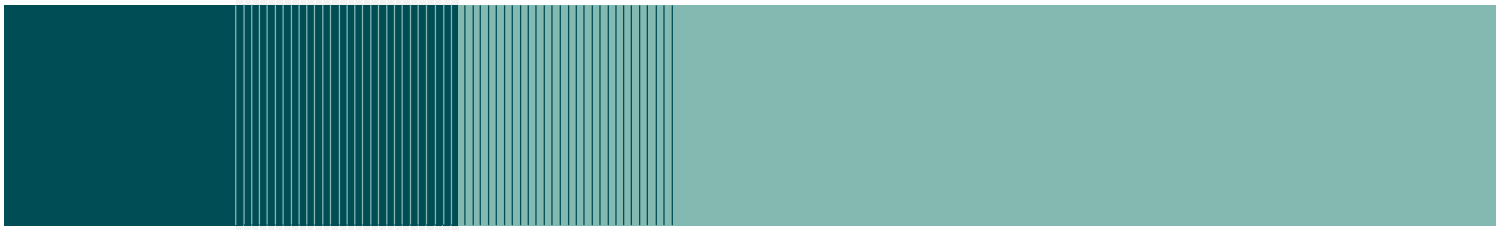
15 October 2007

Review of Ofgem Benchmarking Studies

Prepared for Wales and West Utilities



NERA
Economic Consulting

Project Team

Graham Shuttleworth
Richard Druce
Marco Schönborn

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com

Contents

1. Overview of Ofgem's Benchmarking Studies

Price cap regulation makes revenues independent from actual costs, and so creates an incentive for regulated firms to reduce costs. However, the level of allowed revenues still has to bear some relation to expected actual costs. During the current gas distribution price control review (GDPCR), Ofgem has applied benchmarking techniques to opex and capex in an attempt to estimate the level of "efficient" costs, as a technique for setting the level of the revenue cap. Benchmarking, if not done or interpreted properly, can create several problems that make it unsuitable for regulation. The Monopolies and Mergers Commission (MMC) highlighted some of these general problems with benchmarking in its 1997 inquiry into Northern Ireland Electricity,[1] in which it rejected the regulator's benchmarking as lacking in robustness (because other benchmarking exercises showed different results).

More specifically, however, there are several problems with both Ofgem's methodology and application of benchmarking, which – if implemented – may cause large divergences between GDNs' allowed revenue and their actual costs over the coming control period, without good reason. Therefore, Ofgem's benchmarking results must be treated with caution and should be adjusted in the light of other information besides the conclusions of its own benchmarking exercise.

In this chapter, we outline several problems which can arise when benchmarking analyses are used to set regulated firms' allowed costs. Such problems can arise when benchmarking has not been conducted properly, when adjustments have not been made to compensate for deficiencies in benchmarking techniques, and when the results are interpreted incorrectly. We also outline various problems and weaknesses in Ofgem's benchmarking analyses, and suggest some adjustments to the results that Ofgem should apply when setting allowed revenues.

1.1. The Lack of Economic Logic

1.1.1. Financeability

For most direct opex activities (for example), Ofgem has benchmarked GDNs' costs with reference to an estimated "benchmark" or "frontier" level of costs, and disallowed any of GDNs' forecast expenditure that is above this level. At the same time, Ofgem has calculated the cost of capital using parameters derived from average capital market data. However, Ofgem's approach means that GDNs with "average efficiency" will earn less than an average return. That outcome is not consistent with the GDNs' need to attract capital from investors who can put money into other sectors. Therefore, if the results of Ofgem's benchmarking studies are not applied correctly, GDNs will be denied a comparable return on capital, after recovery of opex and depreciation costs.

GDNs' ability to finance their activities is reduced by Ofgem's decision not to allow the GDNs any time to converge to the "efficient" (i.e. target) level of costs, but to impose it from the first year of the regulatory period. Ofgem should make an adjustment to its

---

[1] *Northern Ireland Electricity Plc: A report on the reference under Article 15 of the Electricity (Northern Ireland) Order 1992*, Monopolies and Mergers Commission, 1997, paragraph 2.159.

benchmarking results when setting allowed revenues to ensure that GDNs can finance their activities, by applying a "glidepath", for instance.

Incidentally, by not allowing any time for "convergence" (a glidepath), Ofgem is clearly stating that all costs above the target level should be disallowed because they are inefficient. This approach contrasts with decisions in earlier regulatory reviews (by Ofgem and other regulators) in which the regulated companies are only required to meet the target level of costs after some years. In such decisions, it is less clear whether the regulator is disallowing costs because companies ***should not*** be incurring them, or merely adopting a forecast which assumes that the companies ***can*** eliminate certain costs within a certain time. Ofgem's decision unambiguously treats all costs above the target as inefficient – even though Ofgem's benchmarking conclusions are not sufficiently robust to reach this conclusion.

Moreover, Ofgem has carried out its financeability analysis after deduction of the disallowed costs. This approach undermines the purpose of such analysis, since it does not represent a true test of the GDNs' ability to finance their activities. When Ofgem disallows costs, the costs do not go away; Ofgem's decision merely prevents the companies from recovering them in their revenues. Such decisions may cause real problems for the GDNs' ability to finance their activities, but a test applied to a purely notional level of costs will not check whether Ofgem is meeting its statutory duties in this respect.

1.1.2. Economies of Scale

During the GDPCR, Ofgem has stated that it will not allow smaller GDNs an additional allowance for their inability to achieve economies of scale when setting their indirect opex allowances.[2] Ofgem has also instructed its consultants to ignore economies of scale when undertaking benchmarking analyses.

It seems unlikely that Ofgem's approach is sustainable. Given the limited opportunities for consolidation of the industry, Ofgem cannot expect smaller GDNs to achieve the same indirect opex costs per unit as larger GDNs, even if they are well-managed. The only way that the smaller GDNs could achieve the same costs of larger scale operations as NGG is through a re-merger of all four GDNs. Ofgem would not have permitted the sale of the GDNs in the first place if it believed merger was the desired outcome. Now that the sale has taken place, various financial constraints require that each GDN is allowed a revenue that offers a reasonable prospect of cost recovery.

Because smaller GDNs cannot avoid having higher average costs than larger GDNs, it is not sustainable in the long-run for smaller GDNs to bear the higher average costs due to their smaller size without compensating revenue. Any failure to account for economies of scale would hinder smaller GDNs' ability to finance their activities and will not allow the GDNs to attract capital for efficient investment. Hence, looking to the future, allowed revenues should offer a reasonable prospect of recovering total costs, including any costs incurred as a result of diseconomies of scale (if such costs were not demonstrably inefficient or imprudent).

---

[2] Gas Distribution Price Control Review: Initial Proposals, Ofgem (125/07), 29 May 2007, para. 3.21.

1.1.3. Partial Benchmarking

Ofgem has benchmarked indirect opex, direct opex, capex and repex costs separately. For each of these cost categories, it has conducted a combination of "top-down" analysis of the total cost category and "bottom-up" analyses of detail cost items (sub-categories). Ofgem recognises that the sum of bottom-up benchmarks for separate cost categories would ignore the possibility of trade-offs between different types of expenditure, and so underestimate the feasible level of total expenditure. Therefore, for direct opex Ofgem raised the "efficiency" targets emerging from its bottom-up analysis by 6.2%, representing the average difference between the upper quartile targets given by its top-down and bottom-up benchmarking.[3] However, Ofgem did not apply any similar adjustment to the other cost categories. Nor has Ofgem made any similar adjustment for tradeoffs between repex, capex or indirect opex, even though it set separate targets for these categories.

Ofgem has thus identified a problem with its method with reference to direct opex, and acknowledged that its results must be treated with caution. It has also suggested one possible solution that adjusts its results in the light of other information. However, Ofgem has not applied that solution to every case where the problem arises. Taken together, Ofgem's current benchmarking targets may be infeasible. Moreover, if Ofgem wishes to use the results of its bottom-up analyses to set allowed costs, it must adjust them to account for trade-offs between input categories. Ofgem could avoid this problem by identifying the exact source of inefficient expenditure (e.g. inflated prices for materials, lax project management techniques leading to cost over-runs, etc), or minimise it by benchmarking broader cost categories (which Ofgem calls "top-down" benchmarking).

1.1.4. Choice of Benchmarking Variables

Benchmarking analyses require an assessment of what are the most relevant cost-drivers. In general, when estimating a cost model for benchmarking, regulators should justify the use of a cost-driver based on an engineering analysis of network costs, and statistical evidence that the cost-driver in question does affect costs. Benchmarking models should also, as far as possible, control for differences in operating conditions between firms. However, because benchmarking studies inevitably cannot control for all such differences, their results must be treated with caution and adjusted in the light of other information on the level of costs that firms are able to achieve.

Where Ofgem has conducted benchmarking regressions, it has used "Composite Scale Variables" (CSVs) is an attempt to control for the size of the network. Ofgem's CSV variables are in most cases crude measures of network size, which Ofgem has made no attempt to justify on engineering grounds. In particular, the weightings and cost elements included in the CSVs are highly arbitrary (see Appendix A). One particular example of Ofgem's failure to justify its choice of CSV regards maintenance opex. In this case, PB/Rune conducted a benchmarking regression using a CSV, but Ofgem's simplified this variable without offering a robust explanation as to why (UPs, para. 3.23):

---

[3] *Gas Distribution Price Control Review Updated Proposals* (denoted "UPs" elsewhere in this document), Ofgem (226/07), 24 September 2007, para. 3.61.

> "*We consider that this revised approach to maintenance is more robust and gives results that are reasonable in the context of historical expenditure for each GDN.*"

Another example where Ofgem's benchmarking approach lacks economic logic relates to the choice of benchmarking variable for indirect opex. For the indirect opex benchmarking, Ofgem uses GDNs' revenue as a driver for costs in a number of activities (UPs, tables A7.3-A7.10). This approach makes GDNs appear less "efficient" in indirect opex, if (compared to their comparators) they have reduced costs in other categories and therefore have lower revenues.[4] Such perverse effects arise because Ofgem's consultants have adopted a number of ad hoc benchmarks without giving proper consideration to the economic rationale for their approach.

The results of benchmarking can be highly sensitive to the choice of benchmarking variables used in regressions or other analyses. Therefore, Ofgem needs to justify its choice of benchmarking variables, using engineering and statistical evidence to support its choice. It should also consider any submissions from GDNs which suggest alternative variables or highlight unintentional implications of its choice of benchmarking variables. Where GDNs highlight problems with its analyses, Ofgem must treat results with caution and should provide reasons for adopting one interpretation of the available evidence rather than any other.

1.2. Robustness

Techniques for setting revenues for regulated companies must be objective, and any use of statistical analysis can only be justified if the results are robust. If different analyses produce results that differ widely, the regulator's use of one technique rather than another becomes a matter of arbitrary choice, not objective support for a reasoned decision. Benchmarking analyses are not robust whenever arbitrary choices or unjustified assumptions affect the conclusions. Such choices and assumptions may include the choice of a limited subset of explanatory variables, the weights given to different variables when constructing a composite workload variable, and the interpretation of unexplained (residual) costs as due to inefficiency rather than other factors.

Within the terms of benchmarking, robustness may be understood as requiring analysis to meet the following criteria:

1. The efficiency scores of a company should not fluctuate widely over short periods of time;
2. Adding or removing a few observations should not alter the results significantly; and
3. The results should not vary significantly due to subjective choices over the specification of a model (choosing one variable instead of others, setting the weights of composite variables, etc) or interpretation of the results (how much of the unexplained costs is attributed to inefficiency).

---

[4] Such problems are specific to regulated companies. Their revenues reflect their costs and so fall if they reduce their costs. Companies operating in competitive markets will not experience the same phenomenon. Hence, even if it were "standard practice" to compare costs to total revenues in competitive industries, the suitability of a revenue driver for regulated industries has to be justified anew.

Where benchmarking studies do not meet the above criteria, the results must be treated with caution and adjusted in the light of other information. Any decision to adopt one set of results rather than another should be explained.

1.2.1. Workload "CSV" Variables

Ofgem has not provided any economic explanation of its decision to adopt simple linear regressions using CSVs as the explanatory variable and, moreover, has not justified its choice of the weights used to construct these CSVs. The PB/Rune report on capex/repex does not even state for most categories of capex/repex expenditure what weights were used for the activities included in workload variables. In most cases, the report simply says that workload volumes were weighted by some unspecified unit costs.

Because the assumptions underlying Ofgem's choice of explanatory variables in benchmarking regressions are arbitrary and unjustified, other benchmarking models, such as those proposed by GDNs, are equally valid and should be accorded equal weight. If Ofgem wishes to choose one model rather than another, some explanation is required.

1.2.2. Stability of Results

The results of Ofgem's benchmarking regressions have shown a lack of robustness because the results they produce are not stable. Their instability was demonstrated when PB/Rune updated its benchmarking regressions with 2006/07 data rather than 2005/06 data. For some areas of expenditure (e.g. mains and service repex), merely updating its regressions implied that GDNs had achieved large efficiency savings of 10% or more. Productivity improvements of this magnitude are highly unlikely to have occurred in one year and reflect spurious effects caused by differences between data in different years.

Any work based on regression is likely to be vulnerable to this problem, since there are too few data points for proper statistical analysis. Any regressions with only eight observations will be unstable and unreliable. Because of this limited dataset, the results from regression analysis must be treated with caution and adjusted in the light of other information besides the regression results.

1.2.3. Agreement Between Benchmarking Techniques

If benchmarking analyses are accurate, they will produce similar results for the same group of companies. GDNs' "efficiency" rankings resulting from the top-down and bottom-up direct opex benchmarking analyses show poor correlation (see Table 1.1). Because the top-down and bottom-up methods appear to give different rankings of the companies, one or both of the methods must be inaccurate, which suggests that Ofgem's benchmarking has not been conducted properly.

Table 1.1
Summary of Ofgem Results

	2006-07	Top Down Analysis			Bottom Up Analysis		
	Opex	Upper Quartile Target	Implied Inefficiency	Rank	Upper Quartile Target	Implied Inefficiency	Rank
	(£m)	(£m)	(£m)		(£m)	(£m)	
EE	107	99	8	5	89	17	8
Lo	78	65	13	8	67	11	3
NW	90	78	12	7	74	16	6
WM	61	62	-1	2	52	9	2
NGN	74	77	-3	1	69	5	1
Sc	64	59	5	4	47	17	7
So	110	100	9	6	96	13	5
WWU	75	75	0	3	64	12	4
TOTAL	**659**	**615**	**44**		**559**	**100**	

The LECG indirect opex benchmarking study also suffers from this problem, as we discuss below (see Table 3.1). The GDN rankings from LECG's benchmarking of total indirect opex and the bottom-up benchmarking of the separate indirect opex activities do not agree, and the "efficiency" targets differ considerably between the two methods.

Note however that agreement between a selection of benchmarking methods does *not* mean that they must be accurate. If all the techniques are biased in the same way (e.g. by omitting the same explanatory factor), they may produce similar – but inaccurate – results. This standard therefore acts only as a filter – which Ofgem's results fail to pass.

1.2.4. Cost Standardisation

Benchmarking requires that the cost data of all firms in the sample are comparable, i.e. that they are prepared using the same accounting rules. This is not the case in LECG's indirect opex benchmarking. The problem arises because Scottish and Southern Energy (SSE) provides support services to Scotia Gas Networks (SGN) at marginal cost. Therefore, SGN's support service costs are not comparable with those of the other GDNs.

Ofgem has acknowledged the existence of this bias and tries to compensate for it by adjusting the standard used to set the benchmark. However, the adjustment is non-standard and the effects are unpredictable, since SGN retains some influence over the adjusted benchmarks. Ofgem also adjusts the overall results of the bottom-up benchmarking by reference to a standard set by the top-down benchmarking. However, Ofgem does not consider whether that standard is biased downwards by the inclusion of SGN.

The biases due to inclusion of SGN's data cannot be solved by ad hoc adjustments to the standards used to set benchmarks. Instead, Ofgem should either have adjusted the data for SGN to put it on a comparable basis, or else have omitted SGN altogether.

When standardising the costs for use in benchmarking, Ofgem has not taken account of the quality of output, which affects the comparability of GDNs' costs. For instance, where GDNs fail to meet quality of service targets, Ofgem should presume that their "efficient" costs would be higher if they were meeting these targets. Therefore, the actual costs of GDNs that underperform against quality of service targets should not determine the

benchmark level of costs. Instead, Ofgem should either adjust the cost data of underperforming GDNs to put it on a comparable basis, or else omit underperforming GDNs from benchmarking analyses altogether.

1.3. The Burden of Proof

Ofgem has asserted that costs in excess of those defined by the benchmark are due to "inefficiency" and should be disallowed. In fact, costs may lie above an "efficiency frontier" (or any other standard) because operating conditions differ in ways that Ofgem's benchmarking regressions do not capture, and not solely due to "inefficiency". This problem means that benchmarking results must always be treated with caution.

Each GDN may be unable to identify the specific factor that causes its costs to lie above a benchmark defined by the specific conditions of other companies, because of the difficulty of understanding both the position of other companies and the calculation of the benchmark. GDNs' inability to explain cost differences is not therefore evidence that the differences are due to inefficiency. However, Ofgem has disallowed such costs as inefficient, without offering objective grounds for doing so.

For instance, PB/Rune defined a range within which they believe the unit costs of LTS and storage capex should lie. The report stated that GDNs should be able to achieve unit costs towards "the lower end" of that range, unless there were specific "route difficulties" for a particular project. However, PB/Rune did not define how route difficulties should be measured, or what effect they would have on unit costs, relative to pipelines at the lower end of the range. In order to avoid disallowances of certain projects, therefore, GDNs would have to guess what represented a standard level of "route difficulty", demonstrate how a project exceeded that standard level, and estimate the additional costs associated with the incremental "route difficulty".

Additionally, LECG's indirect opex benchmarking report acknowledges – to some extent – that benchmarking can mistake unobserved differences between firms for inefficiency. For instance, LECG states that "*variations in GDN performance may be due to factors such as difference in organisational structure. For example, for FTE based metrics, a higher cost per FTE may be due to more outsourcing rather than inefficiency.*"[5] However, despite this acknowledgment early on in the report, LECG's approach does not allay the concern that unmeasured differences between firms may bias GDNs' efficiency scores.

In conducting all of its benchmarking analyses, Ofgem should have considered how its results may have been affected by unobserved or immeasurable differences between GDNs (or external comparators). Ofgem should consider any evidence from GDNs that explains why GDNs have costs above a benchmark level. However, Ofgem should also explain what evidence justifies disallowing costs, when they may be due to unobservable factors other than differences in efficiency.

---

[5] Office of Gas and Electricity Markets: Update assessment of indirect opex based upon 2006/07 actual performance (Confidential Draft), LECG, 28/07/2007, para. 2.18.

1.4. The Rate of Transition

When applying the results of benchmarking to set allowed costs for regulated companies, regulators must specify the time by which companies should reach the estimated "efficient" level of costs. The implied rate of productivity growth required by the firm to reach this level of costs is another source of subjectivity. To remove subjectivity, regulators should apply a robust methodology, by extrapolating trends in productivity growth from historic data.

In other British regulated sectors where regulators have attempted to estimate the "efficient" level of costs through benchmarking, they have allowed companies time to achieve the benchmark, by applying a "glidepath" from the current level of costs to the benchmarked level. Applying a glidepath allows firms additional revenue to finance the one-off costs of reducing ongoing opex costs and reflects the observation that reducing costs takes time. Providing no such glidepath implies that the regulator is confident that any "excess" costs are being incurred unnecessarily and inefficiently (as discussed in section 1.3 above). However, contrary to Ofgem's statements, allowing a glidepath does *not* reduce the incentives for firms to cut costs, or create perverse incentives, since companies still benefit from cutting costs relative to the revenue target.

In the Initial Proposals,[6] Ofgem indicated that it would not apply a glidepath for opex, but did not make a corresponding statement for capex. PB/Rune's June 2007 report[7] states firmly that they expect networks to take some time to close the gap between actual and target costs:

> *"We expect Networks behind the upper quartile to improve and close the gap and we have set the Network the target of closing 70% of the cost gap to the upper quartile over the five years to 2012/13."*[8]

In its updated September 2007 report,[9] PB/Rune's recommendations for work management, emergency and maintenance costs all assume that companies will be able to close 70% of an efficiency gap by 2012/13.[10] If Ofgem decides not to offer such a glidepath, but to impose the benchmarked level of costs immediately, it will be acting against the recommendations of its consultants. Unfortunately, it is not possible to say for certain how Ofgem has departed from its consultants' recommendations, since the calculations are not reported in full in the Updated Proposals. (See for example paragraph 1.13 in Appendix 7.)

Ofgem's analysis has also failed to provide rigorous and objective justifications for the assumptions about the future scope for productivity growth. For instance, statements like the

---

6 Gas Distribution Price Control Review: Initial Proposals, Ofgem (125/07), 29 May 2007, para. 3.20.

7 PB Power/Rune Associates (2007a), "Gas Distribution Price Control Review – Five Year Control – (Opex) – Report 1 – Wales & West Network", 18 June 2007. One report was produced for each GDN. For the sake of simplicity we concentrate here on the report for Wales and West Utilities. The report is available at http://www.ofgem.gov.uk/Pages/MoreInformation.aspx?docid=175&refer=Networks/GasDistr/GDPCR7-13.

8 PB/Rune (2007a), p. 59.

9 PB Power/Rune Associates (2007b), "Gas Distribution Price Control Review – Five Year Control – (Opex) – Updated Report – (Confidential)", Reference number 62533 2006 36621 (Update v1.1), plus associated appendices, 18 September 2007.

10 PB/Rune (2007b), Tables A1-6, A1-13 and A1-25.

quotation below, provided without any supporting evidence, offer no objective justification for setting productivity targets.

> "*In response to initial proposals a number of GDN owners raised concerns regarding the productivity assumptions applied to connections, mains reinforcement and repex. We have reviewed these in light of additional information and have revised the productivity assumptions for both mains reinforcement and connections to 1.5 per cent per annum. We consider that a 2 per cent per annum productivity assumption is still appropriate for repex based on the top end of the range of assumptions put forward by the GDNs.*"[11]

Ofgem has not discussed why the figure at the "top end" of the range is appropriate, or why a figure quoted by one company would be appropriate to others. Even within the terms of Ofgem's own framework, the "top end" estimate of productivity growth might include substantial "catch-up" (which is not available to all GDNs), as well as "frontier shift" (which applies to all GDNs).

1.5. Conclusions

If conducted improperly, benchmarking faces several problems when used for regulatory purposes. In several areas of Ofgem's benchmarking analysis, its findings are not robust and its approach does not provide sufficient justification for disallowing GDNs' costs. In particular:

- Ofgem's benchmarking analysis uses regressions which are unreliable because they use too few data points, and functional forms that it has not justified;
- Ofgem's benchmarking analyses produce results which are not consistent between methods or across time (i.e. from updating with new data from 2006/07);
- Ofgem has not allowed for the inability of its benchmarking techniques to distinguish "inefficiency" from unmeasured differences between firms;
- Ofgem has not fully corrected for the infeasible "efficiency" targets that arise from benchmarking different cost categories separately (i.e. partial benchmarking); and
- Ofgem has not applied a consistent policy regarding "efficiency" catch-up, and has not allowed GDNs the cost of improving efficiency in its most recent proposals for GDNs' allowed costs.

When not done properly, benchmarking is not sufficiently reliable for use in regulatory applications. As we have discussed, due to errors in Ofgem's analysis and approach, the claim that it has correctly identified the "efficient" level of costs is unjustified in reality. The result of Ofgem's analyses is a set of revenue allowances that lack any objective justification and which are little more than random outputs of a non-transparent procedure that nobody understands completely.

[11] Gas Distribution Price Control Review: Updated Proposals, Ofgem (226/07), 24 September 2007, para. 4.30.

The next few months provide an opportunity for Ofgem tochange its approach or adjust its results in the light of the shortcomings in the analysis and gaps in the evidence discussed above. Any revised proposals should rely only on techniques which produce robust results, and not on subjective or unjustified assumptions. Because of the problems we have highlighted, Ofgem's consultants' results should be treated with caution and adjusted in the light of other information that GDNs provide.

2. Review of the PB/Rune Capex Report

In this chapter, we review in detail the PB/Rune capex/repex benchmarking study, dated September 2007. Ofgem commissioned PB/Rune to conduct a capex/repex benchmarking study, which it completed in June 2007. The September report is an update to the June report. We have principally reviewed the September update, but as the methodology is in most instances unchanged, we refer to both reports in this chapter.

2.1. LTS and Storage Capex

2.1.1. Demand Assumptions

In its Updated Proposals (UPs 4.21), Ofgem suggests that WWU has overstated demand, through the way it has used NTS and xoserve demand forecasts. PB/Rune states (A1, page 6, para. 8)[12] that "*further investigation of the assumptions and methodology used in the different forecasts would be beneficial."* As PB/Rune suggests, neither Ofgem nor its consultants have evaluated which of these demand forecasts is (1) most likely to be accurate, and (2) most appropriate for calculating capex allowances. Ofgem should therefore examine the assumptions underlying the two and assess which is appropriate for making capex forecasts.

2.1.2. The Effect on Demand from UNC Mod 90

PB/Rune suggests that its capex recommendations are based on a scenario in which the current "enduring" NTS offtake arrangements persist, there is no effect on demand from the introduction of Uniform Network Code (UNC) Modification 90 (section 3.3 para. 6-7, A9.1.1 and A1.2.5), and NTS flexibility capacity is available at zero cost wherever GDNs need it, unless there is evidence of a local constraint (this is "scenario 1"):

> "*The expenditure proposals are based on the Scenario 1 in relation to demand and the availability of NTS diurnal storage.*" (A1.2.5)

However, PB/Rune recognises the uncertainty over the effect of introducing Mod 90 on consumer behaviour, and hence demand. PB/Rune recommends that GDNs undertake modelling to estimate the impact of introducing this modification.

Because PB/Rune's main capex recommendations are based on "scenario 1" assumptions, which assume that current interruptions arrangements persist, and because Mod 90 has been approved and will be implemented in April 2008, these recommendations are inherently unrealistic.

We have not appraised the impact of Mod 90 on GDNs' operating conditions, but we understand it may result in some currently interruptible users "going firm". If so, GDNs may have to increase capacity to meet their 1-in-20 peak demand obligations. Accordingly, "scenario 2," under which Mod 90 is implemented, has higher demand assumptions and leads PB/Rune to recommend higher capex allowances.

---

[12] Unless stated, all bracketed references in this chapter refer to: *Ofgem Gas Distribution Price Control review Five Year Control (Capex/Repex) Update Report v1.1.1*, Parsons Brinckerhoff Ltd /Rune Associates, 24 September 2007.

Although some uncertainty may exist regarding the precise impact of Mod 90, the selected capex targets should contain some extra allowance for the increased capex it will require GDNs to undertake (A9).

2.1.3. The Availability of NTS Diurnal Storage

PB/Rune's recommendations for LTS and storage capex allowances are substantially lower than WWU's forecasts. This appears to be mainly because of PB/Rune's assumption that sufficient diurnal storage will be available from the NTS at zero cost to meet the forecast increases in demand for this product (A17, pages 6-7). However, PB/Rune cannot rule out localised shortages in diurnal storage, as it does not have data on the demands of NTS directly connected customers (A17, page 7).

GDNs may have prepared their capex projects on the assumption that diurnal storage would provide an adequate justification, even though they were required for a variety of other reasons as well. In the discussion on Mod 116, it was established that "flexibility" capacity is a joint product with peak or "flat" capacity, i.e. that the investment cost of a pipeline project could be allocated either to flexibility (diurnal storage) or to peak capacity. Ofgem should therefore give the GDNs a chance to reassess their projects, to see if they can be justified by their ability to meet a need for services other than diurnal storage.

2.1.4. Benchmark Unit Costs

To appraise LDZ investment project costs, PB/Rune estimated benchmark unit costs in £ million/kilometre of pipeline for different diameters of pipe, and used the median cost from their sample at each diameter.[13] PB/Rune states that it adjusted these unit costs where it has specific evidence of "route difficulties" for individual projects (page 16, para. 6). That is, the burden of proof lies with the company to demonstrate that their costs may be higher than average in some cases for reasons other than "inefficiency."

When assessing projects in the latest version of the report, PB/Rune has applied a range of pipeline construction unit costs, and stated that a pipeline without route difficulties should be towards the lower end of the range (table 3-4). The fact that PB/Rune has accounted for "route difficulties" recognises that these simple benchmarks cannot account for all those factors that influence project costs. Therefore, where PB/Rune has assessed any of WWU's projects using these ranges of unit cost benchmarks, GDNs will have to make Ofgem aware of any possible "route difficulties". However, in practice, it may be difficult – if not virtually impossible – to assess the "ordinary" level of "route difficulties" that PB/Rune's benchmark has taken into account and the incremental cost of a "higher" level of "route difficulties".

PB/Rune has also estimated a range of costs for providing increased diurnal storage as a result of interruptible users "going firm." It considers only diurnal storage projects with costs within this range to be "economic." The range of costs is from £50 to £100 per mcm of diurnal storage "sterilised". PB/Rune's cost estimates come from the same dataset as it used for the unit cost analysis.[14] However, projects that increase diurnal storage also provide

---

[13] PB/Rune, June 2006, *Ofgem Gas Distribution Price Control Review Five-Year Control (Capex/Repex) Report 1: Wales & West Network*, Appendix 6.

[14] PB/Rune, June 2006, Appendix 5.

additional benefits, because diurnal storage is a by-product of gas pipeline capacity constructed to meet peak demand requirements, as we discuss above. Therefore, disallowing projects as "uneconomic" because they have a high unit cost per unit of diurnal storage they provide, under-accounts for the other benefits that such projects provide.

2.2. Connections Capex

2.2.1. Benchmarking Regressions

PB/Rune used comparative regression analysis to estimate the "efficient" level of total capex for a given level of connections. Its regression model uses Ofgem's upper-quartile adjusted COLS regression method, which suffers from various pitfalls, as we have described elsewhere in this report. In particular, however, this technique cannot statistically distinguish "inefficiency" from unmeasured differences between firms. It also uses only eight data points (one per GDN), and so cannot be considered robust.

A further problem with this particular exercise is the method PB/Rune has used to construct the workload driver. PB/Rune used weighted average workload as the explanatory variable, and total cost as the dependent variable. PB/Rune constructed the workload variable by weighting outputs by their "representational" unit costs (section 4.2.1, page 19):

$$CSV = \sum Un * Vn$$

where
Un is the "representational" unit cost of activity n, and
Vn is the workload volume for activity n.

It appears from table 8A-7 of PB/Rune's June report that the representational unit costs for connections capex are calculated as the average unit costs for all the GDNs for the activities in this category of capex. However, this assumption needs to be reviewed, as different companies may face different unit costs due to differences in circumstances between them. If PB/Rune overlooks these differences, the results will be biased against companies that face higher unit costs for reasons other than inefficiency.

In areas of capex other than connections, PB/Rune has regressed costs on CSVs, but has not justified its choices of weights within each CSV, and in some cases has not even reported its choices. PB/Rune has not reported the data used in these regressions, so others cannot replicate or review the analysis. This approach is not compatible with objective and transparent regulation.

2.2.2. Theoretical Properties of CSVs

When estimating benchmarking regressions using CSVs, Ofgem and its consultants impose certain econometric (i.e. statistical) restrictions on the functions they estimate. The weights are almost entirely arbitrary in almost all the benchmarking regressions, but embody *a priori* restrictions on the relative impact of certain network characteristics on estimated costs. A more flexible approach would be to estimate the weights assigned to each explanatory variable by including each variable separately within a regression. However, with only eight observations, regressions including several explanatory variables are not feasible and would

expose the unreliability of the analysis. However, using CSVs merely replaces statistical unreliability with – equally unreliable – assumptions. We describe in more detail this theoretical problem with CSVs in Appendix A.

2.2.3. Regression Error

The connections capex benchmarking regression equations shown in section A8.2.3 of the June report do not correspond to the data shown in that section. We have not been able to verify any other regression equations, as this is the only regression equation reproduced in the June report and the September report provides neither the estimated regression equations nor the data with which to conduct the regressions.

2.2.4. Adjustments to WWU's BPQ Forecasts

PB/Rune recommended that the volume of connections capex in WWU's BPQ should be accepted (page 18, para. 2). Hence, the downward adjustments in PB/Rune's recommendations are due to cost (i.e. "efficiency") adjustments to WWU's forecast connections capex. WWU's forecasts contain an increase in costs per connection of 3.5-3.8% per annum.[15] PB/Rune's efficiency adjustment assumes that costs per connection decrease over time, according to certain assumptions regarding the scope for ongoing productivity which it set out in its first report (page 18, para. 1):[16]

> "*WWU has made significant organisational changes during 2005/06 to deliver efficiency improvements in the management and execution of all connections processes and activities. WWU has not quantified precisely the level of improvement expected but we are of the view that efficiency savings within the range 5% in 2007/08 reducing to 2% in 2012/13 are appropriate. This range has been smoothed in the analysis process to 3% year on year over the forecast period.*"

This statement from PB/Rune's June report on capex follows PB/Rune's acknowledgment that WWU is the third most efficient GDN in its total connections capex benchmarking regression (after allowing for regional factors).[17] PB/Rune has not presented any evidence to show that these ongoing productivity targets are feasible or consistent with the productivity improvements achieved historically.

The connections capex allowances recommended in the June report of PB/Rune follow a "glidepath" which assumes that GDNs will be able to close 70% of the "efficiency" gap by 2012/13.[18] If Ofgem has rejected this recommendation, or caused PB/Rune to change the basis of its recommendation, it is acting arbitrarily against the advice of its consultants as to what is feasible. Ofgem may claim (1) that any "excess costs" are attributable to inefficiency and (2) any "inefficient" costs should be disallowed. However, the analysis provided in the PB/Rune reports is insufficient to establish that any costs are due to inefficiency (as opposed to other factors), so Ofgem's first claim is no more than an assertion. The second claim is

---

15 PB/Rune, June 2006, table 4-2.

16 PB/Rune, June 2006, section 4.4.4.1.

17 PB/Rune, June 2006, page 41, paragraph 3.

18 PB/Rune, June 2006, page 41, section 4.4.3.1, paragraph 2.

also impossible to uphold, partly because Ofgem has not established which costs are inefficient, but also because the stock market offers an average rate of return to firms of average efficiency. By creating a regime where only firms of superlative efficiency can achieve the same rate of return, Ofgem is damaging the ability of GDNs to finance their activities.

2.3. Mains and Governor Capex

To assess the efficient unit cost level of mains replacement capex, PB/Rune appears to have conducted benchmarking regressions similar to those used for the connections capex, where PB/Rune constructed a workload driver as follows:

> "*A basket of work approach has been used to produce a weighted average of a number of different work elements (pipe sizes). The driver is calculated by multiplying the work volume by a nominal unit cost for the activity. The approach is not sensitive to the actual level of these nominal unit costs, but works on the relative costs between work types.*"[19]

This analysis is subject to the same criticisms as listed above, which undermine the reliability of the results:

- The results depend on assumptions about relative unit costs which have not been published;
- Regressions and comparisons with eight data points are uninformative;
- PB/Rune's June recommendations assume that companies will close the gap with the benchmark cost level to 30% by 2012/13, but this recommendation seems to be downgraded in the September report without good reason.

In addition, PB/Rune states that:

> "*we are of the opinion that there is scope for improvement driven by optimized management of operations and review of period contract arrangements. It is considered that 2% year on year performance improvement is appropriate for this activity.*" [20]

PB/Rune has not provided evidence to support this assertion.

2.4. Other Operational Capex

Under the category of plant and equipment, PB/Rune notes that WWU has proposed to spend £3.5 million to comply with the Dangerous Substances & Explosive Atmospheres Regulations (DSEARs). PB/Rune acknowledges that "*WWU is right to identify this requirement,*" but noted in its June report that it did not have enough information at that time "*to determine whether this work is necessary and justified.*" It also noted that this issue "*will*

---

[19] PB/Rune, June 2006, section 5.3.4.2, paragraph 1.

[20] PB/Rune, June 2006, section 5.4.4.1.

be followed up before the final determination of the allowances for the review period."[21] The updated report appears not to follow up on this issue, and Ofgem's UPs do not mention expenditure on complying with the DSEARs.

To deal with such uncertainties, Ofgem should have been developing "off-ramps" or "re-openers" – closely defined descriptions of a possible change in circumstances and the adjustment to revenues (or the method of calculating an adjustment to revenues) when those events occur. Otherwise, GDNs will be left facing a risk which they cannot manage, for no good reason.

2.5. Mains and Service Repex

The benchmark level of unit costs for mains and service repex comes from an upper quartile COLS regression comparing costs with weighted average workload. This method suffers from the methodological and statistical problems discussed above, rendering the results unreliable.

The main change between PB/Rune's June and September reports arises from a reduction in HSE programme repex expenditure. PB/Rune justifies this reduction because its benchmark unit costs for mains replacement work have fallen (see Table 2.1 below). The conclusions that "efficient" unit costs have fallen arises from PB/Rune's updated regression analysis, which uses 2006/07 rather than 2005/06 data. Despite this large reduction in recommended costs for WWU, both the 2005/06 and 2006/07 regressions show that WWU is among the most "efficient" companies (page 35, para. 4, and page 38, para. 1). Hence (PB/Rune concludes – page 38, para. 2), the estimated "efficient" frontier has shifted.

It is highly unlikely that productivity could have improved in just one year by between 10 and 20%, which is the magnitude of the change suggested by the data reproduced in Table 2.1 below. It is far more likely that the movement in estimated efficient costs is due to random data errors, a non-robust and mis-specified regression model, and a failure to account for all the factors that affect GDNs' costs. This problem is common to most of the updated regressions, but is most evident for mains and service repex.

Table 2.1
Comparison of "HSE Programmes" Repex: September vs. June Reports

	PB Rune September Report (£ million 2005/06 prices)	PB Rune June Report (£ million 2005/06 prices)	Difference
2005/06	29.9	33.1	£3.2m (10%)
2006/07	32.7	33.2	£0.5m (2%)
2008/09	30.2	35	£4.8m (14%)
2009/10	30.9	38.5	£7.6m (20%)
2010/11	32.4	40.6	£8.2m (20%)
2011/12	33.5	41.9	£8.4m (20%)
2012/13	33.7	43.3	£9.6m (22%)

Source: June report, section 8.5.4, September report, section A1.7.7

[21] PB/Rune, June 2006, section 6.4.2.2.

2.6. Conclusions

In this section, we have identified several problems with PB/Rune's capex benchmarking study. The main methodological problems with this study are common to all the benchmarking studies Ofgem has conducted during the GDPCR:

- PB/Rune's approach does not account for unobserved or immeasurable differences between benchmarked firms, which mean that Ofgem has mistakenly attributed all these differences to "inefficiency";
- The sample size used for PB/Rune's benchmarking regressions is small, meaning the results are not robust;
- The results of the benchmarking analyses are unstable between years, and PB/Rune mistakes this inherent instability in results for productivity improvements, rather than an indicator of unreliability; and
- Many of PB/Rune's assumptions (e.g. regarding the scope for ongoing productivity) are unjustified and arbitrary, and many are hidden.

Benchmarking studies that produce non-robust results and apply arbitrary assumptions in this manner are unreliable. Therefore, they do not give Ofgem a reliable basis for disallowing any of the GDNs' costs.

3. Review of the LECG Indirect Opex Report

Ofgem commissioned LECG to benchmark business support operating costs (i.e. indirect opex) for the GDPCR. The Phase 2 LECG report, submitted to GDNs in August 2007, is an update to LECG's Phase 1 report, using 2006/07 outturn costs (where available) rather than forecast costs that were used in the Phase 1 study (1.3).[22] LECG was also commissioned to check the consistency of the data used for its study between GDNs. LECG did not change its methodology between the Phase 1 and Phase 2 reports (1.5). Where LECG used cost forecasts in both Phase 1 and Phase 2 reports, GDNs were allowed to update their forecasts between the two phases (1.7). LECG states that WWU (and SGN) changed their data submissions significantly, and suggests that Ofgem consider whether the changes were "appropriate" (1.9).

Although Ofgem commissioned this report from external consultants, the report makes clear that Ofgem instructed LECG as to the methodology to be used (e.g. 1.5, 2.8, 2.19 etc), and that Ofgem instructed LECG not to update its methodology between the two reports. The LECG report does not therefore represent any kind of independent estimate that takes account of the criticisms that the GDNs made of LECG's Phase 1 analysis.

3.1. Using Revenues to Indicate Size

For many of the cost categories LECG benchmarked, it used cost (in that category) as a share of revenue to compare firms' "efficiencies." For this purpose, LECG estimated a notional revenue based on the "natural RAV", rather than the actual ("sculpted") RAV applied to the GDNs in the GDN sale, which adjusted each GDN's RAV to fit with the national tariff structure. LECG's method of controlling for firms' size can produce misleading results in a benchmarking study.

To illustrate this problem, suppose that in one year ("year 0") three firms have costs equal to their revenues of £50, £100, and £150 million, and that each spends 10% of their total costs on IT (£5m, £10m and £150m, respectively). Assume that none of the firms can reduce their IT costs, as they are already "efficient." However, suppose that a few years later the two firms with lower revenues have cut *other* cost items, so that their total costs have fallen, from £50m to £45m, and from £100m to £90m respectively. These firms will now be spending 11.1% (£5m/£45m) and 10.3% (£10m/£90m) of their respective revenues on IT, and will appear inefficient relative to the other firm which has not reduced costs at all, and which is still spending "only" 10% of its revenues on IT (£15m/£150m). Using revenue as an indicator of size causes the firm which has *not* cut costs in other areas to appear most efficient in IT costs.

A similar problem arises where LECG used total opex as a measure of size, as it has done for procurement costs. In this case, savings in other areas of opex can make a GDN look less efficient, even though its spend on procurement has not changed.

Hence, the fact that there are differences in firms' expenditure on individual opex items (as a proportion of revenue/total opex) does not provide proof that a firm is inefficient in that

[22] Throughout this chapter, all bracketed references refer to the LECG report.

particular area. Moreover, LECG's benchmarking approach may be punishing GDNs for past efficiency savings in other areas. This problem demonstrates a lack of economic logic in LECG's approach.

3.2. Partial Benchmarking

The LECG report acknowledges the possibility that benchmarking individual cost categories separately may result in efficiency targets which are infeasible:

> "*The benchmarking of total support services costs has some merit of taking into account potential trade offs between support service cost categories. This explains the lower efficiency savings for three of the GDNs under the total support services costs approach.*"(1.28).

We have previously explained to WWU why "partial benchmarking" can lead to infeasible "efficiency" targets, and WWU has communicated this problem to Ofgem.

Because of the partial benchmarking problem, LECG is wrong to suggest that Ofgem place less weight on benchmarking of total support services than on the benchmarking of individual categories. However, the LECG report does not explicitly prescribe a method for setting revenues based on the findings of its own benchmarking studies.

The table below shows that the results of the top-down analysis produce are much more narrowly dispersed than the results of the bottom-up analysis. Moreover, the top-down analysis shows NGG achieving economies of scale and coming 2nd behind only SGN (see comments below), rather than 4th in the bottom-up analysis. These observations make the top-down analysis (for all its faults) more credible than the bottom-up analysis.

Table 3.1
Comparison of LECG Top-Down and Bottom-Up Approaches
(High Savings Scenario)

Ownership Group	No. GDNs	"Efficient Cost" 2006/07 (£mn pa)				Actual Cost 2006/07 (£mn pa)
		Bottom-Up Analysis		Top-Down Analysis		
		per GDN	Rank	per GDN	Rank	per GDN
NGG	4	21.01	4	17.36	2	26.34
NGN	1	15.91	2	17.54	3	18.52
SGN	2	14.74	1	16.16	1	16.16
WWU	1	15.94	3	18.69	4	21.69
Average		*16.90*		*17.44*		*20.68*
Range		*6.27*		*2.53*		*10.18*

Source: LECG Report, Table 19 and Table 3.

Even if Ofgem does consider that the bottom-up analysis yields more reliable results, this approach does not give consistent treatment to both direct and indirect opex. For direct opex, Ofgem has adjusted for the infeasible targets resulting from bottom-up benchmarking by uplifting the "upper quartile" results of the bottom-up analyses to the average level of the results from the top-down regression. However, for indirect opex, Ofgem has made no such adjustment. If Ofgem, were to apply the same adjustment as for direct opex, from Table 3.1,

it appears that it should "uplift" indirect opex allowances by 3.2%. Without any adjustment to compensate for the use of partial benchmarking, Ofgem may be setting indirect opex targets which – when combined – are not feasible for any one GDN.

3.3. Coverage of Benchmarking

LECG notes (1.20) that although WWU's costs are lower in the Phase 2 report than in the Phase 1 report, its target for efficiency savings has increased. LECG explains that this outcome arises because (1) the costs of the benchmark GDN are also lower, and (2) it was only possible to benchmark a subset of total costs in some areas, such as Human Resources (HR). This incomplete coverage means that LECG's targets cannot be relied upon to determine overall "efficient" costs for the purpose of setting allowed revenues.

3.4. GDN-Specific Factors

LECG's indirect opex benchmarking report does not account for unobserved or immeasurable differences between the GDNs (and the external comparator firms it uses). LECG's report does – to some extent – acknowledge that its benchmarking approach can mistake unobserved differences between firms for inefficiency. For instance, it states that "*variations in GDN performance may be due to factors such as difference in organisational structure. For example, for FTE based metrics, a higher cost per FTE may be due to more outsourcing rather than inefficiency.*" (2.18) However, its approach throughout the report was to ignore the possibility that its approach cannot possibly account for firm-level differences. This appears to be because Ofgem has instructed LECG to ignore factors affecting individual GDNs (2.27). These may include regional differences in input prices, economies of scale (see above), trade-offs between input categories, and many more besides. Ofgem's instructions mean that LECG's results cannot be (and perhaps are not intended to be) taken at face value, but that undermines the value of LECG's analysis.

Ofgem's instructions to LECG (see section 1.3) place a burden on the companies with costs above the benchmark to prove that their higher costs are due to genuine differences between them and the benchmark company, rather than due to inefficiency. In most cases, GDNs will not be able to provide this evidence to Ofgem because they have limited knowledge of the other benchmarked firms. A better approach would have identified in more detail the actual drivers of cost, or the precise sources of the alleged inefficiency that the GDNs need to address. At the very least, the analysis should have allowed for differences between GDNs.

Not only did LECG fail to account for any unobserved or immeasurable differences between the GDNs, it also failed to account for observed differences in GDNs' performance against standards of service. Better performance against service quality standards requires GDNs to incur relatively more indirect opex costs than if they under-perform against targets. LECG's analysis does not account for this factor, and so expenditure devoted to improving service quality may be mistaken for "inefficiency".

3.5. Stability of Results

With reference to the PB/Rune capex and direct opex reports, Ofgem cannot assume that large changes in benchmarking results caused by updating analyses with more recent data is a signal that the "efficient frontier" has shifted. As we discuss in section 3.4, LECG's

benchmarking analysis cannot disentangle unobserved or immeasurable differences between GDNs from "inefficiency." Therefore, large changes in the results from year-to-year are as likely to be due to random data errors and misspecification of the benchmarking models.

LECG highlights an example of this problem, noting in paragraph 1.20 that the efficiency savings demanded of WWU have increased, despite a decrease in WWU's total support costs. This counterintuitive movement in LECG's recommended allowed costs indicates that LECG's results may not be robust. None of LECG's explanations for why this change in the efficiency target occurred deal with the possibility of misspecification.

3.6. The Effects of Weather

LECG recognises that GDNs' revenues change with the weather through the price control formula, whilst the "efficient" spend on support services does not (2.19). However, it appears from LECG's report that Ofgem continued to require the use of revenue as a metric for network size, despite its dependence on weather. Ofgem's apparent rationale is that weather patterns affect all GDNs equally, so this should not affect results.

This explanation is not applicable when LECG benchmarks GDNs against "external" non-GDN companies, as well as other GDNs. LECG recognises this point (2.20) and suggests that external comparators should be excluded from the analysis using revenue as a measure of size. Ofgem has omitted external benchmarks in certain cases (HR, finance, audit and regulation), perhaps for this reason (Updated proposals paragraph 3.33).

In any case, weather patterns do differ between different parts of the country. LECG's omission of non-GDNs from the sample does not correct for any regional differences in weather patterns. This is one example of why LECG's approach (which is largely follows Ofgem's instructions in this case) does not account for differences between benchmarked firms that are not captured by its simple benchmarks. This "unobserved heterogeneity" is indistinguishable from the estimated efficiency scores.

3.7. SSE Provision Support Service to SGN

The LECG report suggests that the SGN cost data used for its benchmarking does not include some of the support costs that SSE provides to SGN. This bias arises because of the arrangement between SGN and SSE. SGN does not pay for support services provided by SSE when SSE has spare capacity, for instance (3.8). As a result, the prices charged to SGN for some support services are not full "arms length" prices, and appear to reflect SSE's marginal cost of provision, rather than the total cost of providing support services that other GDNs face.

LECG has not fully accounted for SGN's arrangement with SSE, which undermines the results of its benchmarking analyses whenever the benchmark level of costs is affected by SGN's costs.

Ofgem has accepted the need to overcome the bias in SGN's ability to use services provided by SSE at marginal cost (UPs, 3.35). When SGN is the least-cost company, Ofgem claims to set the benchmark equal to the cost of the second group (out of the four), rather than at the upper quartile. In fact, Ofgem has not applied this rule, as Table 3.2 shows (see column

"Ofgem Rule"). Instead, Ofgem actually picked the second cheapest company for every item, including not just items where SGN was the cheapest (most items) but also items where SGN was the second cheapest (e.g. Regulation, Property Management 2), or even third cheapest (e.g. Property Management 1).

Table 3.2
Ofgem Benchmarks for Indirect Opex

	Efficiency Scores				Benchmark		Possible Benchmarks			
	NGG	NGN	SGN	WWU	Ofgem Updated	Ofgem Rule	2nd GDN	Upper Quartile	1st GDN (Excl. SGN)	2nd GDN (Excl. SGN)
Information Systems	1.85	2.19	1.50	2.38	1.85	1.85	1.85	1.76	1.85	2.19
Finance & Audit	1.17	0.85	0.42	1.19	0.85	0.85	0.85	0.74	0.85	1.17
Regulation	0.20	0.33	0.27	0.27	0.27	0.25	0.27	0.25	0.20	0.27
Insurance	1.54	1.12	0.95	1.04	1.04	1.04	1.04	1.02	1.04	1.12
Property Management 1	3.80	2.10	3.00	2.70	2.70	2.55	2.70	2.55	2.10	2.70
Property Management 2	29.90	11.00	16.90	36.90	16.90	15.43	16.90	15.43	11.00	29.90
Coporate Center and Communications	1.50	0.96	0.34	1.12	0.96	0.96	0.96	0.81	0.96	1.12
Human Resources	1.32	0.75	0.46	0.48	0.48	0.48	0.48	0.48	0.48	0.75
Legal	0.18	0.27	0.17	0.22	0.18	0.18	0.18	0.18	0.18	0.22
Procurement and Logistics	0.87	0.57	0.42	1.09	0.57	0.57	0.57	0.53	0.57	0.87

Ofgem benchmarks
Matching figures

In any case, if SGN's data is really non-comparable, it should not be included under any rule, including cases where Ofgem does actually calculate an upper quartile and SGN is the second group (out of the four). In such cases, SGN would still be biasing downwards the upper quartile, as we explained in our response to the Updated Proposals.

3.8. Economies of Scale

During the GDPCR, Ofgem has stated that it will not allow for economies of scale when setting the IDNs indirect opex allowances (e.g. IPs 3.21). Ofgem has also instructed LECG to ignore any possible economies of scale in indirect opex from the ownership of more (or larger) networks (2.27), thereby ruling out any proper discussion of the costs that a GDN would actually be able to achieve. As we outlined in section 3.4 above, Ofgem instructed LECG to ignore all differences between firms, including the effect of economies of scale. This approach penalises smaller GDNs like WWU for not achieving the same economies of scale as NGG, for example, which owns four networks. Given the limited opportunities for consolidation of the industry, Ofgem should not expect smaller GDNs to be able to achieve the same indirect opex unit costs as larger GDNs.

The results we presented in Table 3.1 (above) show that – as well as not accounting for economies of scale – LECG has recommended a higher indirect opex allowance "per network owned" for NGG than for some individual GDNs. It is difficult to explain this result either in terms of (ignoring) economies of scale or in terms of relative efficiencies. This outcome undermines confidence in the robustness or reliability of the methodology that LECG has chosen (or been asked by Ofgem) to adopt.

3.9. Conclusions

Like Ofgem's other GDPCR benchmarking studies, LECG's indirect opex benchmarking work is not robust and lacks economic logic. That is, LECG has conducted partial

benchmarking, for which Ofgem has made no correction. Also, LECG has made no correction for observable differences between firms (following Ofgem's instructions).

We have also identified some problems specific to LECG's analysis. The use of expenditure as a proportion of cost to benchmark GDNs, for instance, penalises firms for past reductions in costs in non-benchmarked cost areas. Also, the way in which Ofgem has corrected for the bias caused by the support services that SSE provides to SGN is inconsistent. Ofgem's approach allows SGN's artificially low costs to affect the estimated frontier.

4. Review of the PB/Rune Direct Opex Report

4.1. Introduction

On 18 June 2007 Parsons Brinckerhoff Ltd and Rune Associates (from here on referred to as PB/Rune or "the authors") published a benchmarking study of gas distribution companies for Ofgem (PB/Rune, 2007a).[23] On 18 September 2007 PB/Rune published an update of their study at the request of Ofgem (PB/Rune, 2007b).[24] The data that PB/Rune used in their original benchmarking analyses was actual data for 2005/06 or forecast data for 2006/07. The Gas Distribution Networks (GDNs) submitted their actual expenditure for 2006/07 in June 2007 and Ofgem asked PB/Rune to update their report "taking into account the latest information provided".[25] PB/Rune explain that their updated report is "to be read in conjunction with the initial reports, and provides a commentary on the changes to our conclusions".[26]

Wales and West Utilities (WWU) asked us to review the PB/Rune's updated report (PB/Rune, 2007b). This follows a review of an earlier version of the PB/Rune report that we undertook in March.[27] Where the methodology used in the latest PB/Rune report (2007b) is the same as in earlier versions, our methodological criticisms of the earlier PB/Rune report (2007a) remain valid. However, in some cases, PB/Rune decided to change the methodology in the light of the new data. The executive summary of the latest report contains the following comment on the effect of updating the earlier analysis:[28]

> "*As can be expected whilst there is a reasonable level of consistency between our original and updated regression lines, in some areas there are material changes to the nature of the regression conclusions.* ***It is our view that such changes are inevitable given the sample size and the changing nature of all the GDN businesses.***"[29]

The PB/Rune report provides little information on the statistical analysis contained within it, which makes it difficult to evaluate it in full. However, it is clear that PB/Rune uses many subjective judgements and non-standard methods to construct benchmarks based on eight

---

[23] PB Power/Rune Associates (2007a), "Gas Distribution Price Control Review – Five Year Control – (Opex) – Report 1 – Wales & West Network", 18 June 2007. One report was produced for each GDN. For the sake of simplicity we concentrate here on the report for Wales and West Utilities. The report is available at http://www.ofgem.gov.uk/Pages/MoreInformation.aspx?docid=175&refer=Networks/GasDistr/GDPCR7-13.

[24] PB Power/Rune Associates (2007b), "Gas Distribution Price Control Review – Five Year Control – (Opex) – Updated Report – (Confidential)", Reference number 62533 2006 36621 (Update v1.1), 18 September 2007.

[25] PB Power/Rune Associates (2007b), p. 5.

[26] PB Power/Rune Associates (2007b), p. 5.

[27] NERA (2007), "Gas Distribution Price Control Review: Reports on Costs – A draft report for West and Wales Utilities", 9 March 2007. The NERA (2007) report reviewed: PB Power/Rune Associates (2007c), "Gas Distribution Price Control Review – Five Year Control – (Opex) – Draft Report 1 – Wales & West Network – Confidential", 27 February 2007.

[28] The original data relied on 2005/06 actual expenditure or 2006/07 forecasts, whereby the updated report relies on 2006/07 actual expenditure.

[29] PB Power/Rune Associates (2007b), p. 5. Emphasis added.

data points. Such work will inevitably lack robustness, as PB/Rune recognise, and cannot therefore be relied upon as a true indication of relative efficiency.

4.2. Methodology

PB/Rune use two techniques to evaluate the operating expenditure of GDNs:[30]

- Regression analysis "where the workload is sufficiently well defined to obtain reliable regression results"; and
- "Bespoke review" by their consultants.

The authors do not provide a rigorous criterion of what constitutes "sufficiently well defined", nor what "reliable" regression results are. They appear to base this decision entirely on the R^2 (the goodness of fit) of the regressions.[31] However, even regressions that meet this standard do not necessarily produce statistically significant, reliable or robust results, as the authors admit themselves (see Chapter 4.1).

4.2.1. Robustness

Based on normal standards for economic analysis, we believe that "Robustness" should be understood in the following ways:

1. The efficiency scores of a company should not fluctuate widely over short periods of time.
2. Adding or removing a few observations should not alter the results significantly;
3. The results should not vary significantly due to subjective choices over the specification of a model (choosing one variable instead of others, setting the weights of composite variables, etc) or interpretation of the results (how much of the unexplained costs is attributed to inefficiency).

By PB/Rune's own admittance using an updated dataset with just one year time difference has led to "material changes" in some of the results. This observation by itself significantly reduces the reliability of the results, because it fails by standard number 1 above.

In the section on repairs, the authors "considered treating the Scottish network as an "outlier" and removing it from the analysis".[32] Doing so increased the R^2 from 0.7655 to 0.9383. Despite identifying R^2 as the key indicator of robustness, the authors concluded that "for the purpose of this report" they would use the results of the 2006/07 analysis including all eight networks,[33] without providing any further reasoning for this decision. There are no other references to outliers in the rest of the report.[34]

---

[30] PB/Rune (2007b), p. 6.

[31] PB/Rune (2007a), p. 13.

[32] PB/Rune (2007b), p. 21.

[33] PB/Rune (2007b), p. 21.

[34] PB/Rune (2007b). In the original report, the authors did exclude an outlier in one instance. See PB/Rune (2007a), p. 86.

PB/Rune relied extensively on composite variables in their regression analysis. The reason is that with such a small sample size, a regression model with multiple variables would have lacked any statistical significance. A composite variable is a means of imposing coefficients on a regression which the statistics themselves cannot provide.[35] The use of a composite variable implies that it is not possible to test the statistical significance of the individual cost drivers that make up the composite variable. To ensure robustness, a small change in the weights, or replacing one cost driver in the composite variable by a similar cost driver, should not alter the results of the regression significantly. However, as far as we know, PB/Rune did not perform any such checks on robustness, and has not provided any reasoned explanation of the weights that were finally chosen.

4.2.2. Defining the Frontier

In our earlier review of the PB/Rune report, we criticised three aspects of the authors' definition of the frontier that companies have to reach.

§ The upper quartile frontier

The authors claim that it is "usual"[36] to set a standard so that that the "efficient" frontier equals the costs of the most efficient company. In fact, regulatory practice varies widely. There is no real intellectual basis for the claim that regulated companies should be allowed only the costs of a "frontier company" and no theory of competitive markets supports it.

A competitive market will normally drive out of existence companies whose costs are very high, but their high costs may be due to either management inefficiency or other factors (like use of outdated technology or an expensive location). Those that remain will be a mixture of cheap and expensive plants, which are more or less efficient. In a competitive market, the price will be set by the cost of the most expensive unit produced to meet demand. There is no way to say whether this unit will come from an efficient firm or not.

Moreover, it is not possible to identify the efficient frontier applicable to any particular company with certainty. Errors in the data of the other companies and the omission of explanatory factors will lead to understatement of the costs that a particular company can be expected to achieve.

To counter such criticisms regulators sometimes adopt standards that are less "demanding" than the frontier. PB/Rune adopts the upper quartile, but choosing this standard remains an arbitrary decision, which is not based on sound economic reasoning, and it remains sensitive to the omission of company-specific factors. Given the limited size of the data set, even the upper quartile does not represent a robust estimate of the frontier.

§ Shifting and swivelling the regression line

---

[35] If C are costs and W is a composite variable equal to rX+sY+tZ, then the regression equation C=a+bW is formally equivalent to C=a+b(rX+sY+tZ)=a+brX+bsY+btZ. Fixing r, s and t (i.e. the relative weights in the composite variable) therefore pre-specifies the relative impact of each variable.

[36] PB/Rune (2007a), p. 12.

Having carried out a regression, PB/Rune changes the *slope* of the regression line (the value *b* in $a+b.X$) until it passes through the "upper quartile" point, rather than changing the *intercept* (the value *a* in $a+bX$). The adjustment of the intercept lies behind a number of recognised methods of comparison (including COLS and SFA), but adjusting the slope is a non-standard method.[37]

§ Partial Benchmarking

PB/Rune carried out separate benchmarking analyses for different direct opex cost items, a practice sometimes described as "partial benchmarking". Such analyses ignore the possibility of "substitution" between these cost items, as well as between opex and capex. Ofgem adjusted the results from partial benchmarking by adding an uplift derived by comparing the partial benchmarking of individual cost items with the – also partial – benchmarking of total opex. PB/Rune does not appear to have taken into account any "substitutability" between cost categories.

The authors undertake a normalisation process of the GDNs' expenditure submissions to ensure that all cost items are on a like-for-like basis. Compared with the earlier report, some cost items have been reallocated, which affected the performance of individual companies.[38] The sensitivity of results to these changes shows the risks of benchmarking a small number of companies at a high level of disaggregation.

PB/Rune should have allowed for substitutability between factors of production by introducing additional control variables to account for these effects in the data. However, this is impossible with such a small sample size. In such conditions a benchmarking analysis at disaggregated level cannot possibly give robust or reliable results.

§ Convergence

PB/Rune do not provide any backing for their assumptions on the rate of convergence that companies need to achieve to catch up with the "frontier". The original report contains a statement in the chapter on emergency opex saying that:

> "*The GDNs are not expected to close any gap immediately. The convergence adjustment provides a glide path of cost to the Benchmark performance. The gap is reduced to 30% in 2012/13.*"[39]

Ofgem has since indicated in the Updated Proposals that a glide path is unnecessary or undesirable. However, if the consultants believe that a glide path is reasonable because of some constraint on the rate of cost reduction, they should have made that constraint clear to Ofgem.

---

[37] Offer used something like this method in 1995, after having decided that each electricity distribution company faced the same "fixed cost", but was criticised heavily by the industry for adopting such an approach and has not used it since.

[38] See the statements in PB/Rune (2007b) on pages 11, 26 and 37.

[39] PB/Rune (2007a), p. 47.

4.2.3. Bespoke Reviews

PB/Rune's reports contain many "bespoke reviews", involving purely subjective assumptions and statements which are difficult or impossible to evaluate. The instances are too many to enumerate, but the section on LTS non-routine maintenance opex (Chapter 6.3.7) contains a typical example on AGI Painting:

> *"We have made adjustments to the GDNs' forecast costs as not all of the GDNs' costs were at the steady state level during the base year 2006/07."*

It is not clear how PB/Rune determined that a "steady state" level should be, as there is no other explanation in their report, so we cannot possibly understand or evaluate this process.

4.2.4. Price Effects

PB/Rune assume that some prices will grow at rates that differ from general RPI inflation (hence "real" price effects). PB/Rune also assume that some prices differ across regions. The affected prices are:[40]

- **Contractor prices:** These are assumed to grow at 2.25% per annum above inflation and are subject to a regional factor.
- **Materials prices:** These are assumed to grow at 1% per annum above inflation and are not subject to a regional factor.
- **Direct labour prices:** There are assumed to grow at 1% per annum above inflation and are subject to a regional factor.

The authors did not change their estimates of these effects from the earlier report. The authors refer occasionally to past trends to justify their assumptions, but the numbers used in the analysis are little more than guesses. If they were unsure of these figures, PB/Rune should have experimented with a range of values to ensure robustness of results.

§ Real price effects

PB/Rune do not update their estimates of real price effects from those discussed on pages 17 to 18 of PB/Rune (2007a). In the earlier report, the authors based their forecast of contractor prices on past trends in two price indices published by the DTI: (1) the Price Adjustment Formulae for Construction Contracts Indices (the "Baxter" indices) and (2) the Public Sector Construction Works Indices (Road Construction). On page 18 the authors refer to past wage and salary trends to support their assumption of direct labour costs. However, the earlier report does not explain what assumptions lie behind these calculations and the latest report provides no further indication.

§ Regional factors

On direct labour costs, PB/Rune judge that London wages are on average 30% higher than the national average. The authors then argue that "only Southern and London GDNs are

---

[40] PB/Rune (2007b), Table 2-1 (p. 7) and Table 2-2 (p. 8).

affected and that they are not fully exposed to the 30% uplift as the whole of the GDN is not within London and many activities are carried out away from the London location."[41] However, the choice of the appropriate uplift factor (1.10 for London and 1.03 for Southern) seems to be nothing more than an arbitrary judgement.

§ Conclusion

Real price effects and regional price differences in contractor prices, material prices and direct labour prices are to a large part based on subjective judgements by PB/Rune. The authors did not undertake a sensitivity analysis to changes in their assumptions.

4.3. Cost Categories

In the following sections we review the benchmarking analyses of the various cost categories in more detail without repeating the methodological points from Chapter 4.2, rather we evaluate the benchmarking analyses on the basis of the results and implementation.

4.3.1. Work Management

PB/Rune repeats its benchmarking of work management with a regression using a composite variable of network workload or size constructed as follows:[42]

> Composite Cost Driver (CSV) = Average length of mains Í
>
> (0.3 Í No. of PREs / Average no. of PREs
>
> + 0.3 Í No. of repairs / Average no. of repairs
>
> + 0.4 Í length of < 7bar main / Average length of mains).

The authors provide no support for the choice of these weights in the latest report and the earlier report only contained a statement that they were the product of a subjective judgement:

> "*We have made a judgement on the proportion of work management costs associated with each of the activities*".[43]

Thus, the relative weight give to cost drivers is arbitrary and not backed up with arguments or empirical analysis. PB/Rune carried out no robustness checks to test the sensitivity of results to these variables and parameters.

PB/Rune state that they changed the preferred model after analysis the updated data, because "*the best fit is now obtained if regression is undertaken on a logarithmic basis*".[44] It is not clear whether PB/Rune had sufficient grounds to reach this conclusion on the basis of the R-

---

41 PB/Rune (2007a), p. 19.

42 PB/Rune (2007b), p. 9. See also PB/Rune (2007a), p. 28.

43 PB/Rune (2007a), p. 28.

44 PB/Rune (2007b), p. 10.

squared alone.[45] Moreover, this change in functional form would imply that the cost function changed from a linear form to a Cobb-Douglas form within a year, which is unlikely. Such major changes show that the functional form and results are not robust from one year to the next and do not provide a reliable basis for forecasting actual costs, let alone a notional level of "efficient" costs .

4.3.2. Emergency

The benchmarking of emergency expenditure is subject to the same criticism as the work management benchmarking: the selection of the variables that make up the composite variable and the weights are arbitrarily chosen and were not subject to robustness checks. The composite variable (CSV) is made up as follows:[46]

CSV = 0.8 **Í** total no. of PREs/Average no. of GDN PREs

+ 0.2 **Í** no. of repairs / Average no. of GDN repairs.

In PB/Rune (2007b) the authors explain their choice of weights by saying that the cost drivers are "*weighted by a factor reflecting the proportion of costs **driven** by each of these factors.*"[47] However, PB/Rune does not explain how it established what proportion of costs is "driven" by what factor. In the earlier report, which used the same variables and weights, PB/Rune wrote:

> "*The weights reflect the **assumed** proportion of FCO time allocated to responding to PREs and to D2 rechecks together with site monitoring activities. These latter activities are driven by the number of repairs. (see Appendix 3)*"[48]

This confirms that the weights correspond to "assumed proportions" rather than actual ones, so that the weights are subjective judgements by the authors.[49]

4.3.3. Repair

In the benchmark for costs of repairs, PB/Rune chose weighted average workload as the explanatory variable in their regression analysis, in which the volume of each type of repair is weighted by its unit cost.[50] The authors claim in the latest report that the unit costs "*were*

---

[45] If PB/Rune chose their preferred model on the basis of R^2, they would need to make an adjustment to the simple R^2. The reason is that in the linear model and in the log model the dependent variable is different (y and ln(y) respectively) and therefore the R^2s are not directly comparable. We trust that PB/Rune have compared the right values, although they do not explicitly say so. On page 13 of PB/Rune (2007a), the authors write: "We have used appropriate tests to determine whether the linear or the logarithmic linear regression gives the better fit to the data and have used the regression with the better fit."

[46] PB/Rune (2007b), p. 14.

[47] PB/Rune (2007b), p. 13. Emphasis added.

[48] PB/Rune (2007a), p. 40. Emphasis added.

[49] Appendix 3 is empty and is marked as "redacted" (sic). See PB/Rune (2007a), p. 109.

[50] PB/Rune (2007a), p. 54.

derived from tendered rates for each activity analysed."[51] However, they provide no source for this information.

In the earlier report, PB/Rune claimed that "[a] *number of regression options have been explored in analysing repair costs, including a number of different explanatory variables.*"[52] However, there is no explanation in either report why the selected composite variable is the best one.

The workload projections of PB/Rune rely on arbitrary assumptions and are not backed up by arguments. In the earlier report, PB/Rune argued as follows in relation to interference repairs to mains and services:[53]

> "*The Network has forecast a reduction in interference repairs to mains of about 2% per year* ***but no change to the level of repairs to services****…Interference repairs are driven by the amount of construction activity within the Network, but this can be influenced through improved and focussed communication with those undertaking the work. We think a 2% per annum improvement to mains repairs is realistic and achievable and we accept the Network's forecast.* ***We have adjusted forecast repairs to services to show a 1% per annum improvement.***"[54]

The authors do not give any reasons why repairs to services should show an improvement of 1% per annum, or why WWU's estimate of no change is incorrect. PB/Rune do not seem to have identified how WWU's systems should improved, or whether such improvements would be feasible. PB/Rune's grounds for imposing a reduction in the volume of repairs still rely therefore on assertions without supporting evidence.

4.3.4. Maintenance

The earlier report considered maintenance cost items separately, while PB/Rune now take the view that total maintenance costs should be benchmarked together, on the basis of a logarithmic regression, because of the difficulty of standardising cost allocation.[55] However, for the later report, PB/Rune separates maintenance costs into routine and non-routine costs, with only the routine costs being subject to the regression analysis. PB/Rune than adds back "appropriate" non-routine allowances to derive recommendations after benchmarking of the routine costs. It appears that the identification of "non-routine" maintenance costs was carried out by PB/Rune, but the main report does not explain the basis for making this distinction.

---

[51] PB/Rune (2007b), p. 19.

[52] PB/Rune (2007a), p. 54.

[53] This passage comes from the original report, however there are no fundamental changes between the original and the updated report.

[54] PB/Rune (2007a), p. 59. Emphasis added.

[55] PB/Rune (2007b), pp. 23-24.

The routine maintenance costs are regressed against a composite variable. In order to construct such a composite variable the authors identified a number of drivers "as relevant to the individual categories", i.e.:

- Offtakes and PRS numbers;
- Network length (above and below 7 bar);
- Non PE Length;
- District governor numbers;
- Low pressure holder numbers;
- Network throughput;
- Repair numbers.

PB/Rune claim that they undertook "sensitivity tests" with different driver weightings and decided to adopt equal weightings as a result. However, neither sensitivity tests nor the results are shown or explained further.

4.4. Conclusion

The September 2007 PB/Rune report updates an earlier version of the report dated June 2007. The reports differ mainly in the data used, but there are one or two methodological differences between them as well. Updating methodologies might suggest a willingness to learn from GDN comments. However, the specific cases indicate instead an instability in the data and approach.

Most of the changes in methodology derive from attempts to find an equation that better fits new data. Where the update reflects the use of more recent data (2006/07 rather than 2005/06), this need to change methodology indicates that relationships between the data are not stable. Unstable relations provide no basis for forecasting or predicting future costs.

The results of the report are highly dependent on subjective judgements by the authors and largely rely on arbitrary decisions. We have identified the following weak points in PB/Rune's analysis:

- **Lack of robustness:** the results are, by PB/Rune's own admittance, not robust but PB/Rune have not performed or reported adequate checks on robustness.
- **Sample size:** the results of PB/Rune rely on a sample size of eight, which is not sufficient to provide adequate responses – PB/Rune only found it possible to produce statistically significant regressions by limiting the number of variables, but one cannot then conclude that unexplained variations around the regression lines are due only to inefficiency and not to other explanatory variables.
- **Frontier swivel:** in order to derive the "frontier", PB/Rune swivelled the regression line until it passes through the "upper quartile" point, which is a non-standard approach.
- **Partial benchmarking:** By limiting the analysis to certain opex categories, Ofgem forced PB/Rune to ignore the substitutability between cost categories (factors of production) and to produce potentially biased results.

- **Arbitrariness:** most of the results rely on arbitrary decisions, including:
  - The selection of variables to constitute the composite variables and the weights attached to them;
  - The use of the "upper quartile" as the target for future costs; and
  - The speed of convergence to the target.

The results of the PB/Rune reports can therefore not be relied upon.

Failure to provide proper grounds for statements not only puts in doubt the results of analysis. In the Updated Proposal, Ofgem rejects the notion of a glide path on the grounds that it only rewards inefficient companies. However, it is clear that PB/Rune believe some kind of glidepath would be appropriate. If PB/Rune had investigated properly the potential for cost reduction to provide a justification for its rate of convergence, Ofgem could not have reached the same conclusion.

5. Conclusion

The updated consultants' reports issued in conjunction with the Updated Proposals suffer from the same problems as the earlier versions and do not provide a reliable or robust basis for setting revenue allowances.

Ofgem claims at several points in the Updated Proposals that it has found a way to set revenue allowances without reference to the actual costs of each GDN, but instead by reference to "efficient costs". Neither part of this claim bear close scrutiny.

Leaving aside the lack of any theoretical justification for setting revenue allowances equal to "efficient costs", the analysis provided by PB/Rune and LECG is simply not robust enough to identify "efficient costs". Instead, the results that emerge should be viewed with scepticism, being based on poor (non-comparable) data for a small number of observations, subjective choices over assumptions and methodology, and an arbitrary interpretation of all unexplained costs as "inefficiency". On this basis, Ofgem cannot claim to have identified the "efficient costs" of any GDN and its decision to base revenues on these results amount to assigning arbitrary or random levels of revenue to each company. Such an approach is not consistent with transparency or objectivity in regulation and will not provide long-term incentives for efficient behaviour.

Ofgem's consideration of the consultants' reports has highlighted a number of areas where GDNs can respond with better explanations or better data (e.g. over the justification for LTS investment projects disallowed due to the devaluation of diurnal storage). However, a transparent and objective regulatory regime will only emerge if Ofgem takes proper account of evidence and ignores the unreliable products of its consultants.

Appendix A. Composite Scale Variables: Theoretical Difficulties

Suppose a gas distribution network's costs were driven by two variables (i.e. network characteristic), x_1 and x_2, and that the costs of this network are determined by "log-log" cost function, as follows:

$$\ln C = \ln A + \alpha \ln x_1 + \beta \ln x_2$$

where α and β are parameters which determine the change in the firm's costs attributable to an increase in the logs of the variables x_1 and x_2, and A is constant. To estimate this cost function, one could estimate the following regression model:

$$\ln C = \gamma_0 + \gamma_1 \ln x_1 + \gamma_2 \ln x_2 + \varepsilon$$

where the γ terms are regression parameters, and ε is due to random measurement error. Instead, Ofgem has adopted a number of models using a Composite Scale Variable, which combines several network characteristics into one variable, as the only regressor. In the case given here, one would be estimating the following model:

$$\ln C = \delta_0 + \delta_1(w_1 \ln x_1 + w_2 \ln x_2) + \nu$$

where v represents the random measurement error, the δ terms are regression parameters, and the w terms are weights given to network characteristics within the CSV. This version of the model places restrictions on the coefficients, which require that:

$$\alpha = \delta_1 w_1 \text{ and } \beta = \delta_1 w_2$$

which implies that:

$$\alpha = (w_1 / w_2) \cdot \beta$$

Hence, applying the CSV places an arbitrary restriction on the relative size of the true parameters (α and β). Imposing such a restriction usually lacks any economic logic or intuition. It is simply a method of including several explanatory variables in regressions, where only a small sample is available. However, although the resulting parameters may be more significant in statistical terms, this apparent improvement derives solely from the strong assumption about the relative size of different parameters and does not extract any more information from the data.

NERA
Economic Consulting

NERA Economic Consulting
15 Stratford Place
London W1C 1BE
United Kingdom
Tel: +44 20 7659 8500
Fax: +44 20 7659 8501
www.nera.com

NERA UK Limited, registered in England and Wales, No 3974527
Registered Office: 15 Stratford Place, London W1C 1BE