Dŵr Cymru
Welsh Water

Pentwyn Road
Nelson
Treharris
CF46 6LY

Heol Pentwyn
Nelson
Treharris
CF46 6LY

01443 452 300
dwrcymru.com

Ynon Gablinger,
Director
Water 2020
Ofwat
Centre City Tower
7 Hill Street
Birmingham B5 4UA

1 May 2018

By e-mail only: CostAssessment@ofwat.gsi.gov.uk

Dear Ynon,

Cost assessment for PR19: a consultation on econometric cost modelling

Please find attached our responses to the questions set out in your consultation "Cost assessment for PR19: a consultation on econometric cost modelling".

We believe that this has been a very helpful exercise which has provided an opportunity for Ofwat to share its work and that of several companies and their advisers. The exercise will improve general understanding of the strengths and limitations of cost assessment modelling and should facilitate more robust and fair outcomes.

The models, and our approach to assessment

We have offered comments on all of the models in the consultation in the form of the Response Sheet provided by Ofwat, and this is attached. As requested, we have graded each model either "Green" (no concerns or only minor concerns), "Amber" (some concerns), or "Red" (major concerns) and we have recorded our concerns in the "Comments" column. The approach we have adopted is based on our judgement and can be summarised by the following criteria:

a. does the model contain a plausible selection of explanatory variables? Have cost drivers been included which do not make sense, and have important (on *a priori* grounds) explanatory factors been omitted?
b. are the coefficients on the variables sensible (signs, magnitude)?
c. does the model have good statistical properties (significance, RESET test, etc)?
d. is the distribution of model residuals plausible in terms of size and distribution? and
e. are there any issues regarding the definition of the dependent variable?

Please note that we have tried to be succinct in our comments and have not included every single concern that we have, so if anything is not clear or you have any queries please feel free to drop me an e-mail.

In section 2 of the consultation document ("Approach to Modelling") Ofwat set out six "model development and assessment" criteria which together shaped its own approach to modelling. During the course of our review of the models presented in the consultation a number of common themes emerged, which we consider could usefully be translated into an additional "checklist" of "common sense" conditions/criteria to guide Ofwat's further model development. We have noted these down in Appendix 1 of this letter.

Summary of findings

The Response Sheet is somewhat detailed and technical, so we thought it might be helpful to draw out the main conclusions from our review of the models. A summary of the models that we considered merited a "Green" is attached in Appendix 2.

General

The overall quality of the models varies considerably: at one end of the spectrum there are models which appear somewhat contrived and make little or no objective sense, but there are also several which are plausible and perform well. That said, there are no standout models which are of such a high quality that they alone could be relied upon to set allowable costs at PR19. We think that it will be necessary to draw upon a range of models and other evidence to arrive at robust and fair conclusions.

Water

On water, we found that the best of the wholesale and network plus models were of better quality than the equivalents for water resources, water treatment and treated water distribution. We think this has much to do with the scope for substitution between the various functions, especially between water resources and treatment. In general, the better models used a plausible scale variable, a variable or variables to capture sparsity and density of the area served, a variable that reflected either the complexity of treatment or the mix of raw water to be treated and a measure of maintenance activity. Although we have some misgivings about Ofwat's constructed density measure, preferring the weighted density measure which has more intuitive explanatory power.

Wastewater

On wastewater, we found that there were several models of wholesale and network plus expenditure that were of good quality, (though very few managed to pass the statistical criteria considered, especially that of functional form). Typically these use appropriate scale variables (properties or load) and capture the important explanatory factors of sparsity, asset intensity and the tightness of the consent regime (or the degree of treatment). However, (with a small number of exceptions) the models of sewage collection, sewage treatment, and bioresources plus did throw up significant concerns. For bioresources, although many of the models used plausible combinations of variables, we found that the range of residuals was, in general, implausibly wide. There was also extensive use of variables that are more endogenous to management decisions than exogenous.

Household Retail

For household retail, given the comparatively homogenous nature of the function, we were somewhat surprised that the models produced by Ofwat and the industry did not perform better. In particular, the spread of residuals was in general wider than one might expect, and one or two companies were persistent outliers. The better models sought to capture scale, economies of scope as between single- and dual-service customers, deprivation, the size of the average bill, and meter penetration. However, we were unconvinced by many of the new variables that were specifically constructed by companies for the purposes of this modelling exercise. Moreover, we have misgivings about the use of measures that are not truly exogenous, especially variables that purport to be proxies for deprivation such as default risk, property repossessions, and credit risk. It should also be noted that measures that are constructed by third parties using proprietorial data cannot be verified or replicated in the regulatory process, and there is no way of forecasting them forward. As a general rule, therefore, we favour the use of official statistics and measures that are derived from them in a transparent way.

Enhancement

Finally, whilst we acknowledge the positive efforts made by Ofwat to model certain enhancement activities, this is clearly "work in progress" and there is a lot more work to be carried out to produce equations that could be relied upon for setting allowable expenditures.

Aggregation for cost benchmarking

You stated in the <Guide> tab that you were interested in stakeholders' views on the level of aggregation suitable for econometric modelling in the light of our findings.

As we indicated earlier we have concerns about the **water** resources models as a group, and in particular the implausibly wide range of residuals that they produce. We think this probably reflects, amongst other factors, the high degree of substitution between water resources and elements of "network plus". We are attracted by the suggestion of using wholesale water models, which are generally of higher quality, to set a figure for total allowed expenditure which can then be sub-divided *pro rata* according to the respective water resources and network plus shares of a company's expenditure forecasts.

Another possibility that could be examined is to infer a figure for water resources expenditure by taking the difference between the results obtained from wholesale water and network plus models. However, this could produce perverse results if the model for wholesale water were very different from that for network plus: we think it would be useful, in this case, if the two shared several features in common, including the principal explanatory variables.

In any event we think it could be risky to rely upon the results of any of the water resources models. However, it could be useful to examine them as a "cross-check" upon results obtained from other approaches. The usefulness of disaggregated models as cross-checks holds for other areas as well.

Similarly, for **wastewater**, we would tend to avoid relying on the sub-service models, especially sewage collection, sewage treatment and bioresources plus. We think some bioresources models have promising attributes but we would be reluctant to rely on their results in isolation. Rather, we would advocate a "triangulation" approach involving examination of results from wholesale wastewater, inferred differences between predicted expenditure values from wholesale and network-plus models and the results from the best bioresources equations.

For **retail**, there may be some substitutability issues between debt-related costs and other costs, so we would suggest Ofwat uses whole-retail (totex) models in the first instance and, provided these incorporate the right variables and perform reasonably well, use models of debt costs and other costs to cross-check the results. In any event, given that we think there is scope for improvement in the retail models we will continue to work in this area ourselves and offer any findings that we think may be useful in due course.

Conclusion

In conclusion, it is worth noting one caveat to the comments offered above and in the attached Response Sheet. As we understand it, some adjustments have already been made to the dataset since many of the models included in the consultation were prepared, and the update to include 2017-18 will be a further major change. It is likely that some model characteristics will strengthen, but others will weaken, suggesting that it will be necessary to retain an open mind on the final suite of equations to be used for cost assessment. However, we think that this consultation has provided a sound basis for subsequent development, and we look forward to making further contributions as the work matures.

Yours sincerely

Paul Edwards
Head of Economics & Charges

Appendix 1 – Suggested Additions to Ofwat's "Checklist for Model Acceptability"

The following is an additional "checklist" of "common sense" conditions/criteria to guide Ofwat's further model development.

1) In many models the **coefficient on the scale variable** is materially in excess of unity, and often more than 1.1, apparently implying significant "diseconomies of scale". We do not think that this is plausible: it would not be appropriate to estimate allowed expenditure on the assumption that a company could cut itself in half and reduce overall costs as a result. As a general rule, therefore, if models with coefficients in excess of unity are to be developed, consideration should be given either to constraining the coefficient on the scale variable to unity or examining a unit cost alternative;

2) We noted that in a number of water models companies had added together different sub-measures of "**average pumping head**" to obtain an aggregate figure. As you know, this is invalid because the different sub-measures now use different denominators. As a general rule, therefore, models should not use variables which are derived by adding together values of average pumping head for different parts of the value chain;

3) Companies adopt different approaches to the treatment of **grants and contributions** in defining their explanatory variables. Grants and contributions can be both material and "lumpy", and their inclusion in the definition of expenditure can have a distortionary effect. As a general rule, therefore, we think that expenditure should be defined net of the grants and contributions. However, this should be calculated appropriately for the costs being modelled such that base expenditure models should not be net of grants and contributions that are associated with growth and new connections. The exception would be in the particular circumstances where the model has been designed specifically to cover enhancement expenditure that is related to growth and new development, and explanatory variables that aim to capture the size of that growth are included. ;

4) Care should be given to avoid potential inconsistencies in the treatment of **"customer numbers and/or properties"** in the modelling of retail costs. Since WOCs have only single service customers and WASCs have both single and dual service (and the ratio of one to the other varies a great deal), the use of variables such as "average household bill" and "average revenue per customer" has to make sense alongside other choices such as the denominator in the explanatory variable and/or the definition of the scale variable; and

5) There appears to have been a tendency across the industry to use "all the available data", by which we mean all of the years for which values for relevant variables exist. We think the **choice of how many years' data** to use is a more balanced one: certainly the addition of extra "observations" can help strengthen modelling results, but it is also the case that the longer the dataset the greater the scope for influence from unobserved cost drivers that may impact companies unevenly, creating a risk of spurious results. As a general rule, therefore, any model that is estimated using, several years' data should hold for the last three and the last two years as well.

Appendix 2 – Summary of "Green" models

The following tables summarise the models that we have identified as "Green – Minor or no concerns with the proposed model". There are 57 Green models[1] which is around 15% of all models reviewed. They have a good coverage of cost drivers and we believe that these models could form the basis of a suite of models suitable to be used as part of wider assessment framework to inform cost allowances for 2020-25. The explanatory variables are grouped into the cost driver categories discussed in section 3.2 of CEPA's report published alongside the consultation.

	Nr of models	Green Models		Explanatory factors included in Green models
		Nr	%	
Total	**392**	**57**	**15%**	
Wholesale water	**151**	**21**	**14%**	
Water resources	16	6	38%	Scale - Reservoir capacity Sparsity / density - Sources / DI System characteristics - Propn of DI from reservoirs / boreholes / rivers
Water treatment	10	0	0%	
Water resources plus	12	1	8%	Scale - Conn properties Sparsity / density - Weighted average density System characteristics - Propn of DI from boreholes
Treated water distribution	15	2	13%	Scale - Conn properties, length of mains Sparsity / density – density System characteristics - SRV and Water towers per km of main Levels of activity - % mains refurbished / relined, length of mains laid since 1981
Network plus	48	5	10%	Scale - Conn properties, population served Sparsity / density - props / area, props / area sq, mains per conn prop, mains per conn prop sq, System characteristics - % DI from boreholes / rivers, % treated surface water, % water treated Band 2 and below, % water treated at band 4 and above, % water treated band 5 and above, Levels of activity - % mains refurbished / relined, length of mains laid before 1980
Wholesale water	50	7	14%	Scale - Conn properties, length of mains, population served Sparsity / density - props per km of main, props per km of main squared, mains per conn prop, mains per conn prop sq., km of main per conn prop System characteristics - sources / DI, % DI from reservoirs / rivers, % treated surface water, % water treated at band 1 and below, % water treated at band 2 and below, % treated at band 5 and above, Levels of activity - % mains renewed / relined, length of mains laid before 1980

[1] These models are all work in progress and could change over the next 18 months. Clearly, at this stage 400 models are not manageable so we have taken a very high level approach to reviewing the models to get a more manageable set for further development and review. We believe that 57 models merit progressing further and a further 173 Amber models may cut the mustard with a bit of work or better data. The grading of the models does not imply endorsement as this would require an understanding of the full process and context in which they are to be used and this has not yet been published (for example we have made little comment on the suitability of explanatory factors with regard to their suitability for forecasting data for 2020-25)

	Nr of models	Green Models		Explanatory factors included in Green models
		Nr	%	
Wholesale wastewater	**159**	**31**	**19%**	
Bioresources	37	2	5%	Scale - sludge produced Sparsity / density - % area >2,000 people per km sq, % load in STW band 1-3
Bioresources plus	7	0	0%	
Sewage treatment	13	0	0%	
Sewage collection	12	3	25%	Scale - length of sewers, pumping station capacity Sparsity / density - property density System characteristics - volume of waste collected per pe, regional wages
Network plus	49	12	24%	Scale - Conn props, load, nr of STWs Sparsity / density - Propn of area >4,000 people per sq.km, Propn of area >2,000 people per sq.km, Propn of area <250 people per sq.km, density, load sq., Nr STW squared, load x nr STW, Nr STW per property System characteristics - pumping station capacity per km of sewer, length of sewer / load, nr large tertiary works, nr CSO per km of sewer, ratio of non-res to res population Quality - propn load BOD<10 NH3<1, sum of tight consents
Wholesale wastewater	41	14	34%	Scale - Conn props Sparsity / density - Propn of area >4,000 people per sq.km, Propn of area >2,000 people per sq.km, Propn of area <250 people per sq.km, Nr STW per property System characteristics - pumping station capacity per km of sewer, nr CSO per km of sewer, ratio of non-res to res population Quality - propn load BOD<10 NH3<1 Levels of activity - % growth in number of properties
Retail	**70**	**5**	**7%**	
Bad debt	27	0	0%	
Other expenditure	17	1	6%	Scale - Total customers System characteristics - % dual customers, % metered customers
Totex	26	4	15%	Scale - unique accounts System characteristics - % dual customers, % metered customers, Bill ratio, deprivation measure, IMD score, Income IMD score, average wholesale bill
Enhancement	**12**	**0**	**0%**	
Lead	3	0	0%	
New developments	2	0	0%	
First time sewerage	3	0	0%	
Growth	4	0	0%	