DPCR5 Initial Proposals

Appendices – Cost Assessment

E.ON Central Networks

The photograph on the front cover was taken by Andy Icke, an Environment Advisor for Central Networks. It was taken in the Cotswolds, Worcestershire.

This paper is provided in support of our comments on Ofgem's opex cost assessment methodology (section G, question 1 in the main document).

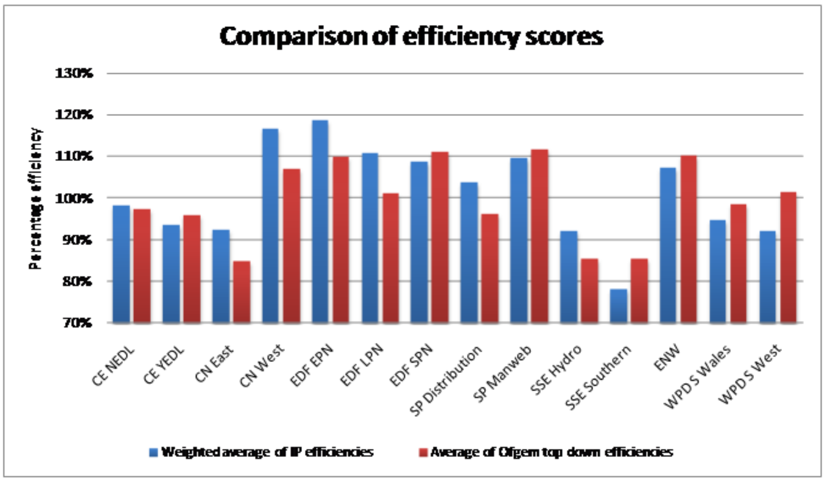
Appendix 1:

Errors in Ofgem's methodology and the need for top-down benchmarking

Top-down Benchmarking

Top-down benchmarking is significantly more robust than bottom-up benchmarking and therefore should be the primary mechanism for benchmarking with bottom-up benchmarking in a supporting role. While Ofgem has suggested its has taken the top-down analysis into account the efficiencies used to set allowances are driven almost entirely by the values obtained from bottom-up benchmarking.

Ofgem suggests it acknowledges the trade-offs between the network operating costs and indirect costs but feel that these are less significant than the benefits of benchmarking at a more detailed level. We had hoped that analysis at a greater level of detail would improve the accuracy by improving the relevance of the drivers used. However, we have found that the results are disappointing and it is likely that "noise" from reporting differences, differences in operation, unusual costs that are not classified as atypical etc have more of an impact at this level. Comparisons at this activity level are therefore *less* reliable, such that any benefit from including more sophisticated drivers is outweighed by the additional issues created by analysing costs at a lower level.



We have created a weighted average efficiency score for the DNOs based on the efficiency scores for NOC and Indirects as stated in the initial proposals. Comparing these values to the top-down efficiency scores shows that, as with the May Paper, the range in efficiency scores is exaggerated by using bottom up benchmarks. This favours WPD and SSE Southern while disadvantaging CN East, CN West, EDF EPN, EDF LPN, SP Distribution and SSE Hydro.

This is evidenced in Appendix C1 which highlights the many errors and issues with each of the bottom-up benchmarks used. While the bottom-up benchmarking can give a general indication of relative efficiency it does not give robust models, as demonstrated by the results of the statistical tests given in table 8 of the Appendix to the initial proposals 94a/09 . This shows that only the top-down regression passes all the statistical tests and has an R squared value above 0.8. Many of the core regressions fail more than one statistical test and with R squared values as low as 0.3 are simply not robust enough to be used in setting cost baselines. While we acknowledge that Ofgem has tried to take a range of bottom-up results into account by including variations, the problem remains that many of these will also be poor quality models and the inherent problems cannot be eliminated by taking averages or creating weighted composite drivers.

As well as being better statistical models, additional top-down benchmarks would allow a wider view of efficiency to be taken which can not be replicated at the bottom-up level, such as taking previous capex into account and normalising for opex/capex boundary issues.

Our analysis has suggested that the range in efficiencies between the DNOs is actually relatively small. Ofgem's top-down benchmarks suggest that the upper quartile efficiency may only be three or four percent ahead of the average efficiency. However Ofgem's efficiency ranges for network operating costs and indirect costs considered separately are far wider. These suggest differences between "best" and "worst" DNOs that are not credible. Similarly benchmarking analysis often results in differences in efficiency between DNOs within the same ownership group, where the management structure, policies, working methods and ultimately efficiency are likely to be the same. The extent of the differences appears to be exaggerated in Ofgem's analysis compared to our findings. While we believe outsourcing differences may play some part, the degree of the observed inconsistency suggests that the benchmarking is subject to significant distortion.

The table below shows compares the difference in efficiency scores between DNOs within the same group for CN's top down methods and Ofgem's bottom up analysis.

	Top-down Benchmarks from CN's paper				Ofgem's efficiency scores from the initial proposals	
Efficiency range within DNO Group	Opex plus faults vs CSV	Totex vs Total costs composite driver	Opex plus faults incl long term capex	Average of top-down benchmarks	NOC	Indirects
CE range	1%	1%	9%	3%	3%	12%
CN range	6%	8%	12%	8%	34%	18%
EDF range	12%	4%	18%	12%	23%	21%
SP range	25%	8%	17%	16%	3%	8%
SSE range	5%	4%	3%	2%	18%	25%
WPD range	0%	0%	0%	0%	7%	0%
Average Range for all DNO groups	8%	4%	10%	7%	15%	14%

Ofgem's analysis suggests higher differences in efficiency between DNOs within the same ownership group, e.g.34% difference for CN DNOs for NOC efficiency. Taking an average of the differences across all the DNO groups it can be seen that Ofgem's methods result in an average range of 14-15% which is double the average range between DNOs from top-down benchmarks at 7%.

Errors in the Faults Benchmarking

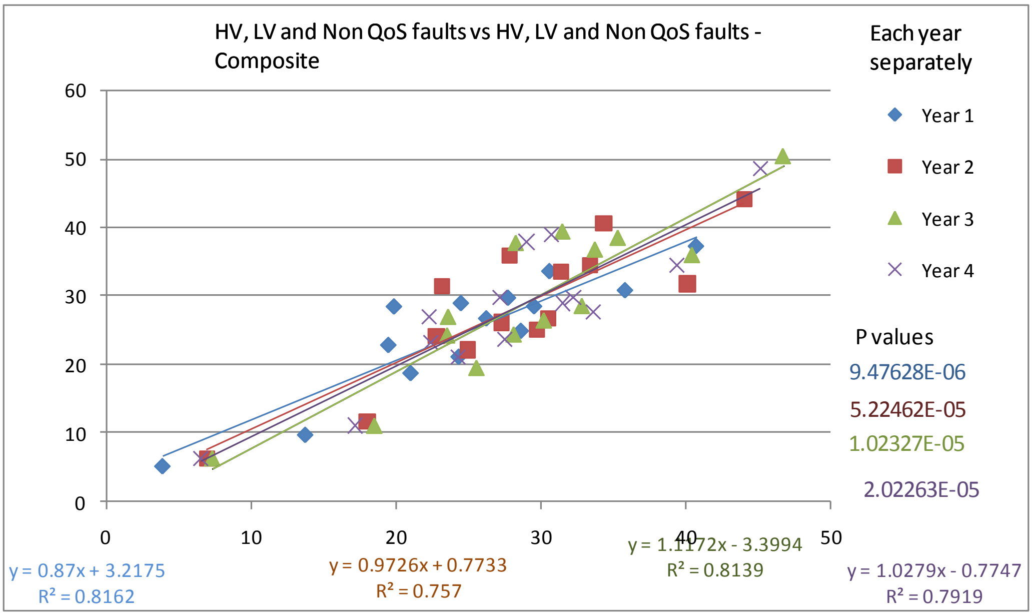
A significant numbers of faults have been excluded relating to both switchgear faults and those classified as "other". These are not reported on a consistent basis between DNOs, as for some DNOs this accounts for 32% of the total HV and LV faults but for other DNOs this is as low as 4%. This leads to considerable distortion in the driver values. As well as CN, this affects CE and WPD significantly. Details have recently been provided to Ofgem including the following proposed revisions to the data to counter this distortion:

- Faults classified as "other" should be pro-rata'd according to the split of known faults between Overhead and Underground.
- HV Switchgear faults are identifiable in the IIS Medium Term Performance report as either Overhead or Underground and should be allocated to benchmarks accordingly.
- The ratio of these faults can be used to pro-rata HV Switchgear costs between Overhead and Underground benchmarks.
- LV Switchgear fault numbers are most likely to relate to fuse failures and therefore costs and fault numbers should be reallocated. For our analysis we have added all of these to the Underground cable faults benchmark.

Additionally, EDF's fault costs are understated as they are net of cable damage recoveries. Cable recovery costs have been added to "other faults" costs rather than being allocated to LV or HV faults and so have not been included in Ofgem's calculation.

Even once the technical errors are corrected we believe that while the overhead faults benchmark is reasonable, the underground faults benchmark, which accounts for 65% of the NOC costs benchmarked, is still not sufficiently robust (see Appendix C1). Additionally the exclusion of Non Quality of Supply (Non QoS) fault costs from the benchmarking causes distortion. Therefore we propose a composite faults benchmark which includes all the costs for HV, LV & service faults, condition related cable replacement and Non QoS faults regressed using multivariate regression against the number of overhead faults, the number of underground faults and the number of customers.

This regression (below) gives statistically good and consistent results over time.



While the decision was taken previously to include severe weather events in the benchmarking, it appears that they have been excluded from both the costs and the fault numbers in this category. While this is not necessarily a problem, it does mean that the allowance setting mechanism needs to apply an uplift to account for the average level of severe weather events as well as the 1 in 20 year storm scenario.

Errors in the I&M Benchmarking

The driver values for the various DNOs are clearly wrong. CNE and CNW are similar networks with similar values for CSV and MEAV. We apply the same policies to our assets and have costs of the same order going forward suggesting that the maintenance workload created by these networks is of a similar order. Ofgem's I&M driver for CNW in the latest version of the regressions is 98,814. This is less than half that for CNE at 201,988. We believe the correct value for CNW is in the order of 190,000. Underestimating the CNW driver by a factor of two would result in a very poor efficiency score which may go some way to explaining Ofgem's NOC efficiency score for CNW of 135%. None of our network operating cost models have found CNW to have an efficiency score above 111%.

We welcome Ofgem's commitment to share the calculation of this metric with us and look forward to helping resolve this issue.

Errors in the Tree Cutting Benchmarking

The benchmark is skewed by some unrealistically high values for spans cut. Scottish Power's tree contractors work on a "spans managed" basis which may be inspection or may be cutting. It is not comparable to the spans cut values from other DNOs and should therefore be excluded as this is not comparing like with like. Similarly it appears that SSE reports a number representing the spans that are clear of trees after the tree cutting work has taken place. So, for example, if CN clears trees from a span emanating from a pole with a transformer on it, it is reported as one span cleared. We understand SSE may, if the transformer feeds say six spans, report six spans cleared even if only one span had trees that needed clearing. This would explain SSE's position as a low cost outlier with an apparent cost per span cut which is significantly lower than the range offered by tree clearance contractors.

There are sudden and significant increases in the spans cut values for 2008-9 for some DNOs that warrant further investigation. This regression is poor from a statistical point of view and the apparent unit cost changes considerably between years to such an extent that it can not be considered robust.

Calculation of NOC efficiency score

All these issues identified above suggest that the tree cutting regression is unlikely to result in efficiency estimates that are meaningful to any degree. Given that Ofgem has quite reasonably used the ESQCR re-opener benchmarking work to set the baseline costs for tree cutting, a question arises as to whether a tree cutting benchmark should be used at all.

It is not clear from a logical perspective that it is necessary or advisable to use a poor quality model for tree cutting efficiency within the NOCs efficiency score that then dictates the baseline costs for faults and I&M. We believe that tree cutting boundary issues and trade offs are more likely to be with capex activities so there seems to be no requirement to use the tree cutting benchmark at all.

Therefore we recommend that the NOCs efficiency score is calculated on the basis of the faults and I&M activities as these are the activities to which the efficiency score is applied. This combined efficiency score should be calculated by adding together the regression costs for these two models and comparing the result to the actual costs. Creating a composite driver for all NOC costs is neither required nor recommended as we believe this also creates distortion.

Indirect Costs

We believe that the regressions which consider indirect costs by groups are not sufficiently robust to be included in the mechanism to set baseline costs. While the regression for Group 2 costs is reasonable the regressions for Group 1 and Group 3 costs are not robust. Therefore the evaluation of indirect cost efficiency should use the single group regression.

Where costs are used to indicate the general workload as a driver for indirect costs these should be gross costs and not exclude customer contributions. There is some doubt as to whether this is the case.

There is some debate as to whether using the actual costs to represent the various measures of workload creates distortion. This adds greater uncertainty to the results and supports the use of the hybrid adjustment for indirect costs.

Efficiency Calculation – Problems with using the Fixed Effects Model

We do not believe the fixed effects model is appropriate for determining the efficiency estimate. We have found that where costs are increasing over time this introduces bias against larger DNOs (see Appendix C2). This method also appears to consistently produce upper quartile values that are artificially low which will have an impact on setting the baseline costs. An example of this is given below, which shows the efficiency estimates for indirect costs regressed as a single group against Ofgem's driver. The value of 88% efficiency is actually lower than achieved by the frontier performer for 2008-9 regressed alone so use of this value creates an impossible standard for DNOs.

Model – Single group of indirect costs	Upper Quartile Value
Fixed effects – fixed slope	88%
Pooled data – no fixed effects	94%
Single year regression for 2008-9	94%

We propose an alternative method which both captures the greater degree of accuracy that should result from using panel data, while still taking the impact of costs in a single year into account. This is achieved by taking the average of the estimates obtained by using the pooled model without fixed effects and the result from a regression using 2008-9 data alone.

Hybrid baseline setting mechanism for Indirect cost baselines

Our benchmarking work has found that as costs are disaggregated to lower levels, the quality of the models becomes poorer. This cherry picking effect that comes into play when using disaggregated costs to set allowances has already been demonstrated by the Gas Distribution review in 2007 where selecting upper quartile costs for each cost category resulted in allowances lower than any company could achieve. Therefore it is clearly not appropriate to set allowances at the upper quartile level, even when costs are only disaggregated into two groups as these are based on aggregating analysis at a lower level.

We acknowledge Ofgem's hybrid approach for allowance setting for network operating costs is a step in the right direction, given the low level of confidence in the results of the bottom-up benchmarking. We believe the hybrid approach should also be applied to indirect costs as:

- the benchmarking models for Group 1 indirect costs and Group 3 indirect costs are no more robust than the bottom-up models for network operating costs;
- it reduces the impact of the cherry picking issue; and
- it is more appropriate given the overall lack of robustness inherent with bottom-up benchmarks.

Incorporation of Top-down Results in determining baseline costs

We believe that top-down results need to be used to determine baseline costs, either on their own or in conjunction with bottom-up analysis, because these methods do not suffer the same problems of poor model specification and account for reporting differences and trade-offs between cost categories. Ofgem's opex efficiency assessment in the initial proposals strongly reflects the bottom-up analysis. This bottom-up analysis is currently affected by data errors and flawed methodology, but even with corrections and improvements the regressions may still be poor. This is due to the greater impact of errors, reporting differences and boundary issues that is inherent in bottom-up analysis and is mirrored in the results of the statistical tests that assess the model quality. This puts undue reliance on analysis which is not sufficiently robust and ignores the opportunity to take a more rounded view.

We believe that the best method of incorporating the results from the top-down benchmarking is simply to take an average of the top-down scores to determine the efficiency of each DNO and the upper quartile value (96%). This overall efficiency score can then be applied to both NOC and Indirect costs to determine the value of baseline costs for 2010-11 at upper quartile efficiency.

We are conscious that Ofgem is unlikely to accept significant changes at this stage and so we also offer another proposal which combines both top-down and bottom-up efficiency scores.

Firstly the NOC and indirect efficiency scores should be calculated having made the corrections to the data and methodology proposed earlier in this section. We believe that after the necessary corrections have been made to the benchmarks and methodology, initial efficiency scores for NOC and for indirects should be calculated as shown below.

$$Initial\ NOC\ efficiency = \frac{Actual\ I\&M\ costs + Actual\ fault\ costs + Non\ QoS\ costs}{I\&M\ regression\ costs + combined\ fault\ model\ regression\ costs}$$

Regression costs are averaged from pooled model + 2008-9 regression

$$Initial\ indirect\ efficiency = efficiency\ score\ from\ single\ group\ of\ indirects\ using\ average\ of\ scores\ from\ pooled\ model + 2008 - 9\ regression$$

The average efficiency from the six top-down models should be calculated as a representative top-down score.

$$Overall\ top\ down\ efficiency\ score = average\ of\ efficiency\ scores\ from\ 6\ top\ down\ models$$

This score should then be averaged with that for Network Operating Costs and also separately averaged with that for Indirect Costs.

$$Amended\ NOC\ efficiency = \frac{Overall\ top\ down\ efficiency\ score + Initial\ NOC\ efficiency}{2}$$

$$\text{Amended indirect efficiency} = \frac{\text{Overall top down efficiency score} + \text{Initial indirect efficiency}}{2}$$

The upper quartile values for use in the hybrid baseline setting mechanism should then be determined from the amended sets of efficiency scores.

We believe that while the averaging process will counteract some of the errors and distortions in the benchmarking of indirect costs it will not be totally resolved and therefore the hybrid allowance mechanism is also appropriate for indirect costs.

Cost indexation and adjustments

Indexation of costs during DPCR5

Having determined appropriate cost baselines using 2008-9 costs Ofgem then makes adjustments for each year to the end of DPCR5. Our comments on these adjustments are given below.

RPE and Efficiency assumption	Ofgem appears to have misinterpreted the efficiency improvements referred to in the FBPQ as supporting the high end of CEPA's range at 1% per annum. While CN has included significant efficiency improvements within its plan, these do not represent frontier shift. Similarly WPD did not assert that it could achieve 1% year on year improvement from its current cost position. Its stance was that it is able to offset the 1% increase in labour costs that arises from its incremental salary structure by making a balancing 1% efficiency improvement.
Indexation of indirect costs to follow capex work	We support this but it should be 1:3 not 1:4 ratio as proposed to Ofgem by the ENA workgroup.
NOC increase driver	We support this adjustment. Given the large disparity between the DNOs' views of appropriate capex spend and the values within the initial proposals it is sensible to assume that the network will become more expensive to inspect, maintain and repair.

Miscellaneous

Cost Adjustments

Our comments on the various adjustments that have been applied to the costs for benchmarking is given below.

Labour & contractor rates	We continue to believe this is only appropriate for the region within the M25. This adjustment is wrongly applied to indirect expenditure categories that do not need to be undertaken within DNOs' operational areas.
Recognition of indirect costs	We do not believe it is appropriate to simply remove these costs for a single DNO group. While we understand that work is ongoing to enable a greater comparison of the indirect costs associated with contractors costs, it is unlikely that the information recently gathered by the DNOs will be of sufficient quality and comparability to enable a robust cost adjustment to be determined. For this reason as an alternative way of normalising for the impact of different degrees of outsourcing we recommend a totex top-down

	benchmarking method be used as one of the top-down benchmarks.
Cable replacement	We agree that it is necessary to include cable replacement costs and lengths within the faults analysis to overcome the identified reporting inconsistencies. However, we found that in multivariate regression of fault costs against both fault numbers and the length of cable replaced was not a statistically significant driver therefore the 18% weighting on this driver is excessive. Ofgem's analysis is likely to have given incorrect results due to the problem in the fault numbers used.
Urbanity	Work is still ongoing to create a cross-check for this adjustment. We believe that it is appropriate to make a general adjustment for urbanity as we are aware of the increased cost of working around the central business district of Birmingham. However, further analysis is required to determine if the specific adjustment for LPN that relates to the increased number of cable tunnels and sites with forced ventilation is appropriate.
Interconnected network / Sparcity	We believe that these factors may also affect cost but are not able to determine whether the adjustment is set at an appropriate level.

Excluded costs

We generally agree with the reasoning behind the costs which have been excluded with the exception of Non QoS faults as described earlier.

There appears to be a problem with the costs for rising mains and laterals in that it removes costs reported in table NL9, where they were not included.

We generally agree with the method used to determine the baseline costs for the excluded costs, however there are certain exceptions which are outlined below.

Dismantlement

Ofgem has suggested that they would use an average of actual costs. However this value is currently artificially low as the sum of the values for 2007-8 to 2009-10 are divided by five, rather than three. As dismantlement costs were not originally separated in the RRP for 2005-6 and 2006-7 these have not been separately identified. We will endeavour to separate these costs, but it is important that this calculation reflects the data available and makes a reasonable estimate.

Unmetered electricity

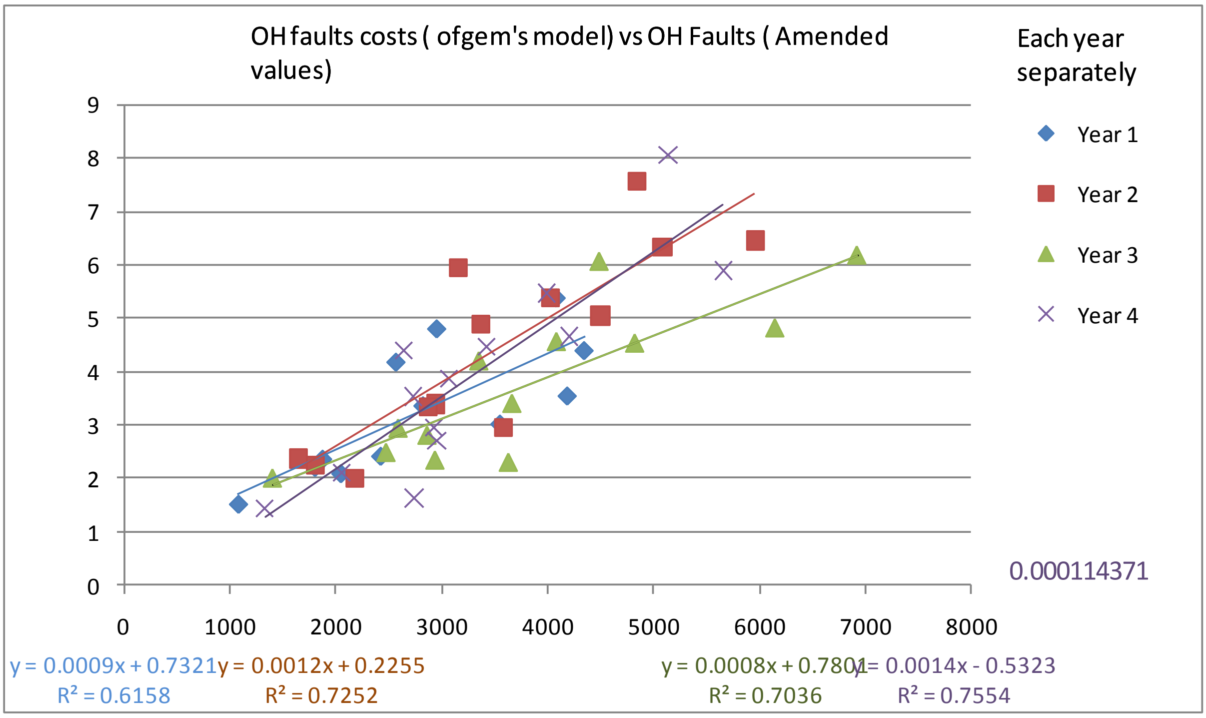
Ofgem's view of the costs of this for DPCR5 is £6.6m lower than CN has suggested in the FBPQ. We are not aware of the reasoning for this reduction and ask Ofgem to provide this.

Appendix 2

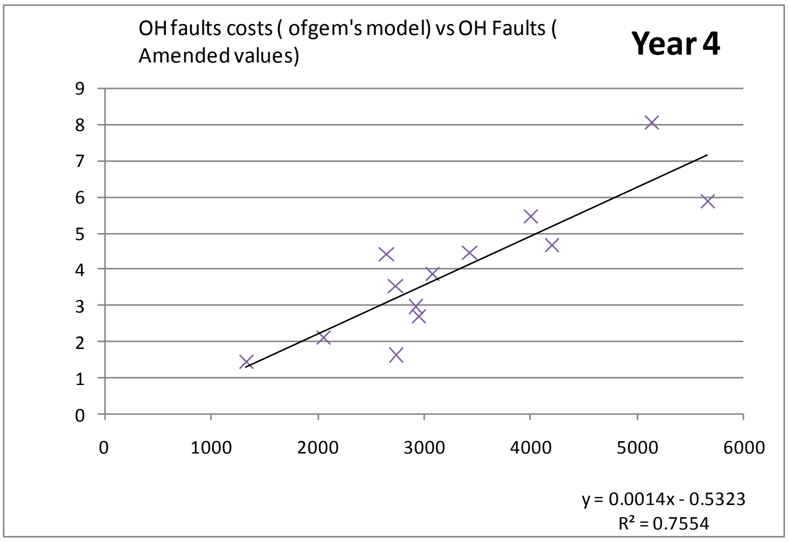
Detailed evaluation of the operational cost assessment regressions

OH Faults

Ofgem's fault numbers ignore both switchgear faults and those classified as "other". These need to be allocated to OH and UG categories as per CN's proposal. The chart below shows the regression including the additional fault numbers.



EDF LPN is excluded as it does not have an overhead network. The model gives consistently high R squared values each year. While the slope for 2007-8 is markedly different, the slope for 2008-9 is consistent with the general pattern.

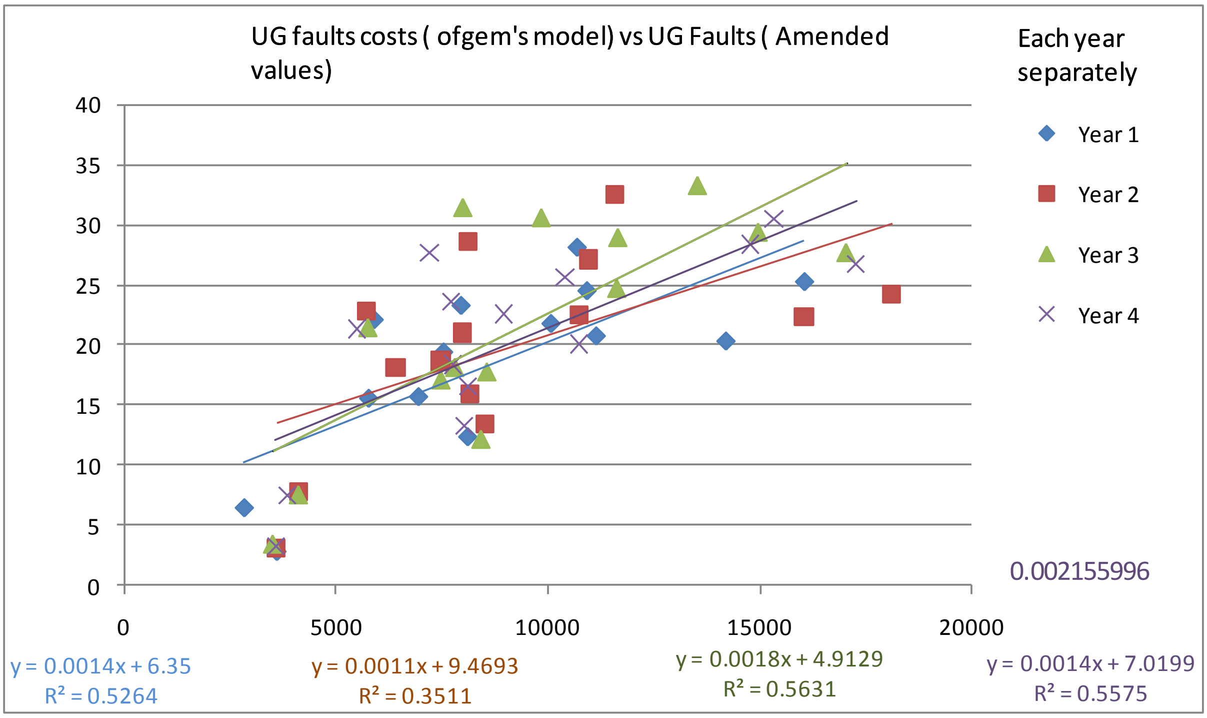


Considering only 2008-9 (left), ENW is a low cost DNO but otherwise the regression appears to be representative. Using pooled and single year models suggests that both CNE and CNW have efficiency scores of 111%.

UG Faults

UG faults account for 85% of the benchmarked fault costs and 65% of the network operating costs benchmarked; therefore the quality of this regression has a large impact on Ofgem's evaluation of NOC efficiency and efficiency as a whole. Ofgem's fault numbers exclude a large number of faults that have been classified as "other" or as switchgear faults. Most of these should be allocated to UG faults and therefore their omission has a significant effect on the UG efficiency value. Fault numbers have a dominant weighting within the composite driver, which should actually place even less weighting on the length of cable replaced.

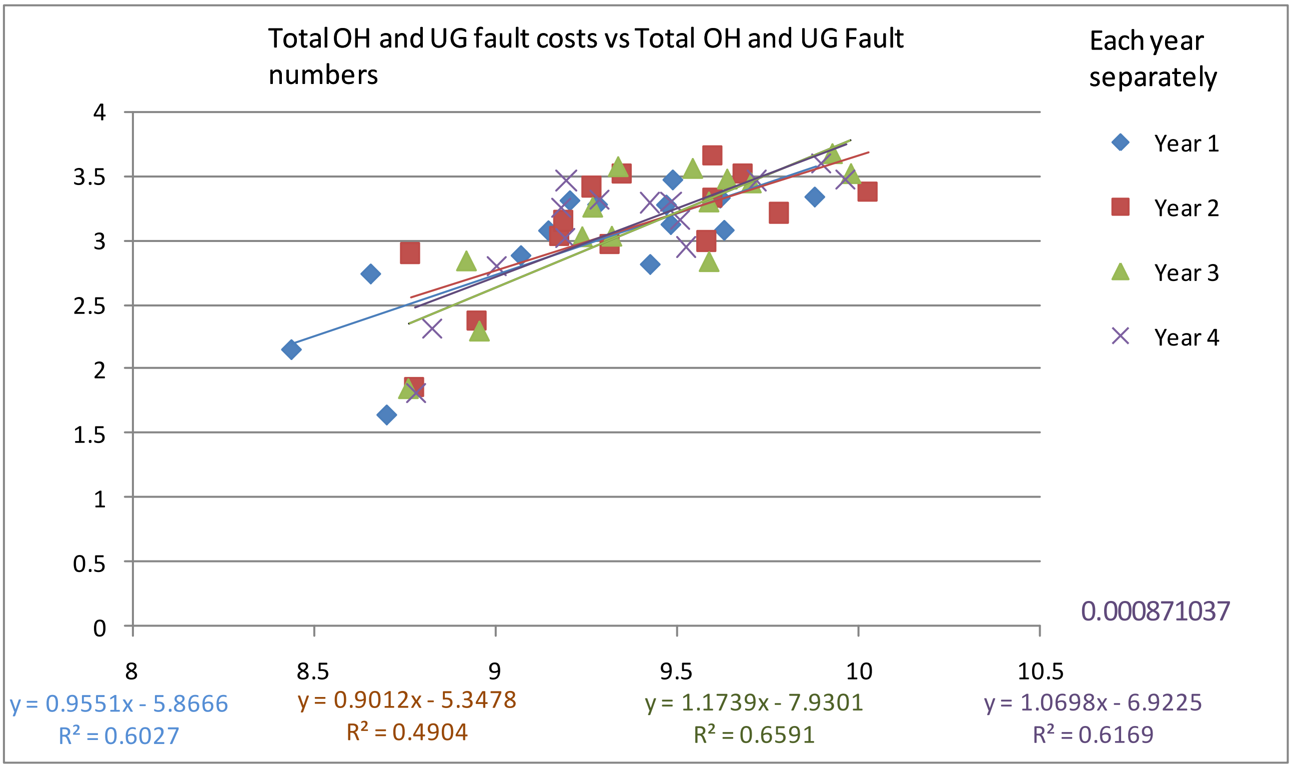
The regression below uses the amended fault numbers and shows that even with the correction of this significant error the model has consistently low R squared values. This suggests that using the output of this model to estimate efficiency is inappropriate.



We have investigated the impact of including the length of cable replaced within a multivariate panel data model. Due to the low weighting ascribed to the cable replacement element this does not significantly alter the efficiency scores or the R squared values.

Combined Faults Model

It is possible that there are differences in fault reporting that would result in costs and fault numbers not being consistently defined as overhead or underground. Faults classified as "other" is an example of a source of error. Therefore as a cross check we have combined all HV and LV faults and costs together as a single driver. This gives a reasonable model which is better quality than that for UG faults but not as good as that for OH faults.



Efficiency scores	Pooled	Fixed effects	Single year
CE NEDL	103.6%	103.3%	103.6%
CE YEDL	100.3%	100.0%	99.5%
CN East	97.8%	97.5%	97.2%
CN West	104.7%	104.4%	104.3%
EDF EPN	99.0%	98.7%	97.9%
EDF LPN	103.1%	102.7%	103.4%
EDF SPN	119.0%	118.6%	118.9%
SP Distribution	110.1%	109.8%	109.9%
SP Manweb	112.3%	112.0%	112.3%
SSE Hydro	73.2%	72.9%	73.8%
SSE Southern	93.9%	93.6%	92.9%
ENW	103.0%	102.6%	102.4%
WPD S Wales	91.4%	91.1%	92.1%
WPD S West	90.8%	90.5%	90.3%

	Pooled	Fixed	Single
max	1.1896	1.1857	1.1891
lower quartile	1.0442	1.0410	1.0408
upper quartile	0.9486	0.9459	0.9395
frontier	0.7316	0.7290	0.7377

This suggests CN efficiency in terms of faults is around average overall and that upper quartile efficiency is at around 94%.

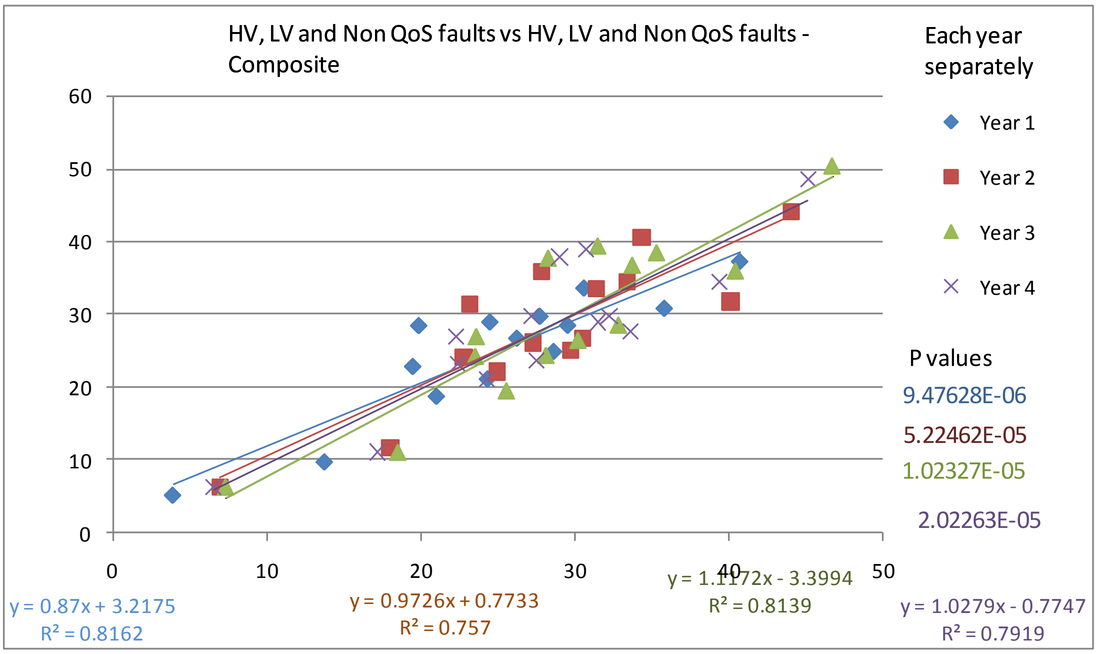
Again this is significantly different to Ofgem's value, but this combined model is statistically more robust than the UG faults element in Ofgem's assessment.

Combined Faults Model including Non Q of S faults

Non Q of S fault costs have been excluded from the benchmarking of network operating costs. CN was seen to be very low cost in this area whereas other DNOs have very high costs in this category. We agree that there is not a sufficiently robust benchmark for these costs; however we also believe that there are cost allocation differences between DNOs that may be affecting the faults benchmarking that will not be picked up if non Q of S faults are simply excluded.

We have created an overall faults benchmark (below) that includes all HV and LV faults plus non Q of S faults and uses a composite driver of customer numbers and total HV & LV fault numbers.

This produces the most stable results year on year of any of the faults benchmarks and suggests that including non Q of S fault costs reduces the apparent differences in efficiency between the DNOs. We believe that this model should be used in place of the separate OH and UG benchmarks.



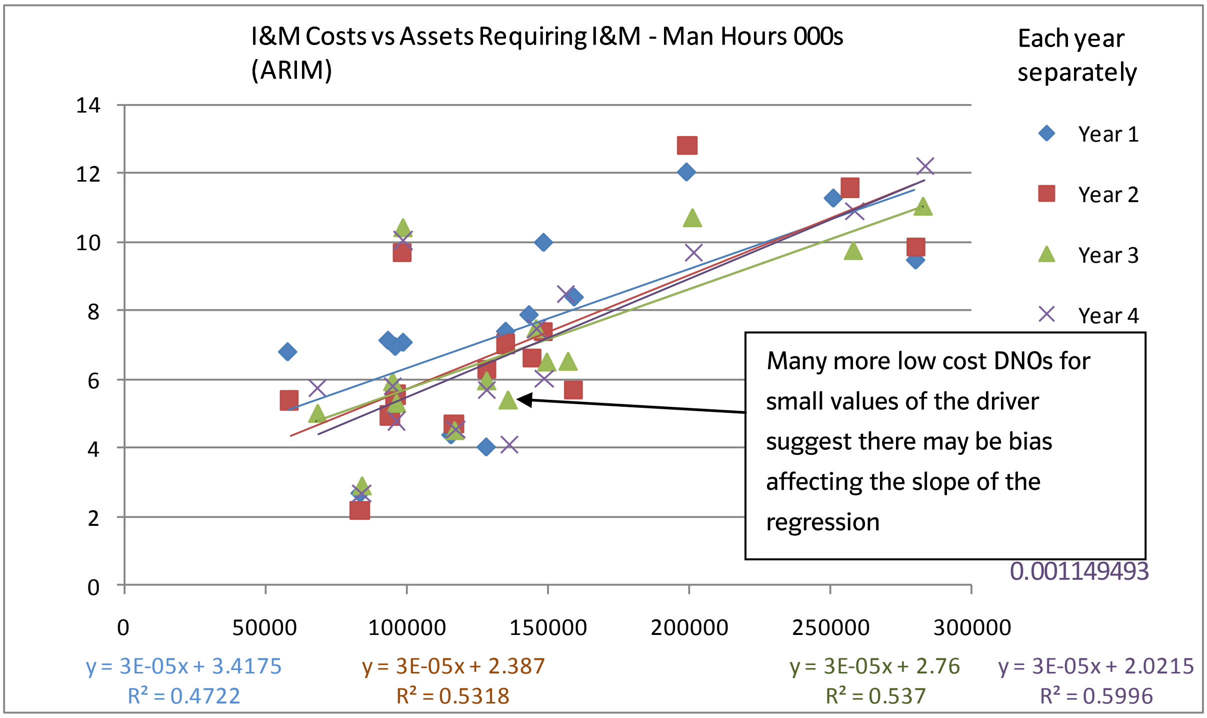
Efficiency	Pooled	Fixed	Single
CE NEDL	104%	104%	105%
CE YEDL	127%	127%	127%
CN East	82%	82%	82%
CN West	93%	93%	92%
EDF EPN	108%	108%	106%
EDF LPN	86%	86%	86%
EDF SPN	131%	131%	130%
SP	110%	110%	110%
SP Manweb	121%	121%	122%
SSE Hydro	98%	98%	107%
SSE Southern	88%	88%	87%
ENW	92%	92%	92%
WPD S Wales	65%	65%	66%
WPD S West	87%	87%	87%

	Pooled	Fixed	Single
Max	1.3060	1.3060	1.3044
lower	1.0912	1.0912	1.0909
upper	0.8700	0.8700	0.8708
Frontier	0.6494	0.6494	0.6608

These results show CN's efficiency to be better than average once non Q of S faults have been taken into account.

I&M

We do not understand the derivation of Ofgem's I&M driver, but it is clear that the value for CNW is wrong as it suggests that the network requires only half the manhours of work that CNE network requires. Ofgem's regression (shown below) is therefore significantly flawed and the efficiency values are incorrect. This may partly explain why as well has having a low R squared value the model fails tests for normality and heteroskedasticity, confirming that the model is unacceptable.

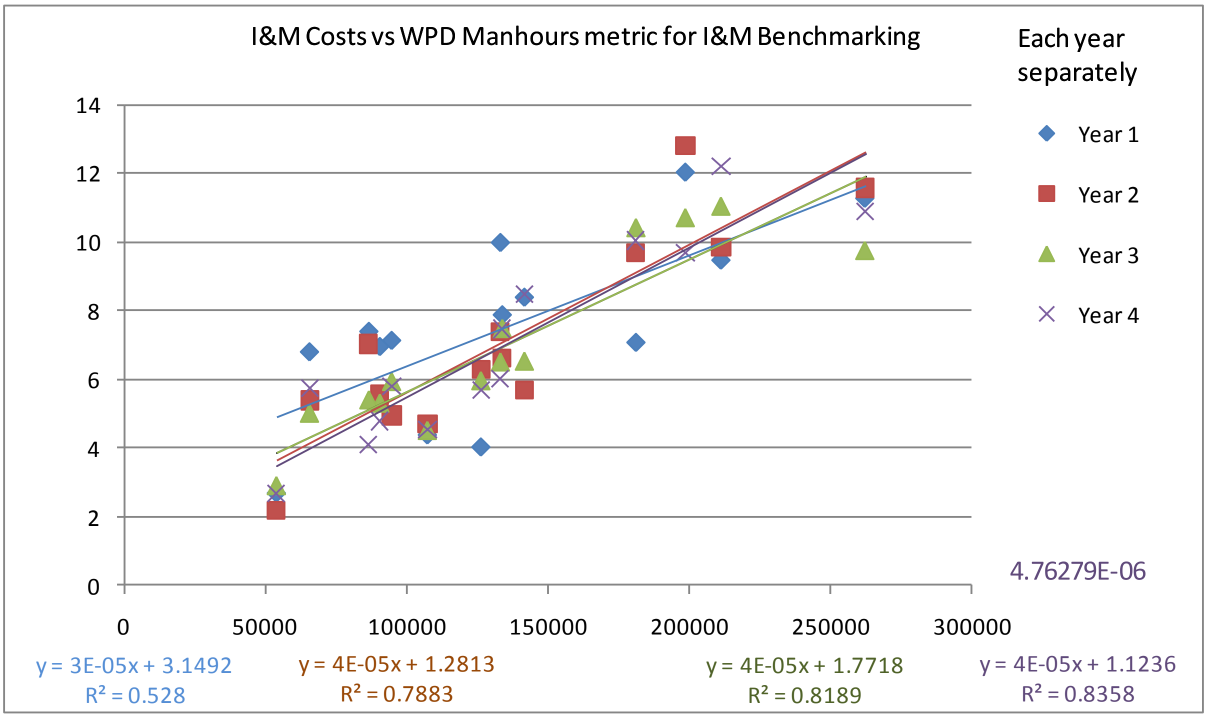


	Pooled	Fixed effects	Single year	Average
CE NEDL	71.3%	73.5%	74.5%	73.1%
CE YEDL	84.5%	87.0%	87.6%	86.4%
CN East	107.7%	110.0%	107.8%	108.5%
CN West	174.5%	180.5%	185.1%	180.0%
EDF EPN	100.9%	102.6%	99.4%	101.0%
EDF LPN	119.4%	124.3%	130.8%	124.8%
EDF SPN	81.8%	83.9%	83.8%	83.1%
SP Distribution	84.1%	87.0%	89.3%	86.8%
SP Manweb	58.9%	60.6%	60.8%	60.1%
SSE Hydro	50.3%	52.2%	54.2%	52.2%
SSE Southern	105.6%	107.3%	103.4%	105.4%
ENW	112.2%	115.1%	114.5%	113.9%
WPD S Wales	102.6%	106.2%	109.1%	106.0%
WPD S West	102.9%	105.6%	105.5%	104.7%
max	174.5%	180.5%	185.1%	
lower quartile	107.2%	109.3%	108.8%	
upper quartile	82.3%	84.7%	84.7%	
frontier	50.3%	52.2%	54.2%	

CNE and CNW operate the same policies under the same management with similar networks. The astonishing difference in efficiency is due to the error in Ofgem's driver.

The fact that this has not been corrected undermines confidence in Ofgem's analysis as a whole.

To determine the likely form of the regression with a realistic manhours metric, a value developed by WPD is used as a driver. The regression (below) appears significantly more realistic and aside from year 1 gives consistently good regressions.



	Pooled	Fixed effects	Single year	Average
CE NEDL	74.5%	76.5%	78.0%	76.3%
CE YEDL	83.0%	84.9%	85.4%	84.4%
CN East	100.1%	101.8%	99.0%	100.3%
CN West	111.8%	113.8%	111.4%	112.3%
EDF EPN	89.3%	90.4%	86.7%	88.8%
EDF LPN	129.7%	134.4%	144.0%	136.0%
EDF SPN	84.4%	86.3%	86.4%	85.7%
SP Distribution	88.3%	91.0%	94.2%	91.2%
SP Manweb	77.7%	80.1%	83.3%	80.4%
SSE Hydro	67.2%	70.0%	76.6%	71.3%
SSE Southern	119.9%	121.8%	118.1%	120.0%
ENW	114.3%	116.8%	116.3%	115.8%
WPD S Wales	103.9%	106.9%	110.2%	107.0%
WPD S West	104.7%	107.0%	107.1%	106.3%
max	129.7%	134.4%	144.0%	
lower quartile	110.0%	112.1%	111.1%	
upper quartile	83.3%	85.3%	85.6%	
frontier	67.2%	70.0%	76.6%	

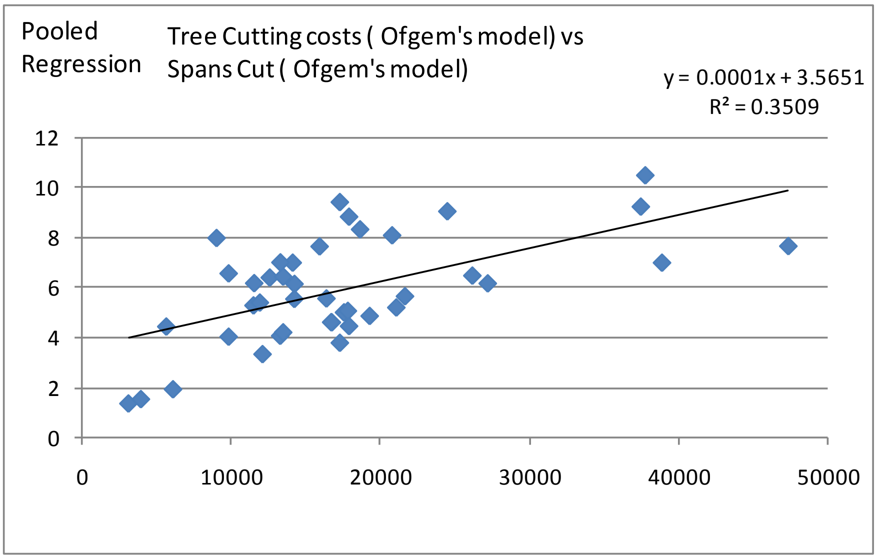
This regression suggests that the efficiencies for CNE and CNW are similar and no worse than 113%.

We believe, from work with EDF, that the WPD manhours metric can be improved further by including manhours relating to substations. This then reduces the apparent spread of efficiency.

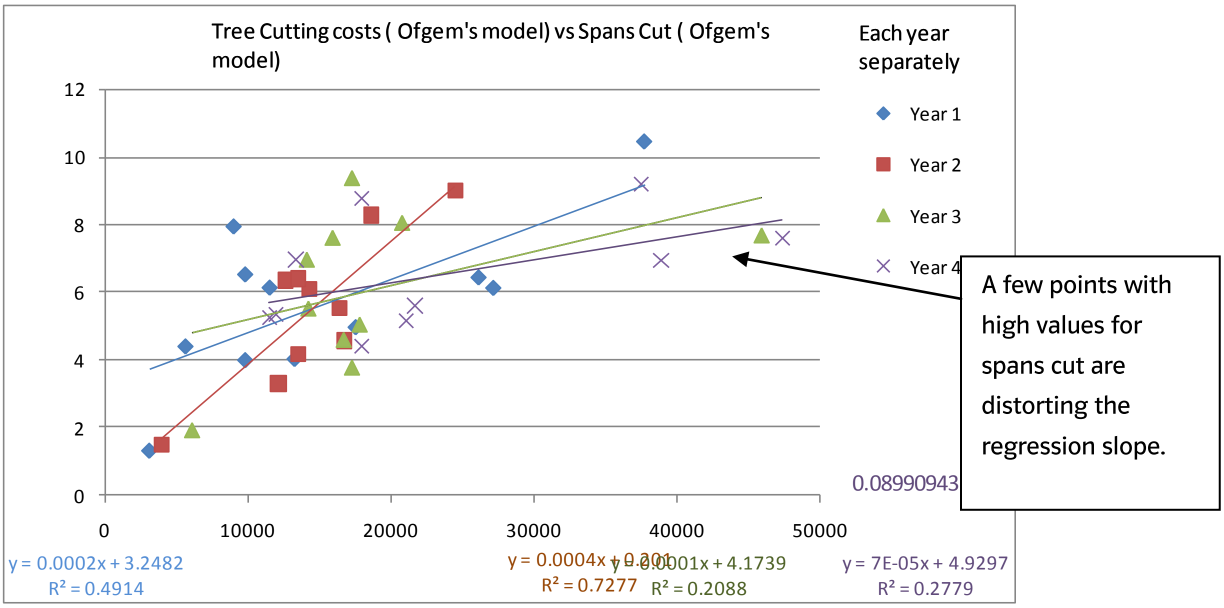
Tree cutting

Ofgem's model fails tests for model specification and heteroskedasticity yet fails to identify outliers. As our analysis has clearly identified outliers in SP and SSE, we feel this casts doubt on Ofgem's work.

In the model below EDF LPN are excluded from the regression as they do not require a tree cutting program. Both Scottish Power DNOs must be excluded as they report spans managed rather than spans cut – an entirely different and non-comparable number.



The pooled regression (left) excludes only EDF LPN and SP yet still has a very poor R squared value. It also shows that the points tend to cluster away from the regression line rather than near it. The problems with this regression become more apparent when each year is analysed separately (below).

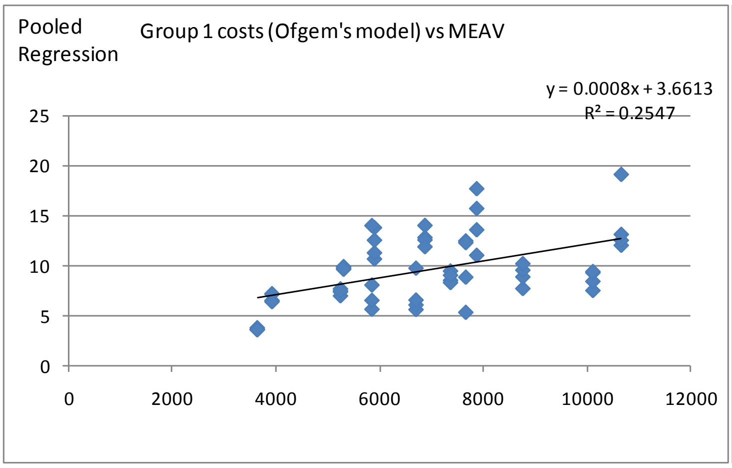


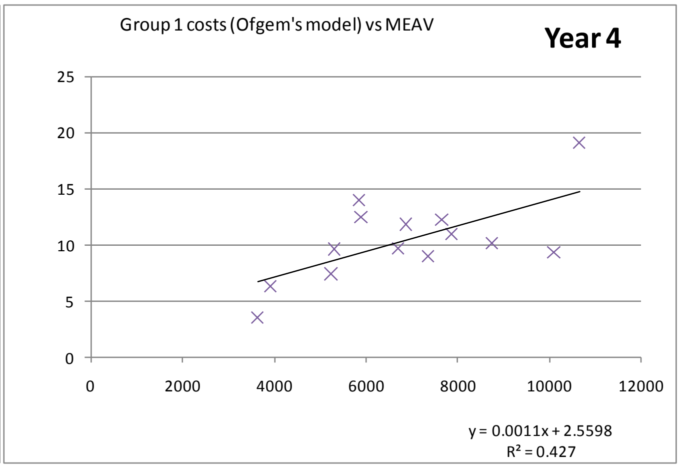
The only year with a reasonable R squared value is that for year 2 (2006-7) which shows a significantly higher slope, and lower intercept than that of other years. The R squared value for 2008-9 is 0.28 which suggests that this relationship is totally unreliable as a predictor of efficiency. Pooling data can improve regression accuracy but it cannot fix data issues as significant as the ones affecting this model. We have recently found out that SSE is reporting spans cut on a basis which would tend to overestimate the spans cut which would explain its position as a low cost outlier with an apparent cost per span cut which is not anywhere near the range offered by tree clearance contractors. Given the very poor quality of this model it is inappropriate to include this in the benchmarking to determine an efficiency score to apply to faults and I&M. Ofgem appears to have set cost baselines for DPCR5 for tree cutting with reference to the ESQCR re-opener benchmarking which seems to avoid the issues around the 2008-9 data in particular. Though we would like to clarify the details of the values used, this seems a more appropriate approach.

Group 1 Indirect Costs

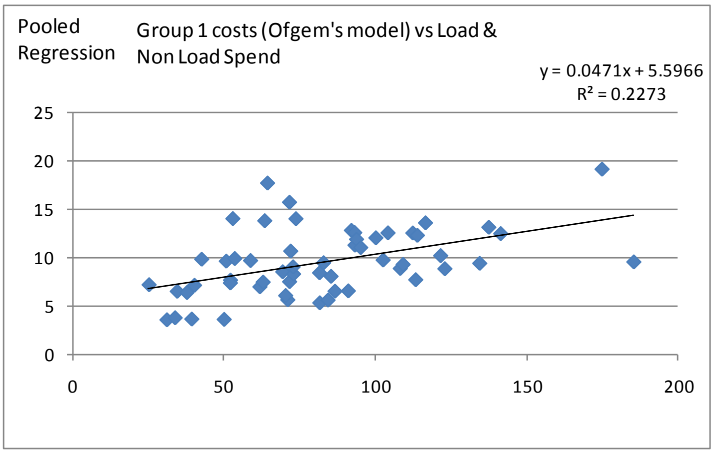
These are regressed against a composite of MEAV and Capex costs. The driver has been set with equal proportions rather than being left for determination within the multivariate regression. However our analysis showed that removing the weighting restrictions did not improve the quality of the model. We believe that the quality of this regression may be limited by failing to include EMCS costs in this category.

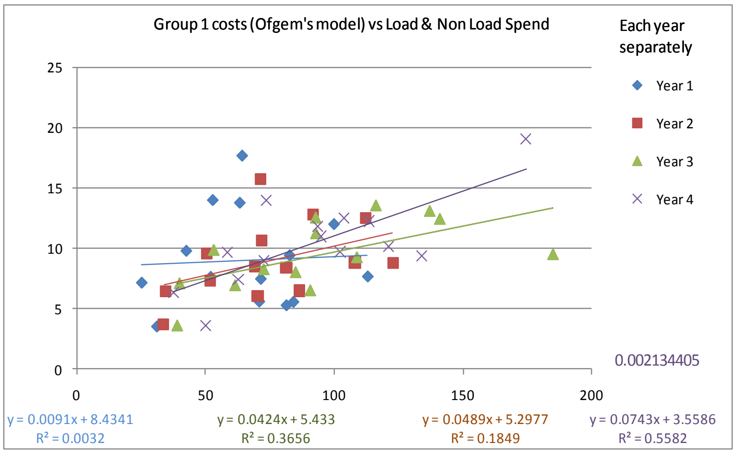
Considering the drivers separately it can be seen that MEAV as a single driver gives low R squared values (0.25). The relationship is significantly different for 2005-6 while the other years have results that are similar to each other, suggesting that the inclusion of the 2005-6 data is likely to worsen the quality of the model. Looking at 2008-9 on its own produces a better R squared than the pooled model, though even this value is still low.





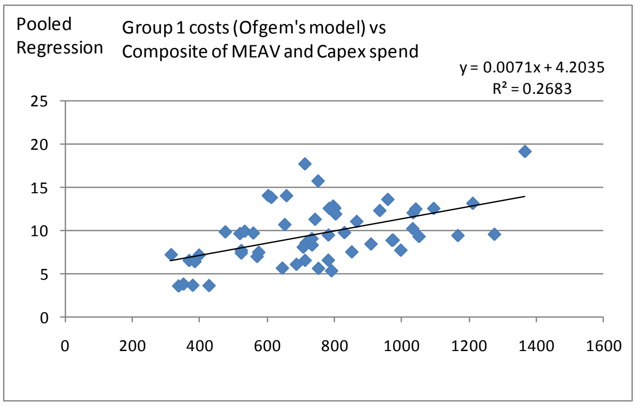
Considering the Load and Non Load Spend as an alternative driver also gives very poor results for the pooled model with significant variation between years. Once again 2008-9 appears to give the "best" result and suggests a steeper slope than the pooled regression.

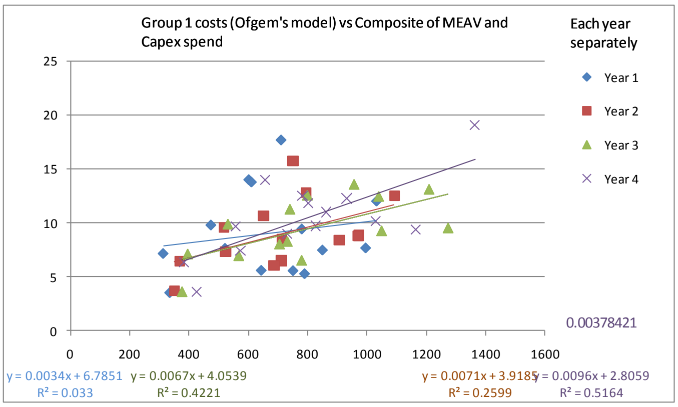




Using the Composite Driver

Given that neither of the drivers within the composite gives a very strong relationship with the costs it is not surprising then that the model using the composite variable (below) also gives poor R squared values, that are well below levels expect for robust analysis. Analysing the data separately for each year suggests that the costs in 2008-9 are higher than the pooled average and shows the year 2008-9 as having the "best" fit of all the individual years.





	Pooled	Fixed effects	Single year
max	1.579473812	1.476489578	1.533548984
lower quartile	1.196086465	1.122695775	1.168928234
upper quartile	0.90419357	0.840150872	0.89562838
frontier	0.496087796	0.456464929	0.518908146

The efficiency values from this model suggest that

- The best DNO, which has an efficiency of 52%, would have costs that are around **half** the average costs.
- The worst DNO, which has an efficiency of 1.53%, has costs that are very much above average costs.

This huge range is far more likely to result from the way in which these costs are accounted for than any actual differences in efficiency.

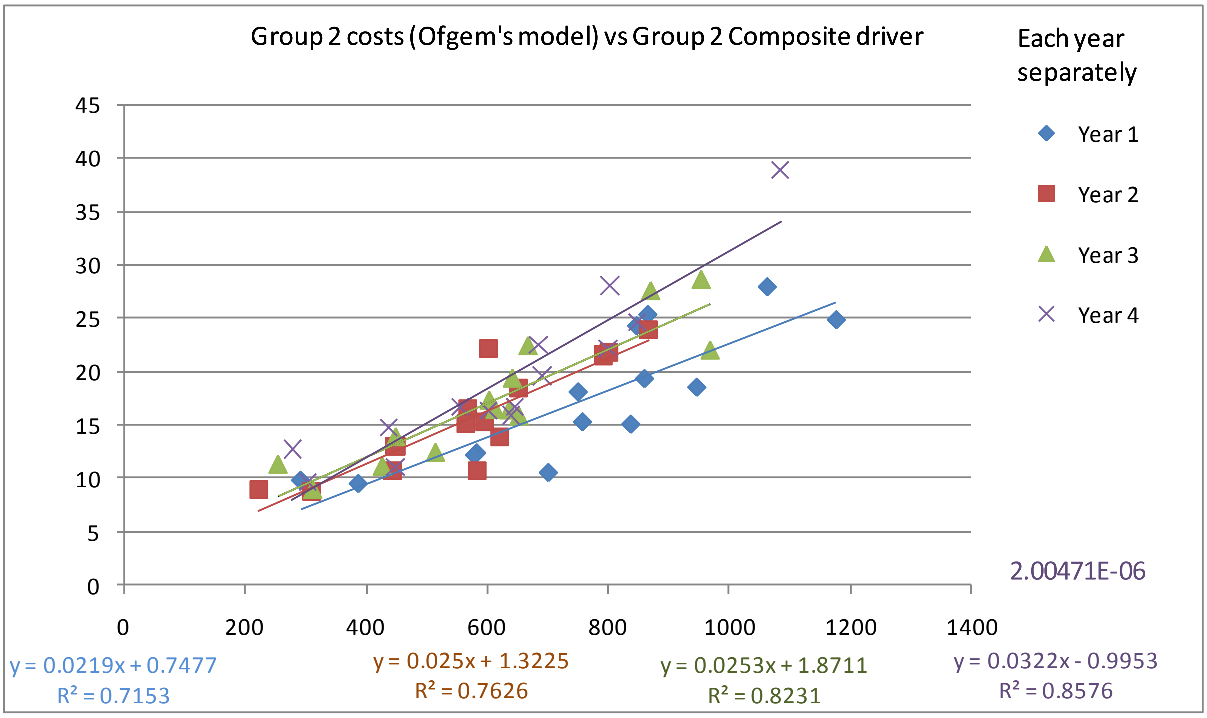
Given the very low value of R squared and the fact that this model fails the test for heteroskedasticity, the model can not be considered to be robust. Therefore we should not include the results of this model, which may also mean that it is not possible to include the results for Group 2 and Group 3 indirects either. Instead, the regression of all indirects as a single group should be used.

Group 2 Indirect costs

Here Ofgem has used a composite of MEAV and direct costs (labour and contractors) as a metric to represent total workload. This is weighted with 54% on direct costs and 46% to MEAV suggesting that the weightings have not been forced.

We agree that these are logical drivers and this regression gives credible results as shown below where each year is considered separately.

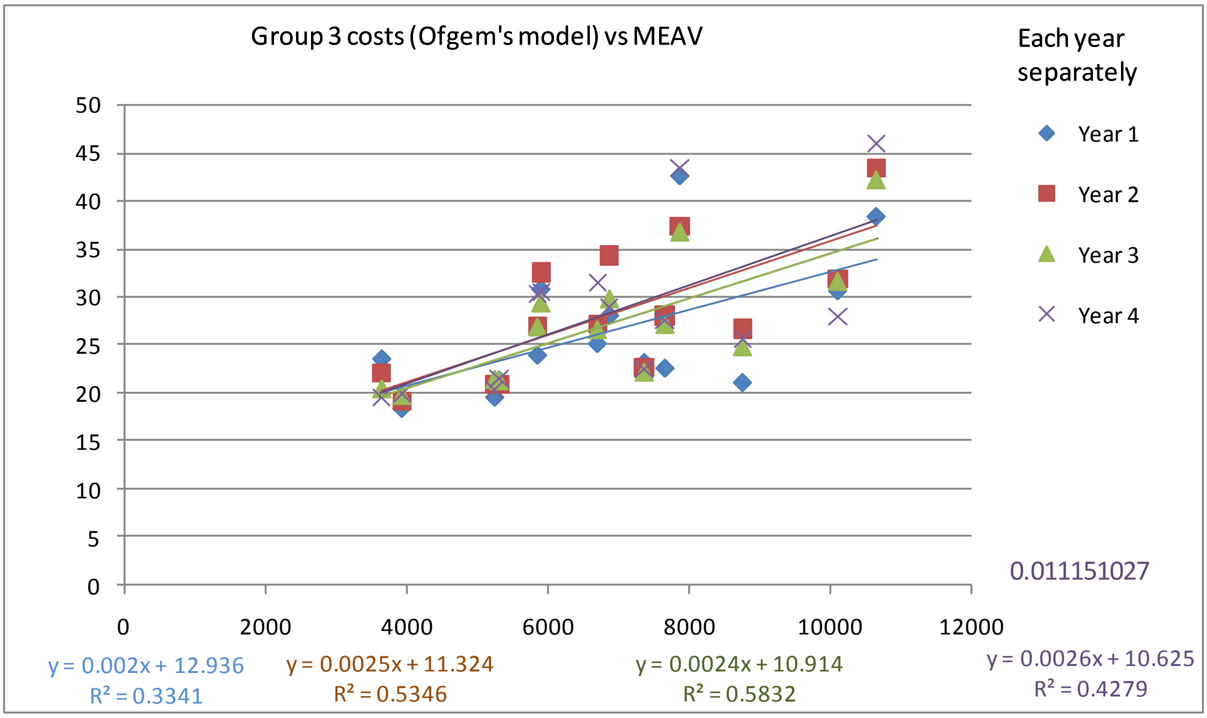
Ofgem's composite – MEAV and Direct costs (Labour and contractors)



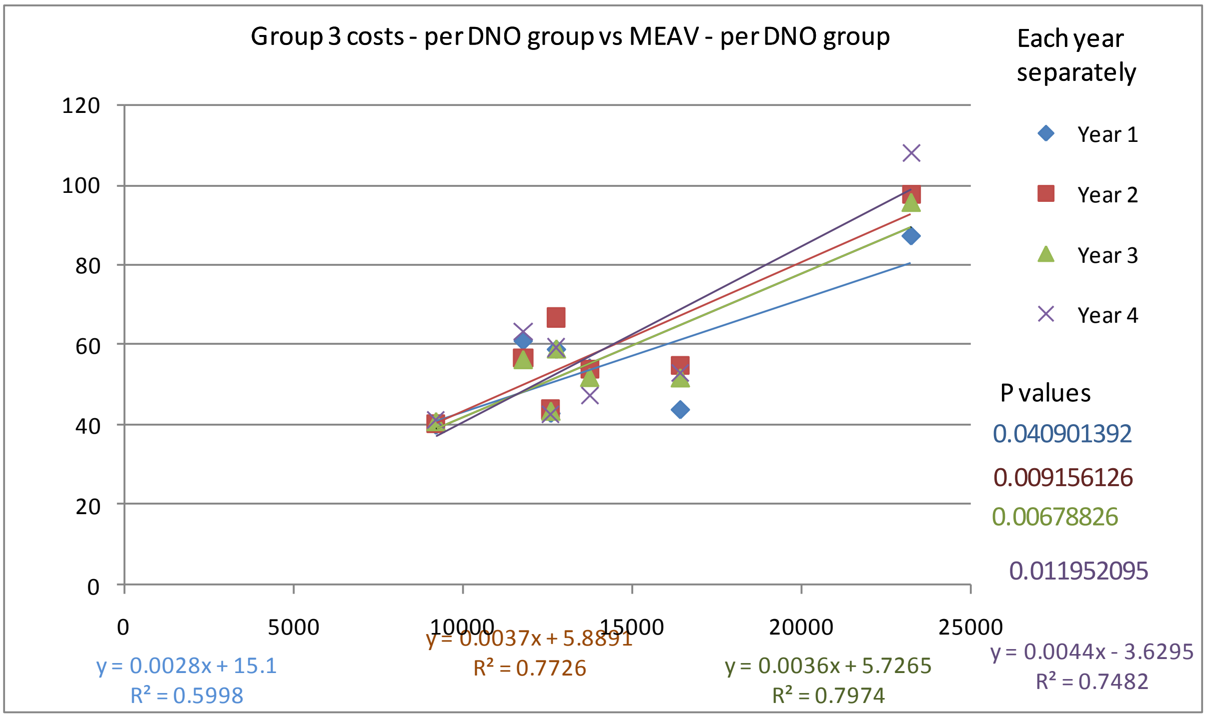
It can be seen that the slope of the regression increases over time. Given that 2008-9 has the highest R squared value of any of the years it should be considered as a more accurate reflection of current costs than will be given by the fixed effects model. The fixed effects model will result in an unreasonably low upper quartile value and bias the efficiency scores against larger DNOs and in favour of smaller DNOs.

While we agree that this regression is not of poor quality, to carry out analysis of indirect costs by group means that the poor quality regressions for group 1 and group 3 must be combined with this regression. Unfortunately they are so poor that this renders comparison by individual groups unacceptable.

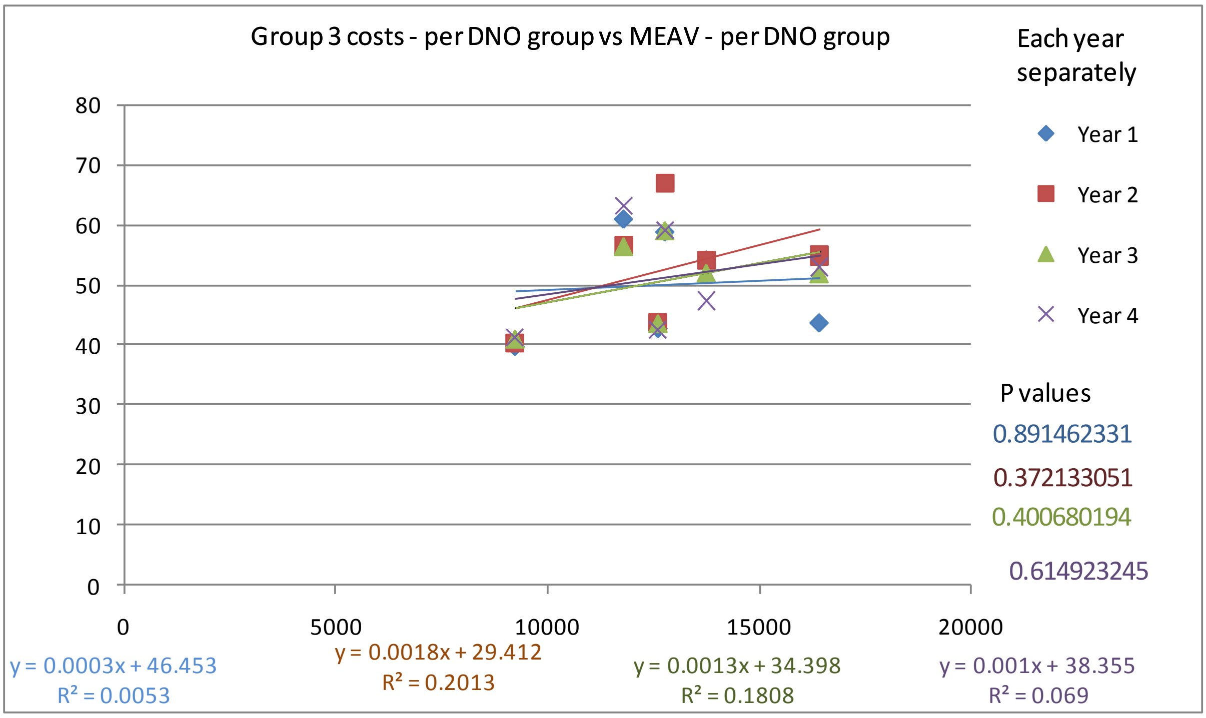
Group 3 Indirect Costs



Ofgem's analysis regresses these costs against MEAV, which produces low R squared values for each year. This model is shown in a linear form above and suggests R squared values are significantly lower than that quoted by Ofgem of 0.67. Ofgem's higher value results from carrying out the regression on a DNO group basis which as there are fewer data points will have a higher r squared value. This can be seen on the group regression below.



This shows a strong relationship for group 3 on the DNO basis. However, this relationship is caused almost entirely by the EDF DNOs on the far right of the chart. Their extreme positioning gives them a lot of influence over the position of the regression line and without them included the strong pattern completely disappears as shown below.

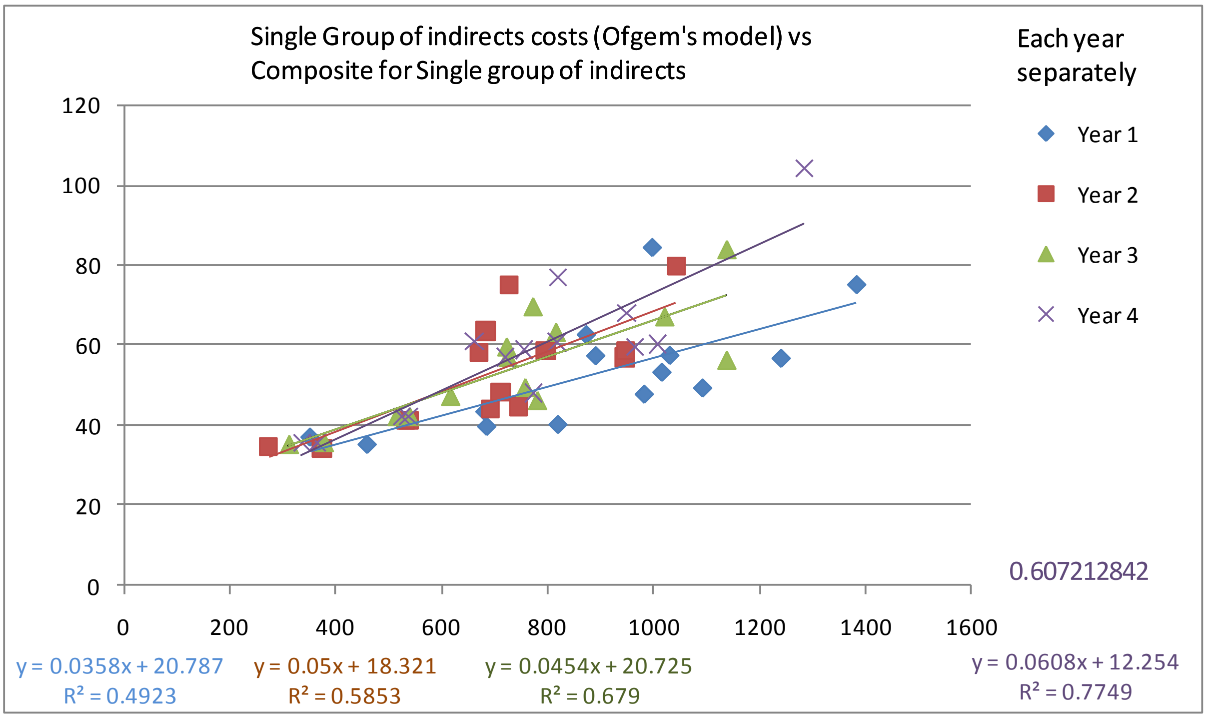


We believe that ENW's singleton status should be recognised within the benchmarking methodology and that regression of group 3 costs by DNO group would initially appear to provide a solution to this. However, we consider the problems with the quality of the regression are significant. In addition to the regression quality being entirely dependent on EDF as shown above, the model fails the test for model specification. Ofgem does not identify any DNOs as outliers whereas visual inspection would suggest that many points are extremely far away from the regression lines. Once again the increase in regression slope over time suggests that a fixed effects model is not appropriate. We do not believe the model is sufficiently robust to be used in combination with other regressions to determine efficiency.

Single group of Indirect Costs

Ofgem has regressed the indirect costs against a composite variable of MEAV and Direct Costs which is similar to that used for group 2 costs but with the drivers forced to 50% weighting each.

We do not agree with Ofgem's reasoning behind restricting the weightings, however in this case the efficiency results are similar. However the fixed effects model exaggerates the position of the upper quartile as outlined in Appendix C2.



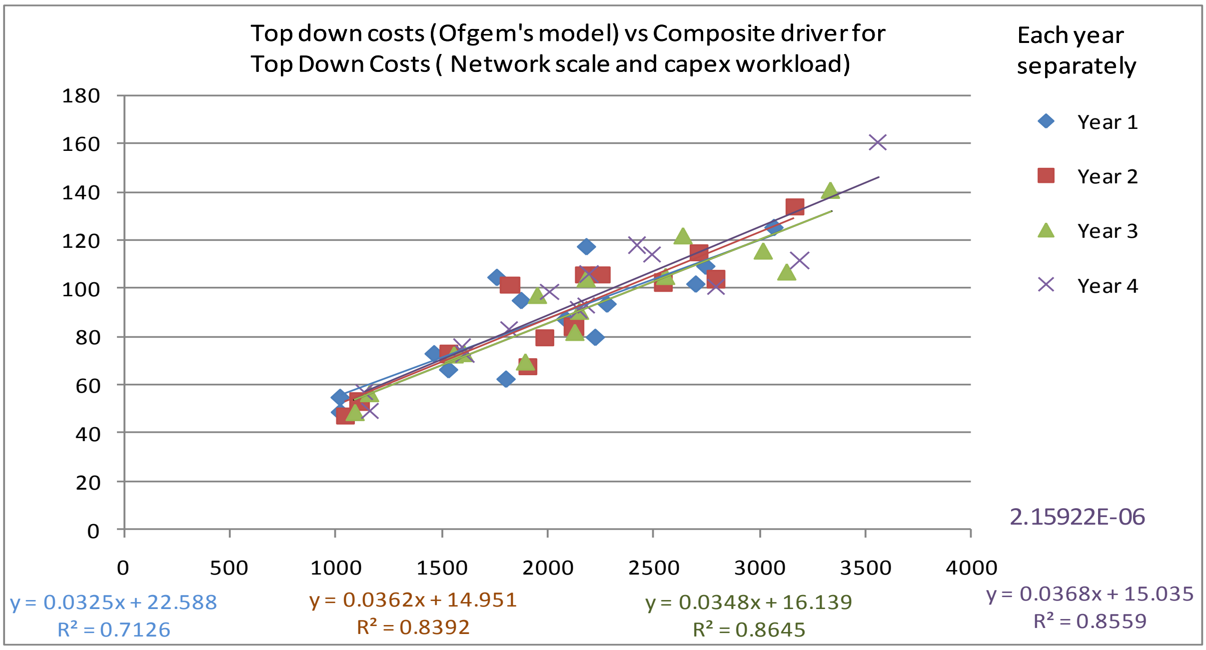
		Fixed	Single year
Max	1.319582492	1.253800048	1.280414235
lower	1.158441535	1.068786171	1.077025297
upper	0.942111822	0.884242904	0.936694438
Frontier	0.833776678	0.812368871	0.909409733

The frontier value for a single year 2008-9 is lower than the upper quartile value Ofgem are using, i.e. they are using a value that is not attainable by the frontier company in 2008-9.

As 2008-9 has a significantly better R squared value than other years it would make sense to put more emphasis on the result from this year and to abandon the fixed effects model and instead take an average of the efficiencies from the 2008-9 single year model and the pooled model.

Top-down Regression – Core version

The core top-down regression uses a composite of MEAV and the labour and contractors elements of Load related and non load related capex. We can see that both drivers are significant individually. Ofgem's regression is pictured below (without logging the costs or driver) and shows consistently high R squared values and consistent positions of the regression lines when years are regressed separately. This regression does not fail any of statistical tests in contrast to the bottom-up models. Given that this is the best model from a logical perspective and from the test results this and other top-down benchmarks need to have a primary role in setting cost baselines.



	Pooled	Fixed effects	Single year
CE NEDL	0.9848	0.9637	0.9757
CE YEDL	0.9817	0.9648	0.9651
CN East	0.8711	0.8589	0.8512
CN West	1.0861	1.0694	1.0639
EDF EPN	1.1274	1.1144	1.0966
EDF LPN	1.0244	1.0043	1.0114
EDF SPN	1.1215	1.1025	1.1019
SP Distribution	0.9851	0.9683	0.9680
SP Manweb	1.1195	1.0992	1.1024
SSE Hydro	0.8490	0.8261	0.8502
SSE Southern	0.8617	0.8509	0.8399
ENW	1.1515	1.1334	1.1287
WPD S Wales	0.9931	0.9659	0.9952
WPD S West	1.0322	1.0099	1.0230
	Pooled	**Fixed effects**	**Single year**
max	1.1515	1.1334	1.1287
lower quartile	1.1111	1.0917	1.0884
upper quartile	0.9825	0.9640	0.9658
frontier	0.8490	0.8261	0.8399

The efficiency of CN as a whole is approaching the industry average. The upper quartile values at around 0.97 suggest that in reality there is significantly less difference between the best and worst performers than is suggested by the lower level regressions. The upper quartile value of 0.84 for indirect costs as used in the initial proposals is unlikely in this context.

Using multivariate regression shows that the fixed effects model depresses the upper quartile value and should not be used.

Ofgem's view including average non load capex also gives consistently good results and suggests an upper quartile value of 0.98. We believe all the top-down benchmarks demonstrate a level of robustness that is not seen in the bottom-up analysis and should therefore be used either as a substitute for or in conjunction with the bottom-up benchmarks

Appendix 3

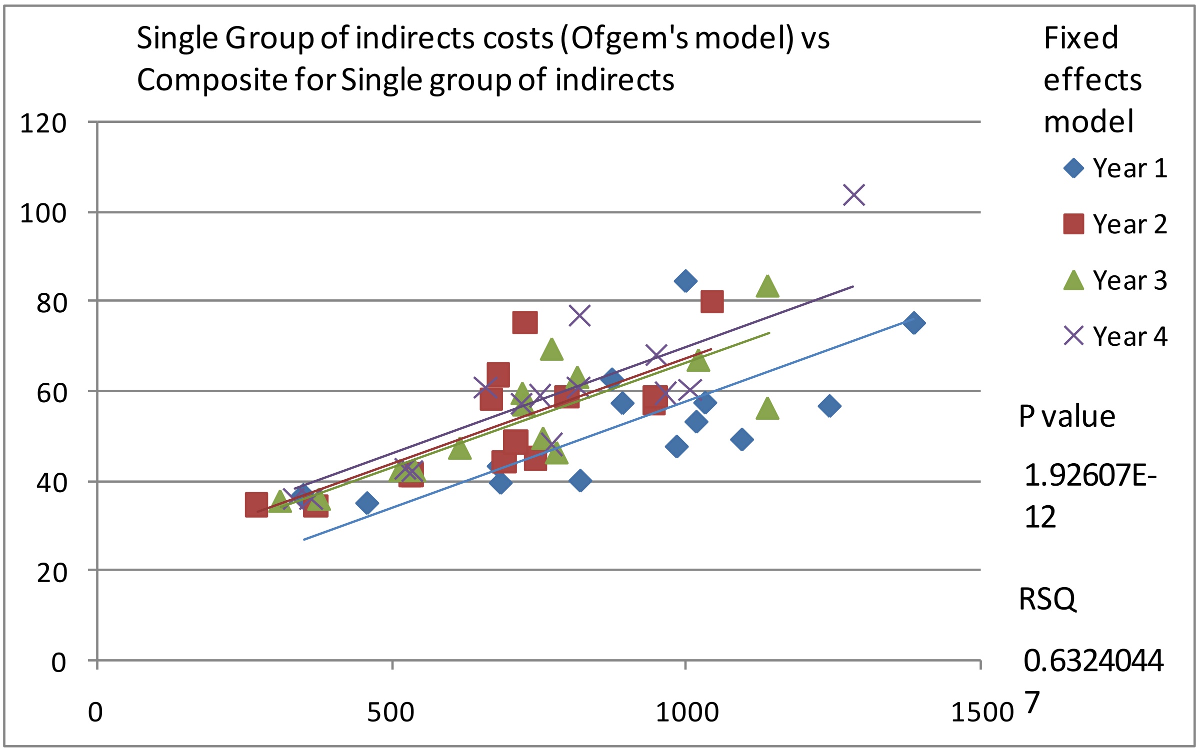
Evaluation of Overall Methodology Robustness

Fixed effects model

In creating a cost model for benchmarking it is reasonable to include a mechanism to cater for the changes that affect costs for each year, even after inflation has been removed. To this end Ofgem proposes using a fixed effects model, which adds in a year specific constant. However we believe that this creates bias against larger DNOs in the case where the variable costs of work are increasing over time and that a better way of taking year specific effects into account while still gaining the benefits of a pooled data model is to use the average efficiency scores of a pooled model and the regression for the current year. The reasoning supporting our conclusion is given below.

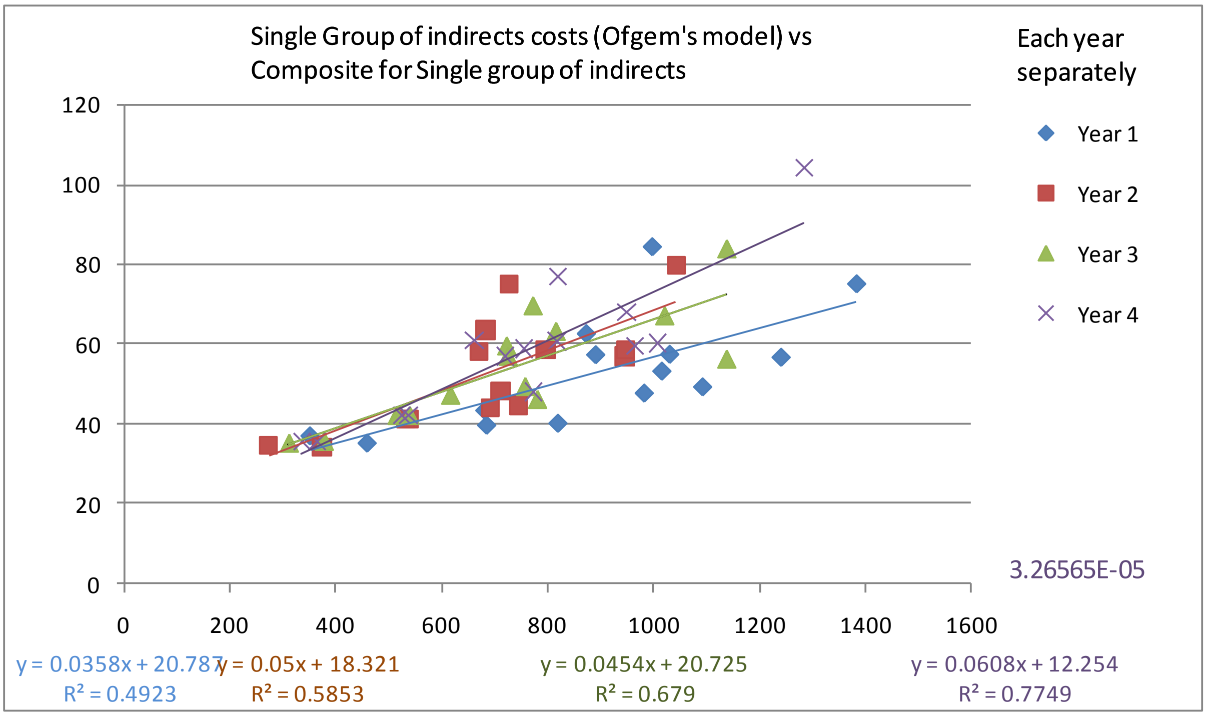
Ofgem's fixed effects model has the effect of moving the intercept of the regression while the slope remains the same for each year. However this suggests a model whereby fixed costs change but variable costs remain the same, which is not likely to represent the true way in which costs are influenced over time. This approach is flawed and biases the results against those DNOs with a larger value for the cost driver as demonstrated in the following paragraphs.

The chart below shows the fixed effects model for indirect costs considered as a single group against its composite driver as provided by Ofgem.

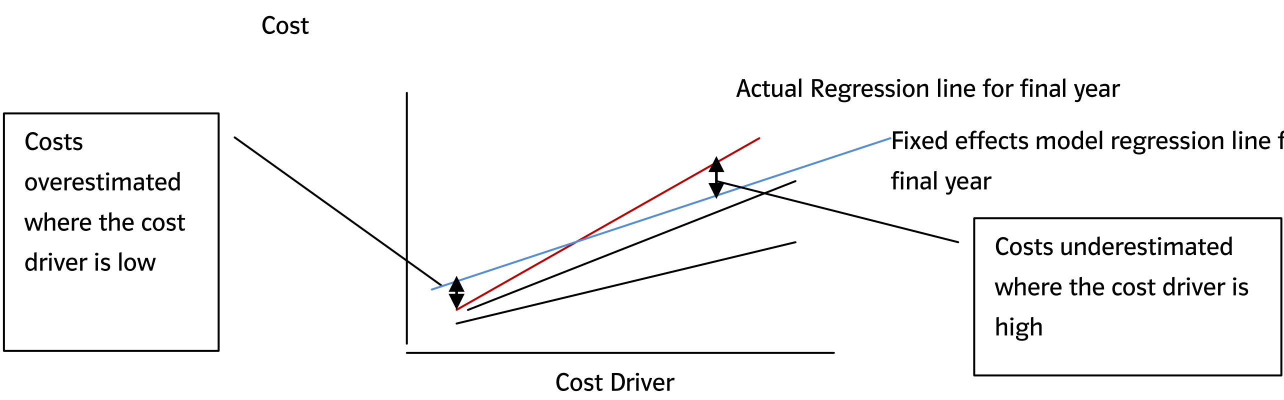


The fixed effects model is constructed so that the lines are all parallel. It can be seen that the year 1 value is lower than years 2 to 4 which are very close together.

A different picture emerges if rather than using the fixed effects model, each year is considered individually as shown below.

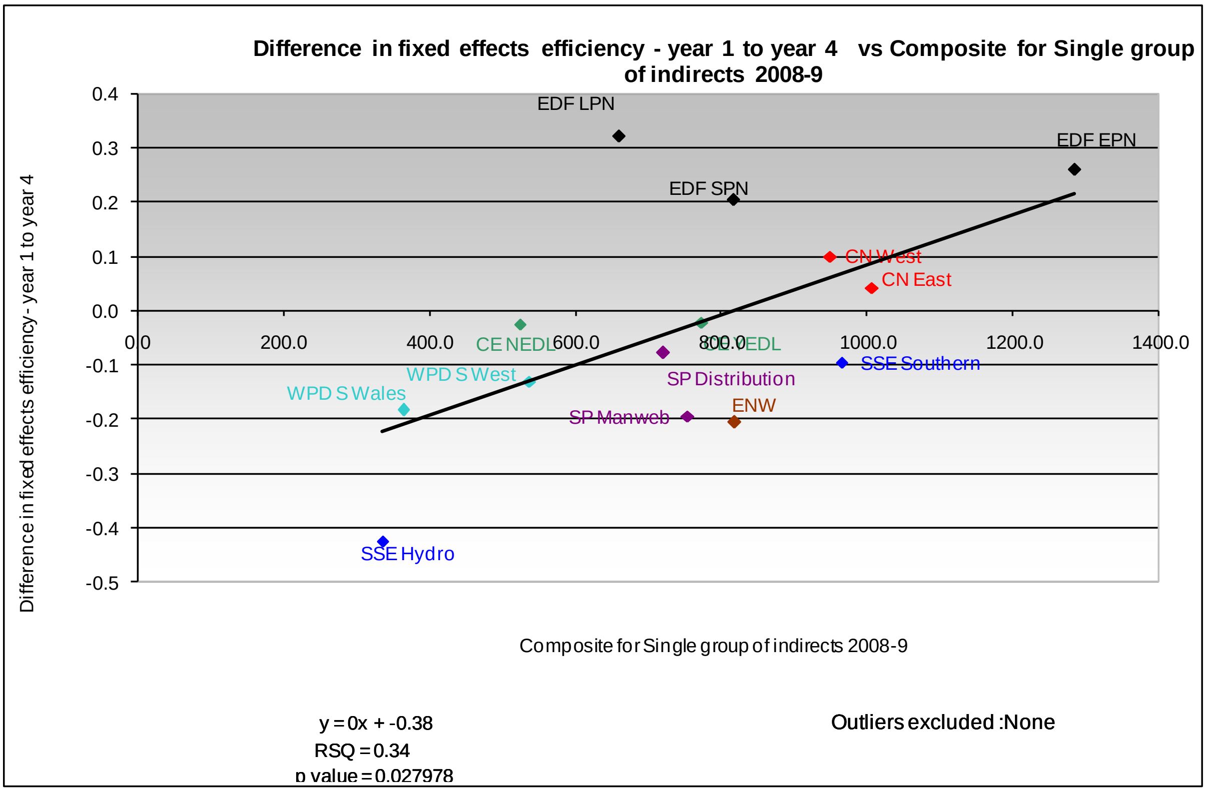


As nothing is fixed by this model, it shows variation in both the intercept and the slope. However, while year 2 and 3 are very similar, it shows a general trend of tilting (ie: changing slope), rather than parallel lines. This suggests that the unit cost of activities is increasing over time. Forcing the slope to remain at a fixed value will result in a slope that is artificially low, but with an intercept that is artificially high. This means that the efficiencies calculated using the fixed effects model are inherently biased against the DNOs with a higher value of the cost driver as shown below for the situation where variable costs are increasing over time.



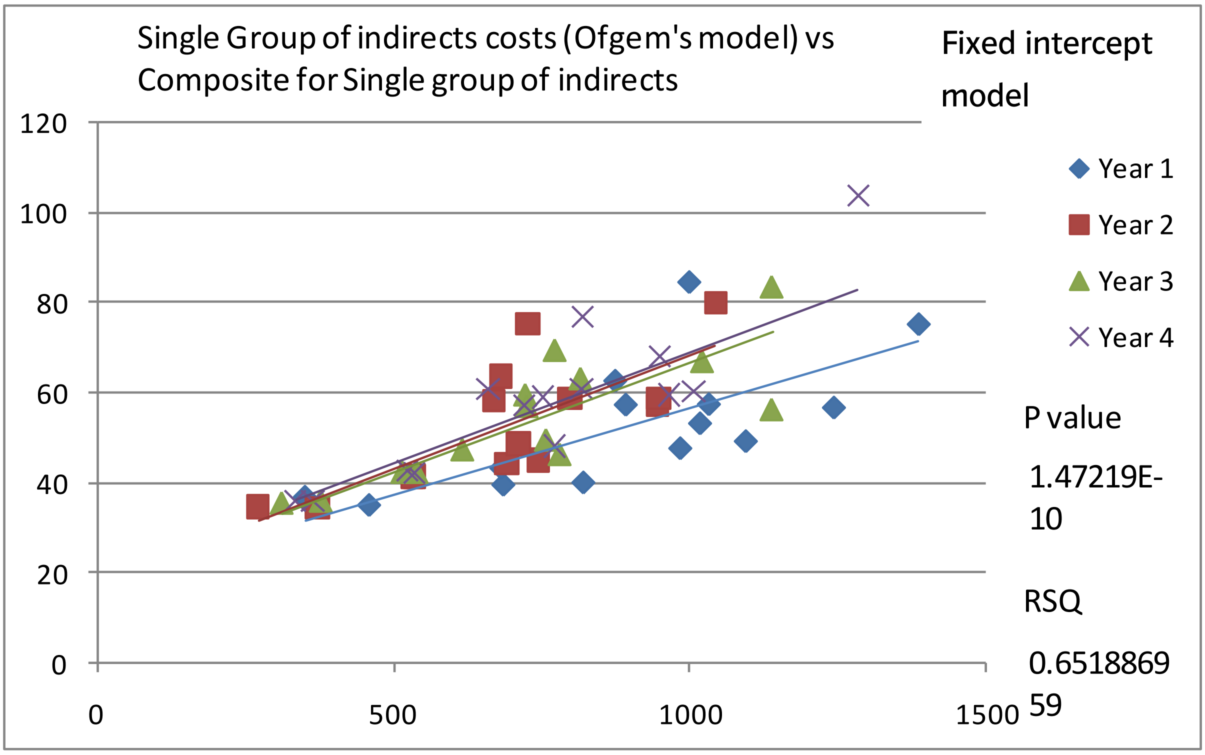
This can be seen in practice when we examine the change in efficiency scores obtained by the fixed effects model for indirect costs. The efficiency scores for year 1 and year 4 are compared and the difference in efficiency scores is then plotted against the cost driver in question.

This shows that "larger" DNOs efficiencies appear to worsen over time whereas those for smaller DNOs improve.

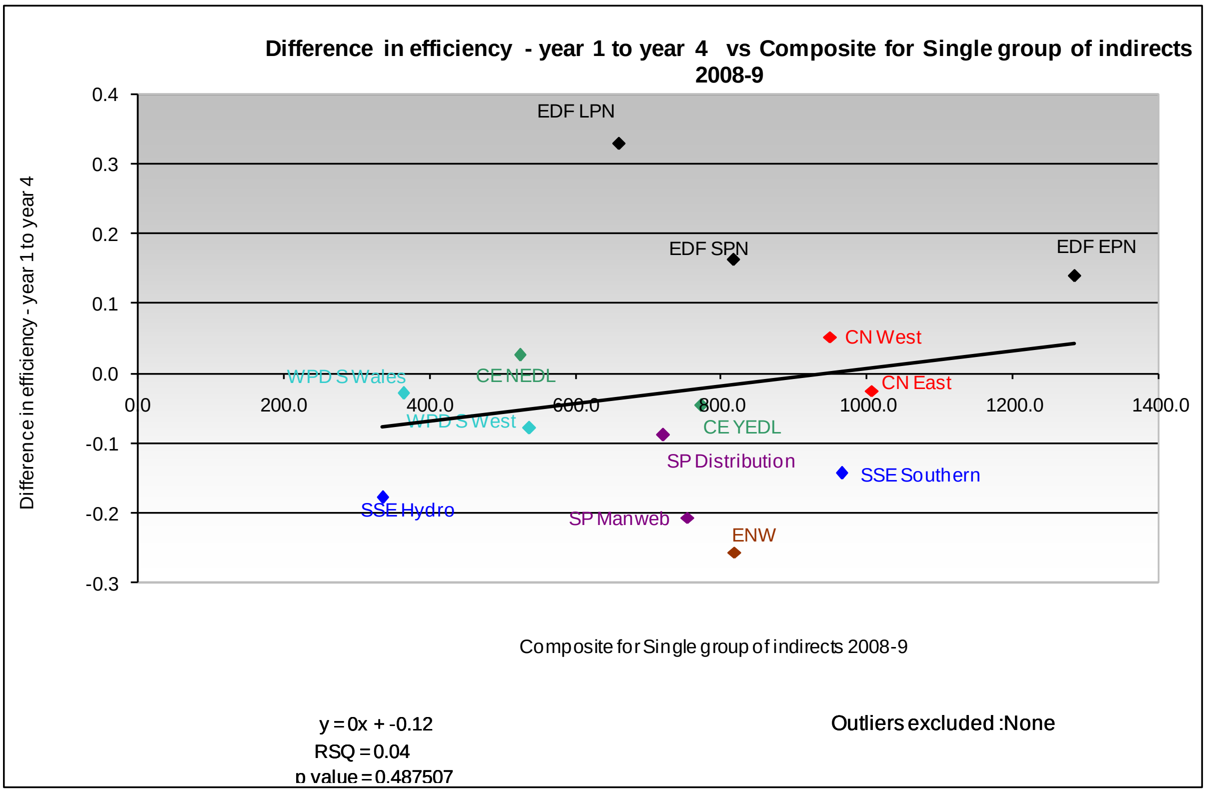


The low p value for this chart suggests that this relationship has not occurred by chance.

A way to confirm that this version of the fixed effects model is the cause of the problem is to compare results with a variation of the fixed effects model which keeps the intercept constant and instead allows the slope to change as shown below. This solves most of the bias problem and is more reflective of how we believe costs vary in reality.



Testing for bias:



The high p value and low r squared imply the vast majority of the bias against those DNOs with a larger cost driver has been eliminated.

While this is clearly an improvement on the model with a fixed slope, there are still similar issues to the fixed effects model and the model is very complex when multiple cost drivers are taken into account.

Therefore we believe the best alternative to the fixed effects model is to consider both a pooled regression of all four years (which allows multiple years of data to be taken into account) and the single year regression (which allows for year specific variations). We believe taking an average of the efficiencies from both of these models would provide a more accurate efficiency score.

Complexity

The DPCR5 benchmarking analysis is an order of magnitude more complex than that carried out in DPCR4. This introduces a greater risk that DNOs will be unfairly disadvantaged as it is incredibly difficult for the DNOs to satisfy themselves that the data is correct and the process is fair. This is due to:

- the large number of spreadsheets and amount of data that has been included,
- the move from single year data to panel data for four years,
- the increased number of regressions,
- the increased complexity of cost adjustments and exclusions,
- the lack of supporting explanatory information to allow replication of the results,
- the omission of information such as the numerical results of the statistical tests, the calculation of the I&M driver, the calculation of the number of overhead and underground faults.

While further meetings have provided some clarity on this issue and as an industry we now have a better understanding of the models, there are still some questions that remain unanswered. While the work required to understand the analysis and its implication has increased significantly, the timetable for the price control remains fixed and there is a high risk that further errors will not be identified until the negotiations are complete. This needs to be included in evaluating the degree of uncertainty that can be applied to the efficiency estimates from the bottom-up benchmarking.

Composite variables vs. multivariate regression

While we agree on the need to consider multiple cost drivers, Ofgem's proposed methodology for creating composite variables is flawed. Combining cost drivers into composites the way Ofgem has proposed is unnecessary and adds an extra layer of complication. It means the drivers have to be standardised and the weightings made to add up to 1.

It also means that when driver which has been deemed to be the "primary driver" is not mathematically the most important, the weightings are arbitrarily changed instead of a deeper investigation being made into why this is the case. Such investigation may reveal a problem with the data itself, or an incorrect interpretation of the situation, or a missing cost driver, or something else. Forcing the weightings suggests that such issues have been ignored and is likely to result in inaccurate efficiency values. Doubts as to the accuracy of the results are raised further by the lack of clear reasoning behind the choice of thresholds for the weightings, nor is it clear why a different threshold was chosen for group 3 indirect costs (66% where 50% has been used elsewhere – even if it did not end up being needed in the final calculations).

Multivariate regression would appear to be the solution as it enables multiple drivers to be included without additional complexity and without forcing weightings. This is a generally approved technique which does not require driver weightings to be standardised as drivers are excluded if they are not mathematically significant.

We agree that it can be useful to see points on a chart and this can be achieved by simply using the "efficient costs" calculated using multivariate regression in place of the cost driver.

This also solves one of the potential problems of top-down regression. While it can be argued that a single cost driver will not accurately reflect total costs, a top-down regression using multiple drivers does not have this problem. The drivers can also be statistically tested to see which do and do not affect costs, and those which do not can be dropped from the regression.

Bias in model selection - exclusion of top-down models

In paragraph 1.4 of the Allowed Revenue – Cost Assessment Appendices (94a/09) Ofgem states that:

> *The results for the top-down analysis are very different to the results of the single group and groups analysis. Our view is that the results for single group and groups are more in line with our expectation of the results given our knowledge of the DNOs. The top-down analysis results are worse for ENW, WPD S Wales, WPD S West and SSE Southern. The top-down analysis results are better for CN West, EDFE LPN, EDFE EPN and SP Distribution.*

This is effectively creating a self-fulfilling prophecy, in that the companies expected to be more efficient will *of course* be found to be more efficient if the results are selected to meet these expectations.

In a well conducted study the analysis should be specified in terms of regressions, drivers, cost adjustments etc, before the data has been gathered and analysed. The robustness of Ofgem's analysis is seriously called into question when it appears to choose its results based on what it wants to see.

Ofgem suggests that the only reason it has given less priority to top-down benchmarking is because of this confirmation bias. This is a completely incorrect reason to rely on bottom-up benchmarking and top-down needs to be seriously considered in its analysis. Indeed, the many flaws in the bottom-up analysis (detailed in Appendix C1) suggest that more credence should be given to the top-down method instead.

Statistical validity of the models

The results of the statistical tests carried out on the models are presented In table 8 of the appendices. We are disappointed that Ofgem has not, at this time, shared the numerical values calculated, instead simply using a yes/no system to reveal results.

We also note that from the May paper to the initial proposals, the Breusch Godfrey test for serial correlation, which was failed by almost all the models, has not been repeated. No explanation has been given as to why this is now omitted and therefore we assume that the models continue to fail this test. While serial correlation may not be considered the most important issue, it does bias standard errors, which in turn can affect t tests for significance of slope.

Of the tests that have been conducted, the vast majority of the bottom-up regressions fail at least one (whereas top-down does not appear to have any problems).

Group 1, UG faults, I&M and tree cutting all fail the White test for heteroskedasticity. As stated in the May paper, this implies 'the standard errors of the coefficients are biased'.

Groups 2, 3, UG faults and tree cutting all fail the Ramsey RESET test for model specification. The May paper suggested some of the issues this might identify: omitted cost drivers, incorrect functional form and correlation between the driver and the residuals.

I&M in 2008-09 fails the Jarque-bera normality test. As was said in the May paper, this is not essential, but 'it is an indication of a well behaved model'.

We believe that such a high number of problems with the different regressions indicates either a generic flaw (such as widespread reporting differences) and/or a high number of specific errors (such as the absent faults, incorrect I&M driver etc). The demonstrable lack of robustness for the bottom-up benchmarking makes it an inappropriate methodology on which to set allowances.

In the Initial Proposals, however, the failure of the statistical tests is given little attention. While there is a suggestion that Ofgem will 'investigate further', there is no indication of what that will entail. In the references Ofgem make to the test results they effectively minimise the scale of the problem by effectively giving equal weight to the test failures as having identified outliers.

Ofgem has also used standardised residuals to identify potential outliers. How this affects the modelling, if at all, is not discussed in the paper, so a critique of that cannot be supplied at this time. However, we find the lack of outliers for tree cutting costs against spans cut to be incorrect. Scottish Power report spans cut on a different basis to the other DNOs – 'spans managed' - but because they are two DNOs affecting the regression line in a similar position, their standardised residual is low enough to pass the tests. They should be excluded from any tree cutting/spans cut regression unless they are reporting on the same basis. We have found that when both cost and driver are used in logarithmic form this makes it harder to spot outliers and therefore the models should be tested for outliers with both cost and cost drivers in their original values.

The R squared values for goodness of fit have also been included by Ofgem. While this is not the be all and end all of regression analysis, it should still be considered. An R squared of 0.6 effectively means the cost driver explains 60% of the variation in costs. An R squared of 0.3 suggests only 30% is explained. While an R squared of 0.3 (as is given for the group 1 regression in table 8) indicates some relationship between the driver and the cost, it implies far more of the variation is left unexplained. We believe this indicates that using that regression as a basis to set cost allowances is fundamentally flawed, when so much variation is unexplained. We do not think it likely that all of the unexplained variation is due to inefficiency. We believe Ofgem should aim to consider only those with an R squared value of 0.7-0.75 or higher for use in allowance setting. Top-down regression once again is an improvement on all of the bottom-up models, with an R squared of 84%.

The statistics tests, reproduced below, counter Ofgem's statement that the analysis is robust. The logical action that follows examining the test results is to put a higher degree of emphasis on top-down modelling and to either put less emphasis on the bottom-up modelling or exclude it altogether.

Table 8 - Results of statistical tests on core regressions

	Jarque-bera test				**Ramsey RESET test**	**F-test**	**White test**	**Standardised residuals**	**R^2**
	Normality test				**Model specification**	**Homogeneity test**	**Heteroscadasticity test**	**Outlier test**	**Goodness of fit**
	2005-06	2006-07	2007-08	2008-09					
DNOs LogLog Model									
TopDown	No	No	No	No	No	No	No	SPM	0.84
Sgroup	No	No	No	No	No	No	No	ENW	0.67
Group1	No	No	No	No	No	No	Yes	-	0.30
Group2	No	No	No	No	Yes	No	No	LPN	0.76
Group3	No	No	No	No	Yes	No	No	-	0.67
Overhead Faults	No	No	No	No	No	No	No	-	0.68
Underground Faults	No	No	No	No	Yes	No	Yes	SSEH	0.58
I and M	No	No	No	Yes	No	No	Yes	CNW, SSEH	0.40
Tree Curring	No	No	No	No	Yes	No	Yes	-	0.50