Scrutinising Ofwat's PR19 Final Determination Models on Sewerage Collection: Economic Invalidity, Triangulation, and Non-transparency of Modelling Principles

Third party submission to the CMA on Botex cost assessment

May 2020

Dr Maria Nieswand

Professor David Saal

Centre for Productivity and Performance
School of Business and Economics
Loughborough University
Loughborough, LE11 3TU
UK

Contact Details

[redacted]

[redacted]

Table of Contents

Preamble

After Ofwat published their PR19 Final Determinations (Ofwat, 2019a), an unprecedented four water and wastewater companies have appealed and now submitted statements of case to the Competition & Markets Authority (CMA). The statements of case were submitted in April 2020 and Ofwat responded to them in early May 2020.

Subsequently, the CMA invited third parties to submit comments on the issues raised in the References from Ofwat and the Main Party Submissions.

Given the short period of time between the availability of Ofwat's References and CMA's submission deadline of May 11, 2020, in this document we solely focus on scrutinising Ofwat's PR19 model specifications for sewerage collection.

We demonstrate that Ofwat's sewerage collection model SWC1 still suffers from economic invalidity, that Ofwat fails to appropriately interpret the coefficient for the scale variable "length of sewers" in SWC1, and that Ofwat incorrectly claims Anglian Water's argument in this regard would disregard the proper interpretation of the model.

We further argue that, due to Ofwat's approach to "triangulation" and the frail nature of its sewerage collection model, the flaws of this model are further transferred to the mid-level and aggregate cost base determined by Ofwat.

Scrutinising the sewerage collection model specifications has also led us to raise concerns regarding the transparency of Ofwat's modelling approach as the accepted level of statistical significance is not specified but used in arguments pro or against certain variables and/ or model specifications.

We emphasise that the issues raised here are illustrative of flaws in Ofwat's sewage collection modelling and should not be considered comprehensive; e.g. they restate only a few key issues with Ofwat's sewage collection models, as necessitated by the limited time we had for this response. They give, however, strong indications for the necessity of reconsidering Ofwat's PR19 final determinations and the suitability of the cost assessment approach and models they have employed.

We also finally note that the absence of detailed scrutiny of the remainder of Ofwat's Wholesale Water and Wastewater models in this document, should by no means taken as indicative of our endorsement of those models. Instead, our opinion with regard to the suitability of Ofwat's overall modelling approach and its models, are best evidenced via consideration of the six submitted reports that we have summarised in a separate document that we have also submitted to the CMA today (Saal & Nieswand 2020).

Evaluating Sewage Collection Models in Ofwat's Final Determination

In December 2019, Ofwat published their final model specifications for wholesale wastewater activities in Table A2.2. of the "PR19 final determinations – Securing cost technical efficiency appendix" (Ofwat, 2019a, page 163). In their report, Ofwat presented two model specifications on sewerage collection activities, i.e. SWC1 and SWC2. The regression results of these models are presented in Appendix 1.

Referring to Saal and Nieswand (2019), in their CMA redetermination statement of case, Anglian Water claimed that both Ofwat's water and wastewater models fail to account for Anglian's operating characteristics, e.g., demographic, topographic and environmental characteristics (Anglian Water, 2020, p. 115, sec 560-562) and criticised Ofwat for triangulating models that "fail either statistical, economic or engineering criteria" (Anglian Water, 2020, p. 119, sec 583).

To underpin their critique, Anglian derived that the first sewerage collection model (SWC1) does neither appropriately control for density as claimed by Ofwat, nor does the model specification meet economic expectations as the coefficient for sewer length, which Ofwat treats as its scale (output) variable, can be shown to be negative (Anglian Water, 2020, p. 119, sec 583-587). The critique expressed by Anglian corresponds with Saal and Nieswand's evaluation of this model (Saal and Nieswand, 2019, p. 53-57) and replicates their criticism on SWC1 and SWC2.

Therefore, Anglian requests that the CMA "remedies the statistical shortcomings of Ofwat's models" (Anglian Water, 2020, p. 131, sec 633).

Subsequently, in May 2020, Ofwat referred to this criticism in their responses to common issues in companies' statements of case by stating:

> "Anglian Water claims that we adopted models that do not follow economic and engineering rationale.[21] In particular, Anglian Water claims that one of our sewage collection models provides a counter-intuitive elasticity for the variable 'length of sewers'." (Ofwat, 2020, p. 18, sec 3.19)

Ofwat dissent with Anglian's view and justify their approach for wastewater model SWC1:

> "Based on this specification, Anglian Water suggests that the effect of sewer length on costs is given by: $\delta - \beta - \gamma$. Results based on the parameters of the sewage collection model presented in our final determinations would, therefore, suggest that sewer length has a negative effect on costs, which contradicts engineering and economic rationale." (Ofwat, 2020, p. 18, sec 3.21)
>
> "We consider that Anglian Water's argument disregards the proper interpretation of the model. In our model, properties/length is a measure of density and capacity/length is a measure of energy intensity per kilometre. The purpose of the

model is that is that δ is the elasticity of length, that is, it captures what happens to costs as a water company becomes bigger, holding the other variables, density and energy intensity per kilometre, constant. This is a reasonable question to ask of our model given that as the length of sewers changes across companies, so does the pumping capacity and the number of properties. To ask the question, what happens to costs when length only increases, means that we are asking the question what happens if we increase length and at the same time decrease the density variable and energy intensity. We do not consider that this is an appropriate question to ask of the model." (Ofwat, 2020, p. 18, sec 3.22)

"For this reason[,] we consider that it is appropriate to interpret δ the elasticity of length, β as the elasticity of energy intensity and γ the elasticity of density. All these parameters have the expected sign." (Ofwat, 2020, p. 18, sec 3.23)

Given Ofwat's response, we would like to clarify Ofwat's misunderstanding regarding ours and Anglian's view and show that Ofwat's interpretation of the estimated coefficient for 'length of sewers' remains incorrect. Consequently, their model specification SWC1 remains economically invalid. Note that we do not dispute with Ofwat's intension to include measures of density and energy (really pumping) intensity, but instead dispute the validity of its modelling approach, variable selections and interpretations.

1.1 Sewerage collection model SWC1 contradicts economic theory by having a negative cost elasticity of sewer length.

In model specification SWC1, Ofwat models BOTEX in logs (of sewerage collection) as a function of sewer length in logs, pumping capacity per sewer length in logs, and the number of properties per sewer length in logs.

According to Ofwat, their intension is to model density by the ratio of properties per sewer length (in logs) and energy intensity per kilometre by capacity per sewer length (in logs; Ofwat, 2020, p. 18, sec 3.22). Note that we do not argue with whether or not these are relevant cost drivers or controls in a sewerage collection model. However, we argue that the definition of the variables in conjunction with the model specification yield an economically invalid model and Ofwat fail to interpret the estimated coefficient for 'length of sewers' correctly.

Formally, SWC1 can be expressed as

$$\log(BOTEX) = \alpha_0 + \delta \log(sewer\ length) + \beta \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + \gamma \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \quad (1)$$

Inserting the regression results published by Ofwat (Ofwat, 2020, p. 163, Table A2.2), this is:

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + 0.998 * \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \quad (2)$$

The mathematically fundamental logarithm quotient rule states that:

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y). \quad (3)$$

Applying the logarithmic quotient rule yields the mathematically identical model:

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * [\log(pumping\ capacity) - \log(sewer\ length)] + 0.998 * [\log(number\ of\ properties) - \log(sewer\ length)] \quad (4)$$

By expanding this equation, we obtain

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * \log(pumping\ capacity) - 0.317 * \log(sewer\ length) + 0.998 * \log(number\ of\ properties) - 0.998 * \log(sewer\ length)] \quad (5)$$

Thus, given that all coefficients in SWC1 are statistically significant, the elasticity of sewer length equals $0.839 - 0.317 - 0.998 = -0.476$.

This is equivalent to stating that the elasticity of sewer length in SWC1 expressed in Equation (5) is $\delta - \beta - \gamma = -0.476$. SWC1 suggests that, ceteris paribus (i.e. keeping all other explanatory variables equal), increasing sewer length by 1% reduces BOTEX by 0.476%.

A comprehensive derivation of the mathematical equivalence of equations (2) and (5) can be found in Appendix 2.

From this, we can conclude that:

a) By mathematical and econometric means, the elasticity of length of sewers in model SWC1 has a negative sign. The negative coefficient of -0.476 is significantly different from zero only at confidence level 14%.
b) Thus, the model specification SWC1 contradicts economic theory and Ofwat's own expectation regarding the sign of this coefficient. Ofwat claims that the correct coefficient has a positive sign and equals δ = 0.839, which is coherent with their expectation expressed in Ofwat (2020, p. 13, sec 3.3 in conjunction with p. 19, sec 3.21).
c) Therefore, Ofwat fail to determine the correct estimate of the elasticity of sewer length when only considering δ. Ofwat's interpretation of the elasticity of sewer length ($\delta = 0.839$.), disregards mathematical (e.g., the logarithm quotient rule) and econometric fundamentals if they refuse to take the coefficients of properties per sewer length and pumping capacity per sewer length into account.
d) By doing so, Ofwat misinterprets the sewer length coefficient in SWC1 and ask exactly the question "[…] what happens if we increase length and at the same time decrease the density variable and energy intensity.", which Ofwat themselves do not consider an appropriate question to ask of the model (Ofwat, 2020, p. 18, sec 3.22).
e) Further, Ofwat's claim that "Anglian Water's argument disregards the proper interpretation of the model" (Ofwat, 2020, p. 18, sec 3.22) is incorrect. **However, Ofwat are incorrect: this is not a matter of interpretation, this is a matter of mathematics**. As shown below in Section 1.2 it can be unequivocally both mathematically and empirically demonstrated via alternative estimation that Ofwat's models include a negative relationship between 'length of sewers' and BOTEX.

1.2 Estimating model specification SWC1 with mathematical identical formulation

We re-estimate Ofwat's SWC1 model (referred to as SWC1):

$$\log(BOTEX) = \alpha_0 + \delta \log(sewer\ length) + \beta \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + \gamma \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \tag{6}$$

and its mathematical equivalent (referred to as SWC1 restated):

$$\log(BOTEX) = \alpha_0 + \delta^* \log(sewer\ length) + \beta \log(pumping\ capacity) + \gamma \log(number\ of\ properties) \tag{7}$$

where $\delta^* = \delta - \beta - \gamma$,

we obtain the following regression results:

Table 1: Re-estimating Ofwat's SWC1 model

	SWC1	**SWC1 restated**
	re1	re2
	b/p	b/p
lnsewerlength	0.839***	-0.476
	{0.000}	{0.144}
lnpumpingcapperlength	0.317*	
	{0.080}	
lndensity	0.998***	
	{0.005}	
lnpumpingcap		0.317*
		{0.080}
lnproperties		0.998***
		{0.005}
_cons	-8.124***	-8.124***
	{0.000}	{0.000}
Econometric_model	Random Effects	Random Effects
depvar	lnrealbotexswc	lnrealbotexswc
N	80	80
vce	cluster	cluster
R_squared	0.931	0.931
VIF_statistic		
RESET_P_value	0.152	0.152

It is straightforward to see that:

a) Both models, i.e. “SWC1” and “SWC1 restated” are mathematically identical. Thus, the elasticities and standard errors for ln(pumpingcapperlength) in model “SWC1” and ln(pumpingcap) in model “SWC1 restated” are identical.
b) Further, the elasticities and standard errors for ln(density) in model “SWC1” and ln(properties) in model “SWC1 restated” are identical. The same applies to the constant.
c) Obviously, due to being mathematically identical R squared and the Ramsey RESET p-values are identical in both models.
d) The only difference between the two models are the coefficients for sewer length. In SWC1 the coefficient equals 0.839 as reported by Ofwat (2019a, p. 163, Table A2.2). That is exactly the value that Ofwat interprets as the elasticity of ‘length of sewers’ disregarding the mathematical relationship between the remaining variables in SWC1.
e) However, as derived in the previous section, the elasticity of ‘length of sewers’ equals -0.476 in model “SWC1 restated”. By definition $-0.476 = 0.839 - 0.317 - 0.998$.

f) As the coefficient for ‘length of sewers” is statistically insignificant in model “SWC1 restated” and given the mathematical identity to model “SWC1”, it can be concluded that the coefficient for ‘length of sewers” in model SWC1 only appears to be statistically significant at a confidence level of 1% when wrongly disregarding the other coefficients of this model.
g) Therefore, Ofwat’s model SWC1 truly violates economic theory, Ofwat’s own expectation regarding the signs of coefficients, and includes a scale variable (‘length of sewers’), which is statistically significant only at a confidence level of 14.4%. This level would not survive academic scrutiny.
h) Ofwat do not actually understand their models as appropriate consideration of the underlying cost elasticities. “SWC1 restated” reveals that contrary to Ofwat’s interpretation of SWC1 the number of properties served (elasticity of 0.998) is the primary cost driver in its model, pumping is an important second driver (elasticity of 0.317), while in contrast to its interpretation of lengths of main as a primary scale driver (output) its actual estimated elasticity is -0.476 and statistically insignificant.

In addition to the regression outcomes, we can also compare the company-specific performances obtained by each of the two models. Table 2 shows the actual and predicted costs as well as the ratio of actual and predicted costs for both models. Note that the actual and predicted costs for each company are not only identical between the two models but also exactly match with Ofwat’s figures for their model SWC1.[1] Hence, the “efficiency score for 5 years” are identical.

Even though, the company-specific performance measures remain unchanged when estimating “SWC1 restated” instead of “SWC1”, doesn’t mean the model specification SWC1 is correct. The conclusion to draw from this exercise is that the cost assessment for sewerage collection depends on at least one invalid model and therefore must be reconsidered.

[1] These are published in FM_WWW2_FD.

Table 2: Company performance for models SWC1 and SWC1 restated

Actual costs	SWC1	SWC1 restated
	sumactualcostxbre1	sumactualcostxbre2
ANH	794	794
NES	438	438
NWT	900	900
SRN	639	639
SVT	856	856
SWB	284	284
TMS	1575	1575
WSH	501	501
WSX	331	331
YKY	656	656
Predicted costs	**SWC1**	**SWC1 restated**
	sumpredictedcostxbre1	sumpredictedcostxbre2
ANH	758	758
NES	398	398
NWT	819	819
SRN	755	755
SVT	1003	1003
SWB	296	296
TMS	1410	1410
WSH	455	455
WSX	372	372
YKY	633	633
Efficiency scores 5 years	**SWC1**	**SWC1 restated**
ANH	1.0478	1.0478
NES	1.1020	1.1020
NWT	1.0995	1.0995
SRN	0.8466	0.8466
SVT	0.8537	0.8537
SWB	0.9574	0.9574
TMS	1.1164	1.1164
WSH	1.1022	1.1022
WSX	0.8908	0.8908
YKY	1.0362	1.0362
3rd best	**0.8908**	**0.8908**

1.3 **The model for sewerage collection is not robust across specifications. Both model specifications SWC1 and SWC2 indicate a contradictory impact of sewer length on BOTEX.**

Model specification SWC2 differs to SWC1 in regard to the right-hand-side-variables and aims to explain BOTEX in logs by sewer length in logs, pumping capacity per sewer length in logs, and the weighted average density in logs.

Formally, this is:

$$\log(BOTEX) = \alpha_0 + \delta \ \log(sewer\ length) + \beta \ \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + \eta \ \log(weighted\ average\ density) \quad (8)$$

Applying the logarithmic quotient rule and inserting the estimated coefficients (see Appendix 1), this model is equivalent to

$$\log(BOTEX) = -6.416 + 0.896 * \log(sewer\ length) + 0.606 * \log(pumping\ capacity) - \ 0.606 * \log\ (sewer\ length) + 0.178 * \log(weigthed\ average\ density) \quad (9)$$

Thus, the elasticity of sewer length in SWC2 equals $0.896 - 0.606 = 0.290$, which should be considered rather low if this is indeed Ofwat's primary scale (output) variable, and relative to the much larger effective elasticity of Botex to pumping capacity.

Given that the correctly determined elasticity for 'length of sewers' is -0.476 in SWC1, the difference to the elasticity for 'length of sewers' of -0.476 in SWC1 is quiet substantial.

This means that

a) The model specifications are not robust and contradicts Ofwat's own aspiration expressed in Ofwat (2020, p. 3, sec 1.3, p. 4, sec 1.10).
b) The absence of robustness between model specifications formed in other cases Ofwat's argument for dismissing other variables, e.g., average pumping head (Ofwat, 2020, p. 17, sec 3.16), which we note we do not concur with. It also formed the argument Ofwat puts forward against an integrated wholesale wastewater model (Ofwat, 2020, p. 19, sec 3.18 and 3.19), which we similarly do not concur with.

1.4 This example clearly illustrates that Ofwat has not taken third-party feedback regarding the flaws in sewerage collection modelling into account with due diligence.

The issue of the negative elasticity of sewer length has already been raised in March 2019 by Saal and Nieswand (2019, p.53) when assessing the cost modelling for the Initial Assessment of Business Plans (IAP). These models were published in Ofwat (2019b, p. 17, Table 3; see Appendix 3 in this document).

The issue remained in the draft determination (Ofwat, 2019c, p. 112, Table A2.2; see Appendix 4 in this document) still occurs in the final determination (Ofwat, 2019a, p. 163, Table A2.2), and thus, Ofwat has not taken third-party feedback into account with due diligence as claimed (Ofwat, 2020, p. 13, sec 3.2).

1.5 Given Ofwat's inappropriate models (SWC1 and SWC2), it is clear that sewerage collection has not been appropriately modelled, let alone triangulated. Moreover, this issue of inappropriate modelling and poor triangulation further extends to the mid-level and aggregate cost base for wholesale wastewater of which the sewerage collection models form a part.

1.6 Given the demonstrated statistical inappropriateness in SWC1 and the statistically insignificant (by normal confidence interval standards) coefficient for weighted average density in logs in SWC2, Ofwat cannot be considered to have employed robust triangulated measures of density, let alone other operating characteristics. E.g. Ofwat have effectively triangulated sewage collection costs with a model that suggests companies should increase their network length to reduce costs (SWC1) and another model in which its density control measure is statistically insignificant by normal standards.

Evaluating Ofwat's Transparency of Cost Modelling Via a Simple Example

1.7 The econometric modelling approach is not fully transparent regarding the accepted level of statistical significance, and hence, is not transparent.

One of Ofwat's modelling principles is "statistical validity" (Ofwat, 2020, p. 13), which means that Ofwat considers the statistical significance of variables when evaluating their model specifications.

For further detail, Ofwat (Ofwat, 2020, p. 13) refer to the principles of their PR19 methodology published in CEPA (2018) where statistical significance of a parameter is described as "A coefficient is significant when it can be tested that it is different from zero with a certain probability." (CEPA, 2018, p. 42, Table 5.1).

With respect to the probability at which a parameter is considered different from zero, i.e. it has no statistically identifiable impact on the dependent variable, Ofwat states that "We do not consider that the common thresholds of statistical significance (eg 95% significance) need to be strictly followed for our model selection." (Ofwat, 2018, p. 9).

From Ofwat's final determination model SWC2, it can be derived that a significance level of 14.6% (log(weighted average density)) is deemed to be acceptable by Ofwat. However, normal economic modelling practice given this low level of significance would be to consider alternative proxies or models for density, or to develop an alternative modelling approach.

While this is approach is not disputed in general, not specifying an acceptable significance level does not allow us to evaluate Ofwat's decision not to include service quality variables in their models. Ofwat argue, for example, against measures of service quality "because it led to a perverse incentive (e.g. higher allowances for poorer service quality) or did not produce sensible and/or statistically significant results." (Ofwat, 2020, p. 23, sec 3.36). However, to our understanding Ofwat provide no publicly available evidence to evaluate these claims.

References

Anglian Water (2020): PR19 CMA Redetermination. Statement of Case. April 2020.

Ofwat (2018): Cost assessment for PR19: a consultation on econometric cost modelling. March 2018.

Ofwat (2019a): PR19 final determinations. Securing cost efficiency technical appendix. December 2019.

Ofwat (2019b): Supplementary technical appendix: econometric approach. January 2019.

Ofwat (2019c): PR19 draft determinations. Securing cost efficiency technical appendix. July 2019.

Ofwat (2020): Reference of the PR19 final determinations: Cost efficiency – response to common issues in companies' statements of case. May 2020.

Saal and Nieswand (2019): A Review of Ofwat's January 2019 Wholesale Water and Wastewater Botex Cost Assessment Modelling for PR19. March 2019.

Saal and Nieswand (2020): CMA Redetermination of Ofwat's 2019 Final Price Determinations: Third party submission to the CMA on Botex cost assessment. May 2020.

Appendix

Appendix 1: Ofwat's Final Determination wastewater wholesale models

Table 3: Econometric models for wholesale wastewater activities

Model name	SWC1	SWC2	SWT1	SWT2	BR1	BR2	BRP1	BRP2
Dependent variable (log)	Sewage collection		Sewage treatment		Bioresources		Bioresources + Sewage treatment	
Sewer length (log)	0.839***	0.896***						
Load (log)			0.779***	0.773***			0.765***	0.762***
Sludge produced (log)					1.274***	1.265**		
Load treated in size bands 1-3 (%)			0.045***		0.057**		0.038*	
Load treated in size band 6 (%)				-0.013**				-0.011**
Pumping capacity per sewer length (log)	0.317*	0.606***						
Load with ammonia consent below 3mg/l (%)			0.004***	0.004***			0.005***	0.005***
Number of properties per sewer length (log)	0.998**							
Weighted average density (log)		0.178 (0.146)			-0.295**			
Sewage treatment works per number of properties (log)						0.397*		
Constant term	-8.124***	-6.416***	-5.228***	-3.988***	-0.389 (0.648)	0.994*	-4.753***	-3.709***
Overall R-Squared	0.93	0.88	0.88	0.87	0.82	0.79	0.92	0.92
Number of observations	80	80	80	80	80	80	80	80

Source: Ofwat, 2019a, page 163, Table A2.2 Econometric models for wholesale wastewater activities.

Appendix 2: Comprehensive derivation of the coefficient for 'length of sewers' in SWC1

Formally, SWC1 can be expressed as:

$$\log(BOTEX) = \beta_0 + \beta_1 \log(sewer\ length) + \beta_2 \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + \beta_3 \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \quad (10)$$

The mathematically fundamental logarithm quotient rule states that:

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y). \quad (11)$$

The randomly chosen numerical examples in **Error! Reference source not found.** demonstrates this. For x=10 and y=20, ln(x)-ln(y)= ln(x/y)=-0.6931 etc.

Table 4: Numerical examples of logarithm quotient rule

x	y	ln(x)	ln(y)	ln(x)-ln(y)	ln(x/y)
10	20	2.30	3.00	-0.6931	-0.6931
25	30	3.22	3.40	-0.1823	-0.1823
299	310	5.70	5.74	-0.0361	-0.0361

Applying the logarithm quotient rule to model SWC1 in Equation (1), yields a mathematical identical formulation of SWC1:

$$\begin{aligned}\log(BOTEX) = \beta_0 &+ (\beta_1) \log(sewer\ length) \\ &+ \beta_2 [\log(pumping\ capacity) - \log(sewer\ length)] \\ &+ \beta_3 [\log(number\ of\ properties) - \log(sewer\ length)]\end{aligned} \quad (12)$$

It should be obvious that estimating Equations (1) and (3) must yield identical results as they are mathematically identical.

The econometric formulation of Equation (3) is

$$\begin{aligned}\log(BOTEX) = \beta_0 &+ \beta_1 \log(sewer\ length) \\ &+ \beta_2[\log(pumping\ capacity) - \log(sewer\ length)] \\ &+ \beta_3[\log(number\ of\ properties) - \log(sewer\ length)] \\ &+ \varepsilon\end{aligned} \quad (13)$$

where ε represents the statistical error term.

Expanding Equation (4) yields

$$\log(BOTEX) = \beta_0 + \beta_1 \log(sewer\ length) + \beta_2 \log(pumping\ capacity) - \beta_2 \log(sewer\ length) + \beta_3 \log(number\ of\ properties) - \beta_3 \log(sewer\ length) + \varepsilon \tag{14}$$

Note that the variable sewer length appears three times in Equations (4) and (5) and has coefficients $\beta_1, \beta_2, and\ \beta_3$ associated with it. Therefore, all three coefficients must be taken into account when evaluating the impact of sewer length on BOTEX. Hence, the overall cost elasticity for sewer length in this model is $\beta_1 - \beta_2 - \beta_3$. It is worthwhile to emphasise that the cost elasticities for pumping capacity and the number of properties remain unaffected by that.

To demonstrate how this relates to the regression presented in **Error! Reference source not found.**, i.e. Ofwat's models for final determination, we restate Equation (1) and change the coefficient names to be consistent with our previous report (Saal and Nieswand, 2019), Anglian's redetermination statement of case (Anglian, 2020) and Ofwat's response to common issues (Ofwat, 2020):

$$\log(BOTEX) = \alpha_0 + \delta \log(sewer\ length) + \beta \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + \gamma \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \tag{15}$$

The regression results[2] for this model are

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * \log\left(\frac{pumping\ capacity}{sewer\ length}\right) + 0.998 * \log\left(\frac{number\ of\ properties}{sewer\ length}\right) \tag{16}$$

Applying the logarithmic quotient rule implies:

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * [\log(pumping\ capacity) - \log(sewer\ length)] + 0.998 * [\log(number\ of\ properties) - \log(sewer\ length)] \tag{17}$$

[2] The regression results are obtained by random effect estimation with clustered standard errors.

By expanding this equation, we obtain

$$\log(BOTEX) = -8.124 + 0.839 * \log(sewer\ length) + 0.317 * \log(pumping\ capacity) - \ 0.317 * \log\ (sewer\ length) + 0.998 * \log(number\ of\ properties) - 0.998 * \ \log\ (\ sewer\ length)] \quad (18)$$

Hence, the elasticity of sewer length equals $0.839 - 0.317 - 0.998 = \ -0.476$.

This is equivalent to stating that the elasticity of sewer length in SWC1 expressed in Equation (6) is $\delta - \ \beta - \gamma = \ -0.476$. SWC1 suggests that, ceteris paribus (i.e. keeping all other explanatory variables equal), increasing sewer length by 1% reduces BOTEX by 0.476%.

Appendix 3: Wholesale wastewater models in IAP

Table 3: Econometric models for wholesale wastewater

Model name	SWC1	SWC2	SWT1	SWT2	BR1	BR2	BRP1	BRP2
Dependent variable (log)	Sewage collection		Sewage treatment		Bioresources		Bioresources + Sewage treatment	
Sewer length (log)	0.739***	0.714***						
Load (log)			0.795***	0.779***			0.788***	0.770***
Sludge produced (log)					1.058***	1.183***		
Load treated in size bands 1-3 (%)			0.045**				0.039**	
Load treated in size band 6 (%)				-0.012*				-0.010**
Pumping capacity per sewer length (log)	0.170**	0.346**						
Load with ammonia consent below 3mg/l (%)			0.004***	0.004***			0.005***	0.005***
Number of properties per sewer	1.471***							
Weighted average density (log)		0.256**			-0.280 (0.121)			
Sewage treatment works per number of properties						0.320*		
Constant term	-8.907***	-5.037***	-5.498***	-4.203***	0.749	0.746	-5.107***	-3.944***
Overall R-Squared	0.91	0.82	0.87	0.85	0.80	0.80	0.92	0.92
Number of observations	70	70	70	70	70	70	70	70

Notes: Dependent variable is modelled base costs in 2017/18 prices, using the CPIH adjustment for each level of aggregation. P values expressed in parentheses are based on clustered standard errors at the company level. *, ** and *** denote significance level at 10, 5 and 1 per cent respectively.

Source: Ofwat, 2019b, page 17, Table 3 Econometric models for wholesale wastewater.

Note: This table incorrectly suggests that in SWC1 the variable "Number of properties per sewer" is not in logs. The published Stata do-files, however, confirm that this variable enters the estimation in logs.

Appendix 4: Wholesale wastewater models in PR19 draft determination

Table A2.2: Econometric models for wholesale wastewater activities

Model name	SWC1	SWC2	SWT1	SWT2	BR1	BR2	BRP1	BRP2
Dependent variable (log)	Sewage collection		Sewage treatment		Bioresources		Bioresources + Sewage treatment	
Sewer length (log)	0.819***	0.897***						
Load (log)			0.856***	0.847***			0.840***	0.809***
Sludge produced (log)					1.217***	1.186***		
Load treated in size bands 1-3 (%)			0.058**		0.053***		0.051***	
Load treated in size band 6 (%)				-0.015*				-0.011**
Pumping capacity per sewer length (log)	0.281*	0.619***						
Load with ammonia consent below 3mg/l (%)			0.004***	0.004***			0.005***	0.005***
Number of properties per sewer length (log)	1.186***							
Weighted average density (log)		0.193 (0.109)			-0.235*			
Sewage treatment works per number of properties (log)						0.325*		
Constant term	-8.588***	-6.534***	-6.273***	-4.765***	-0.542	0.775	-5.777**	-4.296***
Overall R-Squared	0.93	0.87	0.87	0.85	0.82	0.80	0.92	0.92
Number of observations	70	70	70	70	70	70	70	70

Source: Ofwat, 2019c, page 112, Table A2.2 Econometric models for wholesale wastewater activities.