Response to the CMA's Approach to Base Cost Benchmarking

Prepared by: Portsmouth Water

Introduction

On 9 October 2025, the CMA published its Provisional Determinations for the PR24 Water appeals. As a water company operating under the same regulatory regime as the disputing companies, Portsmouth Water is an interested third party. This note provides a response to one key element of the Provisional Determinations - the CMA's approach to base cost modelling and, in particular, the CMA's use of the LASSO regression technique.

We focus on this issue as we have material concerns about the CMA's significant reliance on it to determine base cost allowances. In particular:

- the LASSO approach taken by the CMA represents a substantial departure from the Final Determinations without recognising the limitations and sufficiently justifying the CMA's choices;
- the approach produces counterintuitive base cost models;
- the large changes in efficiency scores resulting from the LASSO modelling have not been sufficiently considered or explained;
- the use of LASSO has the potential to establish harmful regulatory precedent; and
- overall, the approach represents a move away from the direction of travel recommended by the Independent Water Commission.

We are not suggesting that the CMA abandon the use of LASSO at this stage of the redetermination and recognise that LASSO provides a useful tool for assessing the wide range of issues raised by the appealing companies in an integrated way. However, we would ask the CMA to acknowledge the limitations of LASSO and the need to apply it with caution. We would also ask the CMA to make clear that its adoption in this case should not be viewed as precedent for how cost modelling should be carried out at future regulatory price reviews. These should be built on a detailed understanding of the engineering and economic drivers of costs and subject to extensive consultation, peer review and challenge and reflect the conclusions of the Independent Water Commission with regards to the over-reliance on modelling.

Issues with the approach used in the Provisional Determinations

In the PR24 Provisional Determination[1], the CMA has used LASSO (Least Absolute Shrinkage and Selection Operator) to consolidate Ofwat's suite of base cost models into a single model for wholesale wastewater and two models for wholesale water.

Ofwat's approach to modelling base costs at PR24 was based on the average modelled costs across a suite of regression models. Each model draws on different cost drivers and functional forms. In contrast, in the Provisional Determinations LASSO is used to select explanatory variables within a single model, allowing the model itself to determine which drivers are retained or excluded based on statistical fit.

The approach therefore places significantly greater emphasis on automated, data-driven variable selection rather than using engineering and economic rationale to inform these choices. It also reduces reliance on multiple models, giving considerably greater weight to the outcomes of a single model specification. We have a number of concerns about the use of LASSO modelling, and we provide detail on each of these below.

The LASSO approach taken by the CMA represents a substantial departure from the PR24 Final Determinations, without sufficient justification or recognition of the limitations of the new approach

The models used by the CMA represent a significant departure from Ofwat's suite of models and established benchmarking approach. The PR24 cost modelling approach was extensively consulted on, and builds on methodologies developed over the course of many price controls. The LASSO approach by contrast has not previously been employed within the context of water price controls in England and Wales, as far as we are aware.

While substantial methodological change is not an issue in principle, and is within the remit of the CMA, the evidential bar for introducing such changes should be high. In our view, a number of the adjustments introduced in the Provisional Determinations have not been sufficiently justified, and the revised approach brings new limitations and drawbacks that have not been adequately recognised. We note two key examples of this:

- **The treatment of input prices together with scale variables within the econometric model (rather than as a post-modelling adjustment) represents a significant methodological change that has not been explored in sufficient detail**. Ofwat's approach involved setting an ex-ante Real Price Effect (RPE) allowance through a post-modelling adjustment. In contrast, the CMA has moved input price variables for energy and labour within the econometric models, interacting them with selected scale variables.[2,3] We note that the Provisional Determinations do not set out the rationale for selecting these specific scale variables, or considered alternatives (e.g. connected properties, sewer length). The CMA itself notes that these inclusions have a material effect on the catch-up efficiency challenge,

[1] CMA, Water PR24 References, Provisional Determinations, Volume 1
[2] The CMA includes a regional wage variable based on ONS Annual Survey of Hours and Earnings (ASHE) regional manufacturing wages, and an energy price index derived from Department of Energy Security and Net Zero (DESNZ) data.
[3] Energy is interacted with the length of mains in wholesale water models and with pumping capacity in wastewater models, while wages are interacted with the length of mains in wholesale water models and with load in wastewater models.

underlining the need for more comprehensive exploration of alternative approaches, sensitivity testing, and engineering/economic input to ensure the chosen approach is robust.

- **The choice of penalty parameter (λ) within the LASSO framework also represents an important regulatory judgement** which should be recognised as such and warrants greater transparency. In the context of benchmarking, this parameter determines how material a cost driver must be in terms of its statistical contribution to be retained in the model. While the CMA's choice follows established statistical practice[4], the Provisional Determinations do not demonstrate sufficient testing of the sensitivity of results to alternative λ values. Given its material impact on the selected model, this decision should be explicitly justified and accompanied by appropriate sensitivity analysis.

The approach produces counterintuitive base cost models

In principle, a drawback of LASSO is that it will select variables based only on their statistical properties and not on the underlying economic or engineering rationale for whether and how those variables will drive costs. In practice, this drawback is illustrated by the fact that the CMA's LASSO modelling has produced outcomes that are not intuitive and lack clear interpretation. For example:

- All the CMA's base cost models include multiple measures of population density whereas each of Ofwat's base models relied on a single density measure (and its associated quadratic term). The preferred models include *all* candidate density drivers simultaneously. For example, both preferred models for wholesale water include properties per kilometre of main, weighted average density at the local authority district (LAD) level, and weighted average density at the middle-layer super output area (MSOA) level at the same time. There is no clear economic or engineering rationale for why all these variables would be needed to explain costs. This is also problematic from a statistical standpoint, as including several highly correlated variables in the same model can produce a high degree of model instability in repeated samples. Moreover, several of the models also include squared terms without a corresponding linear term, raising further questions around model interpretability.

- Similarly, in the water resources model, a correction published by the CMA notes that while the coefficient on average volume per water treatment works (WTW) is retained, its sign is positive. That is, for every 1% increase in average volume per WTW, the CMA's model suggests that costs increase by around 0.08%, holding other scale factors (such as connected properties) constant. In the original Provisional Determinations, the estimated coefficient was negative, and the CMA noted a negative coefficient was a sign in the 'expected direction'. Based on this reasoning, the positive sign suggests there are diseconomies of scale in water treatment, which would be contrary to the economic and engineering logic and is therefore counterintuitive.

These examples highlight the risk that LASSO-based models can prioritise statistical properties over interpretability, undermining confidence in their ability to predict costs on a forward-looking basis, and their regulatory usefulness. In general, LASSO is also highly sensitive to dataset and specification changes. Even small changes in the dataset, such as removing a single year of data or

[4] The CMA choose the largest value of λ for which the cross-validated mean squared error (MSE) remains within one standard error of the minimum.

redefining a variable, can lead to different variables being selected or excluded, resulting in materially different outcomes.

The large changes in efficiency scores resulting from the LASSO modelling have note been sufficiently considered or explained

The CMA's modelling produces large changes in efficiency scores and allowances for a number of non-disputing companies, relative to Ofwat's Final Determination models. These movements have not been fully interrogated, understood or explained in the Provision Determinations. Given the material impact that these changes have on base cost allowances, it is important that there is a clear understanding of what is driving them, and whether those drivers make sense.

For example, Table 1 illustrates the differences in assessed efficiency between Ofwat's Final Determinations and the CMA's redetermination for the four non-disputing companies with the largest changes in ranking. This illustrates that, if applied across the sector, the CMA's models would produce materially different settlements from those previously accepted by non-disputing companies. Given the comparative purpose of cost benchmarking, a full assessment of such efficiency shifts across all companies, not just disputing companies, should be a pre-requisite for any methodological departure of this scale.

Table 1. Changes in relative efficiency for selected companies

Company	Ofwat PR24 Efficiency	CMA Efficiency	Change in rank	% change allowance
Affinity Water	0.97	1.01	+5	-12%
Bristol Water	1.10	0.91	-9	+12%
Thames Water	0.99	1.04	+8	-11%
Yorkshire Water	1.12	0.99	-7	+7%

Source: Ofwat PR24 Final Determination and CMA PR24 Provisional Redetermination

The use of LASSO has the potential to establish harmful regulatory precedent

We are concerned that the CMA's decision to employ LASSO as the primary basis for its cost benchmarking could be seen as a regulatory precedent. The use of LASSO in this context reduces transparency in regulatory decision-making. Model design choices are determined by automated, data-driven selection rather than through engagement and consideration of the underlying economic and engineering drivers of cost. As a result, model selection is dictated by statistical properties (such as model fit) rather than economic and engineering rationale that have underpinned previous regulatory benchmarking exercises. This lack of transparency risks weakening the shared understanding between regulators and companies regarding the rationale for cost differentials.

Such an approach also makes it more difficult for companies to engage meaningfully with the regulator, both in understanding the determinants of their modelled efficiency and in raising legitimate company-specific adjustments to the benchmarking approach. For example, while LASSO is designed to prioritise variables with strong predictive properties across the whole

sample, it may fail to capture cost drivers that are only important in certain regions or operating environments. As a result, the model may exclude variables that are statistically weak at the sector level but nonetheless material for a subset of companies, reducing the scope for those companies to demonstrate legitimate differences in underlying costs. It may also fail to capture cost drivers that may not have been important in the past but are likely to be important in the future (as companies face new challenges driven by climate change and population growth).

Overall, the approach represents a move away from the direction of travel recommended by the Independent Water Commission

The Independent Water Commission (IWC) review, published earlier this year, provided a diagnosis of longstanding issues with the sector's regulatory framework and set out a series of recommendations for reform. With respect to cost assessment, the review pointed to an overreliance on econometric benchmarking, a 'one-size-fits-all' approach and lack of consideration of engineering evidence.[5] The review also emphasised that there are "*fundamental limits to how precise and accurate a benchmarking framework and econometric tools can be*".[6]

To remedy these issues, the IWC recommended that econometric benchmarking be balanced with company-specific evidence and expert supervisory judgement. While we acknowledge that it is not for the CMA to implement the recommendations of the IWC, its approach nevertheless carries important implications for the sector's future regulatory direction which we believe it is important for the CMA to note in its final decisions.

As the IWC recognises, econometric benchmarking is an inherently backward-looking exercise, relying on historical relationships as predictors of future performance.[7] Increasing reliance on statistical methods such as LASSO risks doubling down on this backward-looking focus, with insufficient recognition of forward-looking considerations such as resilience and asset health. This is particularly important given the growing need for sustained investment to address ageing infrastructure and long-term service risks.

The IWC also observed that elements of Ofwat's current regulatory framework may inadvertently discourage companies from submitting business plans based on their actual investment needs, for fear of financial penalties, and that this was exacerbated by Ofwat's Quality and Ambition Assessment (QAA).[8] We agree that the current framework deters companies from putting forward justified and necessary proposals for asset maintenance or enhancement, and reinforces short-term cost-minimising incentives at the expense of long-term resilience and asset health.

Consistent with the IWC's recommendations, we agree that econometric models should therefore form only one part of the regulatory toolkit, complemented by forward-looking assessments and supervisory judgement. In our view, the CMA's use of LASSO represents a step away from the direction of reform advocated by the IWC, diminishing the role of expert supervisory judgement in designing models and interpreting results. We are concerned that following such an approach in future price controls could exacerbate some of the issues identified by the IWC, rather than helping to resolve them.

[5] Independent Water Commission (July 2025) Final Report, paras. 387 & 392
[6] Independent Water Commission (July 2025) Final Report, para. 388
[7] Independent Water Commission (July 2025) Final Report, para. 387
[8] Independent Water Commission (July 2025) Final Report, para. 386

Conclusion

Given the concerns outlined above, we would ask the CMA to recognise explicitly the limitations of LASSO in its Final Determinations. We would ask the CMA to confirm that its use in the context of a CMA redetermination, does not imply it is an appropriate methodology for future price controls, particularly given the conclusions of the IWC with respect to econometric modelling and asset health. Thus, it should not be treated as precedent to be followed by regulators.

6 November 2025