Yorkshire Water
Western House
Halifax Road
Bradford
West Yorkshire
BD6 2SZ

Ofwat
Centre City Tower
7 Hill Street
Birmingham
B5 4UA

By email: CostAssessment@ofwat.gov.uk

12 January 2023

Dear Ofwat

Yorkshire Water BOTEX+ Cost Model Submission

Thank you for the opportunity to submit BOTEX+ econometric cost models to inform the model consultation in the Spring. We have uploaded the requested material for the Water, Wastewater and Residential Retail price controls to your Sharepoint site.

As part of this submission, we have worked with Oxera to explore a range of alternatives to the PR19 models for assessing efficient expenditure in PR24. We consider that the models presented in this submission perform well against your proposed criteria and represent an improvement on (or are comparable to) the PR19 models. We have primarily focussed on capturing key differences between companies on an outturn basis (e.g. scale, density). Where possible, we have sought to capture some of the cost pressures that companies (including YWS) will face in AMP8 (e.g. increased phosphorus complexity) but this will need to be developed further as additional (outturn and business plan) data becomes available.

When companies were invited to submit cost assessment models, it was unclear as to how Ofwat would model bioresources expenditure (specifically in relation to the treatment of financing costs). Ofwat's approach was later finalised in the PR24 methodology document. While we have not yet developed robust bioresources models (and, therefore, do not present bioresources as part of this submission), we are continuing to explore alternatives to the PR19 models, and we look forward to contributing to the discussion on bioresources modelling during the consultation in Spring 2023.

We summarise our general comments and some specifics for each price control below:

General Comments

- We have some reservations with Ofwat's modelled cost definition, specifically on the inclusion of network reinforcement expenditure in the wholesale models. We consider that network reinforcement needs can differ across companies (and over time) for reasons unrelated to general population growth. We consider that, if network reinforcement remains in the modelled cost definition, some form of post-modelling adjustment may be required.

- In line with Ofwat's guidance on endogeneity, the models that we propose do not account for differences in service quality, either between companies or over time. This means that the base models can only fund the level of service achieved in the historical period by efficient companies and cannot disentangle, amongst other things, service improvements that were achieved historically by companies through additional enhancement funds from those achieved through productivity improvement. If Ofwat intends to set performance commitments beyond those achieved by the industry in AMP7, then a post-modelling adjustment may be required. We note that this adjustment may be complicated by the fact that there is no singular measure of service quality, and that companies may perform well (or poorly) across different service measures which may be driven by exogenous factors.

Water

- We consider that Ofwat's WRP models can be improved by modelling 'weighted average complexity' in levels, as opposed to logarithms. Doing so leads to an improvement in model fit, and allows for the coefficient to be more readily interpreted and validated against operational expectations. Relatedly, we are assessing whether the weights used to construct the weighted average complexity variable are aligned with operational intuition, and we encourage Ofwat to ensure that the weights are appropriate in PR24. As an example, we have presented an alternative weighting system that leads to an improvement in model performance, although this alternative will also need to be validated from an operational perspective.

- We find that connected properties can be a valid alternative measure of scale in the treated water distribution models. Using this cost driver leads to a slight improvement in model quality, and may better capture (implicitly) some of the costs associated with population growth. We find equivalent results in the sewage collection models in wholesale wastewater.

- Ofwat has shared several measures of 'weighted average density' in the latest dataset, and additional measures of density can be constructed using the data published in companies' APRs (e.g. properties per lengths of main). We note that properties per lengths of main typically performs well in the treated water distribution models compared to the weighted average density measures, and an equivalent measure of density was used in Ofwat's wholesale wastewater models in PR19. With respect to the choice of specific weighted average density measure, we could not find a strong operational or statistical argument to support one over the other. As such, we have presented models using Ofwat's PR19 measure, and alternatives that use a more granular version of Ofwat's PR19 measure.

Waste Water

- We consider that Ofwat's PR19 models did not sufficiently account for the age of company networks and the corresponding maintenance and renewal requirements in sewage collection. We believe that this factor is exogenous in the short term and we find that the estimated relationship between asset age and BOTEX+ is consistently statistically significant and operationally intuitive (directionally) across specifications. We consider that customers are protected from endogeneity concerns (i.e. companies deliberately undermaintaining assets to perform better in the cost assessment models) by performance commitments related to asset health and service. If maintenance and renewal requirements are omitted from the cost assessment models in PR24, then a post-modelling adjustment may be required to support additional (or reduced) needs.

- Several companies (including YWS) are experiencing or will experience an increase in Phosphorus removal (P-removal) activity through the WINEP. However, Ofwat's PR19 models do not account for P-removal requirements which have a significant operating cost impact. We have explored incorporating P-removal activity directly in the econometric models, although the estimated relationship between P-removal activity and BOTEX+ is often statistically insignificant due to limited variation in historical data. As such, we have proposed models that account for a composite complexity measure that can account for both phosphorus- and ammonia-related treatment complexity. Such data limitations may not exist once new data (e.g. additional outturn data, AMP8 business plans) is available in PR24, such that P-removal (and other relevant treatment complexities) can be modelled directly.

- Unlike at PR19, we find that network plus models can perform relatively well against Ofwat's stated criteria, when compared to the more granular models. In particular, the model fit is typically higher in the network plus models and the range of estimated efficiency scores is lower, perhaps indicating that modelling at the network plus level better accounts for operational trade-offs and cost allocation issues than the more granular models. As such, we consider that network plus models should form part of Ofwat's suite of models when assessing wastewater expenditure in PR24.

Retail

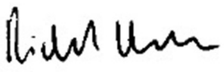
- Based on the current dataset, we agree with Ofwat's provisional assessment that 'bottom-up models' do not perform particularly well when compared to the TOTEX models, especially regarding 'other operating costs'. However, this finding may change once new data is released. Moreover, bottom-up models could provide valid cross-checks to the TOTEX models. As such, we consider that Ofwat should not abandon bottom-up modelling entirely at this stage.

- With respect to the TOTEX models, we find that several drivers of expenditure are statistically insignificant at the standard thresholds. Nonetheless, we include some statistically insignificant cost drivers in the models because it improves the models in other ways (e.g. their inclusion also improves the significance of other coefficients). Moreover, as we find statistically significant economies of scale across model specifications, we have modelled retail costs on an aggregate basis for this submission, which appears to improve some of the diagnostic test results.

- In our analysis, we found that individual companies' performances can be affected by the choice of deprivation measure, yet we found no strong statistical or operational evidence to select a single 'best' deprivation measure. As such, we have constructed 'composite deprivation measures' that could mitigate the risk that companies are over- or under-funded based on a somewhat arbitrary decision on the most appropriate deprivation measure. These composite measures typically perform well in the TOTEX models, and appear to outperform the use of any individual measure of deprivation.

We hope you find this submission useful. If you have any further questions, please send them to regulation@yorkshirewater.co.uk. We look forward to engaging with the Spring 2023 consultation when it is available.

Yours faithfully,



Richard Hepburn

Head of Regulation
Yorkshire Water
richard.hepburn@yorkshirewater.co.uk

Yorkshire Water – WRP, TWD and WW models

Econometric model formula:

1. YWS_WRP1: ln(WRP BOTEX) = $\alpha + \beta_1$ ln(properties$_{it}$) + β_2 (weighted average complexity)$_{it}$ + β_3 ln(weighted average density (LAD)$_{it}$) + β_4 (ln(weighted average density (LAD)$_{it}$))^2 + ε_{it}

2. YWS_WRP2: ln(WRP BOTEX) = $\alpha + \beta_1$ ln(properties$_{it}$) + β_2 (weighted average complexity(log weights))$_{it}$ + β_3 ln(weighted average density (LAD)$_{it}$) + β_4 (ln(weighted average density (LAD)$_{it}$))^2 + ε_{it}

3. YWS_TWD1: ln(TWD BOTEX+) = $\alpha + \beta_1$ ln(lengths of main$_{it}$)) + β_2 ln(booster pumping stations per lengths of main$_{it}$) + $\beta_3$ln(connected properties per lengths of main$_{it}$) + β_4 (ln(connected properties per lengths of main$_{it}$))^2 + ε_{it}

4. YWS_TWD2: ln(TWD BOTEX+) = $\alpha + \beta_1$ ln(properties$_{it}$)) + β_2 ln(booster pumping stations per lengths of main$_{it}$) + β_3 ln(weighted average density (LAD)$_{it}$) + β_4 (ln(weighted average density (LAD)$_{it}$))^2 + ε_{it}

5. YWS_TWD3: ln(TWD BOTEX+) = $\alpha + \beta_1$ ln(lengths of main$_{it}$)) + β_2 ln(booster pumping stations per lengths of main$_{it}$) + β_3 ln(weighted average density(MSOA (pop))$_{it}$) + β_4 (ln(weighted average density(MSOA (pop))$_{it}$))^2 + ε_{it}

6. YWS_WW1: ln(WW BOTEX+) = $\alpha + \beta_1$ ln(properties$_{it}$)) + β_2 (weighted average complexity(log weights)$_{it}$) + β_3 ln(booster pumping stations per lengths of main$_{it}$) + $\beta_4$ln(weighted average density(LAD)$_{it}$) + β_5 (ln(weighted average density(LAD)$_{it}$))^2 + ε_{it}

7. YWS_WW2: ln(WW BOTEX+) = $\alpha + \beta_1$ ln(properties$_{it}$)) + β_2 (weighted average complexity(log weights)$_{it}$) + β_3 ln(booster pumping stations per lengths of main$_{it}$) + $\beta_4$ln(weighted average density(MSOA (pop)$_{it}$) + β_5 (ln(weighted average density(MSOA (pop)$_{it}$))^2 + ε_{it}

Description of the dependent variable

The models have been developed based on Ofwat's cost definitions in the PR24 methodology.

The dependent variables are defined as per Ofwat's consultation analysis files i.e. the sum of:

- Power
- Income treated as negative expenditure
- Bulk Supply
- Renewals expensed in year (infrastructure)
- Renewals expensed in year (non-infrastructure)
- Other operating expenditure excluding renewals
- Maintaining the long-term capability of assets (infrastructure)
- Maintaining the long-term capability of assets (non-infrastructure)
- Addressing low pressure enhancement costs
- Atypical expenditure
- Network reinforcement

It excludes the following cost categories:

- Costs associated with the Traffic Management Act
- Statutory water softening
- NRSWA diversions (non-S185)
- Other non-S195 diversions
- Developer services base cost adjustment

This is also consistent with Ofwat's PR24 methodology.

Description of the explanatory variables

- Connected properties (sum of BN2221 and BN2161).
- Lengths of main (BN1100)
- Properties per lengths of main (Connected properties divided by lengths of main)
- Weighted average complexity, as calculated in Ofwat's analysis files
- Weighted average complexity with logarithmic weights: calculated based on Ofwat's weighted average complexity variable, using logarithmically instead of linearly increasing weights.
- Weighted average density (LAD): Weighted average density suggested in Ofwat's draft methodology, calculated per local authority district with population weights as reported in the published wholesale dataset.
- Weighted average density (MSOA (pop)): Weighted average density, calculated per Middle Super Output Area with population weights as reported in the published wholesale dataset.
- Booster pumping stations per lengths of main (BN11390 divided by lengths of main)

Brief comment on the models

We consider that the models presented in this submission perform well against Ofwat's proposed criteria and represent an improvement on (or are comparable to) the PR19 models. We have primarily focussed on capturing key differences between companies on an outturn basis (e.g. scale, density). Where possible, we have sought to capture some of the cost pressures that companies (including YWS) will face in AMP8 but this will need to be developed further as additional (outturn and business plan) data becomes available.

We note that the models may not sufficiently account for factors that are expected to affect efficient costs in AMP8 for some companies or the entire industry, nor capture the impact of certain exogenous factors on an outturn basis for individual companies.

Water resources plus (WRP)

Relative to the PR19 models, the WRP model fit improves when the weighted average complexity is controlled for in levels instead of logarithms. In addition, the interpretability of the coefficient improves as the coefficient on the weighted average complexity (in levels) estimates the impact of moving 1% of total water treated from complexity band 'x' to complexity band 'y' on predicted costs (in percentages) as the coefficient multiplied by the difference in the complexity bands (y – x). When the variable is modelled in logarithms (as per the PR19 models), the interpretability of the coefficient is less clear[1], making it harder to validate the estimated relationship between complexity and expenditure with operational expectations. As such, we suggest that Ofwat reviews the construction of the weighted complexity measure for PR24.

The weighted average complexity measure has advantages relative to measures capturing the proportion of water treated at different complexity levels or proportion of water from different sources. Hence, we have focussed on it. We however note that the weights used in PR19 to construct the weighted average density measure are somewhat arbitrary. As an alternative weighting method, we propose a weighted average complexity variable where the weights increase logarithmically as opposed to linearly. In this way, the cost impact of (for example) increasing complexity from band 2 to band 3 is larger than the cost impact of increasing complexity from band 5 to band 6. We present this variable as an example for how an alternative weighting system can work within the cost assessment models—we are assessing whether logarithmically increasing weights are aligned with operational expectations, or whether an alternative weighting system would be more appropriate. We ask that Ofwat also validates whether the weights used to construct its weighted average

[1] When modelling in logarithms, the cost impact of moving water between the treatment complexity bands depends on the current level of complexity.

complexity variable are aligned with operational expectations and improves upon the weighting approach where necessary.

Treated water distribution (TWD)

We propose cost models that differ from Ofwat's PR19 models in the following respects.

1. Ofwat controlled for lengths of main as the primary scale variable in its TWD model at PR19, arguing that it performed better than connected properties from a statistical perspective. In the current dataset, length of mains do not outperform connected properties as a scale driver, so we present a model that controls for connected properties as well. Apart from an improved statistical performance, connected properties may also better capture some of the costs associated with network reinforcement activity.
2. We consider that an alternative, asset-based measure of density (i.e. connected properties per lengths of main) can perform well in the TWD model: when replacing Ofwat's density measure in the PR19 TWD model with properties per lengths of main, the model fit marginally improves and the p-value of the RESET test increases (indicating improved performance). Ofwat proposed connected properties per lengths of main for wholesale water in the PR19 modelling consultation for treated water distribution[2] and used an equivalent measure to assess sewage collection costs in wholesale wastewater.[3]
3. With the new dataset, Ofwat published alternative measures for the weighted average density cost driver that differ with respect to: (i) the level of granularity;[4] and (ii) the weighting approach.[5] We note that the models perform similarly (e.g. with respect to model fit, range of efficiency scores and other diagnostics), regardless of the choice of weighted average density measure. However, as MSOA measures are more granular than LAD measures, one may expect that MSOA measures more accurately reflect the density of a company's operating environment. Therefore, we present a selection of models using MSOA measures of density in this consultation. We are still assessing whether population-weighted or area-weighted density is a more operationally appropriate driver of companies' costs. For this consultation, we present models with population-weighted density, in line with PR19 precedent.

[2] For example, see Ofwat (2018). Cost assessment for PR19: a consultation on econometric cost modelling. Appendix 1 – Modelling results. March 2018. p. 13 and p. 53.
[3] See Ofwat (2019), 'PR19 Final Determinations: Securing cost efficiency technical appendix', December, Table A2.2.
[4] At PR19, Ofwat's constructed the weighted average density measure using local authority district (LAD) data. Some of the new measures use more granular population distribution data at the Middle Super Output Area (MSOA) level.
[5] When aggregating the population density estimates for each statistical area within a company's operating region, Ofwat weighted the statistical area by the population within that area ('population-weighted') at PR19. Some of the new measures are instead weighted by the geographical size of the statistical area ('area-weighted').

Note that we continue to estimate a reasonable 'U-shaped' relationship between density and TWD costs with all of the proposed density measures, in line with statistical evidence (i.e. the coefficient on the squared term is positive and significant across a range of specifications) and PR19 precedent.

Wholesale water (WW)

We have incorporated the insights from the WRP and TWD modelling to present examples of feasible models at the WW level. As WW is simply the sum of WRP and TWD, the development needs outlined in the sections above also apply here.

The models do not explicitly account for network reinforcement and other growth enhancement requirements. We note that Ofwat has excluded one growth driver (new connections) from the base modelling dataset, despite this driver being available in the PR19 modelling datasets. We consider that, if Ofwat does not account for network reinforcement or growth requirements explicitly in its cost assessment models, then it should consider post-modelling adjustments. Such an adjustment may not be symmetrical if the industry as a whole is expecting to increase reinforcement or growth activity in AMP8.

General comment on model limitations

As the models are estimated on outturn data and do not explicitly account for measures of service quality, the models will only fund companies to deliver the level of service achieved by the industry average (or the benchmark companies) in the historical period. The analysis of the level of service funded through base expenditure is complicated by the fact that there is no single measure of service, and different companies can perform well on different service measures. Similarly, the cost requirements for improving service quality may depend on the service measure being examined (e.g. it may be more or less expensive to improve internal sewer flooding as opposed to pollution incidents), the level of service already achieved by an individual company (i.e. the marginal cost of improving service may increase as the level of service increases) and exogenous factors (e.g. it may be more or less difficult to improve leakage in densely or sparsely populated areas).

We understand that Ofwat is reluctant to control for service measures explicitly in its cost assessment models. Nonetheless, if unjustified performance commitments are set for PR24 (i.e. beyond those already achieved by the industry), then some form of cost adjustment will be required. Alternatively, in the PR24 final methodology,[6] Ofwat appears open to adjusting some of the performance targets for exogenous factors—this could partly alleviate some of the issues relating to the cost–quality disconnect.

[6] Ofwat (2022), 'Creating tomorrow, together: Our final methodology for PR24 Appendix 9 – Setting expenditure allowances', December, p. 61.

	YWS_WRP1	YWS_WRP2	YWS_TWD1	YWS_TWD2	YWS_TWD3	YWS_WW1	YWS_WW2
Dependent variable	BOTEX (WRP)	BOTEX (WRP)	BOTEX+ (TWD)	BOTEX+ (TWD)	BOTEX+ (TWD)	BOTEX+ (WW)	BOTEX+ (WW)
Connected properties (log)	1.068*** (0)	1.064*** (0)		1.098*** (0)		1.042*** (0)	1.056*** (0)
Lengths of main (log)			1.072*** (0)		1.026*** (0)		
Weighted average complexity	0.118 (0.135)						
Weighted average complexity (log weights)		0.352** (0.031)				0.268*** (0.006)	0.356*** (0)
Booster pumping stations per lengths of main (log)			0.488*** (0.001)	0.540*** (0.006)	0.433*** (0.001)	0.495*** (0.002)	0.341** (0.013)
Weighted average density LAD (log)	-1.361*** (0.005)	-1.383*** (0.002)		-2.606*** (0)			-1.844*** (0)
Weighted average density LAD (log), squared	0.082*** (0.009)	0.085*** (0.004)		0.190*** (0)			0.129*** (0)
Weighted average density MSOA (pop) (log)					-5.561*** (0)	-4.166*** (0.002)	
Weighted average density MSOA (pop) (log), squared					0.393*** (0)	0.268*** (0.001)	
Properties per lengths of main (log)			-14.921*** (0)				
Properties per lengths of main (log), squared			1.898*** (0)				
Constant	-5.917*** (0)	-5.778*** (0)	25.065*** (0)	0.13 (0.935)	15.638*** (0.002)	8.257 (0.118)	-2.366*** (0.002)
Estimation method (OLS or RE)	RE	RE	RE	RE	RE	RE	RE
N (sample size)	187	187	187	187	187	187	187
Model robustness tests							
R2 adjusted	0.91	0.912	0.958	0.956	0.952	0.966	0.972
RESET test	0.329	0.34	0.489	0.157	0.122	0.072	0.158
VIF (max)	200.33	200.739	733.12	208.426	496.844	528.857	208.851
Pooling / Chow test	0.999	0.999	0.903	0.673	0.873	0.971	0.722

Normality of model residuals	0.59	0.53	0.738	0.693	0.014	0.493	0.601
Heteroskedasticity of model residuals	0	0	0.004	0.008	0.046	0	0
Test of pooled OLS versus Random Effects (LM test)	0	0	0	0	0	0	0
Efficiency score distribution (min and max)	50–198%	51–198%	74–138%	80–146%	75–142%	75–151%	79–139%
Sensitivity of estimated coefficients to removal of most and least efficient company	A [weighted average complexity becomes more significant]	A [weighted average complexity becomes more significant]	G	G	G	A [weighted average complexity becomes less, constant becomes more significant]	A [booster pumping station per length of main becomes more significant]
Sensitivity of estimated coefficients to removal of first and last year of the sample	G	G	G	G	G	G	G

Efficiency scores

1. YWS_WRP1

Rank	Company	Efficiency score
1	SSC	50.05%
2	PRT	70.72%
3	ANH	73.78%
4	AFW	81.33%
5	SEW	97.59%
6	HDD	102.56%
7	YKY	104.03%
8	TMS	104.47%
9	NES	106.93%
10	WSH	108.24%
11	BRL	110.23%

12	SVE	110.46%
13	SWB	114.91%
14	NWT	118.87%
15	WSX	120.66%
16	SES	175.43%
17	SRN	197.83%

2. YWS_WRP2

Rank	Company	Efficiency score
1	SSC	51.02%
2	PRT	71.18%
3	ANH	74.48%
4	AFW	82.28%
5	SEW	98.52%
6	HDD	102.29%
7	YKY	103.24%
8	TMS	105.22%
9	NES	106.94%
10	SVE	109.66%
11	WSH	110.94%
12	BRL	112.74%
13	SWB	115.71%
14	NWT	117.81%
15	WSX	121.37%
16	SES	171.20%
17	SRN	198.06%

3. YWS_TWD1

Rank	Company	Efficiency score
1	PRT	73.68%
2	SWB	88.00%
3	NWT	88.17%
4	SRN	91.38%
5	HDD	99.34%
6	NES	101.29%

7	SVE	105.91%
8	AFW	106.92%
9	TMS	108.46%
10	YKY	109.28%
11	WSX	110.46%
12	WSH	113.00%
13	SSC	115.02%
14	SEW	116.05%
15	BRL	131.29%
16	SES	131.72%
17	ANH	137.81%

4. YWS_TWD2

Rank	Company	Efficiency score
1	PRT	79.57%
2	NWT	86.37%
3	SRN	90.67%
4	SWB	94.06%
5	WSX	94.57%
6	SVE	99.98%
7	NES	102.04%
8	AFW	105.14%
9	TMS	110.13%
10	YKY	111.05%
11	HDD	112.58%
12	SSC	117.01%
13	SEW	118.01%
14	WSH	124.86%
15	SES	128.84%
16	BRL	131.50%
17	ANH	145.65%

5. YWS_TWD3

Rank	Company	Efficiency score
1	PRT	74.98%

2	SWB	79.91%
3	SRN	83.25%
4	NWT	89.71%
5	NES	99.94%
6	HDD	104.51%
7	WSX	105.37%
8	TMS	107.04%
9	YKY	109.44%
10	SVE	109.81%
11	WSH	112.74%
12	BRL	114.34%
13	SEW	114.42%
14	AFW	121.90%
15	SES	130.58%
16	ANH	136.20%
17	SSC	142.46%

6. YWS_WW1

Rank	Company	Efficiency score
1	PRT	75.47%
2	SSC	89.61%
3	AFW	91.97%
4	SWB	98.20%
5	NWT	101.34%
6	HDD	102.63%
7	YKY	102.66%
8	NES	103.03%
9	SVE	104.48%
10	BRL	106.23%
11	SEW	106.40%
12	WSH	108.18%
13	TMS	110.62%
14	ANH	115.71%
15	SRN	117.89%
16	WSX	132.03%
17	SES	150.71%

7. YWS_WW2

Rank	Company	Efficiency score
1	PRT	79.44%
2	SSC	87.48%
3	AFW	92.09%
4	NWT	99.37%
5	SWB	99.58%
6	HDD	99.94%
7	NES	100.90%
8	SVE	104.53%
9	YKY	105.14%
10	SEW	107.16%
11	ANH	109.89%
12	TMS	111.61%
13	BRL	111.65%
14	WSH	111.87%
15	WSX	115.04%
16	SRN	127.54%
17	SES	138.80%

Yorkshire Water – SWC, SWT and NPWW models

Econometric model formula:

1. YWS_SWC1: ln(SWC BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$) + β_2 ln(pumping capacity per sewer length$_{it}$)+ β_3 ln(connected properties per sewer length$_{it}$) + β_4 (% mains laid after 2001$_{it}$) + ε_{it}

2. YWS_SWC2: ln(SWC BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$) + β_2 ln(connected properties per sewer length$_{it}$)+ β_3 ln(connected properties per sewer length$_{it}$)^2 + β_4 (% mains laid after 2001$_{it}$) + ε_{it}

3. YWS_SWC3: ln(SWC BOTEX+) = $\alpha + \beta_1$ ln(Connected properties$_{it}$) + β_2 ln(pumping capacity per sewer length$_{it}$)+ β_3 ln(weighted average density$_{it}$)+ β_4 ln(weighted average density$_{it}$)^2 + ε_{it}

4. YWS_SWC4: ln(SWC BOTEX+) = $\alpha + \beta_1$ ln(Connected properties$_{it}$) + β_2 ln(pumping capacity per sewer length$_{it}$)+ β_3 ln(weighted average density$_{it}$)+ β_4 ln(weighted average density$_{it}$)^2 + β_4 (% mains laid after 2001$_{it}$) + ε_{it}

5. YWS_SWT1: ln(SWT BOTEX+) = $\alpha + \beta_1$ ln(Load$_{it}$)) + β_2 (% load with ammonia below 3 mg/l$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + ε_{it}

6. YWS_SWT2: ln(SWT BOTEX+) = $\alpha + \beta_1$ ln(Load$_{it}$)) + β_2 (composite complexity variable$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + ε_{it}

7. YWS_NPWW1: ln(NPWW BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$)) + β_2 ln(load per sewer length$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + β_4 (% load with ammonia below 3 mg/l$_{it}$) + ε_{it}

8. YWS_NPWW2: ln(NPWW BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$)) + β_2 ln(load per sewer length$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + β_4 (% load with ammonia below 3 mg/l$_{it}$) + β_5 (ln(connected properties per sewer length$_{it}$)) + ε_{it}

9. YWS_NPWW3: ln(NPWW BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$)) + β_2 ln(load per sewer length$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + β_4 (% load with ammonia below 3 mg/l$_{it}$) + β_5 (% mains laid after 2001$_{it}$) + ε_{it}

10. YWS_NPWW4: ln(NPWW BOTEX+) = $\alpha + \beta_1$ ln(Sewer length$_{it}$)) + β_2 ln(load per sewer length$_{it}$) + β_3 (weighted average size variable (to band 9)$_{it}$) + β_4 (composite complexity variable$_{it}$) + β_5 (% mains laid after 2001$_{it}$) + ε_{it}

Description of the dependent variable

The dependent variables are defined as per Ofwat's consultation analysis files i.e. the sum of:

- Power
- Income treated as negative expenditure
- Service charges / Discharge
- Bulk Discharge
- Renewals expensed in year (infrastructure)
- Renewals expensed in year (non-infrastructure)
- Other operating expenditure excluding renewals
- Maintaining the long-term capability of assets (infrastructure)
- Maintaining the long-term capability of assets (non-infrastructure)
- Transfer of private sewers and pumping stations
- Atypical expenditure
- Reducing flood risk for properties (OPEX and CAPEX)
- Network reinforcement (OPEX and CAPEX)

It excludes the following cost categories:

- Costs associated with the Traffic Management Act
- Industrial Emissions Directorate
- NRSWA diversions (non-S185)
- Other non-S195 diversions
- Developer services base cost adjustment
- Backcasting adjustment (between bioresources and sewage treatment).

This is consistent with Ofwat's PR24 methodology.

While we have not explored the modelled cost definition in detail as part of this consultation, we note that network reinforcement is typically a 'lumpy' expenditure item that does not correlate well with the cost drivers included in the model. Two companies facing similar population growth rates can have materially different reinforcement requirements, depending on (among other things) where the population growth occurs within their operating regions.[1] We expect that post-modelling adjustments may be

[1] For example, if population growth in a company's operating region is concentrated in areas where there is excess capacity, then the need for additional network reinforcement might be limited. Conversely, if population growth is concentrated in areas that are already capacity constrained, then network reinforcement requirements will be larger.

required for some companies if network reinforcement is included in the modelled cost definition, and the models omit explicit drivers of reinforcement requirements.

Separately, we welcome Ofwat's decision in the final methodology to remove growth at sewage treatment plants from the modelled cost base. We consider that (like reinforcement expenditure above) growth at sewage treatment plants is 'lumpy' and does not correlate well with cost drivers included in the models and underlying issues can be masked at the base cost level. We understand that Arup and Ofwat are developing separate growth models, and we look forward to providing input into this modelling as and when the data and models are shared with the industry.

Description of the explanatory variables

- Connected properties (BN1178).
- Sewer length (sum of BN13535_21 and BN13528)
- Load (STWDP125_21)
- Properties per sewer length (Connected properties divided by sewer length)
- Load per sewer length (Load divided by sewer length)
- Pumping capacity per lengths of sewer (S4029 divided by sewer length)
- Weighted average density LAD (code: WAD_LAD), as reported in the published wholesale dataset
- '% load with ammonia below 3mg/l' (sum of STWDA121 + STWDA122_21 divided by load)
- Composite complexity (sum of '% load with ammonia below 3mg/l' and '% load with phosphorus below 0.5mg/l' (the latter defined as STWDP121_21 divided by total load))
- '% mains laid after 2001' (BB2370 divided by sewer length)
- Weighted average size variable (band 1 to band 9), defined as the weighted sum of the proportion of load treated at different size bands from 1 to 9 (see accompanying analysis files).

Brief comment on the models

The models presented below have operationally intuitive coefficients (directionally) and perform reasonably well against Ofwat's assessment criteria. We note that the magnitude of the relationships between costs and cost drivers (not just the direction) will also require validation to the extent possible. We c continue to examine this and request that Ofwat also ensures that the magnitude of the coefficients is operationally intuitive.

We emphasise that the models are selected based on the data currently available—alternative models may perform better once new data becomes available (e.g. additional years of AMP7, AMP8 business plan information, adjustments to modelled cost definitions) or if new estimation approaches are considered. In particular, the modelling consultation invites companies to submit models estimated using Random Effects, but we consider that alternative estimation approaches should not be excluded from Ofwat's suite of models ex ante.

We also note that the models may not sufficiently account for factors that are expected to affect efficient costs in AMP8 for some companies or the entire industry, nor capture the impact of certain exogenous factors on an outturn basis for individual companies.

Note that the models presented in this submission relate entirely to the network plus price control. When the modelling consultation was shared with companies, it was still unclear as to how Ofwat would model bioresources expenditure (specifically in relation to the inclusion of financing costs). In light of Ofwat's final methodology for PR24, we are continuing to build on Ofwat's PR19 bioresources models and we will engage on bioresources modelling when the consultation is published in spring.

Sewage collection (SWC)

The SWC models that we propose differ from the PR19 models in the following respects.

First, we consider that the PR19 models do not sufficiently account for the differing maintenance requirements that companies may have as a result of (inter alia) the age of their networks. We note that the wastewater dataset does not contain explicit measures of maintenance activity and, therefore, maintenance activity cannot be explicitly included in the wastewater models. However, it is well-established that maintenance needs for specific assets generally increases as assets age,[2] and one asset age variable exists in the wastewater dataset (specifically, the length of sewers laid after 2001).

Incorporating a measure of asset age into the PR19 model (SWC1) improves model fit relative to the PR19 model, and the estimated relationship between asset age and expenditure is operationally intuitive and statistically significant. Indeed, this finding is true across a range of model specifications. Therefore, failing to account for maintenance need could over- or under-fund individual companies.

We note that the CMA argued against the use of asset age variables in the PR19 appeal, stating that the variable was endogenous (i.e. companies could 'under-maintain' or 'under-

[2] We note that this is generally the case across network companies. For example, see DNO working group (2017), 'DNO common network asset indices methodology', January p.32.

renew' their networks in order to receive higher funding).[3] However, the extent to which companies could 'game' the model in this way is likely to be limited, given that companies have established service level and asset health requirements. Moreover, the extent to which a company can influence the age of its assets in the short term (and medium term) is likely to be limited, such that the cost driver is effectively exogenous throughout the course of a regulatory period. In this context, the variable passes Ofwat's criterion of short-term exogeneity.[4]

In any case, the CMA also argued that Ofwat should adopt a new framework for tracking and funding maintenance requirements, using more forward-looking measures. Therefore, we consider that some measure of maintenance activity that reflects maintenance of optimal asset health should be incorporated into Ofwat's econometric models, or other mechanisms should be in place to fund companies for increased maintenance requirements (or reduce funding where maintenance requirements are expected to fall).

Second, the evidence suggests that there is a 'U-shaped' relationship between density and expenditure, regardless of whether weighted average density or properties per length of sewer is the chosen density measure. During the PR19 appeal, the CMA amended Ofwat's SWC2 model (that originally modelled a log-linear relationship between density and expenditure) to include a squared density term, arguing that a squared density term makes operational and engineering sense[5] and is supported by statistical evidence.[6]

As the coefficient on pumping capacity per network length became statistically insignificant when adding a squared density term in model SWC1, we drop this cost driver in SWC2.

Third, we consider that connected properties can be a valid measure of scale when assessing SWC expenditure. Indeed, connected properties may (in theory) capture some aspects of the model costs better, particularly with respect to network reinforcement expenditure. As noted above, even if the model controls for connected properties, individual companies may require post-modelling adjustments to reflect different costs associated with the location of population growth.

[3] Models that explicitly control for maintenance activity (e.g. length of sewers maintained) may also suffer from endogeneity issues. In particular, while maintenance activity is a clearly operational driver of expenditure, companies could 'over-maintain' their networks or forecast excessive maintenance volumes in their business plans (that may not be ultimately delivered) in order to achieve higher allowances in AMP8.

[4] Ofwat stated in its draft methodology that it is 'open to considering drivers that are only endogenous in the long term as the risk of perverse incentives is lower.' See Ofwat (2022), 'Creating tomorrow, together: Our final methodology for PR24 Appendix 9 Setting expenditure allowances', December, p.8.

[5] CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.177.

[6] CMA (2021), 'Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited price determinations Final report', March, para. 4.178.

We find that controlling for connected properties as opposed to network length leads to a moderate increase in model fit, and a narrower range of estimated efficiency scores.

Sewage Treatment (SWT)

On the published dataset, the performance of Ofwat's PR19 SWT models worsens. In particular, the coefficients on the STW-level economies of scale variables ('% load treated in size bands 3 and below' and '% load treated in size band 6 and above') become less significant.

Our proposed SWT models differ from Ofwat's PR19 models with respect to modelling of STW-level economies of scale. Specifically, we:

- include more granular data from Ofwat's 'Large STW' dataset to separate size band 6 into four size bands (meaning that there are nine size bands in total);
- we construct a 'weighted average size' variable using the proportion of load treated at size bands one to nine. This variable is constructed similarly to how Ofwat constructs the weighted average complexity variable in wholesale wastewater i.e. the proportion of load treated in size band 1, plus two times the proportion of load treated in size band 2 and so on.

We consider that this economies of scale variable is superior to the variables that Ofwat used at PR19 because: (i) it can better reflect the varying levels of scale within size band 6 through the use of more granular data; (ii) constructing a weighted average size variable may limit the risk that companies are over- or under-funded based on a (somewhat arbitrary) decision relating to size thresholds. However, as with the weighted average complexity variable in wholesale water, we consider that the weights used to construct the weighted average size variable should be validated from an operational perspective (which we are examining).

Moreover, we note that some companies (including YWS) will experience a large increase in expenditure in AMP8 due to an increase in P-removal activity associated with the WINEP programme. We note that Ofwat's PR19 models do not explicitly account for P-removal activity, which will underfund affected companies. We have explored SWT models that include P-removal as separate cost drivers. However, the estimated relationship between P-removal activity and expenditure was often statistically insignificant or operationally unintuitive. This is likely to be because of limited variation on this variable in the historical dataset (only three companies have P-removal activity at above 1% of total load in the historical dataset).

For this consultation, to overcome current data challenges, we have proposed models that control for a composite complexity variable, defined as the sum of Ofwat's ammonia-related complexity variable and the proportion of load treated with P-consents below 0.5

mg/l. This assumes that the cost associated with increasing ammonia-related complexity is the same as the cost associated with increasing phosphorus-related complexity—an assumption that may not be needed when additional data on P-removal activity becomes available (e.g. with additional years of outturn data and business plan information), or could be amended in light of operational evidence (i.e. a weighted sum could be more appropriate if P-removal is more expensive than ammonia removal). We note that presently there is only a minor difference in model quality between models that control for Ofwat's complexity variable and this new composite measure, and that the models with only ammonia-related treatment suffer from omitted factors. To the extent that other forms of wastewater treatment complexity (e.g. UV treatment) are also material drivers of SWT expenditure, they may suffer from a similar issue and be treated in a similar way.

Network plus (NPWW)

At PR19, Ofwat did not consider network plus models in wholesale wastewater (NPWW), arguing that the models it considered performed poorly. With the additional data, we consider that it is possible to develop NPWW models that perform at least as well as the SWC and SWT models. The model fit improves and the range of estimated efficiency scores narrows when compared to the SWC and SWT models, which could be driven by the ability of the NPWW to account for operational trade-offs between SWC and SWT and possible cost allocation issues.

As NPWW is the sum of SWC and SWT, the insights from the SWC and SWT models also apply to NPWW. For NPWW, we control for sewer length as the primary scale variable (directly related to SWC costs), and control for load per sewer length to account for varying levels of scale at SWT. The remaining cost drivers capture (to varying degrees) density, economies of scale at the treatment plant level, maintenance requirements and treatment complexity.[7] A key distinction in the NPWW models is that we currently do not control for a squared density term due to the fact that it is statistically insignificant or operationally intuitive across a range of specifications.

We note that, while the sign of the coefficients are robust to the data sensitivities that Ofwat is suggesting (i.e. removing the first and last years of data and removing the most and least efficient companies), the statistical significance of the estimated coefficients does change. As such, we consider that NPWW models should be considered alongside SWC and SWT models as part of a broad modelling suite at PR24.

General comment on model limitations

[7] We present models that control for Ofwat's ammonia-related complexity variable to allow for a comparison of model performance with and without the composite complexity variable. Models that control for Ofwat's ammonia-related complexity variable will not sufficiently account for the costs associated with P-removal activity and could be biased against companies experiencing an increase in P-removal activity.

As the models are estimated on outturn data and do not explicitly account for measures of service quality, the models will only fund companies to deliver the level of service achieved by the industry average (or the benchmark companies) in the historical period. Moreover, the analysis of the level of service funded through base expenditure is complicated given that there is no single measure of service, and different companies can perform well on different service measures. Similarly, the cost requirements for improving service quality depend on the service measure being examined (e.g. it may be more or less expensive to improve internal sewer flooding as opposed to pollution incidents), the level of service already achieved by a particular company (i.e. the marginal cost of improving service may increase as the level of service increases) and associated exogenous factors (e.g. it may be more or less difficult to improve internal sewer flooding in densely or sparsely populated areas).

We understand that Ofwat is reluctant to control for service measures explicitly in its cost assessment models. Nonetheless, if unjustified performance commitments are set for PR24 (i.e. beyond those already achieved by the industry), then some form of cost adjustment will be required. Alternatively, in the PR24 final methodology, Ofwat appears open to adjusting some of the performance targets[8] for exogenous factors, which could partly address some of the issues relating to the cost–quality disconnect identified above.

	YWS_SWC1	YWS_SWC2	YWS_SWC3	YWS_SWC4
Dependent variable	BOTEX+ (SWC)	BOTEX+ (SWC)	BOTEX+ (SWC)	BOTEX+ (SWC)
Sewer length (log)	0.778*** (0)	0.669*** (0)		
Connected properties (log)			0.792*** (0)	0.717*** (0)
Pumping capacity per sewer length (log)	0.293** (0.0409)		0.400*** (0.000131)	0.296*** (0)
Properties per sewer length (log)	1.104*** (1.84e-07)	-17.14*** (0.000289)		
Properties per sewer length (log), squared		2.478*** (8.54e-05)		
Weighted average density (log)			-1.848*** (0.00453)	-1.722*** (1.72e-06)
Weighted average density (log), squared			0.129*** (0.00286)	0.126*** (5.21e-08)
% mains laid after 2001	-0.861* (0.0532)	-1.619*** (0)		-1.643*** (0)
Constant	-7.833*** (0)	27.06*** (0.00200)	-0.334 (0.896)	0.155 (0.909)
Estimation method (OLS or RE)	RE	RE	RE	RE
N (sample size)	187	187	187	187
Model robustness				

[8] See Ofwat (2022), 'Creating tomorrow, together: Our final methodology for PR24 Appendix 9 Setting expenditure allowances', December, p. 61.

R2 adjusted	0.924	0.932	0.924	0.940
RESET test	0.016	0.972	0.418	0.13
VIF (max)	2.422	2264	403.9	408.5
Pooling / Chow test	0.731	0.773	0.807	0.329
Normality of model residuals	0.404	0.245	0.0559	0.0604
Heteroskedasticity of model residuals	0.823	0.491	0.256	0.713
Test of pooled OLS versus Random Effects (LM test)	0.000242	0.167	2.18e-05	0.121
Efficiency score distribution (min and max)	89% to 116%	86% to 108%	92% to 111%	90% to 108%
Sensitivity of estimated coefficients to removal of most and least efficient company	A [pumping capacity becomes less significant; asset age becomes more significant]	G	G	G
Sensitivity of estimated coefficients to removal of first and last year of the sample	A [asset age becomes more significant]	G	G	G

	YWS_SWT1	YWS_SWT2
Dependent variable	BOTEX+ (SWT)	BOTEX+ (SWT)
Load (log)	0.743*** (0)	0.735*** (0)
% load with ammonia below 3mg/l	0.00641*** (0)	
Weighted average size variable (band 1 to band 9)	-0.00159*** (2.01e-05)	-0.00160*** (6.13e-06)
Composite complexity		0.00642*** (0)
Constant	-3.917*** (3.36e-06)	-3.820*** (5.77e-06)
Estimation method (OLS or RE)	RE	RE
N (sample size)	187	187
Model robustness		
R2 adjusted	0.885	0.886
RESET test	0.317	0.152
VIF (max)	4.672	4.539
Pooling / Chow test	1	1
Normality of model residuals	0.00718	0.0111
Heteroskedasticity of model residuals	0.726	0.672
Test of pooled OLS versus Random Effects (LM test)	0	0
Efficiency score distribution (min and max)	91% to 143%	88% to 140%

Sensitivity of estimated coefficients to removal of most and least efficient company	G	G
Sensitivity of estimated coefficients to removal of first and last year of the sample	G	G

	YWS_NPWW1	**YWS_NPWW2**	**YWS_NPWW3**	**YWS_NPWW4**
Dependent variable	BOTEX+ (NPWW)	BOTEX+ (NPWW)	BOTEX+ (NPWW)	BOTEX+ (NPWW)
	NPWW1	NPWW2	NPWW3	NPWW4
Sewer length (log)	0.591*** (0)	0.613*** (0)	0.564*** (0)	0.558*** (0)
Load per sewer length (log)	0.988*** (0.00516)	0.592** (0.0124)	1.099*** (0.000363)	1.040*** (0.000297)
Weighted average size variable (to band 9)	-0.00103* (0.0850)	-0.00145** (0.0401)	-0.00108*** (0.00572)	-0.00101*** (0.00566)
% load with ammonia below 3mg/l	0.00589*** (0)	0.00507*** (0)	0.00603*** (0)	
Composite complexity				0.00615*** (0)
Properties per sewer length (log)		0.928** (0.0471)		
% mains laid after 2001			-1.251*** (0.00144)	-1.437*** (0.000228)
Constant	-4.584*** (0)	-4.486*** (0)	-3.550*** (2.53e-06)	-4.372*** (0.000199)
Estimation method (OLS or RE)	RE	RE	RE	RE
N (sample size)	187	187	187	187
Model robustness				
R2 adjusted	0.935	0.943	0.955	0.957
RESET test	0.481	0.33	0.276	0.89
VIF (max)	4.913	11.75	4.977	4.850
Pooling / Chow test	0.997	0.997	0.919	0.903
Normality of model residuals	0.0859	0.0382	0.0118	0.0211
Heteroskedasticity of model residuals	0.449	0.882	0.805	0.828
Test of pooled OLS versus Random Effects (LM test)	0	0	7.70e-07	5.79e-07
Efficiency score distribution (min and max)	88% to 129%	91% to 114%	96% to 125%	93% to 124%

Sensitivity of estimated coefficients to removal of most and least efficient company	A [load per sewer length becomes less significant; weighted average size becomes less significant]	A [load per sewer length becomes less significant; weighted average size becomes less significant; population density becomes less significant]	G	A [weighted average size becomes less significant]
Sensitivity of estimated coefficients to removal of first and last year of the sample	A [load per sewer length becomes less significant; weighted average size becomes less significant]	A [load per sewer length becomes less significant; weighted average size becomes less significant; population density becomes less significant]	A [weighted average size becomes less significant]	A [weighted average size becomes less significant]

Efficiency scores

YWS_SWC1

Rank	Company	Efficiency score
1	NES	89.2%
2	ANH	94.5%
3	SVH	95.0%
4	SRN	96.0%
5	WSX	100.3%
6	YKY	100.9%
7	NWT	101.9%
8	WSH	102.9%
9	SWB	110.4%
10	TMS	115.9%

YWS_SWC2

Rank	Company	Efficiency score
1	NES	86.0%
2	SVH	95.7%
3	YKY	101.0%
4	TMS	101.1%
5	ANH	101.3%
6	NWT	103.1%
7	SWB	103.8%

8	WSX	105.5%
9	WSH	107.1%
10	SRN	107.7%

YWS_SWC3

Rank	Company	Efficiency score
1	WSX	91.6%
2	ANH	93.9%
3	NES	96.4%
4	WSH	98.3%
5	SRN	98.4%
6	TMS	100.0%
7	SVH	100.1%
8	SWB	106.0%
9	YKY	108.8%
10	NWT	110.8%

YWS_SWC4

Rank	Company	Efficiency score
1	NES	89.6%
2	ANH	95.3%
3	SWB	101.8%
4	SVH	101.9%
5	WSH	102.2%
6	TMS	102.3%
7	NWT	102.8%
8	SRN	104.4%
9	WSX	106.1%
10	YKY	108.1%

YWS_SWT1

Rank	Company	Efficiency score
1	SVH	90.5%
2	TMS	92.4%
3	ANH	92.9%
4	WSX	93.9%
5	SWB	95.8%

6	YKY	105.4%
7	NES	105.6%
8	WSH	111.3%
9	NWT	113.2%
10	SRN	142.6%

YWS_SWT2

Rank	Company	Efficiency score
1	SVH	87.7%
2	WSX	92.6%
3	TMS	92.7%
4	ANH	93.4%
5	SWB	95.3%
6	YKY	105.4%
7	NES	105.7%
8	WSH	111.5%
9	NWT	113.6%
10	SRN	140.2%

YWS_NPWW1

Rank	Company	Efficiency score
1	WSX	87.9%
2	SVH	93.4%
3	TMS	95.3%
4	YKY	98.9%
5	SWB	99.4%
6	NES	101.5%
7	ANH	101.5%
8	WSH	102.4%
9	NWT	106.2%
10	SRN	128.8%

YWS_NPWW2

Rank	Company	Efficiency score
1	SVH	90.9%
2	WSX	92.6%
3	TMS	94.8%

4	YKY	97.1%
5	NES	97.7%
6	SWB	100.4%
7	ANH	100.9%
8	WSH	105.1%
9	NWT	110.9%
10	SRN	113.9%

YWS_NPWW3

Rank	Company	Efficiency score
1	YKY	96.5%
2	SVH	96.5%
3	TMS	98.6%
4	NES	98.9%
5	SWB	99.5%
6	ANH	99.8%
7	WSH	100.6%
8	NWT	100.9%
9	WSX	102.6%
10	SRN	124.6%

YWS_NPWW4

Rank	Company	Efficiency score
1	SVH	93.5%
2	YKY	96.4%
3	NES	98.0%
4	SWB	98.9%
5	TMS	99.0%
6	NWT	99.9%
7	ANH	100.0%
8	WSH	101.4%
9	WSX	103.5%
10	SRN	123.7%

Yorkshire Water Services – Residential retail models

Econometric model formula:

1. YWS_RTC1: $\ln(\text{Total operating costs}_{it}) = \alpha + \beta_1 \ln(\text{number of households connected}_{it}) + \beta_2 \ln(\text{average bill size}_{it}) + \beta_3 (\text{\% of metered households}_{it}) + \beta_4 (\text{Total migration}_{it}) + \beta_5 (\text{Composite deprivation metric } 1_{it}) + \varepsilon_{it}$

2. YWS_RTC2: $\ln(\text{Total operating costs}_{it}) = \alpha + \beta_1 \ln(\text{number of households connected}_{it}) + \beta_2 \ln(\text{average bill size}_{it}) + \beta_3 (\text{\% of metered households}_{it}) + \beta_4 (\text{Composite deprivation metric } 1_{it}) + \varepsilon_{it}$

3. YWS_RTC3: $\ln(\text{Total operating costs}_{it}) = \alpha + \beta_1 \ln(\text{number of households connected}_{it}) + \beta_2 \ln(\text{average bill size}_{it}) + \beta_3 (\text{\% of metered households}_{it}) + \beta_4 (\text{Total migration}_{it}) + \beta_5 (\text{Composite deprivation metric } 2_{it}) + \varepsilon_{it}$

4. YWS_RTC4: $\ln(\text{Total operating costs}_{it}) = \alpha + \beta_1 \ln(\text{number of households connected}_{it}) + \beta_2 \ln(\text{average bill size}_{it}) + \beta_3 (\text{\% of metered households}_{it}) + \beta_4 (\text{Total migration}_{it}) + \beta_5 (\text{Composite deprivation metric } 3_{it}) + \varepsilon_{it}$

5. YWS_RTC5: $\ln(\text{Total operating costs}_{it}) = \alpha + \beta_1 \ln(\text{number of households connected}_{it}) + \beta_2 \ln(\text{average bill size}_{it}) + \beta_3 (\text{\% of metered households}_{it}) + \beta_4 (\text{Composite deprivation metric } 3_{it}) + \varepsilon_{it}$

Description of the dependent variable

The dependent variable is defined as in Ofwat's consultation analysis files. The main two differences are:

- Using smoothed doubtful debt (BM9003S instead of BM9003, when available). Use of smoothed debt instead of actual debt may change with new information and will need to be reassessed.
- The dependent variable is not modelled in unit costs as on the current data, the models indicate scale economies. This may change with new information and will need to be reassessed.

The dependent variable used is the sum of:

- Customer services
- Debt management

- Smoothed doubtful debt[1]
- Meter reading
- Other operating expenditure
- Depreciation smoothed over 5 years
- Recharges costs net of recharges income
- Local authority rates + exceptional items when reported separately

Currently our cost definition excludes third party costs and pension deficit repair costs as per Ofwat's PR24 methodology. As in wholesale, we intend to examine the suitability of including/excluding certain items in modelled costs post the model submission.

Description of the explanatory variables

The cost drivers are similar to PR19 with the addition of composite metrics which are intended to capture a rounded view of deprivation on retail costs. The drivers considered are:

- Number of households connected (the sum of R3017, R3019, R3021, R3018, R3020 and R3022)
- Average bill size (rev_t divided by the number of households connected)
- % of metered households (the sum of R3018, R3020 and R3022, divided by the number of households connected)
- Total migration (the sum of internal migration and international migration)
- Composite deprivation metric 1 is computed using the arithmetic average of the three Equifax metrics: Equifax – Insight Postcode Event – % of households with default; Equifax – Credit risk score derived from all Insight data (Score Range is 000-200); and Equifax – Average number of Partial Insight accounts or county court judgements per household. Composite deprivation metric 1 = (*standardized*(eq_lpcf62) + *standardized*(*inverse*(eq_rgc102)) + *standardized*(eq_xpcf2)) * (1/3)
- Composite deprivation metric 2 is computed using the arithmetic average of Income score interpolated and the three Equifax metrics: Equifax – Insight Postcode Event – % of households with default; Equifax – Credit risk score derived from all Insight data (Score Range is 000-200); and Equifax – Average number of Partial Insight accounts or county court judgements per household. Composite deprivation metric 2 = (*standardised*(Incomescore_interpolated) + *standardized*(eq_lpcf62) + *standardized*(*inverse*(eq_rgc102)) + *standardized*(eq_xpcf2)) * (1/4)
- Composite deprivation metric 3 is computed using the first latent factor of performing principal component analysis over the three Equifax metrics: Equifax – Insight Postcode Event – % of households with default; Equifax – Credit risk score derived from all Insight data (Score Range is 000-200); and Equifax – Average number of Partial Insight accounts or county court judgements per household.

[1] This refers to doubtful debts smoothed (Retail household - accounting separation) from the Annual Performance Report data as published in Ofwat's dataset for residential retail (code BM9003S). We have not made any further transformation to this cost item.

Composite deprivation metric 3 = *PCA*(*standardized*(eq_lpcf62), *standardized*(eq_rgc102), *standardized*(eq_xpcf2))

Note: *standardized(a)* refers to the standardisation of the variable a by subtracting its mean and dividing by the standard deviation. This ensures the same scale across variables when computing the average. *Inverse*(a) refers to taking the inverse of variable a. For all observations a_i inverse(a_i) = 1 / a_i. *PCA*(a, b) refers to performing principal component analysis over variables a and b, and obtain the first latent factor.

Brief comment on the models

We have used the same data (including the panel structure) as in Ofwat's published files.

For retail expenditure, we have focused on the development of TOTEX models. When developing bottom-up models (i.e. modelling bad debt related costs and other operating expenditure separately) we found that they performed comparatively poorly (i.e. models had low R-Squared and relevant cost drivers lacked statistical significance). This was particularly the case for 'other operating costs'. However, bottom-up models can provide a helpful understanding of specific cost drivers and their relevance. Also, additional data (e.g. additional years and data adjustments) may improve the statistical performance of the bottom-up models. We therefore consider that bottom-up models can serve a purpose (e.g. as a cross check), even if the results do not feed directly into companies' cost allowances at PR24. We will look to further develop such models post the modelling submission, and propose that Ofwat does not ignore these at this stage of the price review.

The treatment of deprivation and its inclusion in the models is another difference between the PR19 models and our proposed ones. At PR19, Ofwat considered two measures of deprivation and triangulated results between them. The triangulation between competing models has potential to reduce bias, either positive or negative, that individual measures may have on particular companies. Nevertheless, it relies on the bias affecting companies in a specific manner to even out upon triangulation. Alternatively, another option adopted by regulators is the construction of a composite metric that combines competing measures of a particular factor (deprivation in this case) that cannot be individually included in one model because of multicollinearity. Such a composite measure can help to capture the impact of multiple measures in one, which can also moderate volatility in the results from using individual measures separately. Among the available deprivation metrics in Ofwat's dataset, we cannot determine operationally which of them is most directly linked to companies' retail expenditure as they are high-level proxies. Although some deprivation measures performed better than others in the models explored, companies' historical performance are sensitive to the measure included in the models. Therefore, if we were to focus on a specific measure of deprivation, there is a risk that it could result in over- or under-funding individual companies. We have therefore proposed ‘composite deprivation measures’ that can mitigate such risks and is an alternative to triangulation.[2]

On the current data, at the retail TOTEX level, there appears to be strong evidence of economies of scale: the scale driver’s coefficient is significantly smaller than 1 across all specifications. Therefore, we have modelled them on an aggregate basis instead of the unit-cost specification used at PR19. Modelling on an aggregate basis also improves model performance generally.

We note that some of the models below contain cost drivers that are statistically insignificant at standard thresholds (e.g. metering penetration and total migration). Nevertheless, they benefit the models in terms of overall fit (including allowing other important drivers to improve their significance), and their coefficients have the expected sign.

	YWS_RTC1	YWS_RTC2	YWS_RTC3	YWS_RTC4	YWS_RTC5
Dependent variable	Total operating costs (log)	Total operating costs (log)	Total operating costs (log)	Total operating costs (log)	Total operating costs (log)
Number of households connected (log)	0.904*** (0)	0.923*** (0)	0.913*** (0)	0.904*** (0)	0.923*** (0)
Average bill size (log)	0.622*** (0)	0.599*** (0)	0.599*** (0)	0.622*** (0)	0.599*** (0)
% metered households	0.003 (0.237)	0.002 (0.313)	0.003 (0.234)	0.003 (0.231)	0.003 (0.302)
Total migration	0.011 (0.41)		0.011 (0.421)	0.01 (0.413)	
Composite deprivation metric 1	0.086** (0.02)	0.071** (0.032)			
Composite deprivation metric 2			0.083** (0.036)		
Composite deprivation metric 3				0.05** (0.02)	0.041** (0.031)
Constant	0.878*** (0.009)	0.923*** (0.005)	0.884*** (0.009)	0.877*** (0.009)	0.924*** (0.005)

[2] For instance, under the specification of YWS_RTC2, replacing the composite metric by a single deprivation metric yields non-significant coefficients (with p-values ranging from 0.129 to 0.844). On the other hand, the efficiency scores are more consistent across the proposed composite metrics. This underscores the benefits of including a composite deprivation metric.

Estimation method (OLS or RE)	RE	RE	RE	RE	RE
N (sample size)	153	153	153	153	153
R2 adjusted	0.981	0.980	0.980	0.981	0.980
RESET test	0.822	0.708	0.814	0.82	0.708
VIF (max)*	5.209	2.954	4.903	5.193	2.966
Pooling / Chow test*	0.96	0.993	0.956	0.956	0.992
Normality of model residuals*	0.049	0.048	0.071	0.051	0.049
Heteroskedasticity of model residuals*	0.382	0.795	0.297	0.374	0.804
Test of pooled OLS versus Random Effects (LM test)	0	0	0	0	0
Efficiency score distribution (min and max)	Min: 0.8 Max: 1.29	Min: 0.8 Max: 1.31	Min: 0.8 Max: 1.29	Min: 0.8 Max: 1.29	Min: 0.8 Max: 1.31
Sensitivity of estimated coefficients to removal of most and least efficient company	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]	G	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]
Sensitivity of estimated coefficients to removal of first and last year of the sample	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]	A [The coefficient for the composite deprivation metric 1 becomes significant at 99% confidence]	G	A [The coefficient for the composite deprivation metric 3 becomes significant at 99% confidence]	A [The coefficient for the composite deprivation metric 3 becomes significant at 99% confidence]

Note: * These tests are conducted on OLS models.

Efficiency scores YWS_RTC1

Rank	Company	Efficiency score
1	SWB	80.0%
2	SEW	83.6%
3	ANH	87.5%
4	YKY	94.8%
5	WSX	95.9%

6	AFW	96.2%
7	BRL	97.6%
8	SVE	98.4%
9	NES	99.0%
10	PRT	100.2%
11	NWT	100.5%
12	SSC	107.7%
13	HDD	108.8%
14	WSH	116.0%
15	TMS	116.7%
16	SES	124.2%
17	SRN	128.8%

Efficiency scores YWS_RTC2

Rank	Company	Efficiency score
1	SWB	79.7%
2	SEW	84.0%
3	ANH	87.1%
4	YKY	93.5%
5	WSX	94.7%
6	SVE	96.5%
7	NES	97.2%
8	AFW	98.5%
9	PRT	100.0%
10	BRL	100.4%
11	NWT	100.7%
12	SSC	106.0%
13	HDD	106.1%
14	WSH	113.1%
15	TMS	121.9%
16	SES	126.3%
17	SRN	130.8%

Efficiency scores YWS_RTC3

Rank	Company	Efficiency score

1	SWB	79.6%
2	SEW	83.3%
3	ANH	87.7%
4	YKY	94.9%
5	WSX	95.4%
6	BRL	96.8%
7	AFW	96.9%
8	SVE	97.9%
9	NES	99.3%
10	PRT	101.1%
11	NWT	101.7%
12	SSC	106.6%
13	HDD	108.2%
14	WSH	115.2%
15	TMS	118.0%
16	SES	126.0%
17	SRN	128.6%

Efficiency scores YWS_RTC4

Rank	Company	Efficiency score
1	SWB	80.0%
2	SEW	83.7%
3	ANH	87.3%
4	YKY	94.9%
5	WSX	96.0%
6	AFW	96.1%
7	BRL	97.7%
8	SVE	98.4%
9	NES	99.1%
10	PRT	100.3%
11	NWT	100.6%
12	SSC	107.6%
13	HDD	108.9%
14	WSH	116.1%
15	TMS	116.6%
16	SES	124.3%
17	SRN	128.7%

Efficiency scores YWS_RTC5

Rank	Company	Efficiency score
1	SWB	79.7%
2	SEW	84.0%
3	ANH	87.0%
4	YKY	93.6%
5	WSX	94.8%
6	SVE	96.5%
7	NES	97.3%
8	AFW	98.3%
9	PRT	100.1%
10	BRL	100.5%
11	NWT	100.7%
12	SSC	106.0%
13	HDD	106.2%
14	WSH	113.2%
15	TMS	121.8%
16	SES	126.4%
17	SRN	130.6%