

RIIO 2 Tools for Cost Assessment Consultation  
SGN Response  
23<sup>rd</sup> August 2019

## Table of Contents

1	Summary.....	4
2	Approach to econometric analysis .....	6
	Q1: What model estimation options should be considered for our cost assessment and why? .....	6
	Q2: Do you agree with our proposed criteria for developing potential cost pools? If not, what additional criteria do you propose and why? .....	6
	Q3: Should we continue to use the Cobb-Douglas functional form? If not, why? .....	7
	Q4: Do you agree with the proposed model selection criteria and model development phases? .....	7
3	Aggregated econometric analysis.....	8
	Q5: Should the cost driver of the totex regression model be determined by the cost drivers of the 'bottom-up' models, or should the totex regression model account for different explanatory variables? Why? .....	8
	Q6: What could be appropriate cost drivers in middle-up models for opex, capex and repex? Why? .....	9
	Q7: For which opex activities are there trade-offs that support the rationale for testing 'totex and opex plus' modelling? .....	9
	Q8: Are there other particular costs that we should aggregate and test in our analysis? .....	11
4	Disaggregated econometric analysis .....	11
	Q9: Are there trade-offs between opex and capex activities that support the rationale for considering 'opex plus' modelling? .....	11
	Q10: Which cost areas should be assessed using workload drivers as opposed to other cost drivers? Why? .....	11
	Q11: Should repex (or some categories of repex) be excluded from our regression analysis and assessed using other techniques? .....	12
	Q12: Are there other approaches to disaggregated benchmarking that we should consider? .....	12
5	Non-econometric analysis .....	15
	Q13: Should we assess business support costs at a group level in order to address cost allocations across companies within groups? .....	15
	Q14: Which types of business support costs should be benchmarked, and how should they be benchmarked? .....	15
	Q15: Which types of business support costs should be excluded from benchmarking? .....	16
6	Regional factors and company-specific effects .....	17
	Q16: How should we estimate and model the impact of regional factors? .....	17
	Q17: Do you agree with the proposed criteria for justifying regional cost factors that we have outlined? ...	18
7	Real price effects and ongoing efficiency .....	19
	Q18: What RPEs should we account for, how should we gauge materiality, and what criteria should we use for index selection? .....	19
	Q19: What common input and expenditure categories are appropriate for structuring RPEs? .....	20
	Q20: How should we identify an appropriate ongoing efficiency assumption? .....	20
	Q21: How should we determine frontier shift? .....	20
8	Combining the elements of our cost assessment.....	22
	Q22: Should we set the efficiency benchmark at the upper quartile level? .....	22
	Q23: Are there types of expenditure that we should model using only historical or forecast data? .....	22

---

Q24: If we use a combination of aggregated and disaggregated modelling approaches, how should we determine the weight we apply to each?.....	24
--	----

# 1 Summary

SGN welcomes the opportunity to respond to Ofgem's cost assessment consultation. Ofgem has clearly given consideration to a wide range of cost assessment issues.

We are in broad agreement with many of the principles set out in the consultation document, for example the proposed model selection criteria, proposed criteria for developing cost pools and criteria for assessing RPE indices. It is important that Ofgem applies these principles consistently as it continues to develop its cost assessment approach.

A number of questions in the consultation raise complex issues, for example selection of suitable cost drivers, the use of middle-up models, or the assessment of Business Support Costs. In many cases, there are several possible approaches, all with strengths and weaknesses. In these cases, it is often necessary to consider the risks associated with different approaches, carefully test different models, and assess these models against the relevant model selection criteria before a conclusion can be drawn. Further discussion through the CAWG will also be valuable in these cases.

Many of the approaches considered by Ofgem represent adjustments to the GD1 cost assessment approach, but there are a few suggestions that could represent significant departures from the GD1 approach. At a high level, our views on these suggestions are as follows.

- **Opex plus modelling.** We welcome Ofgem's decision to consider aggregating activity-level costs as part of a 'totex and opex plus' approach. The disaggregated models suffer from a number of issues, most importantly that they do not capture trade-offs between cost areas, and can therefore distort incentives to reduce overall costs. In our response we set out groups of costs which, based on an engineering and economic rationale, could be suitable to aggregate.

We note however that Ofgem also appears to consider an only 'opex plus' approach, which does not include totex modelling. Totex modelling is a key part of Ofgem's cost assessment toolkit as it provides an overarching view of companies' value for money. 'Opex plus', if it is used, should be used as an additional cross-check to the totex modelling, not as a replacement. We are therefore strongly in favour of a 'totex and opex plus' approach over an 'opex plus' approach.

- **Within-model regional adjustments.** The question then is whether pre-modelling adjustments or within-model variables are the most appropriate. Broadly, we agree with the pros and cons of each as set out by CEPA in the Appendix 1 (Table 7.1). We agree with Ofgem's prior expectation that density is more likely than labour to be suitable for within-model adjustments (see paragraph 6.39 and 6.40)<sup>1</sup> but ultimately the options will need to be fully tested by Ofgem once the data is available to undertake benchmarking. We expect there will be value in testing both approaches and using each as a cross-check on the other. Weighting together the results of both approaches could also be considered.
- **Determining frontier shift.** Our understanding is that Ofgem is considering using outturn historical cost data to calculate frontier shift. CEPA has suggested that this could be assessed by including a time trend or time dummy variables in Ofgem's econometric analysis. Ofgem should share more detail on this proposal, or any alternative approaches it is considering, as soon as possible to allow all stakeholders to consider them carefully and provide more specific comments. However, we already have a number of concerns based on what Ofgem has indicated so far. Firstly, as a matter of principle, we would advise caution in using outturn data to assess frontier shift. Not all productivity improvements will be repeatable, and just looking at a short period of recent data can lead to misleading productivity estimates. Standard approaches to assessing productivity have relied on external benchmarks such as

---

<sup>1</sup> E.g. because, as Ofgem notes (paragraph 6.22) the urbanity/sparsity adjustments are generally reliant on data provided by the GDNs, while labour adjustment is drawn from an external source.

---

the EU KLEMS data, where a wider set of evidence can be found, and have looked at long-term productivity. Secondly, if Ofgem were to carry out an analysis of outturn costs, we would advise caution in interpreting this estimate as frontier shift. There are a number of reasons why the estimates obtained from this approach to frontier shift might be inaccurate, which we set out in detail in our response. Since the estimates obtained from this approach are unlikely to be accurate, we do not consider it appropriate to use these estimates even as a cross-check of frontier shift.

Finally, we note that the consultation questions do not cover the Business Plan Incentive (BPI). As this incentive cuts across the whole cost assessment approach, we feel it is important that there is detailed discussion on the application of the BPI. The consultation states that Ofgem will be continuing its engagement on the BPI up until the submission of final business plans in December. We urge Ofgem to ensure that these discussions take place, as the BPI stands out as an area where further discussion and development is critical at this stage.

We look forward to continued discussion with Ofgem on cost assessment through the CAWG in coming months.

## 2 Approach to econometric analysis

### Q1: What model estimation options should be considered for our cost assessment and why?

We agree with CEPA's recommendation that Corrected Ordinary Least Squares (COLS) is likely to remain the primary estimation technique for GD2. This is a technique which the industry is familiar with and therefore it enhances transparency and is easy for stakeholders to understand. While we agree that Ofgem could test various alternative options, we expect (in line with CEPA's advice) that the COLS model should remain the primary tool for assessment.

CEPA has recommended testing a Random Effects (RE) model and the use of Stochastic Frontier Analysis (SFA). We agree with CEPA that use of this approach would need to be subject to considerable robustness assessment and sensitivity analysis, and that a high bar in terms of extra benefits over and above COLS should be applied. That said, we agree that it is worth testing, although given the sample size we believe that SFA is likely to be difficult to apply. We also agree that Data Envelopment Analysis (DEA) could be considered but this also suffers from the weaknesses outlined by CEPA. We would of course engage constructively on any model that is considered/developed.

### Q2: Do you agree with our proposed criteria for developing potential cost pools? If not, what additional criteria do you propose and why?

Broadly we agree with the proposed criteria for developing cost pools for GD2. However we would note the following issues.

- **Complementarity:** Ofgem should be clear about how it will determine whether costs are "complementary" i.e. what tests would be applied. CEPA has recommended "testing for year-on-year volatility in expenditure in particular areas/activities that appears to be unrelated and/or correlated with changes in business scale drivers", and "testing the expected consistency of workload drivers between different types/areas of expenditure before costs are grouped together for benchmarking". We think it would be helpful if Ofgem/CEPA could provide some illustrative examples of this analysis using the existing GDNs data for consideration at a future Cost Assessment Working Group (CAWG).
- **Cost boundary complexity:** As we have explained through the CAWG, despite the detail that has been developed around cost reporting guidelines, we still consider that there are a number of cost categories where different business models can lead to different cost allocations across GDNs. Ofgem should set out how it will decide whether a particular cost boundary is too complex or too poorly defined. An effective way to achieve this might be to invite the GDNs to identify and discuss any cost boundary issues as an agenda item at a future CAWG.
- **Risk of inaccurate/biased models:** As we have set out in our business plan appendix, one of the tests Ofgem should apply when evaluating the disaggregated models is a set of simple sense checks around the results. We therefore believe this criteria could include the following tests:
  - 1) checking whether the range of efficiency scores observed in any model can sensibly be attributed to differences in managerial efficiency (we have suggested anything beyond a range of 20% is worth investigating further as a rule of thumb).
  - 2) looking at the efficiency scores for each GDN over time (we would not expect to see very large swings up and down from one year to the next in efficiency scores – such a result would suggest a problem with the model rather than genuine changes in managerial efficiency).

---

CEPA has suggested the principles for developing cost pools should be refined based on the information received in the GD2 business plans (see CEPA Annex 1 page 48). We do not consider that the principles would need refining – they should be applicable irrespective of the data received.

### Q3: Should we continue to use the Cobb-Douglas functional form? If not, why?

We agree that Cobb-Douglas should continue to be used. As Ofgem notes (see paras 2.33 – 2.38) this is a standard approach and will allow Ofgem to observe whether the data suggests there are economies of scale in the sector.

### Q4: Do you agree with the proposed model selection criteria and model development phases?

We agree with Ofgem’s proposed model selection criteria. The outline of economic/technical rationale; transparency; and robustness closely reflects the criteria we outline in our Cost Assessment Technical Annex. We also believe that the statistical tests mentioned are a good selection for measuring robustness.

In our July business plan we also noted that statistical tests are not the sole determinant of whether a model can be considered robust. Pragmatically there are a number of critical tests that also should be evaluated – these are mentioned briefly in CEPA’s paper (see e.g. Table 2.2) but are worth emphasising.

- First, **the data must be thoroughly investigated**. Ofgem should be confident that the costs and cost drivers are being measured in the same way across companies, and assess whether there are any data anomalies or inconsistencies in the way that data is captured. This is likely to give rise to outliers that may then be distorting the regression results or distortions in the time series data or trends. Plotting the data is an important first step in the analysis.
- Second, **the results should be within a reasonable range**. The spread between the highest and lowest performing networks can sometimes be implausibly wide, in a manner that cannot reasonably be attributed to managerial efficiency. Typically we would look to apply a hurdle test here, where any model with a greater than 20% spread of efficiency between the best and worst performing networks should require scrutiny to determine whether the spread is genuine or a modelling error. We would propose 20% as a rule of thumb threshold, as we would not generally expect variations beyond this level to be driven by managerial efficiency alone.
- Third, **the results should be reasonably stable over time**. Very volatile results over time for a given company can suggest an underlying error or modelling issue. We would not expect managerial efficiency to significantly improve/deteriorate (or indeed yo-yo) between consecutive years over the course of a price control. As a result, large year-on-year swings in rankings or efficiency scores would suggest a need for greater scrutiny.

In terms of model development phases, the approach outlined by CEPA seems broadly sensible. It is possible that after phase 2 there remain multiple models that broadly pass the tests – we agree broadly with the prioritisation of model selection criteria set out in CEPA Section 2.1.9 and Table 2.2 (as well as CEPA’s observation that some judgement will need to be exercised given that passing all statistical and logic tests is very challenging in applied work).

### 3 Aggregated econometric analysis

#### Q5: Should the cost driver of the totex regression model be determined by the cost drivers of the 'bottom-up' models, or should the totex regression model account for different explanatory variables? Why?

In paragraph 3.9 Ofgem notes that our response to the SSMC we stated that top-down regressions provide an essential sense check, and mitigate some of the risks associated with bottom-up models. We would like to emphasise that in our view, top-down totex models should be a central part of the cost assessment process and retain significant weight in the overall analysis. As we stated in our response to the SSMC, any move away from at least a 50% weighting of top-down totex models needs firm justification.

However, we acknowledge that there are a number of problems with using cost drivers from 'bottom-up' models to determine the 'top-down' Totex CSV, as discussed in the CAWG and as identified by CEPA. These include:

- The GD1 totex model does not allow for a clear intuitive economic interpretation of the results and the relationship between costs and driver.
- The weightings in the totex CSV across scale and workload variables are essentially arbitrary assumptions – even though they are based on the share of these cost heads in totex, there is no reason in principle to believe that this is the correct weighting for drivers of totex.
- There are clear endogeneity issues with some of the drivers (i.e. workload-based drivers and MEAV)
- Arguably, as noted by CEPA<sup>2</sup>, the use of the GD1 totex CSV means that some of the value in having different modelling approaches (i.e. one totex-based and one disaggregated) is reduced, because in effect both approaches model the same underlying relationship between costs and drivers (albeit the totex model will still capture cost trade-offs and the overall coverage of the cost base was different between the two approaches).

At the same time, it should be stated that the 'bottom-up' cost driver worked in GD1, in the sense that it provided broadly sensible results, and the industry is familiar with the approach. In addition, since we expect there will be changes being discussed in relation to the bottom up models, this might improve the robustness of the totex CSV driver.

Ultimately, we welcome the rigor being developed by CEPA and we think Ofgem should apply the relevant criteria for model selection/robustness testing etc.

CEPA and Ofgem have put forward a number of suggestions for alternatives that could be tested and we have the following comments.

- **Scale vs. workload variables:** We agree with CEPA's view that a distinction should be made between scale variables and workload variables. Ideally these models should be tested separately. In doing this Ofgem should recognise that the analyses are fundamentally different. Using workload drivers in a regression essentially amounts to a unit cost analysis (with an acknowledgement that there will be some fixed cost). Such an approach gives no indication of whether companies are effectively managing the amount of work i.e. doing too much/too little – it only tests whether the companies are efficient *given* the amount of work they choose to undertake.
- **Short-run vs. long-run efficiency test:** Ofgem has suggested (paragraph 2.29) that workload drivers may reward companies with networks in a relatively poor condition but we do not think this is the case – a GDN may have a poorer quality network but still be more expensive per unit of work. In effect, a workload model is a short-run test of efficiency – i.e. have you delivered your chosen workload efficiently? Similarly, a model using MEAV is also a short-run test – are you operating efficiently given the value of the existing asset base? In contrast, totex models with exogenous scale drivers (like

<sup>2</sup> CEPA appendix 1 page 48 and 53



customer numbers) give a more long-run test – i.e. have you scaled your network efficiently over the long-run and selected an efficient workload, given the customer base you serve? All of these questions are potentially relevant, but it is important that Ofgem is clear about what sort of efficiency test is being applied given the model and driver it has selected.

- **Consistency when using density variables:** Ofgem/CEPA are also suggesting other amendments to the models including the possible use of a density variable. As a technical matter it will be important to ensure any final models are internally consistent – for example, if a density variable is calculated by dividing network length over customer numbers (as suggested by CEPA) then this may cause modelling issues if those variables are also used as “cost drivers”.
- **Use of a time trend vs. time dummy variables.** Ofgem/CEPA suggest the use of a time trend instead of year-specific dummies. If time effects are not linear or trending in the same direction, dummy variables may be a more effective way of controlling for this. Ofgem will need to test whether the improved efficiency of the models achieved by using a time trend more than offsets any loss of model “fit”.

## **Q6: What could be appropriate cost drivers in middle-up models for opex, capex and repx? Why?**

We are supportive of Ofgem’s proposal to test middle-up models for GD2. These could present a helpful alternative to the disaggregated models since a number of these are not robust.

Ofgem states that it rejected these middle-up models at RIIO-GD1 because they gave broadly similar results to the totex model (see paragraph 3.11) but we do not think this is a relevant test. The fact that the models gave similar results might actually provide additional support/confirmation of those results. In our view it is potentially reasonable for Ofgem to use both totex models and middle-up models (noting that even at the middle-up level there are still potential trade-offs between opex, capex, and repx that can only be captured by a totex model).

Ofgem should test middle-up cost drivers that were suggested at GD-1 (as described in paragraph 3.10) as well as other options as suggested by CEPA. As with the other proposals these should be subject to the model robustness/testing framework.

## **Q7: For which opex activities are there trade-offs that support the rationale for testing ‘totex and opex plus’ modelling?**

We welcome Ofgem’s decision to consider aggregating activity-level costs. The disaggregated models suffer from a number of issues, most importantly that they do not capture trade-offs between cost areas, and can therefore distort incentives to reduce overall costs. We also consider that a number of the models (as specified at GD1) would fail the Ofgem/CEPA model robustness tests.

Considering a ‘totex and opex plus’ type model is a step in the right direction to addressing these issues. We note however that activities that do not fall under ‘opex plus’ may still be well suited to regression benchmarking (possibly alongside expert review). It should not be assumed that any activity areas that are not captured under ‘opex plus’ will not be regressed.

In determining which activities could be modelled together, CEPA has suggested two statistical tests for complementarity, as set out in paragraph 3.15 of Ofgem’s consultation document:

- testing for year-on-year volatility in expenditure in particular areas/activities that appears to be unrelated and/or correlated with changes in business scale drivers; and
- testing the expected consistency of workload and other cost drivers between different types of expenditure before costs are grouped together for benchmarking.

While the principles of these tests seem broadly sound, we are not clear exactly how these would be applied in practice. Some examples of how Ofgem proposes to apply these tests would be welcome. We also note that, in relation to the first test, trade-offs between some activities may be made over the longer term, and may not be visible within cost data on a year-to-year basis. It may therefore also be worth inspecting data over a longer period and across GDNs.

We have set out below groups of costs that we feel could be suitable to aggregate, and which would be worth testing further. The first two suggestions below are driven by cost boundary complexities (the presence of cost boundary complexity is one of the criteria proposed by CEPA for selecting cost pools).

- **Opex (including BSCs) and repex.** Different companies have different contracting strategies, meaning that there will be differences in the degree to which companies incur overheads directly or through contractors. This leads to different approaches to allocating overheads out of contractor costs. Contractor overheads tend to be bundled into contract rates and can be very difficult to separate out accurately. Repex in particular is likely to be affected by this, meaning that there could be significant boundary complexity between opex and repex, driven by differences in business models and cost allocation approaches. A model combining opex and repex could solve this boundary issue, though we are still of the view that large diameter repex should be removed and assessed separately due to its bespoke nature. There will also need to be careful consideration of how repex should be smoothed over time in order to be combined with opex.  
It could also be worth testing the addition of mains reinforcement into this mix, as mains reinforcement also tends to use a significant amount of contractor labour.
- **Repex and work management. (Also capex and work management.)** The costs associated with network strategy is captured within work management, but is sometimes partly reallocated to capex or repex, where some of the associated work is carried out. This creates a source of boundary complexity, which could again potentially be solved by combining repex (or capex) with work management.

The following aggregation suggestions are driven by areas where we consider there are trade-offs between costs (reflecting another of CEPA's proposed criteria). We note in general that companies are faced with important trade-offs in their asset management strategies, which involves a range of options around whether to repair, refurbish, replace or rebuild assets.

- **Maintenance and capex.** In principle, there is a clear trade-off between opex solutions (maintenance) and capex solutions (asset integrity) within asset management. However, this should be tested carefully as the trade-off may not be this explicit in reality. For example, capital expenditure may be at a level that is keeping overall asset health (and therefore maintenance costs) constant by offsetting the ongoing deterioration of assets, rather than saving on maintenance costs.

In any models where capex and opex costs are aggregated, there will need to be careful consideration of how capex should be smoothed in order to be combined with opex.

- **Emergency, repair and pressure management.** Management of network pressures can reduce leakage, and therefore reduce costs associated with emergency and repairs.
- **Mains reinforcement and demand-side management.** Mains reinforcement and demand-side management are two different solutions to resolving capacity constraints. While costs associated with demand-side management are not very material currently, they could be in future, and Ofgem may wish to consider combining these activities.

As well as identifying suitable groups of costs that can be aggregated, it is also important to carefully consider and test cost drivers that could be used for these models. As explained in Q5 above, simply aggregating the cost drivers of the underlying disaggregated models into a CSV should not be the preferred solution.

We note that later in Ofgem's consultation, in paragraph 4.39, Ofgem considers an only 'opex plus' approach, which does not include totex modelling, and states that this will lead to less aggregative modelling. Totex

modelling is a key part of Ofgem's cost assessment toolkit as it provides an overarching view of companies' value for money. 'Opex plus', if it is used, should be used as an additional cross-check to the totex modelling, not as a replacement. We are therefore strongly in favour of a 'totex and opex plus' approach over an 'opex plus' approach.

## **Q8: Are there other particular costs that we should aggregate and test in our analysis?**

CEPA sets out four cost pooling options in its annex on econometric modelling.<sup>3</sup> These are:

1. aggregated modelling;
2. totex and disaggregated (activity based) modelling;
3. totex and 'opex plus' modelling; and
4. 'opex plus' modelling.

Option 2 is essentially the same approach followed by Ofgem at GD1. We have highlighted various issues with the disaggregated models, but if these issues are addressed then this option could be acceptable.

Options 1 and 3 take a more aggregated view and in our view are therefore preferable. These approaches help to address issues around trade-offs across cost categories and various other weaknesses in the disaggregated models.

As set out in our response to Q7 above, option 4 is least preferable, as it removes any test of overarching value-for-money delivered by GDNs.

In terms of specific cost areas that could be aggregated and tested separately, we think the following areas are worth considering and testing further.

- **Business support costs.** CEPA's table B.1 lists business support costs under opex, and we would welcome testing of whether BSCs would be suitable for inclusion within an 'opex plus' model. As for any models used at GD2, Ofgem's chosen approach should pass the relevant model specification criteria and statistical tests. We have discussed business support costs in more detail in our responses to questions 13-15.
- **Emergency and repairs.** We have previously discussed combining these two cost categories with the CAWG, and we feel that further work would be valuable in testing this.

## **4 Disaggregated econometric analysis**

### **Q9: Are there trade-offs between opex and capex activities that support the rationale for considering 'opex plus' modelling?**

Please see our response to questions 7 and 8.

### **Q10: Which cost areas should be assessed using workload drivers as opposed to other cost drivers? Why?**

As Ofgem and CEPA note, the downside of using workload as a driver is that it can be influenced by the company and therefore potentially distorts incentives. We have also found that there are some categories (notably mains reinforcement and connections) where the workload drivers that were used at GD1 do not give a good fit with the cost data. We have also explained to Ofgem through the CAWG that there is high volatility in the repex

---

<sup>3</sup> CEPA (June 2019) RIIO-GD2 cost assessment – econometric modelling & regional factors, section 5.3

results over time, which is unexpected given that it is essentially a unit cost analysis.

Despite these issues, we think there is a valuable role for analysis based on workload for cost categories where there is a large volume of broadly repetitive work which would be expected to have similar unit cost. For these sorts of activity, the question of whether companies are minimising the unit cost to deliver workload is a relevant and important test of managerial efficiency.

As highlighted by Ofgem (para 4.68) we have noted some ways that the GD1 workload drivers could be improved, for example with repex and CISBOT. We believe that CISBOT workload should be included within the repex workload as there is a clear trade-off in cost, and we believe that CISBOT allows us to deliver our repex programme efficiently by laying less pipe.

In general, any alternative cost driver should pass the robustness criteria set by Ofgem.

### **Q11: Should repex (or some categories of repex) be excluded from our regression analysis and assessed using other techniques?**

We have previously done some analysis, which was presented in the CAWG, on splitting repex into Tier 1 and other Tiers for separate regression benchmarking. The results show that for the wider diameter pipes it is harder to benchmark because the unit cost is more volatile. The data indicates that there is more bespoke work carried out with these wide diameter pipes, and given that there is not much work (data points) as in Tier 1, a unit cost analysis will not be representative of how efficient a project is. Risers should also be removed for separate assessment for similar reasons.

However, given that there is less workload in these categories, the impact of excluding them from the regression might not be significant.

More generally, we have highlighted in the CAWG and in our response to Q12 below that we believe there are a number of issues with the disaggregated repex model as it currently stands. However, we have made some suggestions for how the model can be improved, for example averaging results over GD1 to address the volatility in efficiency scores. Even if these changes are implemented, there may remain some differences in performance that aren't purely explained by differences in efficiency. If this is the case, expert review could help to identify qualitative drivers of these differences, or could help to explain what is driving the regression results. In this case, a weighting between regression results and expert review could be a possible solution. As always with any cost assessment that could be subjective, this must be done in a clear and transparent way, with the process set out and agreed upfront.

### **Q12: Are there other approaches to disaggregated benchmarking that we should consider?**

We have set out in the table below some proposed modifications to specific disaggregated models which were used at GD1.

Repex	<ul style="list-style-type: none"><li>• Include CISBOT workload in repex cost driver;</li><li>• Consider using an average or aggregate of the five year results;</li><li>• Continue to adjust results for innovative processes;</li><li>• Consider removing large diameter pipes from the results (doing an independent engineering assessment on these projects);</li><li>• Update the unit cost assumptions that feed into the synthetic unit costs; and</li><li>• Conduct a sense-check between cost efficiency results and quality of outputs.</li></ul>
-------	--

Capex connections and mains reinforcement:	<ul style="list-style-type: none"> <li>Consider updating the unit cost assumptions;</li> <li>Ensure appropriate adjustments are made to one-off large costs that skew the regression results (such as mains reinforcement costs in London); and</li> <li>Ensure costs are correctly reported (for example the North West reported negative costs in their mains reinforcement).</li> </ul>
Opex	<ul style="list-style-type: none"> <li>Use publicly reported escapes (PRE) as 100% of the emergency cost driver; and</li> <li>Ensure reporting in the repairs model is as consistent as possible in order to use external condition repairs as a cost driver.</li> </ul>

We note that at present, we have significant concerns about the robustness of some models. The RAG rating in the table below summarises our views on the feasibility of resolving the weaknesses in the individual RIIO-GD1 regressions. While many of these weaknesses can be addressed, with the possible exception of the capex regressions, moving to an entirely bottom-up approach would be highly risky as it would remove the overarching cross-check that the totex model provides.

**Table 1: RAG rating for robustness of regression models**

Model	% of totex	RAG rating
<b>Top-down</b>		
Totex	100%	
<b>Bottom-up</b>		
Work management	10%	
Emergency	5%	
Repairs	6%	
Maintenance	6%	
Connections	4%	
Mains reinforcement	1%	
Repex	40%	

The key issues are:

- Work management, emergency, repairs, maintenance (opex regressions):** the current models appear to be relatively good at explaining efficiency, but produce larger ranges in efficiency scores than we would expect can be explained by differences in managerial efficiency alone. Using alternative cost drivers that better explain costs (such as public reported escapes rather than a CSV of customer numbers and external condition reports for the emergency regression), and ensuring that regional adjustments are reviewed and updated can help to resolve this.
- Connections, mains reinforcement (capex regressions):** both capex regressions display infeasibly large efficiency score ranges. The mains reinforcement regression results are also very volatile year-to-year. These results are not credible given that both regressions use workload cost drivers, and are effectively assessing unit costs. A combination of relatively small cost head, highly variable costs, and poor data suggests this cost

area is not suitable for regression analysis. There may be other types of analysis suitable for these cost areas, such as bespoke engineering assessments/judgement.

- **Repex regression:** the repex regression results display significant volatility over time. Again, this undermines its credibility because the regression is essentially a unit cost analysis, and large swings in unit costs from year to year do not appear realistic. However, we believe this regression could be improved by updating the synthetic unit cost driver, and aggregating/averaging data over time to reduce volatility.

These issues will need to be further considered and addressed through the GD2 process.



## 5 Non-econometric analysis

### **Q13: Should we assess business support costs at a group level in order to address cost allocations across companies within groups?**

Ofgem assessed business support costs (BSCs) at a licensee group level at both GD1 and ED1. Because companies generally incur business support costs at a group level, this approach is suitable as it avoids any distortions in benchmarking results that could be caused by different approaches to allocating BSCs to licensees within a group.

However, there are some detailed considerations that should be taken into account when deciding whether to assess BSCs at a group or licensee level, and the final approach taken will need to be compatible with other decisions taken on how BSCs will be assessed at GD2. For example:

- if BSCs are assessed using regression techniques rather than unit cost assessment, using group-level data significantly reduces the sample size available; and
- using group-level versus licensee-level data could have implications on the upper quartile benchmark, and it is worth understanding what these are.

### **Q14: Which types of business support costs should be benchmarked, and how should they be benchmarked?**

It would be beneficial for Ofgem to revisit the GD1 approach to assessing BSCs, and we are open to constructive discussion on this through the CAWG. There are a number of conceptual and methodological questions that need more detailed discussion and testing. As discussed in response to Q4 above, we broadly agree with Ofgem's model selection criteria, and these should be applied when developing the approach to assessing BSCs as well as other cost areas.

Some of the key conceptual issues to explore include the following.

- The types of BSCs to be benchmarked. We are broadly supportive of the types of BSCs that were benchmarked at GD1, with the possible exception of IT and Telecoms (see response to Q15).
- Whether BSCs should be included within "opex plus", if this approach is used. We are open to this idea and it appears to be the approach CEPA is suggesting (see CEPA Appendix B) but it will need further discussion and testing.
- Whether a different modelling technique should be used to benchmark BSCs, such as regression rather than unit cost analysis. Regression analysis would have the benefit that it would take account of any economies of scale that might apply to larger groups that can spread their BSCs across a number of licensees. The current ratio analysis does not account for possible economies of scale.

Some of the more detailed methodological issues include the following.

- Ofgem's GD1 BSC models were not transparent, and the results were difficult to replicate. In contrast, the approach Ofgem describes for RIIO-ED1 seems to be significantly more transparent since it relied on data that all the DNOs had access to (see Ofgem 5.18 – 5.21). However, it does appear that for RIIO-ED1 there were material changes in the methodology between the fast-track and slow-track decisions – if Ofgem is considering significant changes to the BSCs models for RIIO-GD2 it would enhance transparency and robustness if these can be explained to the sector early enough for a proper engagement and review process.
- At GD1, Ofgem made a net to gross adjustment to companies' BSCs, re-allocating BSC costs that companies had allocated out into other cost categories back into BSCs. This was to ensure that companies' BSCs were compared on a like-for-like basis. This adjustment is essential to the robust assessment of BSCs, and is worth re-visiting to ensure it is done as accurately as possible. One potential issue is that companies that make more use of contractors may have more overheads allocated within

their contractor costs, rather than within BSCs. Separating out and reallocating overheads from contractor costs can be challenging, but is important as different business model choices shouldn't impact benchmarking results other than the extent to which they have a genuine impact on efficiency.

- Some of the cost drivers underlying the CSV Ofgem used in its unit cost assessment would not have passed the criteria that Ofgem has set out for good cost drivers. For example employee numbers and spend are within the control of companies and their use within the CSV can create perverse incentives. These cost drivers should be revisited.
- If BSCs are assessed on an aggregate basis, as was done at GD1, Ofgem should revisit the question of what cost driver to use. At GD1 Ofgem aggregated the cost drivers for each individual BSC area into a CSV; however the approach to aggregation appeared to have some weaknesses (e.g. Ofgem appears to have used arithmetic rather than geometric averaging). This should be revisited, and other options, such as scale drivers, should also be considered.

### **Q15: Which types of business support costs should be excluded from benchmarking?**

We broadly agree with benchmarking the same business support cost categories as were benchmarked at GD1, with the possible exception of IT & Telecoms. IT & Telecoms can be difficult to compare across companies at any given point in time for a number of reasons, including:

- different legacy systems;
- different timing of investments in IT & Telecoms; and
- transitional periods while moving to different systems. For example, some companies are moving away from capex-heavy systems to more opex-based cloud systems. This could create accounting or timing differences in costs between GDNs.

Given the materiality of IT & Telecoms, which makes up the largest proportion of business support costs (approximately 40% across all GDNs over GD1), these differences could drive significant differences between companies in the benchmarking assessment for reasons other than differences in efficiency.

As mentioned in the cost assessment consultation, at ED1 Ofgem used expert review (with 50% weighting) alongside unit cost benchmarking to assess IT & Telecoms costs. We think that expert review is a sensible approach to assessing IT & Telecoms, either alongside or instead of a benchmarking model.

We also note that at GD1, insurance costs were excluded from the BSC benchmarking. We agree with Ofgem's reasoning at GD1 for taking this approach: differences in risk appetite and appropriate levels of coverage between companies make it different to ensure a like-for-like comparison. It is also important that Ofgem does not create artificial incentives for GDNs to under-insure.

Finally, at GD1, costs related to training and apprentices were assessed separately from business support costs. GDNs worked with EU Skills to develop their workforce planning models and provided long term estimates of their workforce renewal requirements. Ofgem used these forecasts to form a view of numbers of training and apprentice programmes needed, and applied a unit cost to calculate allowances. Training and apprentices costs should continue to be assessed separately from BSCs due to fact that the underlying drivers of these costs, such as current workforce composition, can vary significantly between GDNs.



## 6 Regional factors and company-specific effects

### Q16: How should we estimate and model the impact of regional factors?

It is clear that regional variations in cost exist and are driven by external factors outside of GDNs' control. There is substantial regulatory precedent and evidence of these factors. We therefore consider that it is essential these factors are addressed in the benchmarking analysis, so as to generate a fair and robust outcome.

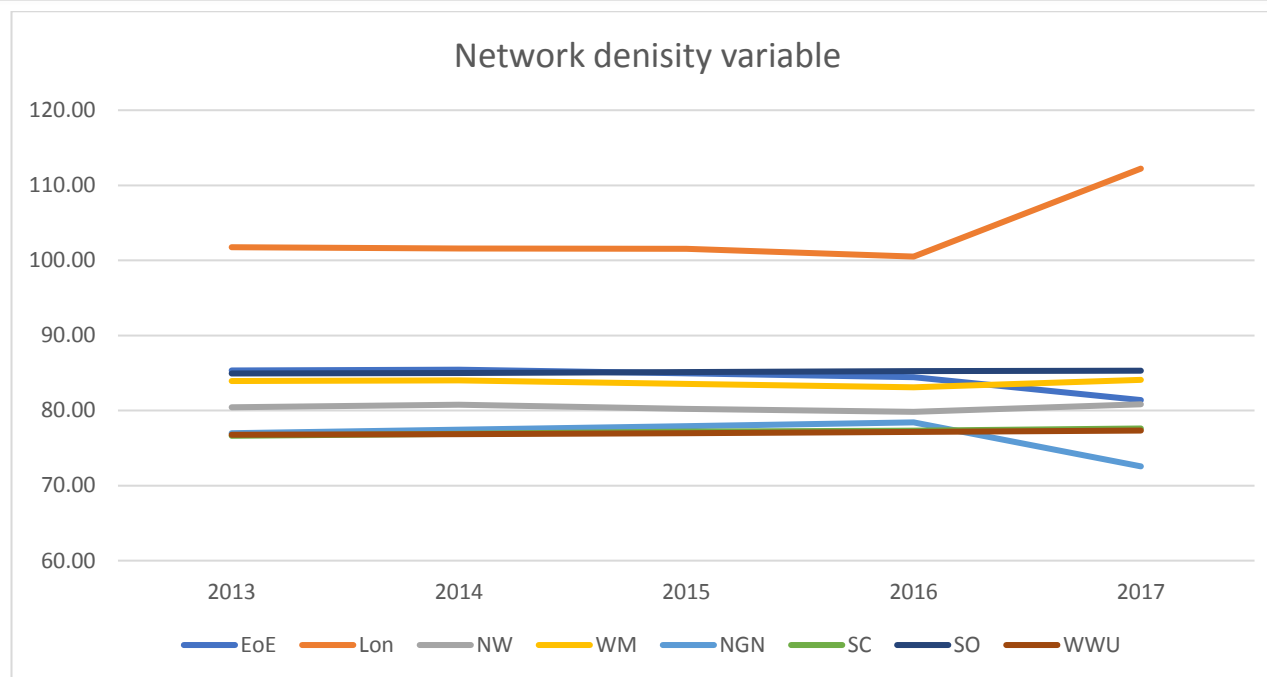
We agree with Ofgem's view that post-modelling adjustments are not appropriate and should not be used (paragraph 6.13). The question then is whether pre-modelling adjustments or within-model variables are the most appropriate. Broadly, we agree with the pros and cons of each as set out by CEPA in the Appendix 1 (Table 7.1). We agree with Ofgem's prior expectation that density is more likely than labour to be suitable for within-model adjustments (see paragraph 6.39 and 6.40)<sup>4</sup> but ultimately the options will need to be fully tested by Ofgem once the data is available to undertake benchmarking. We expect there will be value in testing both approaches and using each as a cross-check on the other. Weighting together the results of both approaches could also be considered.

If Ofgem uses a within-model variable, Ofgem should attempt to evaluate the materiality of this in terms of the effective impact on GDN's cost base in £m terms (i.e. by looking at the difference between allowances with and without the variable). This would allow Ofgem to be able to produce a table similar to that shown in Table 6.1 of the consultation, such that GDNs can understand the impact of the variable, and compare this with their own information on direct valuation of region specific costs. As we have noted in our business plan, SGN has undertaken an extensive investigation into a number of region-specific cost areas and we believe this information should be utilised by Ofgem in the cost assessment process.

We note that the variables that are used in the within-model options should be subject to scrutiny (in line with the general need to check the data). For example, we have noticed that the density variable CEPA appears to have used exhibits some significant movements in the last year of the estimation window which will need to be verified (see chart below).

---

<sup>4</sup> E.g. because, as Ofgem notes (paragraph 6.22) the urbanity/sparsity adjustments are generally reliant on data provided by the GDNs, while labour adjustment is drawn from an external source.



As a technical matter, it will also be important to ensure that the selection of a density variable does not create econometric issues. For example, one of the proposed variables (i.e. customer numbers over network length) is based on data which may also be used as the cost driver in some of Ofgem/CEPA's proposed modelling approaches. Ofgem should seek academic/technical advice on whether this is likely to cause biased or distorted results from the econometrics (in effect, the density variable is equivalent to placing a special restricted form on certain cost drivers).

We would re-iterate our view that imposing symmetrical adjustments is essentially arbitrary and has no economic rationale. Ofgem's consultation suggests some GDNs were supportive of this method but we note that CEPA appears not to have commented on its merits. The modelling adjustments are there to ensure that a fair comparison is made between GDNs – this should be the over-arching driver for modelling choices. Ofgem should reject the proposal for arbitrarily imposing symmetry in its decision document in this consultation.

We also disagree with Ofgem's statement (paragraph 6.19) that pre-modelling adjustment removes the incentives on companies to mitigate these costs. Companies have a strong incentive to control costs during the price control period through the totex incentive mechanism. SGN has invested significant effort in seeking to mitigate regional cost issues e.g. through multi-skilling staff in sparse areas of Scotland.

### Q17: Do you agree with the proposed criteria for justifying regional cost factors that we have outlined?

We understand this question refers to the criteria outlined in paragraph 6.11 of the consultation. The proposed criteria appear sensible in principle and may be useful to guide Ofgem. However, we note that a company affected by regional factors may not be able to provide evidence that they are affected "in a materially different way to others", because the company will not have access to other company's costs and data. Therefore it is likely that Ofgem would need to make that evaluation.

We expect that the urbanity, sparsity and regional wages adjustments that were made at RIIO-GD1 will continue to be justified against the proposed criteria. There is a clear rationale and precedent for these adjustments, and Ofgem should take that into consideration as a starting point. Arguably therefore these criteria should only be applied when considering any new regional adjustments that are requested.

## 7 Real price effects and ongoing efficiency

### Q18: What RPEs should we account for, how should we gauge materiality, and what criteria should we use for index selection?

We agree with CEPA that there is no clear rule for how to assess materiality of RPEs,<sup>5</sup> so we feel that this is an area for careful consideration and discussion.

In gauging materiality, we broadly agree with Ofgem and CEPA that materiality criteria relative to both totex and CPI are relevant. However, we disagree with Ofgem that each RPE needs to be shown to be material relative to both totex and CPI.<sup>6</sup> If an RPE is found to make up a significant proportion of totex, even a small deviation from CPI over time could have a very large absolute impact on costs. Similarly, if an RPE makes up a small proportion of totex but deviates significantly from CPI for a sustained period of time, this could result in a material impact on costs. Therefore these two criteria cannot be considered independently – an RPE should only need to pass one of the two criteria to be considered material overall. Of course, if both criteria are failed then the RPE can be considered immaterial.

This is in accordance with CEPA’s recommendation that Ofgem takes “a risk-based approach focusing on cost categories that represent a relatively large share of totex and/or that would likely face relatively large movements over time.”<sup>7</sup> [Emphasis added.] CEPA suggests the following rule of thumb that combines materiality relative to both totex and CPI:

*“the price index and CPI-based measure should be sufficiently large to change totex by 0.1 percent per year relative to the default approach (and therefore an aggregate 0.5% of annual totex over the five-year price control period). That would mean a cost category covering 5% of totex would need to demonstrate an average historical RPE difference of 2% a year; a category covering 1% would need to exceed 10%.”<sup>8</sup>*

When assessing materiality, there is also a question of what level of aggregation input costs should be assessed at. CEPA notes that categories of cost could “arguably be sub-divided into smaller categories that would individually be classed as immaterial but together may represent a material RPE.”<sup>9</sup> We agree with the level of aggregation set out in Ofgem’s consultation, which is a well-established approach, i.e. direct labour, contract labour, materials, plant and equipment, transport, and other. Assessing materiality on a more granular basis risks setting an overly burdensome materiality threshold.

Section 3.2 of CEPA’s paper on frontier shift sets out an approach to selecting indices to index RPEs. We generally agree with the process set out by CEPA: generating a long list of potential index options based on regulatory precedent, stakeholder suggestions and wider research, and assessing these options against a set of assessment criteria to generate a shortlist. We also agree with the high-level assessment criteria (simplicity, credibility, accuracy, independence, transparency and timeliness), and most of the detailed criteria for index selection set out by CEPA. However, we have comments on a few specific criteria.

- **Reflects movements in the respective input cost category (or a distinct portion thereof) for a notional efficient company in the sector.** Our understanding of Ofgem’s guidance to date is that the existence of RPEs should be based on an assessment of companies’ own input cost movements, rather than those of a notional efficient company.
- **The index has a low or no chance of being manipulated by actions of companies in the sector.** While it is important that no individual company should be able to manipulate a chosen index, for most suitable

<sup>5</sup> CEPA (June 2019) RIIO-GD2 cost assessment – frontier shift, page 15.

<sup>6</sup> Ofgem, RIIO-2 tools for cost assessment, paragraph 7.9.

<sup>7</sup> CEPA (June 2019) RIIO-GD2 cost assessment – frontier shift, page 16.

<sup>8</sup> CEPA (June 2019) RIIO-GD2 cost assessment – frontier shift, page 17.

<sup>9</sup> CEPA (June 2019) RIIO-GD2 cost assessment – frontier shift, page 15.

indices in areas such as construction contractor costs, the GDNs are likely to influence these indices at least to an extent. We feel that the materiality threshold for this criterion should therefore be set relatively high so as not to rule out indices that are a good reflection of GDNs' costs.

- **Series does not capture ongoing efficiency.** We agree that it is important that Ofgem's approach to RPEs and ongoing efficiency should not result in any double counting of ongoing efficiency. We agree with CEPA that further work is needed to determine that Ofgem's treatment of RPEs and ongoing efficiency are consistent.<sup>10</sup>
- **Series is publicly available.** We would challenge that the series must be publicly available in order to be used to index RPEs. If Ofgem identifies a good index that requires a subscription to be accessed, and there are good alternatives that are publicly available, we feel that this shouldn't be a barrier to using that index. While ideally all stakeholders would be able to access the data and replicate the calculations, this is not essential, and index suitability should take priority over this.
- **A forecast comparable to the index is available from the same data provider.** Although the availability of a forecast of the index from the same data provider would be helpful to the understandability and implementation of the index, our view is that this is a nice-to-have rather than an essential criterion. Given the difficulty in finding suitable indices, Ofgem should avoid imposing overly onerous criteria where these are not essential.

## Q19: What common input and expenditure categories are appropriate for structuring RPEs?

We agree with the common input and expenditure categories set out in the consultation. The relevant input categories are direct labour, contract labour, materials, plant and equipment, and other. We think that property costs, particularly rent and utility bills, may be another input category that could be worth testing for RPEs.

The relevant expenditure categories are direct opex, indirect opex, capex, repex mains and repex services. These categories have been used in the past, are well established and well understood within the sector.

## Q20: How should we identify an appropriate ongoing efficiency assumption?

Ofgem has historically assessed ongoing efficiency through growth accounting analysis using the EU KLEMS dataset. This is a standard approach which is well established in regulatory precedent and widely understood. We see no reason to deviate from this approach.

## Q21: How should we determine frontier shift?

Our understanding is that Ofgem is considering using outturn historical cost data to calculate frontier shift. CEPA has suggested that this could be assessed by including a time trend or time dummy variables in Ofgem's econometric analysis. Our view is that it is not clear that this additional assessment of frontier shift is appropriate in principle or in its implementation.

Firstly, as a matter of principle, we would advise caution in using outturn data to assess frontier shift. Not all productivity improvements will be repeatable, and just looking at a short period of recent data (e.g. just GD1 data) can lead to misleading productivity estimates. Standard approaches to assessing productivity have relied on external benchmarks such as the EU KLEMS data, and have looked at long-term productivity.

Secondly, if Ofgem were to carry out an analysis of outturn costs, we would advise caution in interpreting this estimate as frontier shift. There are a number of reasons why the estimates obtained from this approach to frontier shift might be inaccurate:

- An analysis of outturn historical costs will likely capture both the impact of frontier shift, and the impact

---

<sup>10</sup> CEPA (June 2019) RIIO-GD2 cost assessment – frontier shift, pages 13-14.

of catch-up efficiency. CEPA has also acknowledged this issue: *“We consider that the results obtained from the decomposition of historical company cost performance to estimate historical ongoing efficiency (as distinct, for example, from catch-up efficiency and input price pressures) should be interpreted with caution as it is likely to be difficult to isolate such effects with confidence.”*<sup>11</sup>

- Similarly, if there are variables omitted from Ofgem’s analysis which might in fact be important in explaining movements in cost over time (for example, quality changes), the omitted variables will distort the parameter estimate on time trend variables or time dummy variables.
- An analysis of outturn costs will assume that non-repeatable improvements will necessarily apply going forward, and risks overestimating frontier shift. This could include, for example, legislative changes that have resulted in a reduction in GDNs’ costs. We note that this could lead to double counting of productivity across price controls, when productivity targets are rebased at the end of the period.

Since the estimates obtained from this approach are unlikely to be accurate, we do not consider it appropriate to use these estimates even as a cross-check of frontier shift.

Lastly, CEPA notes that, *“If the time trend coefficient is used to produce forecast costs, an estimate of frontier shift will therefore already be in part or fully included in the allowance produced by the model.”*<sup>12</sup> We agree that if Ofgem includes a time trend in its regressions and then extrapolates this trend when forecasting cost allowances, also imposing a productivity adjustment could lead to double counting of productivity. Ofgem should be careful to avoid this.

---

<sup>11</sup>CEPA, RIIO-GD2 cost assessment – frontier shift, p. 10-11.

<sup>12</sup> CEPA, RIIO-GD2 cost assessment – frontier shift, pages 11.

## 8 Combining the elements of our cost assessment

### Q22: Should we set the efficiency benchmark at the upper quartile level?

Ofgem has historically used the upper quartile as the efficiency benchmark. This is long-established practice and it reflects the fact that Ofgem cannot have complete confidence that the models it is estimating are robust. We think it is important to recognise that the benchmarking cannot be perfect, even if it passes the various robustness checks and selection criteria. We agree with the characterisation set out in the CEPA paper (Section 4.2.1).

We agree with Cadent's view (see Ofgem consultation para 8.5 – 8.6) that the RIIO-ED1 approach of aggregating model results before calculating an upper quartile will help to limit the effect of cherry picking. However, it will still be imperative that any mid-level or disaggregated modelling that is used is subject to the proper scrutiny and robustness checks. The upper quartiling approach should not be an excuse to lower the bar on the quality of the underlying analysis.

CEPA has recommended that Ofgem explore the case for removing the glide-path applied for RIIO-GD2 (see Ofgem para 8.9 and CEPA page 10 and Section 4.2.2). The glide-path was part of the IQI mechanism at GD1, and by interpolating between modelled costs and companies' own forecasts, meant that networks were only required to close 75% of the gap between their costs and Ofgem's modelled upper quartile costs. This had three key benefits:

- it allowed companies time to catch up to the efficiency target;
- it recognised the weaknesses of the benchmarking models, but also served as part of the IQI incentive arrangements; and
- it was also a critical part of the IQI incentive, strengthening the incentive to submit good forecasts.

With respect to the second point, in our view it is critical that there is an assessment of the robustness and quality of the data and the analysis *before* any judgement is taken on whether or not to apply a glide-path. We agree with CEPA that if the models continue to exhibit a wide range of efficiency scores or volatility over time, or if there are evident problems in the data or drivers, then this would indicate that the models are less reliable. It would therefore be inappropriate to remove interpolation entirely. As both Ofgem and CEPA have noted, many past price controls have employed a more gradual shift to the frontier even than the 75% interpolation employed at GD1. In particular, Ofgem must be careful not to remove the glide-path and use of the upper-quartile, as this would remove two important balances against potential inaccuracies in the benchmarking models.

Finally, if Ofgem does decide to remove interpolation, this will represent a significant tightening in the regulatory allowance setting process relative to previous controls. It will be important to consider this change in the context of the wider price control framework i.e. financeability tests and, in particular, when considering the proposed 50bps cost of equity adjustment.

### Q23: Are there types of expenditure that we should model using only historical or forecast data?

As Ofgem has noted the RIIO-GD1 approach involved using a blend of historical and forecast models. We broadly agree with CEPA's evaluation of the pros and cons of each (see CEPA Appendix 1 Section 3.2.3).

- Using historical data has the advantage that it reflects real outcomes that capture on-the-ground, achieved efficiency. In contrast, forecast data is inherently uncertain, and there is no guarantee that



the implied productive process (i.e. the modelled link between forecast costs and forecast drivers) is actually achievable.

- Considering forecasts, however, can be helpful if company business plans are projecting step changes or structural changes in the categories and drivers of costs (e.g. if new activities are included in the cost base for the GD2 period).

We note that there is a distinction to be made between totex and disaggregated modelling. We think that benchmarking forecasts of disaggregated cost heads increases the potential for data allocation distortions (since forecast costs are potentially easier to switch between categories, exacerbating the allocation risk).

We strongly disagree with CEPA's suggestion that the choice of history/forecast data should be determined by whichever approach predicts the lowest modelled cost (see CEPA Appendix 1 page 32 and 33). The choice of sample should ultimately be determined by employing the same model selection criteria as Ofgem has proposed elsewhere. Just because a model generates lower allowances does not make it more reliable – this is true for essentially any modelling choice Ofgem makes (e.g. choice of cost driver, technique, sample period etc).

We also think it is important for Ofgem to acknowledge that historical/forecast data has two potential uses under Ofgem's modelling framework.

- First, there is the question about what data to use to run the regression and **estimate parameters** i.e. the regression sample. In principle either history, or forecast, or a full run of all history/forecast data could be included in the panel that is used in the regressions.

We agree with CEPA's assessment that Ofgem needs to consider whether there are changes in cost allocations over time or structural breaks in the data. For example, the move to cloud-based IT systems could lead to a change in the drivers of IT costs, and is likely to mean that more IT costs are categorised as opex rather than capex. This means that historical IT costs may not be representative of forward-looking costs. Adding to this complexity, companies are at different stages in adopting cloud technology, and their forecast IT costs will reflect this.

The data templates will be essential to this process of assessing the data for structural breaks and differences in forecasts. The templates will need to be completed carefully, and should be the focus of discussions through the CAWG around how to address any breaks or inconsistencies.

- Second, there is the question about how Ofgem then uses the estimated parameters **to predict modelled costs and calculate efficiency scores**. At GD1 Ofgem calculated efficiency scores in a single reference year. In the historical model this was 2011/12 and in the forecast model this was 2013/14. Efficiency scores were calculated by using the estimated regression parameters to project modelled costs in those reference years, and comparing the modelled costs to actual/forecast costs in that year.

In principle we think that it might be possible to separate these questions entirely. For example, it might be possible to **estimate parameters** using only historical data but use those parameters to **calculate efficiency scores** in forecast years (or indeed other combinations of historical/forecast data).<sup>13</sup> Ultimately Ofgem should have a clear economic rationale for its choice of historical/forecast years for both estimating parameters and calculating efficiency scores.

Finally, if forecast data is going to be used in the benchmarking, the GDNs must be given sufficient time in advance of the draft determination to have prior sight of the data and to understand Ofgem's direction of travel on the modelling framework. This will ensure that the inherently uncertain forecast data can be properly scrutinised by the other GDNs for errors/inconsistencies, which will help to get around some of the weaknesses of using forecasts.

<sup>13</sup>

CEPA has pointed out that this was Ofwat's approach in PR19 – see CEPA Appendix 1 footnote 43.

---

## **Q24: If we use a combination of aggregated and disaggregated modelling approaches, how should we determine the weight we apply to each?**

In principle, a fully aggregated approach to benchmarking modelling is desirable, as it removes any distortions caused by differences in cost allocation across companies, and it ensures that companies' decisions on trade-offs only affect benchmarking results to the extent that they minimise overall costs. This means that an aggregated approach tests overall value-for-money, and encourages networks to optimise not only within cost categories, but when making decisions on how to allocate spending across cost categories.

However, in practice, disaggregated models can be helpful in providing a cross-check to aggregated modelling results, they can help to identify specific cost areas that are driving inefficiency, and they can sometimes give better statistical results because more variation is often seen disaggregated costs and cost drivers, which can make these variations easier to model.

If Ofgem continues to see value in using both aggregated and disaggregated models, we recommend ensuring first that the disaggregated models are suitably robust and accurate. If this can be achieved, we propose that an equal weighting continue to be applied to each.