Ofwat's consultation of assessing base costs at PR24

Hafren Dyfrdwy response

3 February 2022

RHAGOROL O'R TAP
WONDERFUL ON TAP

HAFREN DYFRDWY
severn dee

Executive summary

We welcome the breadth of this first consultation and the opportunity to share our thoughts on the emerging approach to cost assessment. Overall, the PR19 cost assessment was effective – as the CMA has noted – and so it makes sense to evolve rather than seek a radically different approach. We recognise that thinking will continue to develop and mature over the coming months as we move towards pragmatic solutions for delivering robust models. Alongside this, we are conscious that certain areas may need more detailed/ specialist thinking and testing. Moreover, we are keen to play our part in helping evolve the approach to cost modelling for PR24 and our initial thinking on the key areas follows.

Separate consideration of Hafren Dyfrdwy wastewater costs. Our wastewater business is a clear outlier given its size and scale. Logically, it would not be sensible to include it as a separate observation when constructing econometric models given that our size, scale and nature of our waste business is highly likely to result in significant outlier effects. We think a pragmatic way forward here would be to retain the PR19 approach that used aggregated data with Severn Trent and then split the cost allowance in proportion to the two business plan submissions, in a way that reflects the makeup of the HDD business and its more-expensive-to-run smaller assets.

Specifications. We have identified several areas where there look to be very worthwhile improvements available for the base models. The nature of the area we serve means that assessments of population density and network complexity are likely to be particularly material for us.

Scope. This should focus as much as possible on base opex and capital maintenance (Botex), with enhancement cost drivers put through different assessments. We note that assessing growth costs is an ongoing challenge and suggest areas worth a further look include (i) approaches outside of econometric models and (ii) increasing explanatory power within these for a subset of growth costs.

Adjustments. We welcome the extra focus on how cost adjustments are assessed. Making these symmetrical could help manage model misspecifications. Materiality will also be important and we expect there may be cases for them to be non-symmetrical, such as areas with a lack of historic data.

Cost service relationship. As thinking in this area develops, a key focus will be defining the capital maintenance and enhancement expenditure that will flow into the models. The performance levels inherently assumed in the models would then follow the logic of the cost models. That said, we are aware that previous reviews have found performance and efficiency to have a very loose relationship that tends to the average – a strong sign that the models do not account for external factors that impact performance.

Future capital maintenance/asset health. We see a clear risk that the current approach is yet to reasonably account for future pressures on asset health, so it makes sense to raise understanding in this area. One fix we think is worth a look is forecast, rather than historic, efficiency ratios. We also suggest adding asset health metrics in the models, along with appropriate ways to mitigate endogeneity concerns.

Contents

1. Principles and development of a base cost assessment approach

1.1 Principles of cost assessment

We are broadly supportive of the suite of principles identified in the consultation document.

Engineering logic should come first when specifying models and selecting explanatory variables (in line with principle 1). Statistical validity is important and can be used to select between models and variables that are supported by engineering logic, but it should not serve as the primary determinant of which variables to include in the models. In ***Chapter 2.3*** we discuss some variables that we believe could be changed from the PR19 modelling suite.

Statistical validity is an important aspect of model selection and will support robustness (as per principle 4). Variables should not be selected purely based on statistical significance, nor should they necessarily be excluded if they are not statistically significant but give a sensible coefficient and are strongly supported by engineering logic. Tests describing the 'goodness of fit' of models can be used to select the models with the optimal predictive power from a pool of models that are supported by engineering logic.

We are in favour of retaining the Random Effects estimator for PR24, given it worked at PR19. It keeps the approach consistent and there doesn't seem to be a superior alternative.

As per principle 6, models need to work in their wider cost assessment context. This includes making sure that efficiency benchmarks are set on a sound basis that reflects the risk and uncertainty of models. This will need to be reflected when discharging principle 5.

1.2 Modelling Hafren Dyfrdwy at PR24

Our primary concern is around how Hafren Dyfrdwy should be modelled at PR24. We are of the view that HDD should be treated as its own entity where this is feasible. In water, this certainly looks possible given our size relative to the smaller Water Only Companies. In waste, this appears to be substantially more challenging as we are a material outlier given that we are significantly smaller than any other company in the sector.

At PR19, the waste models combined Severn Trent England (SVE) and Hafren Dyfrdwy (HDD) into a single company. ***Figure 1*** below shows a set of regressions that follow this single company approach for the 2018-19 and 2019-20 financial years. For contrast, ***Figure 2*** then goes on to show a set of regressions for the same period with SVE and HDD treated as their own separate entities. We can see that the coefficients change substantially between the two approaches, and that the second approach reduces the model quality significantly according to the AIC/BIC values. This is particularly true for the sewage treatment specifications, where the non-scale variables all become insignificant.

A further challenge with the second approach is that the coefficient for the percentage of wastewater treated at small works (pctbands13) becomes negative, which is counterintuitive. This is likely a result of HDD having significantly lower total expenditure than any other company in the sector at the same time as having a significantly higher proportion of wastewater treated at these more-expensive-to-run smaller sites.

Figure 1 – Regression coefficients (p-values) for SWC and SWT specifications with SWT and SWT combined into a single company, mirroring the SWT19 process. These use a panel of data from 2011-12 to 2019-20.

Variable	SWC1	SWC2	SWT1	SWT2
lnsewerlength	0.8390 (0.000)	0.8649 (0.000)	-	-
lnpumpingcapperlength	0.2875 (0.086)	0.5144 (0.019)	-	-
lndensity	0.9777 (0.008)	-	-	-
lnwedensitywastewater	-	0.1875 (0.125)	-	-
lnload	-	-	0.7794 (0.000)	0.7814 (0.000)
pctbands13	-	-	0.0416 (0.016)	-
pctnh3below3mg	-	-	0.0036 (0.000)	0.0038 (0.000)
pctbands6	-	-	-	-0.0126 (0.025)
constant	-8.0385 (0.000)	-6.1084 (0.000)	-5.2144 (0.000)	-4.1201 (0.000)
R^2	0.9353	0.8925	0.8728	0.8645
AIC	-139.90	-133.75	-100.81	-101.44
BIC	-124.90	-118.75	-85.81	-86.44
RESET	0.0941	0.1069	0.3106	0.2669
MAPE 2018-20	13.18%	17.03%	12.92%	10.61%

Figure 2 – Regression coefficients (p-values) for SWC and SWT specifications with SWT and SWT modelled as separate companies for 2018-19 and 2019-20. These use a panel of data from 2011-12 to 2019-20.

Variable	SWC1	SWC2	SWT1	SWT2
lnsewerlength	1.0622 (0.000)	0.9768 (0.000)	-	-
lnpumpingcapperlength	0.5332 (0.006)	0.5955 (0.002)	-	-
lndensity	0.7889 (0.112)	-	-	-
lnwedensitywastewater	-	0.1908 (0.152)	-	-
lnload	-	-	0.8045 (0.000)	0.9081 (0.000)
pctbands13	-	-	-0.0012 (0.948)	-
pctnh3below3mg	-	-	0.0015 (0.144)	0.0014 (0.113)
pctbands6	-	-	-	-0.0062 (0.176)
constant	-9.8780 (0.000)	-7.3983 (0.000)	-5.3350 (0.000)	-6.1723 (0.000)
R^2	0.9692	0.9642	0.9282	0.9317
AIC	-133.46	-131.55	-77.93	-78.88
BIC	-118.33	-116.42	-62.93	-63.75
RESET	0.1924	0.0817	0.3894	0.4574
MAPE 2018-20	34.11%	20.56%	61.03%	34.84%

As this analysis highlights, there would be a strong logic for following the approach taken at PR19 in waste. This would then see PR24 to model a single company for waste called 'SVH' that consists of both HDD and SVE, and make sure that the smallest waste company in the sector does not have an outsized impact on the outcomes for the sector mode. Forecast costs could then be allocated based on HDD and SVE business plans, again following the methodology used at PR19, in a way that reflects the makeup of the HDD business and its more-expensive-to-run smaller assets.

An alternative, but much more resource-intensive, approach could see HDD excluded entirely from the wastewater modelling process and given its allocation based on a deep dive of the business plan. Given HDD's size, it's unlikely that such an approach would prove to be pragmatic.

1.3 Developing and improving PR24 wholesale base cost models

1.3.1 Appropriately allowing for population density

The recognised engineering logic is that costs are elevated at both extremes of density and sparsity. At PR19, this was the underlying premise for including weighted-average-density and weighted-average-density-squared.

Given our size as a company and that our patch is dominated by ultra-rural and ultra-low-density areas we see that this as an important area to get as right as we can.

It is true that the two terms allow for a U-shaped distribution of weighted-average-density with respect to cost. As the existing variable stands, our principle concern is that the local authority data used to create it is not giving consistent information across all companies, which then leaves the variable unable to fulfil its intended purpose. To us the data used does not appear to satisfy these engineering expectations for three reasons:

- The calculation of weighted average densities is problematic:
  - the calculated weighted average values are mathematically influenced by the size and level of aggregation of local authorities included in a company's density value. As ***Figure 3*** shows, greater disaggregation gives a much clearer, richer and representative view of the types of density that a company actually serves.
  - Weighted averages skew values towards the most populous areas (generally also the densest areas) served, irrespective of whether they are typical of the wider supply area. Consequently, it is unlikely to represent the sparsity of a company's supply area (eg, Powys) away from the urban population centres that dominate the calculated value (in our case Wrexham).
- Providing one value of density per company describes a single point towards the centre of the distribution, whereas density is likely to drive costs at a granular sub-company level. The logic here is that cost pressures are at their greatest in the areas of extreme density and/or sparsity. When averaged out at the company level, these cost pressures become significantly diluted and under-reflected in the model.

Figure 3 – weighted average density increases with disaggregation of areas

Unit	Population	Area	Density (*weighted total*)
Single area			
A	100	1	100
Total	*100*	*1*	*100*
Some disaggregation – eg, large combined authorities			
A	60	0.5	120
B	40	0.5	80
Total	*100*	*1*	*104*
Increased disaggregation – eg, generic local authorities			
A	60	0.5	120
B	30	0.1	300
C	10	0.4	25
Total	*100*	*1*	*164.5*
Most disaggregation – eg, postcode sectors			
A	50	0.1	500
B	30	0.1	300
C	10	0.4	25
D	10	0.4	25
Total	*100*	*1*	*345*

The impact of local authority size and the skew to dense population centres

Mathematically the number and size of local authorities will materially impact the calculated view of density. For example, the more disaggregated the local authorities within a company supply area, the higher the weighted-average-density will be. This is particularly stark for us as we have just three local authorities in our water supply area, and just one for waste.

In overall terms, the weighted-average-density metric is skewed heavily towards the most populous – and in general the densest – areas that a company serves. We have found that, for multiple companies, the 2011 weighted-average-density is significantly above the 75th percentile and, for some companies, is even considered an outlier. As we've touched on above, some of this pattern will also reflect the different sizes of local authorities.

We've found that using more granular postcode sector data gives a much truer picture of the way densities distribute across the supply area. The upshot is that the weighted-average-density then lies consistently within the 75th percentile for all companies, irrespective of the distribution or extent of sparse areas that they serve.

Off the back of this analysis, it very much looks like that the density terms used in the PR19 models are not actually accounting for the increased costs of operating in rural areas. Rather it suggests that the PR19 models have (i) worked on the basis that there are higher costs associated with having the least dense urban areas – a result that does not align with engineering expectations and (ii) targeted a level of density that does not represent most of the area that a company needs to supply. In short, the squared density term is not serving its intended purpose, so should be excluded from the model. We should therefore look at alternatives to account for density extremes, while retaining the linear term in the model.

The impact of considering density only at a company-average level

The current weighted average measure of density gives a single density-value per company. The accompanying squared density term then generates a 'curve' of associated costs, with the aim of acknowledging the costs associated with both sparsity and density. The significant challenge that comes with this approach is the underlying assumption that a company's weighted-average-density applies uniformly across its patch, which then means it cannot account for the costs a company incurs from serving both sparse and dense areas.

One alternative approach considered at PR19 for density was separate metrics for the proportion of local authorities above or below a certain density or sparsity threshold. These metrics could then have been used alongside, or in place of, an average density figure. One advantage from of this approach is that it starts to acknowledge the distribution of densities across a company's patch, in line with the way that engineering logic says cost pressures should show themselves.

The difficulty with this density threshold data is its ability to disaggregate local authorities in a way that truly reflects the density that a company has to contend with. This is particularly acute for a company that is as small us. On our water patch, we have just three local authorities, which means that the only values we could use for the variables are zero, 0.33, 0.66 or 1. In other words, they are giving a blurred view of density as they're unlikely to accurately reflect the share of either our company area or population beyond these thresholds.

To explore a clearer and more representative view of density, we have used more granular postcode sector as we believe that it is more likely to reveal more robust and appropriate thresholds. ***Figure 4*** shows a table of moments for each company based on the postcode sector distribution.

Figure 4 – Table of the moments of the postcode sector density distribution

Company	Kurtosis	Skewness	Standard Deviation	Mean
WSH	8.882	2.583	1756	1192
SVE	8.659	2.061	2449	2288
HDD	7.213	2.532	1522	890
SWB	5.723	2.072	2081	1513
NWT	5.577	1.531	2585	2994
SRN	4.936	1.873	2966	2839
PRT	3.545	1.871	3370	3162
TMS	3.441	1.325	5954	6699
SEW	3.009	1.56	1652	1491
ANH	3.005	1.586	1638	1319
YKY	2.677	1.484	2160	2113
NES	2.454	1.393	2279	2356
WSX	1.306	1.484	1466	1095
BRL	0.949	2.072	2753	2568
AFW	0.282	0.972	2348	2608
SSC	-0.164	0.576	1793	2254
SES	-0.326	0.84	2151	2378

- **Kurtosis**. A company with high kurtosis – the extent to which the data distribution has tails – suggests that there are a high number of density outliers. This is particularly problematic for companies like us as well as SVE, NWT, SWB and WSH. The supply areas for these companies all look to have a lot of rural areas alongside very dense outlier regions or vice versa – circumstances where the weighted-average-density measure will not account for density outliers.
- **Skewness**. This provides another view of this outlier issue – high skewness indicates that a company's postcode sectors are skewed heavily towards lower densities. Where high skewness is combined with high kurtosis, a company has to contend with mostly very rural areas, but with a significant number of very dense outliers as well.

Conclusion

Our considered view is that the squared density term should be removed as it is unable to capture the costs for companies that serve both highly dense and extremely sparse areas. We recognise that this move would leave a misspecification with no explanatory power that is linked to the extremes of density/sparsity (which has strong engineering logic). We also recognise that using weighted-average-density alongside thresholds of high density and sparsity is problematic, because high density is colinear with weighted-average-density and because the current sparsity values at the local authority level do not fully expose areas of extreme sparsity.

While we have yet to find the perfect solution, we consider that the following areas have merit for improving on what we have now:

- Our analysis at postcode sector level is promising and provides a much richer picture. As a next step, we could seek ways to improve data granularity of data and then include relevant variables in the models.

- If collating granular data is unlikely, then we could assume just a central view of density in the models using the weighted average, but not the weighted average squared. We could then use symmetrical adjustments for sparsity and density, such as the kurtosis – or the tailedness – of the data).

1.3.2 Appropriately allowing for network complexity

Network complexity

There has been a lot of debate as to whether Average Pumping Head (APH) or Boosters per Length (BPL) is the correct variable to account for water network complexity. Looking ahead to PR24, we believe that there is notable merit to be had from including both variables in the TWD and WW model specifications – principally because these variables represent two very different aspects of cost.

Engineering logic

The engineering logic for considering APH and BPL as separate cost drivers is as follows:

- **APH** increases as the need to pump water within the network increases. This is largely driven by the topography of the supply area – because of very flat or very hilly regions. A single booster pumping station with a high lift and capacity could be equivalent to multiple smaller pumping stations delivering an equivalent volume and lift. Crucially, both scenarios would incur significant energy costs.
- **BPL** increases with the number of booster pumping stations. In practice, there may be many low-capacity/lift-booster pumping stations in a particular area of network, such as those serving small populations in hilly areas like Powys, that contribute significantly to our costs given our size and the absence of the large urban centres that other companies serve. Where booster pumping stations are needed, they are usually matched with corresponding pressure-release valves. Simplistically, a booster lifts the water on one side of a hill, with the pressure release valve used to manage pressure, velocity and the risk of bursts downstream.

Where boosters lift relatively small volumes of water, they will not individually contribute much to the aggregated APH value. However, these additional assets and corresponding pressure-release valves all require ongoing capital maintenance and create weak points in the network that are prone to leaks and bursts, and hence drive further capital maintenance needs. This therefore creates a capital maintenance cost pressure.

Overall, APH is likely to provide a good proxy for opex (power) costs, while BPL is likely to reflect long-term capital maintenance and asset-heath cost pressures within the network. <u>Our assessment is that by including both these variables in the models will provide a much more complete view of totex costs than is possible from using just one of these variables in isolation</u>.

Statistical logic

Figure 5 below shows a scatter chart of mean log BPL and mean log APH in Treated Water Distribution. From a visual inspection, there is very little correlation between the two variables (Pearson correlation of -0.0898). Consequently, there is no problem with including them both in the same regression.

Figure 5 – the relationship between Boosters Per Length and Average Pumping Head in Treated Water Distribution

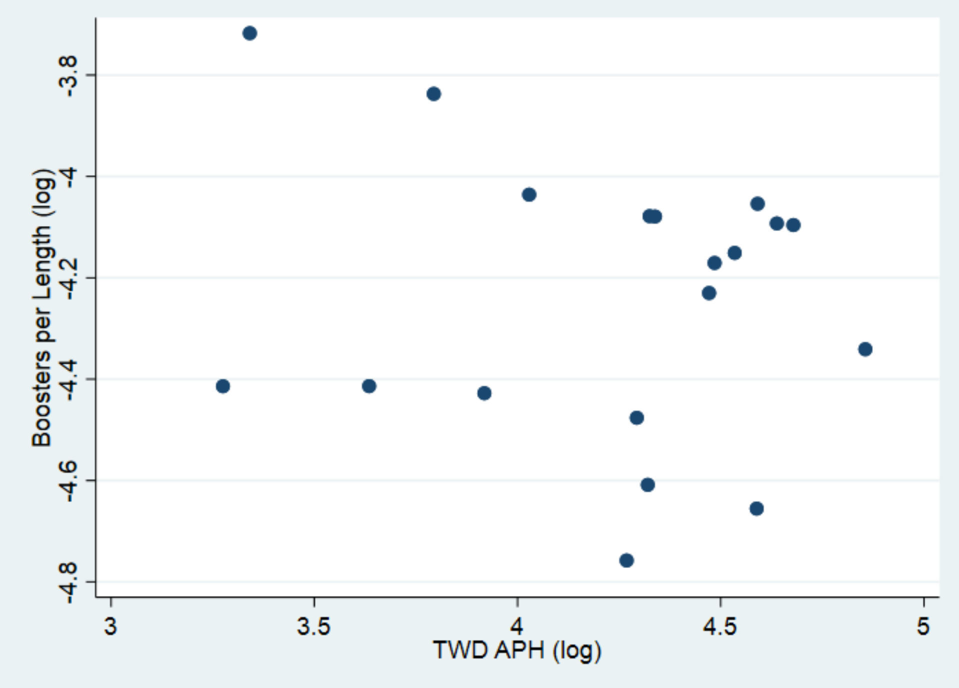
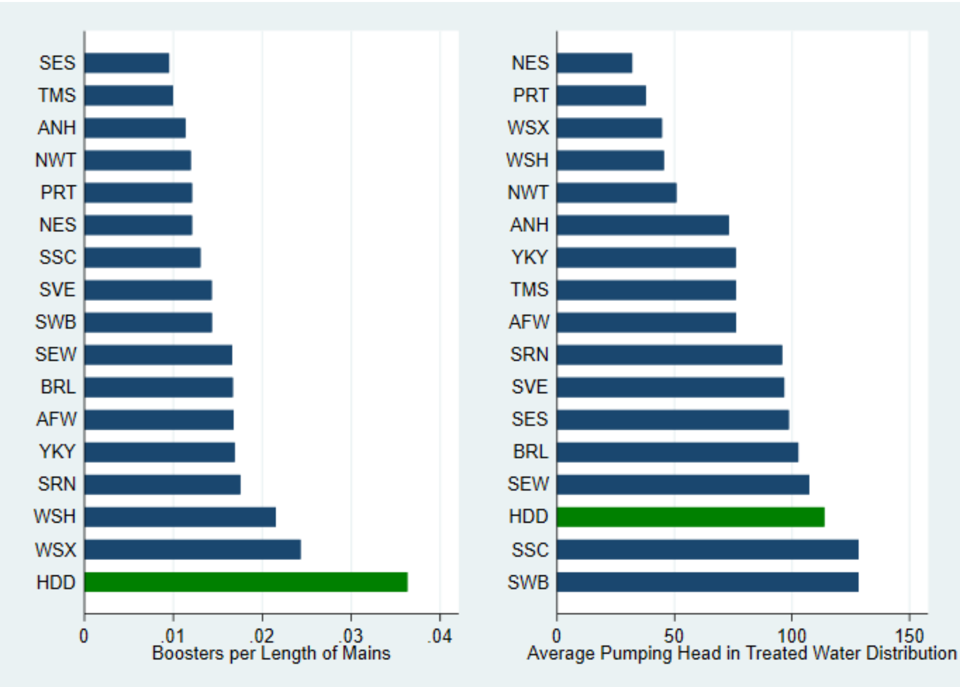


Figure 6 shows that HDD has a significantly higher number of boosters per length of mains than the rest of the industry, as well as a high APH in treated water distribution. If the PR24 models were to include APH but not boosters per length, the models would miss a material and significant cost element for HDD.

Figure 6 – Boosters Per Length of Mains and Average Pumping Head in Treated Water Distribution (panel data 2011-12 to 2019-20)



Regression analysis for the TWD specification – as shown in ***Figure 7*** below – finds that both the BPL and APH variables are significant. By considering the R^2, AIC, BIC, and MAPE metrics together, we can see that the joint specifications perform better than separately. We also see strong signs of robustness as the coefficients only change marginally between the joint and the single network complexity variable specifications. Overall, we consider this as strong evidence for including both variables, whilst acknowledging that the quality of APH data has known challenges and is subject to review.

Figure 7 – Regression coefficients (p-values) with just BPL, APH + BPL, just APH for the TWD specification (panel data 2011-12 to 2019-20)

Variable	Original TWD Model	TWD with APH + BPL	TWD with just APH
lnwedensitywater	-3.1969 (0.000)	-3.4112 (0.000)	-4.1374 (0.000)
lnwedensitywater2	0.2562 (0.000)	0.2696 (0.000)	0.3117 (0.000)
lnlengthsofmain	1.0620 (0.000)	1.0696 (0.000)	1.0916 (0.000)
lnboostersperlength	0.5504 (0.000)	0.4988 (0.000)	-
lnaph_twd	-	00994 (0.047)	0.1117 (0.042)
constant	6.1092 (0.000)	6.2366 (0.000)	6.8898 (0.000)
R^2	0.9623	0.9652	0.9580
AIC	-100.87	-104.34	-97.46
BIC	-79.43	-79.84	-76.02
RESET	0.0481	0.0403	0.0408
MAPE 2018-20	20.45%	19.39%	20.42%

1.3.3 Wider modelling

Growth

We consider that the scope of the econometric models should conform, as close as possible, to a Botex[1] definition as these are the most repeatable and modellable costs. We would strongly counsel against adding growth costs to these models – to create Botex+ models – as there is very little explanatory power for them.

The difficulty is that growth costs are lumpy, which means that including them is likely to reduce model quality rather than improve it. This is particularly true for Network Reinforcement, which is a function of highly specific local conditions and is unlikely to be able to be captured by company-wide variables.

Nevertheless, we recognise the need for growth to be considered as part of the cost assessment. Our thinking on how this area could be taken forward would be to:

- remove growth from econometric models and consider it separately if robust growth models can be found;
- if such robust models cannot be found for the entirety of growth, we propose removing the elements of growth that are particularly lumpy and cannot be modelled (eg, network reinforcement) and use a deep-dive assessment of these costs instead. The Botex+ model could then include a variable to account for the remaining on-site growth costs; and
- resort to retaining all growth costs in the Botex+ models only if the above methods are not possible.

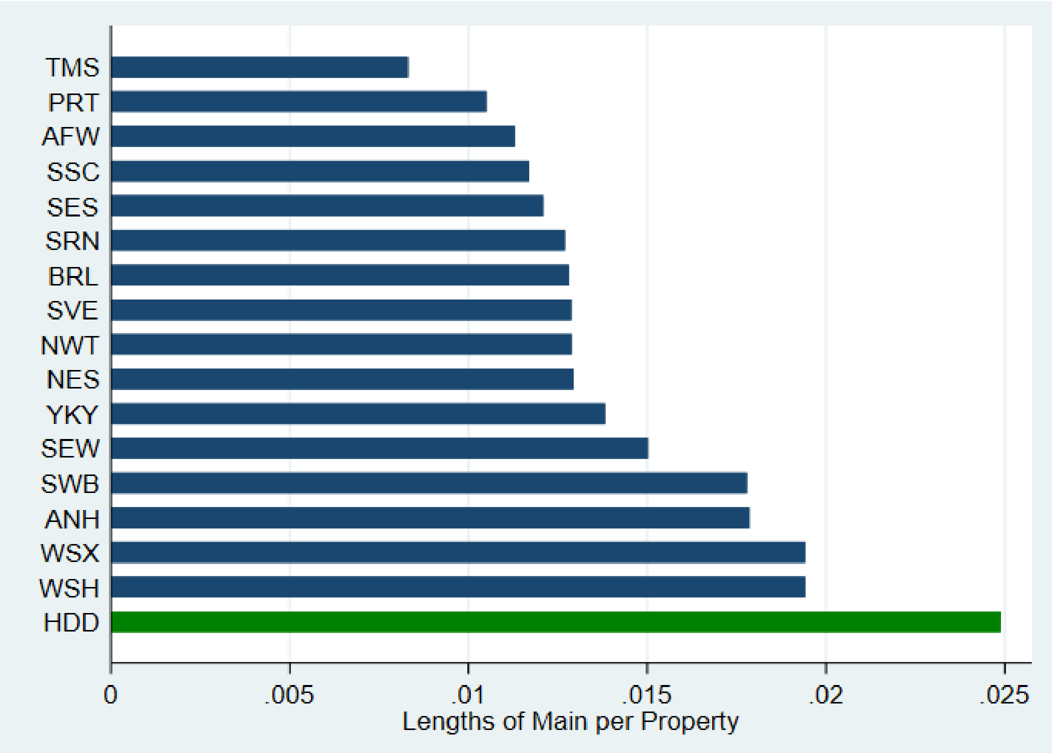
[1] Botex = base operating expenditure and capital maintenance

Composite scale variables

Scale is an essential driver – when we used it in our models as the only explanatory variable, it generally giving an R^2 value of 0.9. This is intuitively logical, as the size of a business – the area and the population it serves – clearly has a close relationship with how much it costs to operate. The challenge is that it is not always clear which scale driver has the greatest impact on cost – in fact, there is often a case for including multiple scale-drivers in the model. This is particularly true for models describing higher levels of aggregation. Given that we're aiming for sensibly simple models, we propose that composite scale variables are used in place of multiple scale-drivers.

Figure 8 shows just how sensitive we are as a company to this issue given our proportionally large network length relative to the number of connected properties. As it stands, the current wholesale water specification does not reflect this and so misses the significant and material impact on our costs from this driver. By contrast, the largest element of wholesale water costs, Treated Water Distribution, uses lengths of main as the scale driver, so it is logical for this variable to also influence on the wholesale water specification.

Figure 8 – average length-of-mains per property over the sample period *(panel data from 2011-12 to 2019-20)*



Water resource cost drivers

Currently, water resources cost drivers are not used in the base models. The upshot is that the water resources costs that are present in the WRP models are driving model noise instead. Nevertheless, we appreciate that finding appropriate drivers will not be straightforward. Going forward, we recommend that exploring this area further in order to identify drivers with a clear underlying engineering logic.

Water treatment complexity

The water treatment complexity we face as a company is one of the highest, if not the highest in the sector, and is a significant driver of our costs. Our concern with the current water treatment complexity metrics is that they may be contributing to model misspecification. Our logic here is that:

- certain treatment processes may be in the wrong treatment complexity band,
- multiple higher-cost treatment processes may not be significantly more expensive than a single higher-cost treatment process, and
- single higher-cost treatment processes may not be significantly more expensive than multiple lower-cost treatment processes.

Accounting for changes over time – indicators and trends

Time indicators and/or time trends serve as an important model control because they account for the cost changes over time that the other variables used in the models do not capture. This is likely to be the case for changes in quality, input prices, or the material conditions companies face.

The advantage of this approach is that the time trend may robustly capture real prices effects within the model. It would then be sensible to discuss whether there is a need to estimate them separately. Removing the real price effect adjustment may be needed so as to avoid double-counting.

Sample period selection

We believe that increasing the amount of data observations in the model is important for model accuracy and therefore predictive capability. Consequently, we support the use of the full dataset back to 2011-12. We also note that adding time trends into these models is a helpful way to make sure older data remains relevant aside from any efficiency improvements over that time.

We've also given thought to whether Ofwat could use business plan forecasts in its models, recognising that this might help to identify changes in costs driven by the challenges of climate change, net zero, greater resilience and population growth. On balance, we believe that using forecast could prove unhelpful for Ofwat as these would, in all likelihood, inject uncertainty into the models and materially reduce their robustness.

2. Cost adjustment claims

We welcome the commitment to retain a cost adjustment claims process, which we see as a critical safety-valve to the cost adjustment process. We think that it very much helps the process of arriving at econometric base models that are sensibly simple by recognising that such models will not be able to capture all drivers of cost that effect each company.

In managing the inherent risk associated with sensibly simple models, we are conscious that an effective adjustment process is needed. We believe that one must-have for this process will be administering it in a way that allows companies to apply and reasonably expect success where they proportionate cases to put forward.

We think it is sensible to consider adjustment claims for base and enhancement expenditure as distinct from each other. This will allow the assessment of claims to be appropriately tailored to the different circumstances of base and enhancement expenditure.

We support the premise of symmetrical adjustments where they relate to addressing weaknesses in model specification and think that they can play a key role in improving the wider cost assessment approach (principle 5). This can be a critical way of improving the coverage of cost assessment where model improvements are neither possible nor desirable. We also think it is important to set materiality tests that need passing before applying any counteracting symmetrical adjustments to other companies.

We also strongly believe that there still needs to be a realistic opportunity to make and receive non-symmetrical adjustments. These would be appropriate (i) where costs relate to one or two companies – such as those in Wales – and any negative cost adjustment would be trivial or zero, or (ii) where the sector had not incurred the identified costs in the past.

Our considered view is that we should not rule out potential claims that either have not been raised before or have been rejected historically. Holding companies to historic levels of understanding and decision making does not feel appropriate given (i) the complexities of understanding and interpreting econometric information, (ii) improvements in the ability to capture and analyse data, and (iii) the extra information, knowledge and insight that is now available.

Implicit allowance

The implicit allowance (IA) sets out the extent to which Ofwat's default cost assessment process partially (or fully) accounts for cost adjustment claims. On base, this will relate to the coverage of the expenditure identified within the final suite of base econometric models.

We accept that calculating implicit allowances can be complex. At PR19, we did at times find it difficult to follow and learn from used to calculate the implicit allowances. With this in mind, we would encourage that a good time ahead of the PR24 cost adjustment claim submissions, Ofwat sets out (i) some principles and examples of how IAs should be calculated in typical cost adjustment areas, and (ii) how this interacts with the 'need for cost adjustment' assessment gate. This will then allow companies to refine and prioritise potential claims accordingly. Similarly, it appears sensible to encourage companies to set out their cost adjustment claims both net and gross of the IA and the assumptions used in arriving at a material net cost adjustment claim.

Timing of cost adjustment claim submissions

We would cautiously welcome the opportunity to submit cost adjustment claims ahead of submitting the final business plan. This is particularly relevant if symmetrical claims are likely to be used regularly.

In this case, the benefits of early visibility of the cost models will only be realised if there is equivalent early visibility of likely symmetrical claims – given that both model adjustments and symmetrical adjustments will inform the likely cost benchmark used to assess companies' plans.

We see that the timing of an early submission will be important, as will be feedback, early sight of and visibility of the type of claims being proposed. It's crucial that this is on a timescale that gives companies an appropriate opportunity to respond and incorporate any feedback. If an equivalent timescale to PR19 were followed, this would take us to May 2023, which feels too close to the final submission deadline to undertake substantive reassessments. We also recognise that submitting claims overly early will likely reduce their quality and increase the disconnect with the final submitted plan.

Potential cost adjustment claims

In **Figure 9** below, we set out a list of potential adjustments that we could develop and put forward at PR24. These largely reflect the issues that we have discussed in this ***Chapter 2*** about model specification. It is also worth highlighting that were the model specifications to change in a way that addresses these issues then we are unlikely to need these adjustments.

We are conscious that, as we develop our business plan and gain greater insight of the cost assessment and modelling for PR24, other potential cost adjustment claims may arise. So, it's worth us noting that this current list is not a long-list that we expect to refine down.

Separately, we are considering the potential scope for enhancement business cases and associated enhancement cost adjustment claims. We are aware that enhancement cost assessment is not within the scope of this consultation and are engaging separately in this area through the CAWG.

Figure 9 – potential cost adjustment claims for PR24

Potential claim subjects	High level basis	Symmetrical/ non-symmetrical	Possible data needs
Density/ sparsity	Areas of high density and sparsity both drive cost, but this is not captured in the weighted-average-density variable currently considered in the base models.	Symmetrical/ non-symmetrical	Population density info at a more granular level (eg, postcode sector data from census)
Water network complexity	Depending on the network complexity explanatory factors used, the operating or capital maintenance costs are not likely to be appropriately accounted for. These are material costs which are not necessarily sensitive to one cost driver	Symmetrical / non-symmetrical	Standardised APH data Booster pumping station data
Water resources costs	The type, location and size of water resources assets impact on the costs that companies incur in both the water resources and network plus price control (through the knock-on impact on treatment complexity). These are controlled largely by the geography and geology of the company rather than scale or population density.	Symmetrical	More granular water resources related cost driver info. For example, expanding the information reported previously at PR19 in Wr2 to account for capex.
Future operating and maintenance impacts of prior enhancement expenditure	Historical costs will not account for the future operation or maintenance of enhancement activity implemented in the previous period (e.g. WINEP driven activity). The future impacts of enhancement interventions are unlikely to be strongly sensitive to model scale drivers. Consequently, over and above any allowance through a time trend or sensitivity to a complexity driver, very little allowance can be assumed.	Non-symmetrical	Whole-life cost information relating to incurred enhancement costs expenditure

3. Capital maintenance and asset health

Ofwat does not believe there is yet a case for changing its approach to funding as a result of future asset health cost pressures. We are much less certain that this will prove to be the case and are concerned about the potential impact on us, as such a small company, should a material asset-health problem arise in future. The emerging future context shows the challenges that our asset base will need to face into and keep operating.

- An increasing need to manage climate change uncertainty, and threats to physical and cyber security.
- Changes to the types, configuration and complexity/expected-life of assets that are driven by increasing expectations and standards on service resilience – for example, the new and novel interventions needed to meet the leakage target for 2050 drive down PCC, all of which will create future operating and maintenance cost pressures.
- Changing expectations and standards on environmental performance and 'waste' – such as maintaining service whilst also delivering against net zero objectives, and meeting societal expectations on leakage, pollution and PCC.

These issues impact on the size and shape of the asset base we rely on, the amount of stress that it will need to handle and the way it will respond to the stress. These all point to the need to take greater account of future asset health pressures when determining the future level of efficient / sustainable capital maintenance investment.

With this in mind, the cost assessment methodology will need to make sure that it adequately provides for capital maintenance and network renewal to account for:

- the physical properties and changes in the asset base over time,
- the external pressures that may increase the amount of stress on the asset base in the future,
- improvements made to manage existing assets more effectively, and
- potential incentives to chase near-term cost efficiencies if cost assessment and asset health deliverables are not closely aligned.

Better accounting of future challenges

The consultation document sets out three potential options for how future capital maintenance requirements might be better accounted for in the cost assessment process:

- Using business plan forecasts as inputs into base econometric models – as we set out in Chapter 2, **we do not support this option** as we consider the use of business plan forecasts will materially reduce both the robustness of models and the accuracy of any efficiency benchmark.
- Setting the catchup efficiency challenge based on business plan forecasts rather than historic outturn costs – **we cautiously support** considering this option further, while noting that it will need managing carefully if it is to deliver as intended.
- Including explanatory factors in the models that relate directly to capital maintenance/asset health activity – while we note Ofwat's concerns about endogeneity, **we think that this option has merit** particularly if (i) it is combined with independent assessments of the endogenous drivers and (ii) it has clear linkage to the outcomes framework to safeguard delivery.

We do not consider that separate modelling of capital maintenance is a realistic way forward as this would run counter to the totex assessment – the foundation stone of whole cost assessment approach.

Asset health measures

In light of the above and Ofwat's concerns on the lack of supporting evidence, we agree with the desire to improve understanding on asset health. Going forwards, it will be helpful to understand how this information is likely to be used as part of the PR24 cost assessment so that appropriate confidence can be provided whilst minimising excessive regulatory burden.

The criteria for assessing potential asset health metrics that Ofwat set out in table 5.1 of the consultation document are sensible and coherent. We explicitly support the acknowledgement that back-casting new asset health metrics may be challenging where they have not been specifically collected in the past.

We think that it is right not to pre-empt the UKWIR work on potential asset health measures. What we can say at this stage is that the four proposed types of asset health measures[2] appear sound.

Regarding the collated examples of potential measures across each of the four types that are presented in table 5.2 of the consultation document, we make the following high-level comments:

- Some of the aggregated measures can be very complex to calculate and can require significant analytical work to report consistently across companies.
- When using service performance measures, it is important to take care that underlying asset health, rather than operational performance, is being tracked.
- It may be sensible to consider the use of metrics that better track changes in the shape and characteristics of the asset base – eg, number/value of short-life EM&I (Electrical, Mechanical and Instrumentation) assets including meters, loggers etc. These could potentially help indicate changes in future maintenance needs.
- There may be merit in focusing asset health measures where (i) there are known variances with respect to company asset bases, such as our high number of DSRs relative to the population we serve, or (ii) where asset renewal is expensive or lumpy, such as trunk main renewal or our DSR maintenance programme.

We note that most of the metrics put forward in table 5.2 have a level of prior reporting – something that is important for generating consistent comparative information. Where prior reporting exists, the definitions in use would mark a logical starting point for setting the clear reporting requirements that these measures will need.

We also recognise that there will be challenges in gathering information that has not previously been reported. In such situations, a series of iterations to refine guidance, and company interpretation of that guidance, will be needed before meaningful comparisons are possible.

[2] Asset Characteristics, Maintenance activity, Asset and service performance measures and Aggregated measures.

4. Cost-service relationship

We agree that the interaction between cost and performance is a fundamental component of the price review. At PR19, econometric cost models essentially assumed that companies were delivering a common level of performance. Where this was not the case and the performance variations resulted in companies incurring different levels of cost, this discrepancy will have contributed to model noise. We see this as an important uncertainty to account for when setting the strength of the catch-up efficiency challenge.

Our reading of section 6 of the consultation document is that there is room to develop the logic further. It appears to us that implementing some of the proposals in the consultation could override the fundamental definitions and expectations of base and enhancement. Where this to happen, it would muddle the interaction between base and enhancement activity and create extra complexity.

One way forward here would be to make sure that the basis of the costs that are included within the scope of base econometric models do indeed reflect their fundamental basis. Appropriate remedies would then only need seeking where the interaction between cost and performance gives rise to an unacceptably large level of model misspecification.

Base costs should relate to maintaining a constant level of base service

We think that care is needed around the assertion that "efficient companies will continue to improve performance over the long term from base expenditure." Our particular concern is that there are companies who may look efficient through the narrow lens of totex outperformance, but have levels of service delivery that suggest that totex savings are coming at the expense of performance.

We agree that base costs should relate to capital maintenance and base operating expenditure – the expenditure needed to maintain a constant level of performance. As the Regulatory Accounting Guidelines set out in RAG4.10 (Table 4J and K):

> "Maintaining the long-term capability of the assets: Capital expenditure on [infrastructure / non-infrastructure assets] excluding third party capex to maintain the long-term capability of assets and to deliver base levels of service. Where projects have drivers both of enhancement and capital maintenance, companies should apply a method of proportional allocation to allocate costs between enhancement and capital maintenance"

We are conscious that service levels might improve over time as companies find new ways of working and responding to ODI incentives. Where this is the case, Ofwat would rightly want to protect the interests of customers by holding companies to those improvements at successive price reviews. We would note here that while companies can choose to improve levels of service, this should not be considered as a zero-sum efficiency improvement – in most cases, companies would need to choose to forego efficiency gains in order to drive performance improvements with the decision ultimately resting on the interaction of cost allowances and ODI incentives.

We also agree that companies who do not meet a defined level of service should be expected to improve – say, a common service level where there are no external divers such as regional geography or geology. This is a situation where low performing companies need to catch up with their peers and one that plays out differently from the assumption that base expenditure will lead to a gradual increase in service.

The cost-service assessment framework – 'What base buys?' and subsequent use of adjustments

We support the aspiration to identify the level of performance associated with modelled base expenditure (ie, what does base buy?). We recognise that this is an area that we're facing into afresh and think the approach outlined so far is a helpful start. We are also of the view that there is a bit of a way to go before we can confidently say that the approach is appropriate to take forward for PR24.

The base econometric models account for all relevant base operating and capital maintenance expenditure and seek to identify the relationship between costs and fundamental cost drivers. This means that they should take account of all performance levels across the sector. At PR19, performance metrics were not included as model drivers due to concerns with endogeneity[3], which has the fundamental consequence that the coefficients will assume an average level of performance / service provision.

Assuming that there is a recognisable relationship between cost and performance, the catchup efficiency challenge should be set relative to the performance of the fourth-placed or UQ company, rather than the UQ performance level. The logic here is that the efficient costs of the UQ company and the efficient costs of delivering UQ performance could be very different.

The consultation document appears to suggest that it is possible to identify the relationship between historical cost efficiency and outcomes performance. Having run some modelling ourselves, we are concerned that this relationship is poorly defined and circumstantial. When reviewed against PR19 data, our analysis shows that performance varies significantly with efficiency. With an appropriate level of tolerance, the performance of the benchmark companies' quicky tends towards the average.

The difficulty looks to be that, even though there are examples of efficient companies with high performance levels, this is not seen consistently enough to model robustly. The results also support the hypothesis that performance is subject to external factors not accounted for in model explanatory factors, rather than having a one-to-one relationship between efficiency and performance.

In taking this forward, we think that it will be essential for the process for identifying the base performance associated with the efficiency benchmark, to build in a level of tolerance – say one or two companies either side of the benchmark. A logical starting point for defining the level of performance assumed in the base econometric models is either the industry average level or the average level that performance commitments have set out for the end of AMP7. Continuing this logic, the identified average level of performance should then project forward on a flat basis – sensitive to any related enhancement expenditure assumed at PR24 – rather than following an ongoing trend.

To help protect the interests of customers, Ofwat could scrutinise material divergences in performance from the average level. It could also exclude underperforming companies when calculating the efficiency benchmark. Where this relates to model specification, this could be through specific/symmetrical adjustments or ODIs. That said, diluting the fundamental principles of base and enhancement expenditure and adding extra complexity to the efficiency methodology could rapidly take us away from a transparent and coherent approach.

[3] It is possible – but not yet proven – that the some of the chosen exogenous cost drivers (such as density or complexity) may indirectly contribute to performance levels. But, the engineering basis for this has yet to be explored in detail. In principle, we consider that the decision on whether performance should focus on a common or bespoke service level should be based on the extent that performance is shown to be driven by exogenous factors outside of company control (bespoke) or not (common).

Using econometric models to identify levels of performance

Theoretically, econometric analysis could identify efficient levels of identified base service. Rather than cost, the dependent variable would be flipped to a particular aspect of performance. This would then be explained through external explanatory factors that impact on performance outside of company control. Given the rationale of common performance levels – identified targets are applicable to all companies – this econometric approach could be justified for non-common performance levels.

Performance modelling would also be subject to major uncertainty. There are also some highly complex performance metrics, such as leakage or PCC, that have many underpinning causal relationships. Realistically, we do not think that these can be adequately developed to an appropriate level of robustness for PR24 that, particularly when coupled with the existing cost modelling uncertainties. We see that the complexity that comes with such composite modelling would significantly reduce the level transparency and would likely move us significantly away from the aspiration for sensibly simple models.

Composite performance variables are another way to include performance cost drivers in the cost models. We see that this has the potential to improve model specification and suggest that using principal component analysis could avoid the need to define weightings (see Chapter 2). Endogeneity issues could also be managed by using independently defined forecasts (see Chapter 4). The challenge here is that aggregating metrics would only improve model predictive performance if the performance metrics were (i) broadly consistent and (ii) sensitive to the same underlying exogenous explanatory factors. We would recommend that, before pursuing such approaches, the causal relationships between cost and cost-drivers for each of the of variables is mapped out and explored in more detail.

5. Residential retail cost assessment

We are supportive of the intention to separate Covid-19 related bad-debt costs from the standard provision. We suggest proceeding carefully here as this might lead to issues with the data and that relevant guidance would help make sure that reporting is consistent across companies.

If this is not possible, or the approach would further complicate data issues, then an indicator variable accounting for these exceptional years – ie, equal to 1 for the pandemic years and 0 otherwise, as discussed in Chapter 2 – may be a better way to solve to this problem rather than smoothing or even making no adjustment at all. We see that smoothing would only serve to spread the noise across multiple years, while making no adjustment would end up keeping the noise for the 'covid years'. By contrast, a time indicators variable would help to eliminate the majority of the noise while keeping years not impacted by Covid-19 consistent.

We would also welcome the move to include a broader range of affordability measures that will be available from the credit ratings agencies – eg, energy bill sizes, housing costs. This is particularly relevant given the cost of living increases that customers are currently facing.