

Comments on ‘Cost assessment for PR19: a consultation on econometric cost modelling’

Dr Karli Glass* and Dr Anthony Glass^{‡,#}

This document is organised into two sections. The first section briefly summarises the content of the Ofwat consultation document in relation to the issues we raise for consideration. The second section discusses three issues to consider as the cost modelling for PR19 evolves.

1 Background in the context of our comments

The issues we raise relate to the following three features of the consultation document.

- a) The appendix of the consultation document reports a large number of models which all use clustered standard errors for statistical inference. The use of clustered standard errors in the consultation document reflects the approach of the Competition and Markets Authority (CMA) to Bristol Water’s PR14 price control appeal.¹
- b) Ofwat note on page 10 of the consultation document that ‘In a small sample...we cannot accommodate many explanatory factors’. The reported models are all therefore parsimonious model specifications, which raises the possibility of omitting important cost drivers.

* School of Business and Economics and Centre for Productivity and Performance, Loughborough University. Email: k.glass@lboro.ac.uk

‡ School of Business and Economics and Centre for Productivity and Performance, Loughborough University. Email: a.j.glass@lboro.ac.uk

The views expressed in this document are those of Karli Glass and Anthony Glass and do not necessarily represent the views of the School of Business and Economics and Centre for Productivity and Performance at Loughborough University.

¹ CMA (2015): ‘Bristol Water plc: A reference under section 12(3)(a) of the Water Industry Act 1991’.

- c) As all the reported models are parsimonious and use clustered standard errors for statistical inference, this raises the issue of the underlying intuition for the selection of these model specifications.

2 Issues to consider

2.1 Standard errors

We recognise that the use of clustered standard errors across all the reported models is in line with the approach of the CMA (2015) and clustered standard errors are beneficial, at least in principle, as they account for correlation in the residuals within a cluster. In practice, the case for using clustered standard errors in the water cost modelling is not clear cut because the sample size is small. On this issue in a recent academic study Ibragimov and Muller (2016) note that when using clustered standard errors, a variable is more likely to be not significant than is actually the case when the sample size is small.² This is supported by evidence of under rejection of the null hypothesis in simulation results for some scenarios. In particular, on page 87 they state with reference to clustered standard errors with small samples that the:

‘...null rejection probability can be substantially below the nominal level α for some values of σ^2 [where σ^2 denotes the variance]’.

2.2 Omitted explanatory variables

As a result of the under-rejection problem with clustered standard errors in small samples together with the parsimonious model specifications in the consultation document, some potentially important explanatory variables may have been omitted. In benchmarking this has implications for the errors and hence the efficiencies. One should therefore exercise caution when omitting variables. For example, when modelling bad debt costs an explanatory variable reflects the probability that a

² Ibragimov, R. and U. K. Muller (2016): ‘Inference with few heterogeneous clusters’. *Review of Economics and Statistics*, vol. 98, pp. 83-96.

customer defaults on paying a water bill. There are a large number of indicators of deprivation and this variable is capturing just one of them. By also including the index of multiple deprivation (IMD) as an explanatory variable a wider range of deprivation indicators would be captured (e.g., crime, health deprivation and disability, education skills and training, and barriers to housing and services).

2.3 Further insights into the current modelling proposals

Ofwat have evidently taken into account a range of different factors and not just used statistical inference to arrive at their models. The statistical inference in the consultation document, however, is based only on clustered standard errors. Given the above evidence for small samples that clustered standard errors under reject the null hypothesis that a coefficient is zero, we propose a simple two-step process to help ensure a more systematic approach to the statistical inference.

It is clear that Ofwat are placing a great emphasis on a high level of transparency in the consultation process by providing underlying data, details of its own models and various consultancy reports. Reporting the results of the following two-step process would also provide greater intuition on how Ofwat arrived at its set of preferred model specifications.

- (i) Due to concerns about omitting valid cost drivers because of the under-rejection problem with clustered standard errors in small samples, we suggest that selection of the preferred model specifications in terms of the choice of explanatory variables should be based on models where the standard errors are not clustered.
- (ii) Run the preferred model specifications from (i) with clustered standard errors. Comparing corresponding estimated models from steps (i) and (ii) will indicate the impact of using clustered standard errors on the significance of the explanatory variables. In theory, whether the standard errors are clustered or not should not affect the parameter estimates and thus the errors and efficiencies. Reporting models with and without clustering of the standard errors will establish whether this is the case in practice.