Appendix 8i:
An in the round assessment of cost adjustment claims

Author: Oxera

YorkshireWater

Independent assessment of Yorkshire Water's historical cost performance and consideration of its AMP7 cost adjustment claims in this context

Prepared for
Yorkshire Water

August 2018

Final

www.oxera.com

Contents

Executive summary		**1**
1	**Introduction**	**3**
2	**Methodology**	**4**
2.1	Dataset	4
2.2	Modelled expenditure	5
2.3	Cost drivers	6
2.4	Modelling approach	11
3	**YKY's efficient BOTEX level—water**	**16**
3.1	Catch-up assessment, YKY models	16
3.2	Catch-up assessment, Ofwat models	17
3.3	Catch-up assessment, value chain models	19
3.4	Summary of results	21
4	**YKY's efficient BOTEX(Growth) level—water**	**22**
4.1	Catch-up assessment, YKY's models	22
4.2	Catch-up assessment, Ofwat's models	23
4.3	Summary of results	24
5	**YKY's efficient BOTEX level—wastewater**	**25**
5.1	Catch-up assessment, YKY models	25
5.2	Catch-up assessment, Ofwat models	26
5.3	Catch-up assessment, value chain models	29
5.4	Consideration of YKY's forward-looking cost adjustment claims	30
5.5	Summary of results	31
6	**YKY's efficient BOTEX(Growth) level—wastewater**	**32**
6.1	Catch-up assessment, YKY's models	32
6.2	Catch-up assessment, Ofwat's models	33
6.3	Summary of results	34
7	**Conclusion**	**36**
A1	**Regression output—wholesale water**	**37**

A1.1 Ofwat's wholesale water BOTEX models 37
A1.2 Ofwat's growth enhancement models 41
A1.3 YKY aggregate BOTEX and BOTEX(Growth) models 42
A1.4 Stochastic Frontier Analysis results 45

A2 Regression output—wholesale wastewater 47

A2.1 Ofwat's wholesale water BOTEX models 47
A2.2 Ofwat's growth enhancement models 50
A2.3 YKY aggregate BOTEX and BOTEX(Growth) models 53
A2.4 Stochastic Frontier Analysis results 56

Figures and tables

Figure 2.1 YKY characteristics, Ofwat's water models 7
Figure 2.2 YKY characteristics, YKY's water models 8
Figure 2.3 YKY characteristics, Ofwat's wastewater models 9
Figure 2.4 YKY characteristics, YKY's wastewater models 10

Table 3.1 YKY's relative ranking, YKY's models 16
Table 3.2 YKY's estimated efficiency, YKY's models 17
Table 3.3 Summary of triangulated results, YKY's models 17
Table 3.4 YKY's rank, Ofwat models 18
Table 3.5 YKY's estimated efficiency, Ofwat's models 18
Table 3.6 YKY's efficient expenditure, POLS and RE (upper-quartile benchmark) 19
Table 3.7 Summary of triangulated results, Ofwat's models 19
Table 3.8 Value chain modelling results—YKY models 20
Table 3.9 Value chain modelling results—Ofwat models 20
Table 3.10 Aggregating results 21
Table 4.1 YKY's ranking, YKY's models 22
Table 4.2 YKY's estimated efficiency, YKY's models 22
Table 4.3 Summary of triangulated results, YKY models 23
Table 4.4 YKY's ranking, Ofwat's models 23
Table 4.5 Aggregating results from disaggregated analysis 24
Table 5.1 YKY's relative ranking, YKY's models 25
Table 5.2 YKY's estimated efficiency, YKY's models 26
Table 5.3 Summary of triangulated results, YKY's models 26
Table 5.4 YKY's rank, Ofwat's models 27
Table 5.5 YKY's estimated efficiency, Ofwat's models 27
Table 5.6 YKY's efficient expenditure, POLS and RE (upper-quartile benchmark) 28
Table 5.7 Summary of triangulated results, Ofwat's models 28
Table 5.8 Value chain modelling results—YKY models 29
Table 5.9 Value chain modelling results—Ofwat models 30
Table 5.10 Aggregating results 30
Table 6.1 YKY's ranking, YKY's models 32
Table 6.2 YKY's estimated efficiency, YKY's models 33

Table 6.3	Summary of triangulated results, YKY models	33
Table 6.4	YKY’s ranking, Ofwat’s first-time sewerage models	33
Table 6.5	YKY’s ranking, Ofwat’s new developments, sewage growth, and reducing sewer flooding risk models	34
Table 6.6	Aggregating results from Ofwat’s BOTEX and disaggregated enhancement analysis	34

Executive summary

Oxera understands that Yorkshire Water (YKY) is considering to submit cost adjustment claims for its wastewater service in its PR19 business plan. The context of an independent assessment of YKY's historical base expenditure,[1] which is a focus of this report, is to consider if YKY's anticipated step-changes in costs over AMP7 can be robustly captured through a simple roll-forward of the historical analysis.[2] To that end, we have been asked to evaluate the need for possible cost adjustment claims for YKY over AMP7. YKY's cost adjustment claims have been assessed and quantified separately by its engineering consultants.[3]

To determine YKY's historical relative efficiency on wholesale BOTEX and BOTEX(Growth), models submitted by YKY as part of Ofwat's modelling consultation were taken as a starting point.[4] Ofwat's published models in its modelling consultation[5] were also considered. The suite of models used to estimate YKY's relative efficiency comprised both aggregate and granular models submitted by YKY and Ofwat.[6]

Based on the last six years of data spanning 2011/12 to 2016/17, multiple model specifications to enable a 'balanced' consideration of industry operational drivers, multiple efficiency estimation approaches, and a wide consideration of aggregation and triangulation possibilities, YKY's efficient BOTEX and BOTEX(Growth) expenditure is estimated as below.

Wholesale water: YKY's historical efficient BOTEX is in the range of £1,137m–£1,150m, with a central estimate of £1,150m (in 2016/17 prices).[7] Given its historical AMP outturn BOTEX of £1,150m,[8] YKY is estimated to be relatively efficient. The historical assessment also suggests that YKY is efficient in BOTEX(Growth).

Wholesale waste: YKY's historical efficient expenditure is in the range of £1,310m–£1,361m, with a central estimate of £1,355m (in 2016/17 prices). Given its historical AMP outturn expenditure of £1,361m,[9] YKY's historical catch-up is estimated to be around 0.4% (c. £6m) over an AMP. The historical assessment also suggests that YKY is broadly efficient in BOTEX(Growth) with an estimated efficiency gap of between 0% and 4% (with a modest gap of about 0.3%[10]).

Therefore, on an outturn basis, YKY is assessed to be broadly efficient on the water service with a modest gap on the wastewater service. We also

[1] Includes BOTEX, which refers to the sum OPEX and capital maintenance expenditure, and BOTEX(Growth), where enhancement expenditure with respect to growth activities are added to BOTEX.
[2] AMP refers to Asset Management Period—a five-year period in which Ofwat sets prices following the submission of company business plans. AMP7 refers to the years 2021–25.
[3] See YKY WWN+01 Cellared properties, YKY WWN+04 Wastewater Growth and YKY BR-01 Bioresources - WINEP enhancement expenditure for details of its cost claims.
[4] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.
[5] Ibid.
[6] As a general point we note that both sets of models may require development in light of further data refinements and industry business plan assumptions over AMP7.
[7] This is based on the median across the efficient estimates from YKY's and Ofwat's models under the different modelling approaches.
[8] Based on average annual expenditure in the modelling period, 2011/12–2016/17 multiplied by five. The figure excludes unmodelled items as per Ofwat's modelled cost definition in the modelling consultation. See section 2.2 for a description of these.
[9] Ibid.
[10] This is based on the median across the efficient estimates from YKY's and Ofwat's models under the different modelling approaches.

reassessed the models considered in this report in light of the 2017/18 data submission in July 2018. While we acknowledge that the 2017/18 data has not gone through Ofwat's quality assurance process and may subsequently be amended, the model results and estimated efficient cost predictions were largely insensitive to the additional year. YKY's position on water and wastewater is also broadly similar to that assessed above, suggesting no significant change in the conclusions drawn in the report.[11]

We note that the historical efficiency position estimated for YKY may be driven, to some extent, by YKY's and the industry's historical levels of activity as well as expenditure profiling over the modelled period. As such, the estimated historical relationship between cost categories and cost drivers may be unrepresentative of future elasticities and can produce inappropriate AMP7 cost predictions if a simple roll-forward of the historical relationship is adopted.

To establish if YKY's proposed cost claims for PR19 are captured through a simple roll-forward of the outturn analysis, we estimated the forecast efficient baseline for AMP7 based on forecast activity levels provided by YKY. The estimated variance between the YKY's business plan expenditure and model forecasts in both YKY's and Ofwat's models is notably large. Indeed, the historical cost assessment models considered in this report appear to not adequately account for the step-increase in YKY's future expenditure.

A driving factor is that the activities/cost drivers used in YKY's and Ofwat's models do not robustly represent YKY's forward-looking expenditure requirements in AMP7 (e.g. expenditure for WINEP and further improvement to service levels). From the model forecasts using a roll-forward approach, it appears likely that YKY's cost adjustment claims for PR19 are incremental. This is particularly so as YKY's cost claims on sewerage relate to factors such as the proportion of cellared properties, number of flooding incidents, and sludge produced, which are not adequately represented in the historical models. We note that if Ofwat were to model companies' business plan assumptions over AMP7, it could provide a reference point, to some extent, on the need and materiality of some of YKY's claims.

The highlighted limitations also apply to YKY's water service AMP7 projections, even if YKY does not intend to submit cost claims in the area.

Hence careful consideration in assessing YKY's and the industry's forward-looking expenditure and YKY's cost claims will be required.

[11] In particular, on water, the efficiency gap of the central estimate remains unchanged at 0% for both BOTEX and BOTEX(Growth), while the corresponding gap for wastewater is approximately 1.3% and 1.2%, respectively.

1 Introduction

Oxera understands that Yorkshire Water (YKY) will be submitting three cost adjustment claims for its wastewater service together with its PR19 draft business plan—namely reducing internal flooding for cellared properties; wastewater growth; and capacity investment in bioresources. These claims involve a step increase in YKY's BOTEX and BOTEX(Growth) for AMP7.

YKY has commissioned Oxera to undertake an overall assessment to evaluate whether these cost adjustment claims could be picked up by a roll-forward of relationships estimated using outturn data. To that end, we have undertaken an independent assessment of YKY's historical BOTEX as well as BOTEX(Growth)[12] performance for its wholesale water and wastewater services using top-down econometric modelling.

The claims themselves are not assessed as part of this report. They have been assessed and quantified separately by YKY's engineering consultants.[13] Instead, the underlying factors of the need for such claims (and the extent to which they are accounted for in the cost assessment models) form part of our wider considerations.

To determine YKY's relative efficiency on wholesale BOTEX and BOTEX(Growth), models submitted by YKY as part of Ofwat's recent modelling consultation were taken as the starting point.[14] Ofwat's published models in this consultation were also considered as an alternative set of evidence. We note that both Ofwat's and YKY's models need further development with additional data and to be reinforced with bottom-up analysis from an operational perspective.

The estimation approaches used in the modelling consultation were examined, alongside alternative methods. Therefore, both YKY's models submitted as part of the consultation and Ofwat's models were used to inform the assessment of YKY's historical efficiency and assess the need of its cost claims.

The report is structured as follows:

- section 2 sets out the methodology used to determine YKY's efficient cost baseline;
- section 3 presents the estimated catch-up in wholesale water BOTEX;
- section 4 presents the estimated catch-up in wholesale water BOTEX(Growth);
- section 5 presents the estimated catch-up in wholesale wastewater BOTEX;
- section 6 presents the estimated catch-up in wholesale wastewater BOTEX(Growth)
- section 7 concludes and suggests limitations to the use of historical relative efficiency assessment models to derive AMP7 forecasts.

[12] BOTEX(Growth) includes BOTEX and growth enhancement expenditure.
[13] See YKY WWN+01 Cellared properties, YKY WWN+04 Wastewater Growth and YKY BR-01 Bioresources - WINEP enhancement expenditure for details of its cost claims.
[14] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.

2 Methodology

In PR14, Ofwat set TOTEX efficiency targets using a combination of TOTEX, BOTEX and enhancement expenditure models. In the PR19 methodology documents, Ofwat indicates that there will be distinct price controls for network plus and water resources in wholesale water.[15] Similarly, the wholesale wastewater price control will be split into bioresources and network plus price controls.

In its most recent consultation on econometric cost modelling,[16] Ofwat has further indicated its intention to distinguish between BOTEX and enhancement expenditure to provide input into the setting of TOTEX efficiency targets. BOTEX models at various levels of aggregation have been published for both wholesale water and wholesale wastewater.[17]

In the consultation, Ofwat developed two water enhancement categories (expenditure associated with meeting lead standards and with new developments and new connections) and two wastewater enhancement categories (expenditure associated with first-time sewerage schemes and with new developments, growth and sewage treatment works and reducing sewer flooding risk for properties).[18]

We note the reservations about enhancement modelling that were also highlighted by some companies in the responses to the consultation. The concerns about modelling enhancement expenditure are brought about by this expenditure (related to some activities) being more idiosyncratic than BOTEX and driven by regional, company-specific factors that are difficult to robustly capture in an econometric framework.[19] Partly due to these factors, YKY submitted BOTEX models and BOTEX(Growth) models at the aggregate level in the consultation.

Therefore, the focus of this report is an assessment of YKY's historical efficient aggregate BOTEX and BOTEX(Growth). The growth-related enhancement models developed by Ofwat for consultation will be assessed in conjunction with Ofwat's aggregate BOTEX models to provide a view of YKY's efficient level of BOTEX(Growth).

In this section, we outline our approach to modelling YKY's efficient historical level of BOTEX and BOTEX(Growth), the dataset used in the modelling (including the modelled expenditure), and the associated cost drivers considered in the models.

2.1 Dataset

Our analysis is based on the dataset published by Ofwat as part of the econometric modelling consultation.[20] For both water and wastewater, the data spans a six-year period from the financial year 2011/12 to the financial year 2016/17, and it therefore includes data from both AMP5 and AMP6.

[15] Ofwat (2017), 'Delivering Water 2020: Our final methodology for the 2019 price review', December, p. 137
[16] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.
[17] Ibid.
[18] Ibid.
[19] For example, see the responses by Thames Water and Welsh Water, available here: https://www.ofwat.gov.uk/consultation/cost-assessment-pr19-consultation-econometric-modelling/#Responses, accessed 20 August.
[20] Available here: https://www.ofwat.gov.uk/consultation/cost-assessment-pr19-consultation-econometric-modelling, accessed 20 August.

The wholesale water dataset is an unbalanced panel consisting of 18 companies in the first five years (2011/12–2015/16) and 17 companies in the final year (due to the merger between South West Water and Bournemouth Water), resulting in a total of 107 observations. The wholesale wastewater dataset is a balanced panel consisting of 10 companies, creating a total of 60 observations.

2.2 Modelled expenditure

In the assessment of YKY's efficient historical level of BOTEX, we have excluded cost items from modelled costs that are either outside management control or could provide perverse incentives with respect to cost reduction, consistent with Ofwat's modelling approach. All expenditure has been deflated to 2016/17 prices using CPIH, and as such, all results presented in this report are also in 2016/17 prices.

In wholesale water, the excluded costs are abstraction charges/discharge consent; business rates; third-party costs; costs associated with the Traffic Management Act; costs associated with statutory water softening; enhancement CAPEX (including infrastructure network reinforcement); pension deficit recovery payments; and atypical costs.

Modelled BOTEX(Growth) in YKY's submitted models includes enhancement expenditure in supply/demand balance; new developments and connections; resilience; and metering over and above modelled aggregate BOTEX. Modelled BOTEX(Growth) in Ofwat models is modelled BOTEX plus enhancement expenditure in new developments.[21]

In wholesale wastewater, the excluded costs are business rates; third-party costs; costs associated with the Traffic Management Act; costs associated with the Industrial Emissions Directive; enhancement CAPEX (including infrastructure network reinforcement); pension deficit recovery payments; and atypical costs.

Modelled BOTEX(Growth) in YKY's submitted models includes enhancement expenditure for first-time sewerage; sludge enhancement (growth); new development and growth; growth at sewage treatment works (excluding sludge treatment); resilience; and reducing sewer flooding risk for properties over and above modelled aggregate BOTEX. Modelled BOTEX(Growth) in Ofwat models is modelled BOTEX plus expenditure in both enhancement categories modelled by Ofwat.[22]

The modelled costs used in Ofwat's models (above) are defined differently to those used in the models submitted by YKY for the consultation. YKY's models have been reassessed using Ofwat's definition of modelled costs.

Unless stated otherwise, both outturn expenditure and efficient predicted expenditure presented throughout the report refer to expenditure over an AMP. Given the modelling period uses four years of data from AMP5 and two years of data from AMP6, this involves taking the average per annum expenditure and multiplying by five to derive expenditure over a five-year period.

[21] Because Ofwat did not develop enhancement models for all growth-related enhancement expenditure for the consultation, the aggregation of BOTEX with enhancement expenditure described above will not coincide with the BOTEX(Growth) modelled in YKY's models.
[22] Similarly to water, Ofwat did not develop enhancement models for all growth-related enhancement expenditure for the consultation. Therefore, the aggregation of BOTEX with enhancement expenditure described above will not coincide with the BOTEX(Growth) modelled in YKY's models.

2.3 Cost drivers

2.3.1 Wholesale water

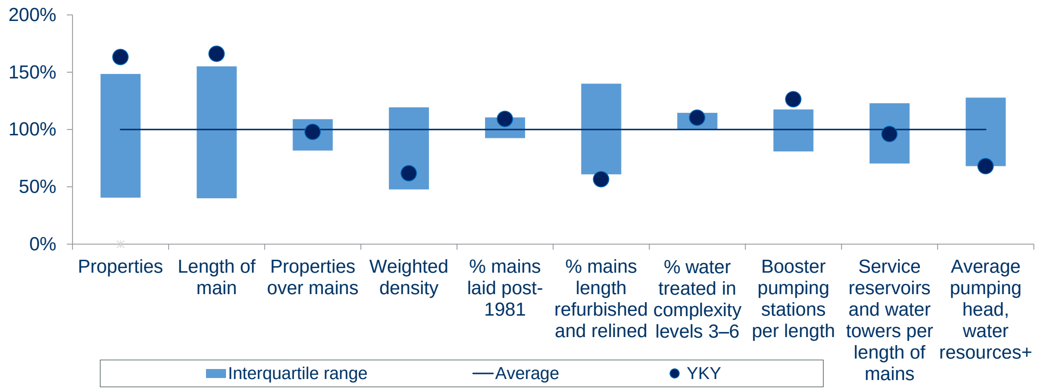
Ofwat included cost drivers in its 12 wholesale water BOTEX models to control for industry characteristics—scale, density, network and treatment complexity, pumping activity, and maintenance-related costs. These are described briefly below.

- **Scale**. The number of connected properties and total length of the network are used to control for the scale of operations. Each variable is used as a scale driver in half of the BOTEX models. When the length of the network is used as the scale driver, the number of connected properties per length of mains is controlled for in three models.
- **Density**. Ofwat considers two density measures and directly controls for density only when the length of the network is used as a scale driver. The number of connected properties per length of mains and the weighted average density measure are controlled for in three models each. The exact construction of the weighted average density measure is not clear.
- **Treatment complexity**. In three of the 12 BOTEX models, Ofwat uses the percentage of water treated in treatment plants of complexity levels 3–6 to control for treatment complexity. Raw water treatment expenditure is expected to be closely related to the treatment requirements of the raw water.
- **Network complexity**. The number of service reservoirs and water towers per length of mains and the number of booster pumping stations per length of mains are used in four Ofwat models each. These variables may also indirectly capture the impact of sparsity on expenditure.
- **Pumping requirements**. In the nine models that do not control for treatment complexity, Ofwat controls for average pumping head in water resources plus.[23] Average pumping head is used as a proxy for the pumping requirements that a company faces.
- **Maintenance**. The proportion of mains renewed or relined and the proportion of mains laid or refurbished after 1981 have been used to control for maintenance activity. The former directly controls for maintenance activity, while the latter controls for it indirectly, on the assumption that older assets require more maintenance.

Figure 2.1 reflects the interquartile range and how YKY compares to the industry average for the cost drivers used in Ofwat's cost assessment models.

[23] The collective name for water resources, raw water distribution and water treatment.

Figure 2.1 YKY characteristics, Ofwat's water models



Source: Oxera analysis.

Figure 2.1 shows that YKY is a relatively large company in terms of the scale drivers used in Ofwat's models. It also has relatively old mains and undergoes relatively little renewal activity. Two network complexity drivers provide possibly competing interpretations of YKY's network complexity—the booster pumping stations per length variable shows YKY to have a relatively complex network, whereas the number of service reservoirs and water towers per length of mains variable shows YKY to have a network of average complexity. In the remaining cost drivers controlled for by Ofwat, YKY comes out to be a typical company.

The six BOTEX models submitted by YKY as part of the consultation also sought to control for industry characteristics with the following variables.

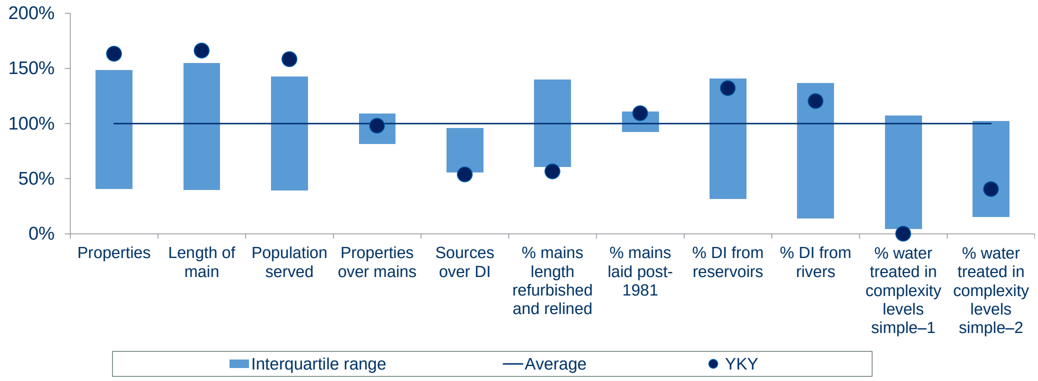
- **Scale**. Because a particular scale driver may capture different dimensions of scale,[24] a range of scale drivers were considered. Three models control of length of the network, two control for connected properties, and one controls for population served.
- **Density**. The number of connected properties per length of mains is controlled for in all six models, three of which also control for the number of connected properties per length of mains squared.[25] This functional form may better suit the a priori expectation that both extreme density and extreme sparsity will increase BOTEX.
- **Diseconomies**. The number of sources per distribution input is used to control for diseconomies of scale at the source and treatment plant level. The expectation is that retrieving less water from more sources will be more costly than retrieving more water from fewer sources.
- **Maintenance**. YKY's models control for the proportion of mains renewed or relined in five out of six models and the proportion of mains laid or refurbished before 1980 in two models. The latter is simply one minus the proportion of mains laid after 1981, and is therefore comparable to the age of assets variable used in Ofwat's models.

[24] This was noted in Ofwat's consultation: Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling, Appendix 1 – Modelling results', March.
[25] A partial-translog model. The translog model subtracts the sample mean from the density variable before squaring to reduce the collinearity between the variables. This does not affect the estimated elasticities or estimated efficiency scores.

- **Treatment complexity**. This is controlled for in all six models. All models control for the proportion of input from reservoirs, and three models control for the proportion of input from rivers. Surface water tends to be more complex to treat, and the source-type variable may also account for different costs of abstraction. However, the relationship between source-type and treatment complexity is imperfect. One model controls for the proportion of water treated in complexity band 1 and below, while two models control for the proportion of water treated in complexity band 2 and below.

Similarly to Figure 2.1, Figure 2.2 shows how YKY compares to the industry in the above cost drivers. YKY's relative position in water source variables supports the view that YKY treats relatively complex water compared to the industry average.

Figure 2.2 YKY characteristics, YKY's water models



Source: Oxera analysis.

2.3.2 Wholesale wastewater

Ofwat submitted for consultation eight BOTEX models, with a view to controlling for operating characteristics across the industry.

- **Scale driver**. Ofwat considered the total load received by wastewater treatment works (WTWs) and connected properties as the drivers capturing the scale of operation. Apart from the first two models, Ofwat controlled for the two measures of scale alternatively across models as each driver may capture an alternative relationship with cost.

- **Diseconomies of WTWs**. Ofwat controlled for the proportion of load treated at small WTWs, defined as bands 1–3, in all of its BOTEX models. The inclusion of this cost driver is intended to capture the possible diseconomies of scale experienced at small WTWs.

- **Density**. All but the first two Ofwat BOTEX models control for density through the number of connected properties per length of sewer. Ofwat suggests that the relationship between this measure of density and cost is positive as 'there are more properties to collect sewage from' per length of sewer.[26]

- **Age of assets**. Half of Ofwat's models control for the proportion of sewers laid after 2001 as a proxy for maintenance and renewals

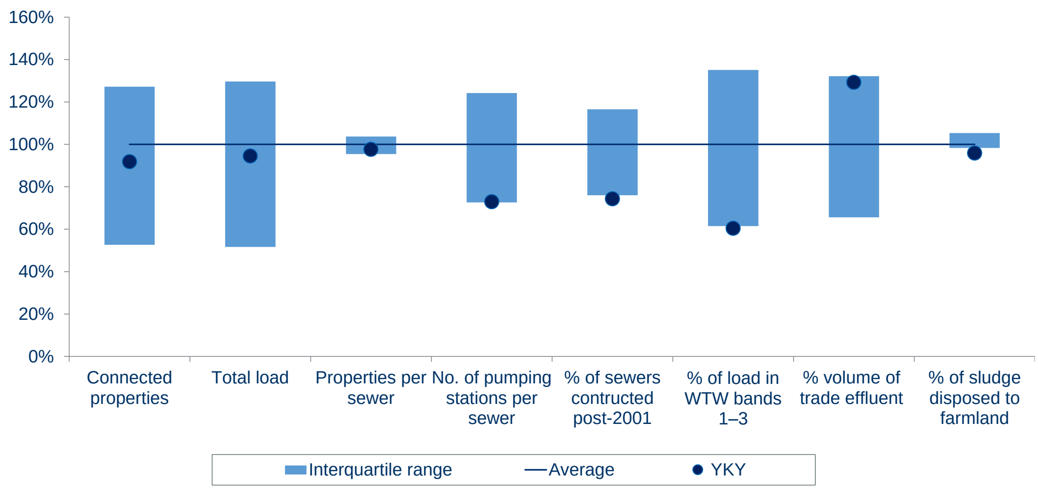
[26] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March, p. 20.

activities. As newer assets are likely to require less maintenance, controlling for this variable serves to explain differences in maintenance expenditure across the industry.

- **Treatment complexity**. Two of eight of Ofwat's BOTEX models control for proportion of load from trade effluent customers as a measure of treatment complexity. As load from industry is likely to contain more pollutants required for treatment, companies with a higher proportion of load from trade effluent customers are likely to incur higher costs for an increased treatment complexity.
- **Sludge disposal**. Two of eight of Ofwat's BOTEX models control for the proportion of sludge disposed to farmland. However, in the dataset used for modelling, the variation of this driver across the industry appears small (with the exception of United Utilities). Therefore, although the expected sign of the relationship between this driver and cost is negative, the impact of including this driver in the model may be driven by outlier observations.
- **Pumping requirements**. Only one model controls for topology, through the number of pumping stations per length of sewer. Companies operating in conditions with higher pumping requirements are expected to incur higher network cost.

Figure 2.3 reflects the interquartile range for the cost drivers used in Ofwat's cost assessment models and how YKY compares to the industry average. The data suggests that YKY is approximately an average-sized company but is atypical (being in upper (or lower) quartile) on the cost drivers capturing operational characteristics adopted by Ofwat.

Figure 2.3 YKY characteristics, Ofwat's wastewater models



Source: Oxera analysis.

The four BOTEX models submitted by YKY during the cost assessment consultation also attempt to capture important operating characteristics.

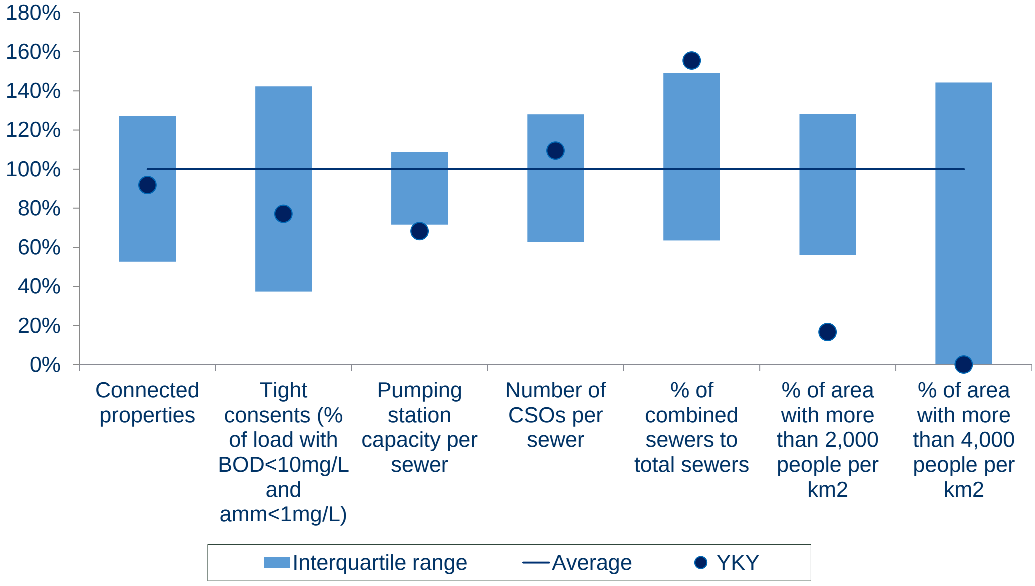
- **Scale driver**. YKY's models control for connected properties as the main scale driver.
- **Treatment complexity**. All of YKY's submitted models control for treatment complexity through tightness of consents—proportion of load subject to BOD (<10mg/l) and ammonia (<1mg/l) consents. As treatment

complexity is likely to increase with the tightness of consents to which a WTW is subjected, the relationship between this driver and cost is expected to be positive. This may be a better driver of treatment complexity as consents are exogenously determined.

- **Pumping requirements**. Pumping station capacity per length of sewer is controlled for in all of YKY's models to capture the relationship between pumping requirements and expenditure. This driver is an alternative measure of pumping compared to the number of pumping stations in Ofwat's models.
- **Maintenance activity**. YKY's models controlled for the number of combined sewer overflows (CSOs) per sewer or the proportion of sewers that are combined sewers as alternative proxies for maintenance activity. Among these, other possible drivers to control for maintenance were presented in the modelling consultation.
- **Density**. YKY's BOTEX models control for (population) density with the granular measure of the proportion of area with more than 2,000 people per km^2 in two of four models and the proportion of area with more than 4,000 people per km^2 in one of four models. As population density increases, cost savings may be made through the use of larger WTWs, which have greater scope for economies of scale. As such, the expected relationship between population density and cost is negative.

Figure 2.4 YKY characteristics, YKY's wastewater models shows that YKY may be considered a 'typical' company in terms of scale (connected properties), tight consents and the number of CSOs. However, it is atypical relative to the upper (or lower) quartile in pumping station capacity per sewer, proportion of combined sewers and (population) density.

Figure 2.4 YKY characteristics, YKY's wastewater models



Source: Oxera analysis.

2.3.3 Growth enhancement

This section focuses on the cost drivers controlled for to explain growth-related enhancements in the water and wastewater models submitted by Ofwat and YKY for consultation.

- **New developments and new connections (water).** Ofwat's enhancement models are parsimonious, controlling for one driver in each model. OE4 controls for population served (smooth) and OE5 controls for the number of new connections (smooth).
- **First-time sewerage schemes.** Ofwat's first-time sewerage models control for various measures capturing activity involving s101A schemes, such as the number of s101A schemes completed, the number of connected properties served by such schemes, and the average number of properties per s101A scheme.
- **New developments, sewage growth and sewage treatment works and reducing sewer flooding risk for properties.** Ofwat's new developments, sewage growth, and reducing sewer flooding risk models control for various measures of scale of operations. In the first two models, Ofwat controls for residential population and total properties (household and non-household) billed for sewerage respectively. The subsequent two models control for the average load received per sewage treatment work incrementally to the scale driver to capture the possible economies of scale at treatment works.
- **YKY's BOTEX(Growth) drivers.** The drivers in YKY's BOTEX(Growth) models consist of BOTEX drivers and enhancement (growth) drivers appended. All wholesale water BOTEX(Growth) models control for enhancements to the supply/demand balance per DI, and two models further control for number of new connections per connected property. All wholesale wastewater BOTEX(Growth) models control for properties growth.

2.4 Modelling approach

2.4.4 Estimation approach

As considered in the modelling consultation, we use the pooled ordinary least squares (pooled OLS, or POLS) method as a starting point to estimate YKY's efficient baseline level of expenditure. The POLS approach assumes that each of the 107 observations in water and 60 observations in wastewater is independent, and ignores the fact that the data on the same set of companies is recorded over time.

Owing to the uncertainty in estimating the cost frontier using POLS (as the method cannot distinguish between statistical errors, inefficiency and company heterogeneity not captured in the model specification), Ofwat and other regulators make ad hoc adjustments to the estimated gap between the companies' actual cost and the benchmark. At PR14, Ofwat adjusted the gap to the upper-quartile (UQ) level of efficiency levels to account for statistical errors.

In contrast, when considering Bristol Water's appeal of PR14, the Competition and Markets Authority (CMA) considered that the average benchmark to be appropriate for Bristol Water based on the cost models it developed in the inquiry.[27] As shown in Oxera (2013),[28] such ad hoc adjustments may overcompensate or undercompensate for specific companies, even when the adjustment is broadly correct across the industry as a whole.

As such, we widen the evidence base to provide an unbiased estimate of YKY's efficient level of outturn expenditure. Part of this requires that multiple model specifications are considered that control for different sets of industry cost drivers to account for heterogeneity in the industry. Different cost drivers could proxy for key operational factors to different extent for companies and we consider that a triangulation of results across appropriate model specification can provide a more balanced view of companies' cost performance. In addition, we consider estimators that can take into account the panel structure of the data and possible unobserved company heterogeneity. To this end, we consider Random Effects (RE) models and panel Stochastic Frontier Analysis (SFA).

While comparable to the POLS estimator, the RE estimator can take into account unobserved firm heterogeneity when estimating model coefficients. RE models were used in the PR14 assessment,[29] and the enhancement expenditure models developed in the consultation process.

Based on the estimated parameters on the cost drivers, the RE estimator gives two efficiency estimates for each company: one is time-varying and the other time-invariant (i.e. the unobserved company-specific effect not captured through cost drivers). Whether the time-invariant prediction represents legitimate differences in efficient expenditure based on unobserved differences in regional operating environments, or permanent differences in efficiency, requires regulatory judgement. As at PR14, we assume that the time-invariant prediction represents permanent differences in efficiency when using this approach and test this assumption under the SFA model described below. To that end, similar to POLS, the RE estimator requires an ad hoc adjustment to derive efficient cost predictions.

SFA can separate the estimated residual into noise and efficiency by imposing assumptions on the distribution of the inefficiency and noise terms. This has particular advantages over the other two estimators considered in this report at the expense of additional assumptions; in particular, no ad hoc adjustment is required since statistical noise (data/modelling errors) is separated from inefficiency. The particular SFA model we use in this report, based on Kumbhakar et al. (2012),[30] is often referred to as the 'four-component model'. This is because it can separate the residual (the estimated gap between companies actual cost and the benchmark) into four components: (i) uncontrollable fixed differences in firms' operating environments; (ii) permanent differences in efficiency; (iii) time-varying noise; and (iv) time-varying efficiency differences.

As SFA can separate noise from inefficiency at a company-specific level, the estimated efficiency scores can be used to inform the choice of

[27] Competition and Markets Authority (2015), 'Bristol Water plc: A reference under section 12(3)(a) of the Water Industry Act 1991 Report', October, paras 4.205–4.245.
[28] Oxera (2013), 'Recommendations on cost assessment approaches for RIIO-ED1', February
[29] Ofwat (2014), 'Basic cost threshold models', April.
[30] Kumbhakar, S. C., Lien, G. and Hardaker, J. B. (2012), 'Technical efficiency in competing panel data models: A study of Norwegian grain farming', *Journal of Productivity Analysis*, **41**:2, September, pp. 1–7.

benchmark in POLS and RE models for a particular company.[31] For example, if an SFA model predicts a company to have a lower efficiency score than UQ-corrected POLS or RE, it may indicate that a UQ benchmark is too lenient and overcompensates for noise for that company. The converse may hold if the SFA model predicts a higher efficiency score than POLS or RE. This makes the choice of UQ (or another benchmark) for a particular company less dependent on ad hoc adjustments and regulatory judgements.

To summarise, our catch-up assessment is based on three main estimation approaches.

- **Pooled OLS approach**: this assumes that each observation represents an independent company, ignoring the fact that the data consists of repeated observations on the same companies over time.
- **Random Effects (RE) approach:** this estimation technique accounts for the panel structure of the data and treats the unobservable individual effects (i.e., company-specific factors) as random. No distributional assumption for the inefficiency component is needed for this approach.
- **SFA (four-component) approach**: further extensions in an SFA panel setting allow for explicit interpretation of the results in terms of uncontrollable company-specific effects, noise in data/modelling errors, persistent inefficiency, and transient inefficiency. Such a decomposition and interpretation is currently not possible using the other approaches.

2.4.5 Triangulation and model selection

Our assessment of YKY's efficient baseline level of expenditure focuses on the following suites of models:[32]

- **YKY-submitted models**. Wholesale BOTEX and BOTEX(Growth) models submitted by YKY in March 2018 as part of the consultation on econometric modelling. The modelled expenditure has been updated to be consistent with Ofwat's definition and the models have been re-estimated using the data provided in the consultation.
- **Ofwat-submitted models**. Wholesale BOTEX and selected enhancement models were presented (separately) in the same consultation document. As Ofwat did not submit BOTEX(Growth) models for consultation, BOTEX(Growth) is assessed by aggregating estimates from the BOTEX and growth-related enhancement models.

In their models, both YKY and Ofwat aim to capture a wide range of industry characteristics. However, some companies' performances could be sensitive to the exact choice of variable used to capture particular characteristic. There can be multiple legitimate variables to capture industry characteristics; given the constraints of econometric analysis, and the dataset available for consideration, only a limited number of cost drivers may be used in any one model. In view of the arguments and limitations

[31] Note that traditional SFA models will rarely predict a company to be 100% efficient. While this aspect of SFA is explored further in Kumbhakar, Parameter and Tsionas (2013), in this report, we present normalised efficiency scores where the frontier company has 100% efficient. See Kumbhakar, Parameter and Tsionas (2013), 'A zero inefficiency stochastic frontier model', Journal of Econometrics, **172**:1, pp. 66–76.

[32] The regression outputs of both YKY's and Ofwat's models is shown in Appendix A1 for wholesale water and A2 for wholesale watsewater.

above, we triangulate results across a range of models to mitigate biases that may arise when focusing on any particular model.

Additionally, we note that Ofwat's models can be developed further such that they capture appropriate industry drivers and the model outputs are aligned with operational, economic and statistical expectations. We have made such suggestions in the March submission on behalf of YKY. In this report, we consider a selection of Ofwat's models for triangulation based on which industry characteristics are considered and the level of statistical uncertainty. The two approaches to model selection are as follows:

Approach 1: select models based on the cost drivers included and their ability to control for key operational characteristics for YKY. Specifically, it is our understanding that Ofwat will engage in further model development to ensure that all models control for key operational characteristics.[33]

Approach 2: select models based on the relative certainty with which they predict either (i) YKY's expenditure or (ii) the benchmark companies' expenditure. Note that the statistical properties of models may change with further model development and updated data. We consider the following selection approaches.

- **Noise to signal ratio (benchmark or efficient companies)**—models are selected if a majority of benchmark companies had a noise to signal ratio less than the median in the industry. A noise to signal ratio is a measure of prediction uncertainty. The notion is that the models that predict the cost of the benchmark companies with reasonable certainty are able to determine the benchmark (i.e. cost frontier) with reasonable certainty.
- **Noise to signal ratio (YKY)—**models are selected if they predict YKY's expenditure with relative certainty, defined as having a noise to signal ratio lower than the median in the industry. The overall allowance for YKY is a function of the cost predictions for it from the models and the benchmark correction (if more stringent than the average). In this approach, we identify models that predict YKY's cost with reasonable certainty even where the benchmark may not be identified accurately to provide an alternative view.

We note that both Ofwat and YKY models presented in this report require further development in light of additional data refinements (2017/18 data undergoing Ofwat's quality assurance as well as companies' business plan data) and reinforced with bottom-up analysis from an operational perspective.

Our triangulation approach specifically consists of taking an average cost prediction across different suites of models and then applying an efficiency correction derived from the appropriate benchmark.

YKY and Ofwat both provided econometric cost models for BOTEX at the price control level, which we examine as a further cross-check on the results from the aggregate BOTEX analysis.[34] When aggregating results from value chain models, as well as results from Ofwat's BOTEX and enhancement models, our approach is to take an average prediction across the various suites of models, sum the relevant predictions to the relevant aggregate

[33] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.
[34] We do not consider models such as water treatment or sewage collection, which do not represent a complete price control.

level, and then apply the appropriate efficiency correction. This is to ensure that potential trade-offs across the value chain (and between BOTEX and BOTEX(Growth)) are captured and that companies are benchmarked to actual achieved leading performance in the industry. Value chain modelling may be able to better capture the different relationships between costs and specific cost drivers at the price control level. It also allows the efficient level of expenditure to be calculated directly, rather than relying on assumptions to apportion the predicted efficient expenditure from aggregate models into the two relevant price controls. However, value chain models can be sensitive to data allocation issues and reporting inconsistences, as noted by several companies in their responses to the March consultation.[35] As such, the core body of evidence focuses on aggregate BOTEX (and aggregate BOTEX(Growth)), and value chain models (aggregated to the overall level) are used only as a cross-check.

[35] For example, see Welsh Water's response to the modelling consultation, available here: https://www.ofwat.gov.uk/consultation/cost-assessment-pr19-consultation-econometric-modelling/#Responses, accessed 20 August.

3 YKY's efficient BOTEX level—water

In this section, we present our top-down assessment of YKY's efficient cost baseline for wholesale water. Both YKY's and Ofwat's econometric cost models have been used to assess YKY's expenditure and potential for cost reduction.

The section is structured as follows:

- section 3.1 provides an assessment of YKY's efficient baseline level of expenditure using models submitted by YKY;
- section 3.2 provides an assessment of YKY's efficient baseline level of expenditure using models submitted by Ofwat;
- section 3.4 summarises the results to derive a final range of estimates.

3.1 Catch-up assessment, YKY models

As previously discussed, we consider that the models submitted by YKY as part of the consultation process capture key industry-wide characteristics, and they therefore form part of our assessment of YKY's relative efficiency.

Table 3.1 shows YKY's rank in the six models used in the historical assessment. It appears to be a relatively efficient company in most models, being ranked between first and seventh out of 17. YKY tends to perform relatively worse in the SFA models compared to POLS and RE.

Table 3.1 YKY's relative ranking, YKY's models

	YKYWW1	YKYWW2	YKYWW3	YKYWW4	YKYWW5	YKYWW6	Average (triangulated)
Rank (pooled OLS)	1	4	1	4	5	6	3
Rank (RE)	1	4	1	5	6	5	2
Rank (SFA)	1	7	2	6	7	4	3

Source: Oxera analysis.

To derive an efficient cost prediction from POLS and RE models, we use the SFA results to inform the appropriate benchmark for YKY. The table below shows YKY's estimated efficiency score under three different benchmarks (upper quartile, upper quintile and upper decile) for both POLS and RE. This is compared to the estimated efficiency from SFA. The results are shown in Table 3.2.

It should be noted that the triangulated results (last column of the table) provide results based on average of cost predictions across models benchmarked to the appropriate level; this is our preferred approach. YKY has an efficiency of 1 if it is estimated to be at least as efficient as the chosen benchmark.

Table 3.2 YKY's estimated efficiency, YKY's models

	YKYWW 1	YKYWW 2	YKYWW 3	YKYWW 4	YKYWW 5	YKYWW 6	Average (triangulated)
OLS							
UQ	1.00	1.00	1.00	1.00	1.00	1.00	1.00
UQi	1.00	1.00	1.00	1.00	0.99	0.99	1.00
UD	1.00	0.92	1.00	0.93	0.95	0.94	1.00
RE							
UQ	1.00	1.00	1.00	1.00	1.00	1.00	1.00
UQi	1.00	1.00	1.00	0.98	0.99	1.00	1.00
UD	1.00	0.93	1.00	0.94	0.94	0.93	1.00
SFA							
Four-component	1.00	0.98	1.00	0.98	0.98	0.98	0.99

Source: Oxera analysis.

The results show that the YKY's position is largely invariant to the choice of benchmark and the three estimation approaches. In this regard, an upper-quartile benchmark, as considered by Ofwat at PR14[36] and Ofgem in the RIIO controls,[37] would suggest that YKY's historical expenditure is efficient.[38] It should be noted that although YKY tends to rank worse in the SFA models (see Table 3.1), the estimated level of inefficiency is modest.

Based on the above assessment, a summary of YKY's performance and the implied level of catch-up is shown in Table 3.3. YKY's historical efficient baseline assessed through YKY's models is estimated to be in the range of £1,137m–£1,150m.

Table 3.3 Summary of triangulated results, YKY's models

	OLS	RE	SFA
Modelled actual AMP costs (£m)	1,150	1,150	1,150
Efficiency-corrected predicted costs (£m)	1,150*	1,150*	1,137
Catch-up target	0%*	0%*	1%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

3.2 Catch-up assessment, Ofwat models

The YKY BOTEX models discussed in section 3.1 represent one piece of evidence for our assessment of an unbiased efficient historical cost baseline for YKY. Ofwat's econometric models provide an alternative assessment of YKY's efficient level of BOTEX, subject to the caveats outlined in section 2.4.5.

Table 3.4 shows YKY's performance across the twelve Ofwat BOTEX models. On these, YKY is only ranked outside the upper quartile in OWW10

[36] Ofwat (2014), 'Cost assessment – Advanced econometric models', March
[37] Ofgem (2012), 'RIIO-GD1: Initial Proposals – Step-by-step guide for the cost efficiency assessment methodology', August, p.13
[38] YKY's efficiency scores across all models and estimation techniques are shown in Appendix A1 of wholesale water and A2 for wholesale wastewater.

and OWW12. There is no large difference in YKY's ranking when the different estimators are considered.

Table 3.4 YKY's rank, Ofwat models

	OLS	RE	SFA
Rank range (out of 17)	1–6	1–7	1–7

Note: The full range of estimated ranks is presented in Table A 4 of Appendix A1.2.

Source: Oxera analysis.

Ofwat's models tend to predict YKY to be broadly efficient, as shown in Table 3.5.

Table 3.5 YKY's estimated efficiency, Ofwat's models

	Range	Average (triangulated)
OLS		
Upper quartile	0.99–1	1.00
Upper quintile	0.98–1	1.00
Upper decile	0.92–1	1.00
RE		
Upper quartile	0.98–1	1.00
Upper quintile	0.97–1	1.00
Upper decile	0.93–1	1.00
SFA		
Four-component	0.93–1	0.99

Note: The full range of estimated efficiencies is presented in Table A 3 of Appendix A1.2.

Source: Oxera analysis.

As discussed, the models presented by Ofwat may underestimate or overestimate YKY's efficient cost baseline. As such, we have triangulated different suites of Ofwat's models based on a statistical or operational criterion, following the approach outlined in section 2.4.5.

The full Ofwat suite was considered for triangulation under Approach 1. Although YKY might be considered atypical in terms of scale and maintenance requirements, all of Ofwat's models control for these factors directly. Since YKY may be considered typical in the remaining variables (see Figure 2.1), it is not clear ex ante which models may overcompensate or undercompensate for YKY's characteristics. Approach 2 is as set out in section 2.4.5.

Table 3.6 shows YKY's efficient cost prediction under the POLS and RE estimator when applying an upper-quartile benchmark under three triangulation approaches. Across the three suites within Ofwat's set of models, YKY is estimated to be an efficient company.

Table 3.6 YKY's efficient expenditure, POLS and RE (upper-quartile benchmark)

	Outturn AMP costs (£m)	POLS efficient cost baseline (£m)	RE efficient cost baseline (£m)
Full suite (Ofwat)	1,150	1,150 (0%)*	1,150 (0%)*
Noise to signal ratio (benchmark companies)[1]	1,150	1,150 (0%)*	1,150 (0%)*
Noise to signal ratio (YKY)[2]	1,150	1,150 (0%)*	1,150 (0%)*

Note: [1] OWW2, 3, 7–12. [2] OWW1–4, 7–12. * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. Numbers in parentheses refer to YKY's catch-up target. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Based on the above assessment, as well as estimates under SFA, a summary of YKY's performance and the implied level of catch-up is shown in Table 3.7. YKY's historical efficient expenditure assessed through Ofwat's model specifications is estimated to be in the range of £1,142m–£1,150m.

Table 3.7 Summary of triangulated results, Ofwat's models

	OLS	RE	SFA
Modelled actual AMP costs (£m)	1,150	1,150	1,150
Efficiency-corrected predicted costs (£m)	1,150*	1,150*	1,142
Catch-up target	0%*	0%*	1%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

3.3 Catch-up assessment, value chain models

Separate modelling of the components of base TOTEX—for example, water resources and network plus separately—could provide an alternative view on YKY's efficiency, and may better capture company-specific factors affecting the cost components.

Although we have already noted reservations regarding value chain modelling on the limited dataset available for consideration and given reporting concerns noted by companies in the consultation response, it can serve as a cross-check to the aggregate analysis shown in the previous sections.

While a rigorous selection of the benchmark was not considered for the value chain models, the upper-quartile benchmark is supported by the evidence presented in sections 3.1 and 3.2.

Table 3.8 Value chain modelling results—YKY models

	Water resources		Network+	
	Range	**Triangulated**	**Range**	**Triangulated**
Modelled actual AMP costs (£m)	120	120	1,030	1,030
OLS				
UQ efficient cost (£m)	103–106	103	1,030–1,030	1,030*
Rank	10–11	10	3–4	3
RE				
UQ efficient cost (£m)	102–105	102	1,030–1030	1,030*
Rank	10–11	10	3–5	4

Note:* YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Table 3.8 shows YKY's performance in the value chain models submitted by YKY. YKY performs significantly better in network plus models compared to water resources models. However, given that the potential trade-offs across the value chain are not accounted for in value chain modelling, it is not clear whether the estimated gaps in the individual controls represent efficiency.

Table 3.9 shows results from Ofwat's value chain models. These support the analysis in section 3.1 that YKY is estimated to be relatively efficient in network plus and relatively inefficient in water resources.

Table 3.9 Value chain modelling results—Ofwat models

	Water resources		Network+	
	Range	**Triangulated**	**Range**	**Triangulated**
Modelled actual AMP costs (£m)	120	120	1,030	1,030
OLS				
UQ efficient cost (£m)	81–103	94	1,030–1,030	1,030*
Rank	11–16	13	1–4	1
RE				
UQ efficient cost (£m)	70–103	88	1,030–1,030	1,030*
Rank	11–16	15	1–5	1

Note:* YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Table 3.10 shows the outcome when the results from value chain models are aggregated to give an estimate of YKY's efficient overall BOTEX expenditure.

Table 3.10 Aggregating results

	YKY disaggregated results	Ofwat disaggregated results
Modelled actual AMP costs (£m)	1,150	1,150
OLS		
UQ efficient cost (£m)	1,150*	1,150*
catch-up target	0%*	0%*
Rank	3	1
RE		
UQ efficient cost (£m)	1,150*	1,150*
Catch-up target	0%*	0%*
Rank	3	2

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Aggregating results from value chain models supports the conclusion from the analysis presented in sections 3.1 and 3.2 that YKY is a relatively efficient company on the dataset considered in the modelling.

3.4 Summary of results

Considering the results from a range of models and specifications developed by YKY and Ofwat and evaluated under different estimation techniques, the analysis presented indicates that YKY's efficient BOTEX is in the range of £1,137m–£1,150m[39] with a central estimate of £1,150m.[40] This compares to an observed expenditure of £1,150m, implying a catch-up of approximately 0%.

[39] Comparing results in Table 3.3 and Table 3.7.
[40] The median of results in Table 3.3 and Table 3.7.

4 YKY's efficient BOTEX(Growth) level—water

In this section we present our top-down assessment of YKY's efficient BOTEX(Growth) baseline for wholesale water. Both YKY's and Ofwat's econometric models have been used in the process.

As discussed earlier in Section 2.4, YKY developed BOTEX(Growth) models for the modelling consultation, thus allowing for a direct assessment of YKY's efficient level of BOTEX(Growth). Ofwat modelled certain enhancement costs separately, and as such, require aggregating results from growth-related enhancement models with results from its BOTEX models, following the triangulation technique outlined in section 2.4.

4.1 Catch-up assessment, YKY's models

As previously discussed, YKY's BOTEX(Growth) models provide an estimate of YKY's historic relative efficiency directly. Table 4.1 shows YKY's relative performance in YKY's BOTEX(Growth) models.

Table 4.1 YKY's ranking, YKY's models

	YKYWW7	YKYWW8	YKYWW9	YKYWW10	Average (triangulated)
POLS	1	1	4	4	2
RE	1	1	3	3	2
SFA	4	6	6	8	5

Source: Oxera analysis.

The table shows that YKY performs relatively well in BOTEX(Growth) models, being more efficient than the upper quartile in all POLS and RE models. However, YKY's position worsens when using SFA to estimate the models.

Table 4.2 shows YKY's estimated efficiency score in YKY's BOTEX(Growth) models.

Table 4.2 YKY's estimated efficiency, YKY's models

	YKYWW1	YKYWW2	YKYWW3	YKYWW4	Average (triangulated)
OLS					
UQ	1.00	1.00	1.00	1.00	1.00
UQi	1.00	1.00	1.00	1.00	1.00
UD	1.00	1.00	0.99	1.00	1.00
RE					
UQ	1.00	1.00	1.00	1.00	1.00
UQi	1.00	1.00	1.00	1.00	1.00
UD	1.00	1.00	1.00	1.00	1.00
SFA					
Four-component	0.99	0.99	0.99	0.99	0.99

Source: Oxera analysis.

The table generally supports the view that YKY is relatively efficient in BOTEX(Growth). Although Table 4.1 shows that YKY performs worse in SFA models, Table 4.2 reveals that the extent of this decline is limited. Specifically, even when YKY is ranked eighth out of 17, the estimated catch-up is approximately 1%. Note that the SFA models predict a lot of noise in the sample, so the estimated efficiency range across the industry is narrow.[41]

Table 4.3 summarises the results from YKY's BOTEX(Growth) models. The analysis supports the view that YKY is generally efficient in BOTEX(Growth), with an estimated catch-up of between 0 and 1%.

Table 4.3 Summary of triangulated results, YKY models

	OLS	RE	SFA
Modelled actual AMP costs (£m)	1,217	1,217	1,217
Efficiency-corrected predicted costs (£m)	1,217*	1,217*	1,205
Catch-up target	0%*	0%*	1%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

4.2 Catch-up assessment, Ofwat's models

Ofwat did not publish BOTEX(Growth) models in the main consultation process, but it did publish two enhancement models in new developments. This is a growth-related enhancement expenditure and as such can be used to inform YKY's level of catch-up, even if it omits some growth enhancement activity.[42]

It should be noted that Ofwat only published results from the RE estimator in the appendix on modelling results. We do not consider SFA for enhancement modelling.

Table 4.4 YKY's ranking, Ofwat's models

	OE4	OE5
Rank (pooled OLS)	1	1
Rank (RE)	1	1

Source: Oxera analysis.

Table 4.4 shows YKY's ranking in Ofwat's enhancement (growth) models. YKY is estimated to be the frontier company in these models. As such, further results from analysis, as shown in previous sections, have been omitted.

[41] The SFA models do estimate time-varying inefficiency close to the 10% statistical significance level. The SFA results for the BOTEX(Growth) models are presented in Table A 12 of Appendix A1.4.
[42] Such as enhancements to the supply/demand balance and resilience.

Table 4.5 Aggregating results from disaggregated analysis

	BOTEX	Enhancement**	BOTEXGrowth)
Modelled actual AMP costs (£m)	1,150	24	1,174
OLS rank	1	1	1
OLS Catch-up target	0%*	-	0%*
RE rank	1	1	1
RE Catch-up target	0%*	-	0%*

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. ** Enhancement expenditure is not assessed in isolation but for the assessment of BOTEX(Growth). Therefore the catch-up target estimate for enhancement modelling is not reported. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Table 4.5 shows that YKY's level of BOTEX(Growth) is efficient in Ofwat's models and supports the analysis presented in Table 4.3.

4.3 Summary of results

The analysis in this section suggests that YKY is broadly efficient in BOTEX(growth). Specifically, the estimated efficiency gap is approximately 0% to 1% in YKY's models and approximately 0% in Ofwat's models. This indicates an overall catch-up in the range of 0% to 1%.

5 YKY's efficient BOTEX level—wastewater

In this section, we present our top-down assessment of YKY's efficient cost baseline for wholesale wastewater. Both YKY's and Ofwat's econometric cost models have been used to assess YKY's expenditure and potential for cost reduction.

This section is structured as follows:

- section 5.1 provides an assessment of YKY's efficient baseline level of expenditure using models submitted by YKY;
- section 5.2 provides an assessment of YKY's efficient baseline level of expenditure using models submitted by Ofwat;
- section 5.4 summarises the results to derive a final range of estimates.

5.1 Catch-up assessment, YKY models

As previously discussed, the models submitted by YKY as part of the consultation process capture key industry-wide characteristics. Subject to the models undergoing further re-evaluation in light of new data, these models form part of our assessment of YKY's relative efficiency.

Table 5.1 reflects the estimated rank based on the historical assessment of relative efficiency in YKY's four models. YKY is estimated to be a relatively efficient company, ranked within the top three companies (out of 10) with the exception of YKYWWW1. The estimated efficiency position for YKY is broadly consistent across the three estimation approaches as set out in section 2.4.4.

Table 5.1 YKY's relative ranking, YKY's models

	YKYWWW1	YKYWWW2	YKYWWW3	YKYWWW4	Average (triangulated)
POLS	2	2	1	1	1
RE	4	2	1	1	2
SFA	4	2	1	1	1

Source: Oxera analysis.

Table 5.2 shows YKY's estimated efficiency score under the three benchmarks that we are considering (upper quartile, upper quintile and upper decile) for both POLS and RE. A comparison against efficiency estimates from SFA is performed to determine the appropriate benchmark for YKY.

Table 5.2 YKY's estimated efficiency, YKY's models

	YKYWWW1	YKYWWW2	YKYWWW3	YKYWWW4	Triangulated
OLS					
Upper quartile	1.00	1.00	1.00	1.00	1.00
Upper quintile	1.00	1.00	1.00	1.00	0.96
Upper decile	0.97	0.99	1.00	1.00	0.89
RE					
Upper quartile	0.99	1.00	1.00	1.00	1.00
Upper quintile	0.96	1.00	1.00	1.00	1.00
Upper decile	0.84	1.00	1.00	1.00	0.96
SFA					
Four-component	1.00	1.00	1.00	1.00	1.00

Note: The efficiency scores are calculated as the ratio of predicted cost divided by modelled costs.

Source: Oxera analysis.

OLS and RE estimate YKY to be efficient under the three benchmarks across all of YKY's models, with the exception of YKYWWW1. In YKYWWW1, YKY is estimated to be broadly efficient under OLS and RE when assuming the upper-quartile benchmark, but is estimated to be about c.3% in OLS and c.16% in RE when the upper-decile benchmark is assumed. However, this benchmark is with respect to the frontier company (i.e. upper decile suggests 10% of 10 companies, which is 1)—i.e. assuming that there is no noise in the models, which is an extreme assumption.

Alternatively, across the board, the results from SFA suggests that the upper-quartile benchmark is likely to be most appropriate when setting the efficiency challenge for YKY.

The efficient historical costs obtained using YKY-submitted models are summarised in Table 5.3. As YKY was estimated to be more efficient than the benchmark with OLS and RE, the efficient expenditure was set to the outturn expenditure. As a result, across the three modelling approaches, YKY's estimated efficient cost baseline is £1,361m, with an implied level of catch-up of 0%.

Table 5.3 Summary of triangulated results, YKY's models

	OLS	RE	SFA
Modelled actual AMP costs (£m)	1,361	1,361	1,361
Efficiency-corrected predicted costs (£m)	1,361*	1,361*	1,361
Catch-up target	0%*	0%*	0%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

5.2 Catch-up assessment, Ofwat models

Similar to our approach in water, Ofwat's BOTEX models provide an alternative view to the estimates suggested by YKY's models in section 5.1, subject to the caveats outlined in section 2.4.5. Table 5.4 reflects a range of

YKY's relative efficiency positions estimated across the models using POLS, RE and SFA.

Table 5.4 YKY's rank, Ofwat's models

	OLS	RE	SFA
Rank range (out of 10)	3–7	4–7	4–7

Note: The full range of estimated ranks is presented in Table A 16 of Appendix A2.2.

Source: Oxera analysis.

Contrary to YKY's models, Ofwat's models suggests that YKY is a relatively inefficient company and is ranked between the upper and lower quartiles (out of 10).

Ofwat controls for key operational factors **alternately** across its eight BOTEX models—in other words, all of Ofwat's BOTEX models exclude certain operationally important drivers for the industry as well as for YKY. This is especially important when assessing YKY's efficient cost baseline, given that YKY appears to be an atypical company relative to the industry average with respect to the drivers in Ofwat's models (see Figure 2.3). As such, we would recommend developing a more robust suite of models for sewerage service that are aligned with operational, economic and statistical expectations.

Consistent with the approach set out in section 2.4.4, the results from SFA are used to inform the appropriate benchmark (see Table 4.5). Across the suite, YKY's efficiency gap is estimated to be approximately 3% with SFA, suggesting that the appropriate efficiency challenge for YKY is between the upper-quartile target and upper-quintile target. As such, there appears to be no strong reason to deviate from the regulatory precedent followed by Ofwat at PR14 and Ofgem in the RIIO controls in considering an upper-quartile benchmark.

In estimating efficiency under SFA, three of Ofwat's models (OWWW5, OWWW7 and OWWW8) were unable to separate noise and inefficiency.[43] A possible reason for this could be that the models are mis-specified, requiring further development.

Table 5.5 YKY's estimated efficiency, Ofwat's models

	Range	Average (triangulated)
OLS		
Upper quartile	0.95–1	0.99
Upper quintile	0.93–0.99	0.94
Upper decile	0.89–0.98	0.90
RE		
Upper quartile	0.93–0.99	0.98
Upper quintile	0.9–0.97	0.95
Upper decile	0.87–0.94	0.94
SFA		
Four-component	0.91–1	0.97

Note: The full range of estimated efficiencies is presented in Table A 15 of Appendix A2.2.

Source: Oxera analysis.

[43] The SFA results for the BOTEX models are presented in Table A 23 in Appendix A2.4.

As discussed, Ofwat's models may underestimate or overestimate YKY's efficient cost baseline. As such, we have triangulated different suites of Ofwat's models based on a statistical or operational criterion, using the approach outlined in section 2.4.5. This approach will need to be updated with additional data and modelling refinement.

As noted, YKY appears to be an atypical company based on the cost drivers used in Ofwat's models (see Figure 2.3). Additionally, since key industry operational characteristics are omitted in each of the eight Ofwat BOTEX models (as discussed earlier), there is no one model that adequately accounts for the atypical nature of YKY in these drivers. Indeed, no subset of Ofwat's suite appears to appropriately capture YKY's operational characteristics; therefore, the full Ofwat suite was considered for triangulation under Approach 1. Approach 2 is as set out in section 2.4.5.

Table 5.6 shows YKY's efficient cost prediction under the three triangulated sets of models. Across the triangulated sets, the historical efficient baseline for YKY is estimated to be between £1,345m and £1,350m, with a median efficiency gap of approximately 1% under POLS. The assessment based on the RE estimator is broadly consistent with the results for POLS. Under RE, the historical efficient baseline for YKY is estimated to be between £1,310m and £1,321m, with a median efficiency gap of approximately 4%.

Table 5.6 YKY's efficient expenditure, POLS and RE (upper-quartile benchmark)

	Outturn AMP costs (£m)	POLS efficiency-corrected predicted costs (£m)	RE efficiency-corrected predicted costs (£m)
Full suite (Ofwat)	1,361	1,350 (1%)	1,310 (4%)
Noise to signal ratio (benchmark companies)[1]	1,361	1,345 (1%)	1,321 (3%)
Noise to signal ratio (YKY)[2]	1,361	1,350 (1%)	1,310 (4%)

Note: [1] OWWW1, 3, 5, 8. [2] OWWW1–8. Numbers in parentheses reflect YKY's catch-up gap. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Table 5.7 incorporates the assessment of YKY's efficient historical cost baseline according to POLS and RE and compares it with the assessment based on SFA.[44] A central estimate of YKY's historical efficient costs across Ofwat's models ranges between £1,310m and £1,350m over an AMP, with a catch-up target between 1% and 4%.

Table 5.7 Summary of triangulated results, Ofwat's models

	OLS	RE	SFA
Modelled AMP costs (£m)	1,361	1,361	1,361
Efficiency-corrected predicted costs (£m)	**1,350**	**1,310**	**1,325**
Catch-up target	1%	4%	3%

Note: All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

[44] The central estimate corresponds to the median value across the different triangulated suites for POLS and RE, whereas the SFA results consider Ofwat's full suite only.

Source: Oxera analysis.

5.3 Catch-up assessment, value chain models

As with wholesale water, value chain modelling is used as a cross-check to the aggregate wholesale analysis above. Similarly, an upper-quartile benchmark is assumed when assessing the value chain results. Table 5.8 and Table 5.9 reflect the assessment of YKY's historical efficient baseline by price control when considering YKY's and Ofwat's models respectively.

Table 5.8 Value chain modelling results—YKY models

	Network+		Bioresources	
	Range	**Triangulated**	**Range**	**Triangulated**
Modelled actual AMP costs (£m)	1,015	1,015	346	346
OLS				
UQ efficient cost (£m)	1,015–1015	1,015*	265–346	300
Rank	1–3	1	2–10	9
RE				
UQ efficient cost (£m)	979–1,015	1,015*	268–343	300
Rank	1–5	2	5–10	9

Note:* YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

In both YKY's and Ofwat's network+ models (see Table 5.8 and Table 5.9 respectively), YKY's historical expenditure is assessed to be an efficient company. On the other hand, YKY is estimated to be relatively inefficient in bioresources. However, as noted earlier, evaluating companies' efficiency position for each price control level in isolation does not account for the operational trade-offs.

Furthermore, it was noted by companies in response to the modelling consultation that the Ofwat value chain models (especially network plus) appear simple and do not capture key operational characteristics of the industry.[45] Therefore, despite YKY's relatively efficient performance in Ofwat's network plus models, these models will require further development.

[45] See, for example, Southern Water's response to the consultation, available here: https://www.ofwat.gov.uk/consultation/cost-assessment-pr19-consultation-econometric-modelling/#Responses, accessed 20 August.

Table 5.9 Value chain modelling results—Ofwat models

	Network+		Bioresources	
	Range	Triangulated	Range	Triangulated
Modelled actual AMP costs (£m)	1,015	1,015	346	346
OLS				
UQ efficient cost (£m)	1,015–1,015	1,015*	294–312	292
Rank	1–3	2	8–8	8
RE				
UQ efficient cost (£m)	1,005–1015	1,015*	274 to 295	274
Rank	2–4	3	8–9	8

Note:* YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

Table 5.10 below reflects YKY's efficiency position when aggregating up the results from the value chain modelling. The results from the aggregation suggest that YKY is an efficient company in both YKY's and Ofwat's models. This is consistent with the assessment in the preceding sections.

This result is unsurprising given YKY's efficiency position in Network+, which comprises a significant portion of the value chain. Furthermore, the results in Table 5.10 reinforce the argument that value chain models may not adequately account for operational trade-offs, and setting efficiency targets at the price control level in isolation may be inappropriate.

Table 5.10 Aggregating results

	YKY disaggregated results	Ofwat disaggregated results
Modelled actual AMP costs (£m)	1,361	1,361
OLS		
UQ efficient cost (£m)	1,361*	1,361*
catch-up target	0%*	0%*
Rank	2	2
RE		
UQ efficient cost (£m)	1,361*	1,361*
Catch-up target	0%*	0%*
Rank	2	3

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

5.4 Consideration of YKY's forward-looking cost adjustment claims

As part of the assessment, YKY's and Ofwat's models were used to provide an initial view of YKY's efficient allowance for AMP7, using activity forecasts provided by YKY.

The models appear to account for base expenditure related to growth in operations, to some extent, since growth in connected properties is explicitly controlled for in the models (see section 2.3.2). On the contrary, the drivers relating to the atypical cost adjustment claims, among other increases in expenditure in AMP7 (e.g. WINEP expenditure), are not robustly captured in the models (over and above the general correlation captured in the cost drivers included in the models). The drivers for these forward-looking cost adjustment claims include the growth in sludge produced—from the increased compliance to remove phosphorous—and the reduction in the number of flooding incidents and/or number of blockages. [46]

Furthermore, the allowance for growth in population estimated by the models is based on the expected average growth of YKY's service and is estimated based on the historic relationship between properties growth and expenditure for the industry. Should the unit cost to provide wastewater service in YKY's operational areas be larger than the historical unit cost estimated by the models (e.g. because of profiling of expenditure or differences in the activities undertaken, service improvements targeted), the models will not adequately account for the expected increase in expenditure for YKY over AMP7.

Additionally, the historical relative efficiency position assessed in this report for YKY through econometric modelling may be driven, to some extent, by YKY's and the industry's historical levels of activity, as well as expenditure profiling over the modelled period. This may be addressed to some extent should Ofwat include companies' business plan data in their cost assessment models.

From the limitations suggested above, a naïve rolling forward of cost allowances for YKY based on outturn relationships is unlikely to account for the step-change in YKY's expenditure for AMP7, and suggests that its cost claims should be assessed incremental to the model forecasts.

The highlighted limitations also apply to YKY's water service AMP7 projections, even if YKY does not intend to submit cost claims in the area.

5.5 Summary of results

The analysis presented in the preceding sections considers a range of modelling specifications developed by YKY and Ofwat, evaluated under different estimation techniques. From the analysis, similarly to wholesale water, YKY is assessed to be broadly efficient on a historical basis, with efficient BOTEX being in the range of £1,310m–£1,361m,[47] with a central estimate of £1,355m.[48] This compares to an observed expenditure of £1,361m, implying a catch-up efficiency target of approximately 0.4%.

[46] See YKY WWN+01 Cellared properties, YKY WWN+04 Wastewater Growth and YKY BR-01 Bioresources - WINEP enhancement expenditure for details of its cost claims.
[47] Comparing results from Table 5.3 and Table 5.7.
[48] A median of results from Table 5.3 and Table 5.7.

6 YKY's efficient BOTEX(Growth) level—wastewater

In this section, we present our top-down assessment of YKY's efficient BOTEX(Growth) baseline for wholesale wastewater. Both YKY's and Ofwat's econometric models have been used in the process.

Similarly to water, YKY developed BOTEX(Growth) models for wastewater that will inform our assessment of YKY's efficient level of BOTEX(Growth). As Ofwat also developed enhancement models separately from BOTEX in wastewater, Ofwat's BOTEX and enhancement expenditure estimates are aggregated as per the triangulation technique described in section 2 to arrive at an efficient BOTEX(Growth) estimate for YKY.

6.1 Catch-up assessment, YKY's models

As reflected in Table 6.1, the assessment of aggregate BOTEX(Growth) in YKY's models suggests that YKY is ranked within the upper quartile under OLS and RE (apart from in YKYWWW1). The results from SFA suggest that YKY is ranked just outside the upper quartile and seventh (out of ten) in YKYWWW1. However, in the four BOTEX(Growth) models, SFA is unable to statistically distinguish between noise and inefficiency in the models.[49]

Table 6.1 YKY's ranking, YKY's models

	YKYWWW5	YKYWWW6	YKYWWW7	YKYWWW8	Average (triangulated)
POLS	2	3	2	2	1
RE	4	3	3	2	3
SFA	7	4	4	3	4

Source: Oxera analysis.

Correspondingly, Table 6.2 suggests that YKY is estimated to be efficient when the upper-quartile benchmark is assumed and slightly inefficient under the upper-quintile and upper-decile benchmarks.

Given that the SFA approach is unable to statistically distinguish between noise and inefficiency, it does not provide insight on which of the three ad hoc adjustments considered could be appropriate for YKY and warrant a more detailed investigation.

Nevertheless, the upper-quartile benchmark may be appropriate, consistent with the assumption in BOTEX models, where the statistical quality of the models are comparatively better (in terms of adjusted R^2 and range of efficiency scores across the industry).

[49] The SFA results for the BOTEX(Growth) models can be found in Table A 24 of Appendix A2.4.

Table 6.2 YKY's estimated efficiency, YKY's models

	YKYWWW 5	YKYWWW 6	YKYWWW 7	YKYWWW8	Average (triangulated)
OLS					
UQ	1.00	1.00	1.00	1.00	1.00
UQi	1.00	0.99	1.00	1.00	1.00
UD	0.97	0.98	0.99	0.98	0.99
RE					
UQ	1.00	1.00	1.00	1.00	1.00
UQi	0.94	0.99	0.97	1.00	0.97
UD	0.86	0.98	0.97	0.98	0.95
SFA					
Four-component	1.00	1.00	1.00	1.00	1.00

Source: Oxera analysis.

YKY's efficient historical BOTEX(Growth) suggested in YKY's models across the various modelling approaches is summarised in Table 6.3. Similarly to its BOTEX position, YKY is estimated to be an efficient company in overall BOTEX(Growth).

Table 6.3 Summary of triangulated results, YKY models

	OLS	RE	SFA
Modelled actual AMP costs (£m)	1,490	1,490	1,490
Efficiency-corrected predicted costs (£m)	**1,490***	**1,490***	**1,486**
Catch-up target	0%*	0%*	0%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

6.2 Catch-up assessment, Ofwat's models

In this section, the results from Ofwat's growth-related enhancement models are presented to inform the assessment of YKY's historic efficient baseline for BOTEX(Growth). The two enhancement categories (first-time sewerage; and new developments, sewage growth, and reducing sewer flooding risk models) modelled and published by Ofwat are considered growth-related enhancements. Ofwat did not publish overall BOTEX(Growth) models for consultation.

Similarly to water, Ofwat published results from the RE estimator in the appendix on modelling results for their enhancement models. Similarly, we do not consider SFA for enhancement modelling.

Table 6.4 YKY's ranking, Ofwat's first-time sewerage models

	OE6	OE7	OE8
Rank (pooled OLS)	5	2	2
Rank (RE)	3	2	2

Source: Oxera analysis.

The estimated efficiency position of YKY in Ofwat's first-time sewerage models suggests that YKY is relatively efficient on first-time sewerage (see

Table 6.4). There are a number of companies—including YKY—that have a relatively low expenditure on first-time sewerage historically. Thus, Ofwat's approach to modelling enhancements appears to estimate an inappropriately large range of efficiency scores across the industry.[50]

Table 6.5 YKY's ranking, Ofwat's new developments, sewage growth, and reducing sewer flooding risk models

	OE9	OE10	OE11	OE12
Rank (pooled OLS)	6	6	6	6
Rank (RE)	6	6	6	6

Source: Oxera analysis.

In Table 6.5, Ofwat's new developments, sewage growth, and reducing sewer flooding risk enhancement models suggest that YKY is estimated to be relatively inefficient in terms of its rank across the industry.

Table 6.6 Aggregating results from Ofwat's BOTEX and disaggregated enhancement analysis

	BOTEX	First-time sewerage enhancement**	Sewage growth enhancement**	BOTEX (Growth)
Modelled actual AMP costs (£m)	1,361	0.76	136	1,498
Rank (OLS)	4	3	6	4
Catch-up target (OLS)	1%	-	-	2%
Rank (RE)	6	2	6	5
Catch-up target (RE)	4%	-	-	4%

Note: * YKY is estimated to be at least as efficient as the benchmark. The catch-up efficiency target has therefore been set at 0%. ** Enhancement expenditure is not assessed in isolation but for the assessment of BOTEX(Growth). Therefore, the catch-up target estimates for enhancement modelling are not reported. All costs are shown in 2016/17 prices. AMP expenditure refers to the average expenditure per annum, multiplied by five.

Source: Oxera analysis.

The results from aggregating estimates from YKY's BOTEX with Ofwat's disaggregated enhancement modelling are presented in Table 6.6 above. The aggregated BOTEX(Growth) results for YKY from Ofwat's models are broadly consistent with the estimated efficiency position for BOTEX, given its significant proportion of BOTEX(Growth).

The analysis of Ofwat's base and enhancement expenditure modelling suggests that YKY is slightly inefficient in BOTEX(Growth) relative to the upper-quartile benchmark, with a efficiency gap estimated to be between 2% and 4%.

6.3 Summary of results

The analysis in the preceding section suggests that YKY is broadly efficient historically in BOTEX(Growth). The estimated efficiency gap is approximately 0% to 0.3% when estimated directly in YKY's aggregate BOTEX(Growth) models, and the estimated efficiency gap is 2% to 4%

[50] The range of efficiency scores are illustrated in Appendix 1a of Ofwat's consultation on econometric cost modelling, available here: https://www.ofwat.gov.uk/wp-content/uploads/2018/03/Appendix-1a-Supplementary-charts-for-models-in-appendix-1_Final.xlsx, accessed 20 August.

when aggregating Ofwat's BOTEX and selected growth enhancement models modelled separately. This indicates an overall catch-up in the range of 0% to 4%, with a central estimate of about 0.3%.[51]

As YKY's BOTEX(Growth) models are very similar to their BOTEX models (the former controlling additionally for properties growth), similar conclusions (i.e. caveats) noted in a mechanistic rolling forward of historical relationship of BOTEX (see section 5.4) can be drawn for BOTEX(Growth) as well.[52]

[51] A median of results from Table 6.3 and Table 6.6.
[52] No forward rolling of Ofwat's enhancement allowance for AMP7 was undertaken given the atypical nature of the expenditure.

7 Conclusion

In this report, we provide an independent assessment of YKY's historical BOTEX as well as BOTEX(Growth) performance for its wholesale water and wastewater services, and the extent to which YKY's cost projections and claims over AMP7 are captured through a roll-forward of the historical analysis.

To determine YKY's relative efficiency on BOTEX and BOTEX(Growth), models submitted by YKY and published by Ofwat as part of Ofwat's modelling consultation were considered. The pooled OLS estimation approach used in the modelling consultation was examined, alongside SFA and RE, which are able to accommodate company heterogeneity and isolate inefficiency from uncertainty/noise. The alternative approaches are used widely by UK and continental European regulators.

Based on the last six years of data and a wider consideration of aggregation and triangulation possibilities, YKY's historical BOTEX and BOTEX(Growth) spend on water is assessed to be broadly efficient; similarly, its outturn wastewater BOTEX is estimated to be efficient, while its wastewater BOTEX(Growth) has a modest gap. These results are robust to the consideration of the 2017/18 data.

To establish if YKY's proposed cost claims for PR19 are captured through a naïve roll-forward of the outturn analysis, we estimated the forecast efficient baseline for AMP7 based on forecast activity levels provided by YKY. The estimated variance between the YKY's business plan expenditure and model forecasts in both YKY's and Ofwat's models is notably large.

A driving factor is that the activities/cost drivers used in YKY's and Ofwat's models do not robustly represent YKY's forward-looking expenditure requirements in AMP7 (e.g. expenditure for WINEP and further improvement to service levels). Additionally, a simple roll forward of companies' efficient baselines may be influenced by activity levels across the modelling period and company-specific expenditure profiling. From the model forecasts using a roll-forward approach, it appears likely that YKY's cost requirements over AMP7 are underestimated, and that YKY's cost adjustment claims for PR19 are incremental. Should Ofwat model companies' business plan assumptions over AMP7, it could provide a reference point, to some extent, on the need and materiality of some of YKY's claims.

In light of the limitations illustrated in this report, careful consideration in assessing YKY's and the industry's forward-looking expenditure and YKY's cost claims will be required.

A1 Regression output—wholesale water

A1.1 Ofwat's wholesale water BOTEX models

The tables below show regression results from Ofwat's models based on the dataset published by Ofwat as part of the consultation on econometric cost modelling in March 2018.[53]

Two key diagnostic tests are shown at the bottom of the regression table using pooled OLS (Table A 1)—the RESET test and the Breusch–Pagan Lagrangian multiplier test. The former is used to detect whether the functional form of the model is correctly specified whereas the latter tests for panel structure in the data.

The higher the p-value of the RESET test, the more likely it is that the model is well-specified. The null p-value of the Breusch–Pagan test recognises a panel structure in the data and therefore justifies the use of panel estimation techniques such as Random Effects (RE).

Results from RE regressions are shown in Table A 2. The diagnostic tests (Hausman and over-identification tests) suggest that a RE panel specification tends to be preferred over an alternative panel specification such as Fixed Effect.

YKY's estimated efficiency scores and ranking in each of Ofwat's models can be found in Table A 3 and Table A 4, respectively.

[53] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.

Table A 1 Regression results—Ofwat's BOTEX models (pooled OLS)

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	OWW9	OWW10	OWW11	OWW12
Dependent variable	Ln(wholesale water base cost)											
Ln(connected properties)	1.109***	1.078***	1.114***	1.053***	1.037***	1.081***						
Ln(length of mains)							1.114***	1.072***	1.114***	1.086***	1.031***	1.082***
% mains renewed and relined	0.177	0.185*	0.191*	0.286**	0.247**	0.276***	0.210*	0.174	0.197*	0.184	0.130	0.165
% mains laid or refurbished after 1981	-0.007*	-0.007	-0.007	-0.005	-0.005	-0.006	-0.008*	-0.006	-0.007*	-0.009*	-0.007	-0.008*
Ln(average pumping head of water resources plus)	0.272***	0.170*	0.199**				0.231**	0.172*	0.196**	0.252**	0.207*	0.231*
Ln(booster pumping stations per length of main)		0.280**			0.392***			0.320*			0.353**	
Ln(service reservoirs and water towers per length of mains)			0.202**			0.336***			0.183			0.165
% water treated in water treatments in band 3–6				0.004	0.003	0.004**						
Ln(density)							0.918***	1.148***	1.071***			
Ln(weighted average density)										0.248***	0.330***	0.290***
Constant	-11.53***	-9.491***	-10.42***	-10.12***	-8.036***	-8.976***	-10.57***	-9.571***	-10.31***	-8.201***	-6.600***	-7.676***
Observations	107	107	107	107	107	107	107	107	107	107	107	107
Estimation technique	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Adjusted R-squared	0.972	0.976	0.975	0.963	0.974	0.973	0.973	0.976	0.974	0.968	0.971	0.969
RESET test	0.372	0.145	0.684	0.0464	0.0467	0.161	0.346	0.162	0.476	0.0252	0.0211	0.0194
Breusch–Pagan Lagrangian multiplier test	0	0	0	0	0	0	0	0	0	0	0	0

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

The table below shows the RE regression results for BOTEX.

Table A 2 Regression results—Ofwat's BOTEX models (RE)

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	OWW9	OWW10	OWW11	OWW12
Dependent variable	Ln(wholesale water base cost)											
Ln(connected properties)	1.110***	1.083***	1.114***	1.056***	1.039***	1.079***						
Ln(length of mains)							1.115***	1.087***	1.116***	1.086***	1.048***	1.085***
% mains renewed and relined	0.214**	0.216**	0.213**	0.251**	0.245**	0.245**	0.216**	0.217**	0.215**	0.201**	0.200**	0.199**
% mains laid or refurbished after 1981	-0.008**	-0.007*	-0.008**	-0.007**	-0.006*	-0.006*	-0.009**	-0.008**	-0.008**	-0.009**	-0.008*	-0.008**
Ln(average pumping head of water resources plus)	0.261***	0.183**	0.208***				0.223***	0.184**	0.205***	0.252***	0.220**	0.243***
Ln(booster pumping stations per length of main)		0.257**			0.382***			0.228			0.287*	
Ln(service reservoirs and water towers per length of mains)			0.181**			0.315***			0.122			0.0980
% water treated in water treatments in band 3–6				0.003*	0.002*	0.003***						
Ln(density)							0.881***	1.037***	0.984***			
Ln(weighted average density)										0.252***	0.318***	0.279***
Constant	-11.48***	-9.704***	-10.55***	-10.01***	-8.042***	-8.949***	-10.37***	-9.660***	-10.23***	-8.251***	-7.020***	-7.978***
Observations	107	107	107	107	107	107	107	107	107	107	107	107
Estimation technique	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE
R-squared (overall)	0.973	0.977	0.976	0.964	0.975	0.974	0.974	0.977	0.976	0.970	0.972	0.970
R-squared (between)	0.983	0.988	0.986	0.975	0.986	0.985	0.985	0.987	0.986	0.980	0.983	0.980
R-squared (within)	0.143	0.138	0.138	0.117	0.112	0.110	0.148	0.140	0.143	0.160	0.158	0.158
Hausman test	0.856	0.786	0.769	0.735	0.710	0.607	0.992	0.872	0.910	0.492	0.485	0.519
Over-identification test	0.697	0.326	0.308	0.554	0.551	0.529	0.969	0.117	0.0213	0.156	0.0278	0.00560

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 3 YKY's estimated efficiency—Ofwat's models

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	OWW9	OWW10	OWW11	OWW12	Average*
OLS													
UQ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00
UQi	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.98	1.00
UD	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	1.00	0.92	1.00
RE													
UQ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.98	1.00
UQi	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	1.00	0.97	1.00
UD	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.99	0.93	1.00
SFA													
Four-component model	0.95	1.00	1.00	0.98	1.00	1.00	0.98	1.00	0.99	0.94	0.95	0.93	0.99

Note: * Average refers to the triangulated efficiency across all twelve models.

Source: Oxera analysis.

Table A 4 YKY's relative position—Ofwat's models

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	OWW9	OWW10	OWW11	OWW12	Average*
POLS	5	1	1	3	1	1	2	1	1	6	4	6	1
RE	5	1	1	3	1	1	2	1	2	6	4	7	2
SFA	5	1	1	3	1	5	2	1	2	7	4	7	2

Note: * Average refers to the triangulated position across all twelve models.

Source: Oxera analysis.

A1.2 Ofwat's growth enhancement models

Ofwat presented two enhancement models in growth-related activity. Specifically, enhancement expenditure related to new developments was modelled as a function of either population served or number of new connected properties. Both costs and cost drivers were smoothed using a three-year moving average.

The results from OLS and RE modelling are presented in Table A 5 and Table A 6 respectively.

Table A 5 Regression results of enhancement models (OLS)

	OE4	**OE5**
Dependent variable	Ln(wholesale water new developments cost)	
Ln(population served, smoothed)	1.034***	
Ln(new connections, smoothed)		1.050***
Constant	-6.291***	-0.264
Observations	70	70
Estimation technique	POLS	POLS
Adjusted R-squared	0.821	0.812
RESET test	0.00267	0.105
Breusch–Pagan Lagrangian multiplier test	0	0

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 6 Regression results of enhancement models (RE)

	OE4	**OE5**
Dependent variable	Ln(wholesale water new developments cost)	
Ln(population served, smoothed)	1.061***	
Ln(new connections, smoothed)		1.040***
Constant	-6.498***	-0.242
Observations	70	70
Estimation technique	RE	RE
R-squared (overall)	0.823	0.815
R-squared (between)	0.849	0.831
R-squared (within)	0.395	0.299
Hausman test	1.86e-07	0.948
Over-identification test	0.00239	0.964

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

A1.3 YKY aggregate BOTEX and BOTEX(Growth) models

Regression results from YKY models submitted in the consultation in March 2018 are shown in the tables below. Table A 7 shows the OLS regression results for BOTEX models. The RESET and Breusch–Pagan test have the expected p-value under OLS.

Table A 7 Regression results of YKY BOTEX models (OLS)

	YKYWW1	YKYWW2	YKYWW3	YKYWW4	YKYWW5	YKYWW6
Ln(Length of mains)	1.036***	1.035***	1.062***			
Ln(connected properties)				1.005***	1.022***	
Ln(population served)						1.015***
Ln(properties over mains)	0.947***	1.017***	0.935***			
Ln(properties over mains), demeaned				-0.0899	-0.0521	-0.263**
Ln(properties over mains), demeaned squared				1.078***	1.148***	0.892**
Sources over DI	0.640***	0.726***	0.625***	0.474***	0.687***	0.609***
Proportion of mains renewed/relined		26.39**	29.38***	31.78***	28.79***	24.51***
Proportio of mains laid before 1980	0.814***		0.894***			
Proportion of DI from reservoirs	0.547***	0.550***	0.564***	0.207**	0.215**	0.259***
Proportion of DI from rivers	0.217	0.274	0.299*			
Proportion of water treated in band 1 and below				-0.832***		
Proportion of water treated in band 2 and below					-0.700***	-0.591***
Constant	-10.08***	-9.999***	-10.48***	-9.430***	-9.654***	-3.350***
Observations	107	107	107	107	107	107
Estimation technique	POLS	POLS	POLS	POLS	POLS	POLS
Adjusted R-squared	0.972	0.971	0.977	0.980	0.981	0.981
RESET test	0.962	0.540	0.539	0.627	0.455	0.204
Breusch–Pagan Lagrangian multiplier test	0	0	4.77e-10	4.58e-05	0.000621	0.000333

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 8 shows the RE regression results for BOTEX models.

Table A 8 Regression results of YKY BOTEX models (RE)

	YKYWW1	YKYWW2	YKYWW3	YKYWW4	YKYWW5	YKYWW6
Ln(Length of mains)	1.039***	1.033***	1.058***			
Ln(connected properties)				1.006***	1.020***	
Ln(population served)						1.015***
Ln(properties over mains)	0.870***	0.990***	0.924***			
Ln(properties over mains), demeaned				-0.128	-0.0923	-0.294***
Ln(properties over mains), demeaned squared				1.071***	1.146***	0.878**
Sources over DI	0.577***	0.722***	0.614***	0.432***	0.615***	0.555***
Proportion of mains renewed/relined		25.04**	24.13***	26.99***	26.91***	24.44***
Proportio of mains laid before 1980	0.992***		0.879***			
Proportion of DI from reservoirs	0.535***	0.545***	0.556***	0.224*	0.223**	0.260***
Proportion of DI from rivers	0.206	0.300*	0.293**			
Proportion of water treated in band 1 and below				-0.672***		
Proportion of water treated in band 2 and below					-0.601***	-0.518***
Constant	-9.880***	-9.860***	-10.36***	-9.427***	-9.628***	-3.351***
Observations	107	107	107	107	107	107
Estimation technique	RE	RE	RE	RE	RE	RE
R-squared (overall)	0.973	0.973	0.978	0.981	0.982	0.983
R-squared (between)	0.986	0.984	0.989	0.992	0.993	0.993
R-squared (within)	0.00121	0.121	0.129	0.108	0.114	0.124
Hausman test	0.419	0.982	0.993	0.527	0.662	0.910
Over-identification test	0.373	0.937	0.780	0	0.228	0.725

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 9 shows the OLS regression results for BOTEX(Growth) models.

Table A 9 Regression results of YKY BOTEX(Growth) models (OLS)

	YKYWW7	YKYWW8	YKYWW9	YKYWW10
	Ln(wholesale water base cost + growth)			
Ln(length of mains)	1.045***	1.043***	1.035***	1.036***
Ln(connected properties per length of mains)	0.848***	0.835***	0.914***	0.870***
Sources over DI	0.432***	0.376***	0.502***	0.355**
Proportion of mains laid before 1980	0.905**	0.843**		
Proportion of DI from reservoirs	0.438***	0.442***	0.419***	0.428***
Proportion of mains renewed/relined			16.20	17.73*
Enhancements to the supply/demand balance over DI	1.989**	1.912**	1.845	1.669
New properties over connected properties		11,614		25,875**
Constant	-9.554***	-9.509***	-9.232***	-9.200***
Observations	107	107	107	107
Estimation technique	POLS	POLS	POLS	POLS
Adjusted R-squared	0.972	0.973	0.968	0.970
RESET test	0.837	0.837	0.0921	0.0368
Breusch–Pagan Lagrangian multiplier test	0	0	0	0

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 10 shows the RE regression results for BOTEX(Growth) models.

Table A 10 Regression results of YKY BOTEX(Growth) models (RE)

	YKYWW7	**YKYWW8**	**YKYWW9**	**YKYWW10**
	Ln(wholesale water base cost + growth)			
Ln(length of mains)	1.048***	1.046***	1.040***	1.040***
Ln(connected properties per length of mains)	0.810***	0.788***	0.886***	0.855***
Sources over DI	0.402***	0.353***	0.473***	0.394**
Proportion of mains laid before 1980	0.996***	0.988***		
Proportion of DI from reservoirs	0.420***	0.430***	0.385**	0.406***
Proportion of mains renewed/relined			19.70**	22.09**
Enhancements to the supply/demand balance over DI	1.766**	1.606*	1.586*	1.304
New properties over connected properties		11,764		19,138*
Constant	-9.460***	-9.419***	-9.149***	-9.142***
Observations	107	107	107	107
Estimation technique	RE	RE	RE	RE
R-squared (overall)	0.974	0.974	0.970	0.971
R-squared (between)	0.986	0.986	0.980	0.981
R-squared (within)	0.0535	0.0698	0.143	0.176
Hausman test	0.631	0.206	0.995	0.917
Over-identification test	0.737	0.243	0.977	0.245

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

A1.4 Stochastic Frontier Analysis results

Table A 11 and Table A 12 show YKY's estimated efficiency scores on wholesale BOTEX and BOTEX(Growth) models with the SFA four-component estimation approach respectively.

The tables also show the p-value for the corresponding likelihood ratio test, which tests for the presence of time-invariant and time-varying inefficiency.

Table A 11 A YKY's estimated efficiency on wholesale BOTEX models (SFA four-component)

Model	YKY's estimated efficiency	Likelihood ratio test (time-varying inefficiency)	Likelihood ratio test (time-constant inefficiency)
YKYWW1	1.00	0.16	0.03
YKYWW2	0.98	0.03	1.00
YKYWW3	1.00	0.03	0.16
YKYWW4	0.98	0.02	1.00
YKYWW5	0.98	0.02	1.00
YKYWW6	0.98	0.02	1.00
OWW1	0.95	0.02	0.00
OWW2	1.00	0.04	0.02
OWW3	1.00	0.04	0.00
OWW4	0.98	0.01	0.00
OWW5	1.00	0.02	1.00
OWW6	1.00	0.02	1.00
OWW7	0.98	0.03	0.18
OWW8	1.00	0.04	0.02
OWW9	0.99	0.03	0.00
OWW10	0.94	0.03	0.00
OWW11	0.95	0.04	0.00
OWW12	0.93	0.04	0.00

Source: Oxera analysis.

Table A 12 YKY's estimated efficiency on wholesale BOTEX(Growth) models (SFA four-component)

Model	YKY's estimated efficiency	Likelihood ratio test (Time-varying inefficiency)	Likelihood ratio test (Time-constant inefficiency)
YKYWW7	0.99	0.21	1.00
YKYWW8	0.99	0.15	1.00
YKYWW9	0.99	0.13	1.00
YKYWW10	0.99	0.12	1.00

Source: Oxera analysis.

A2 Regression output—wholesale wastewater

A2.1 Ofwat's wholesale water BOTEX models

Table A 13 and Table A 14 show regression results from Ofwat's models based on the dataset published by Ofwat as part of the consultation on econometric cost modelling in March 2018.[54] They also reflect key statistical diagnostics for Ofwat's models using the POLS and RE approaches respectively, as outlined in section A1.1.

YKY's estimated efficiency scores and ranking in each of Ofwat's models can be found in Table A 15 and Table A 16 respectively.

[54] Ofwat (2018), 'Cost assessment for PR19: a consultation on econometric cost modelling', March.

Table A 13 Regression results—Ofwat's BOTEX models (pooled OLS)

	OWWW1	OWWW2	OWWW3	OWWW4	OWWW5	OWWW6	OWWW7	OWWW8
	Ln(wholesale wastewater base costs)							
Ln(properties)						0.976***	0.961***	0.975***
Ln(load)	0.877***	0.852***	0.924***	0.910***	0.921***			
% lengths replaced post 2001	-0.0127**	-0.0147**			-0.0133***			-0.0153***
Ln(pumping stations per sewer length)		0.141						
% load treated in STWs bands 1-3	0.0336*	0.0189	0.0610***	0.0520***	0.0483***	0.0660***	0.0546***	0.0504***
% load from trade effluent customers			0.0692***			0.0870***		
% sludge disposed to farmland				-0.00771***			-0.00885***	
Ln(properties per sewer)			1.170***	0.667*	0.742***	1.317***	0.688**	0.775***
Constant	-5.562***	-4.803***	-3.503**	-3.701***	-4.243***	1.078	0.655	0.0854
Observations	60	60	60	60	60	60	60	60
Estimation technique	POLS	POLS	POLS	POLS	POLS	POLS	POLS	POLS
Adjusted R-squared	0.946	0.951	0.958	0.963	0.966	0.963	0.967	0.971
RESET test	0.00136	9.18e-07	0.00979	0.00164	0.00111	0.00194	0.000174	0.0101
Breusch–Pagan Lagrangian multiplier test	1.23e-08	4.53e-09	1.19e-05	2.48e-07	0.0111	0.00871	6.77e-05	0.360

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 14 shows the RE regression results for BOTEX.

Table A 14 Regression results—Ofwat's BOTEX models (RE)

	OWWW1	OWWW2	OWWW3	OWWW4	OWWW5	OWWW6	OWWW7	OWWW8
	Ln(wholesale wastewater base costs)							
Ln(properties)						0.999***	0.974***	0.984***
Ln(load)	0.904***	0.865***	0.957***	0.925***	0.941***			
% lengths replaced post 2001	-0.00469	-0.00900**			-0.00876**			-0.0131***
Ln(pumping stations per sewer length)		0.272***						
% load treated in STWs bands 1-3	0.0384**	0.0152	0.0603***	0.0552***	0.0530***	0.0615***	0.0553***	0.0521***
% load from trade effluent customers			0.0228			0.0386		
% sludge disposed to farmland				-0.00683***			-0.00721***	
Ln(properties per sewer)			1.052**	0.849**	0.867***	1.093***	0.808***	0.826***
Constant	-5.991***	-4.671***	-4.110***	-3.508***	-4.236***	0.456	0.709	0.121
Observations	60	60	60	60	60	60	60	60
Estimation technique	RE	RE	RE	RE	RE	RE	RE	RE
R-squared (overall)	0.943	0.944	0.956	0.964	0.966	0.961	0.968	0.972
R-squared (between)	0.964	0.960	0.975	0.981	0.986	0.981	0.985	0.994
R-squared (within)	0.0804	0.254	0.167	0.288	0.127	0.123	0.272	0.0982
Hausman test	0.352	0.204	0.0208	0.108	0.00224	0.00426	0.0967	0.00133
Over-identification test	2.44e-06	0.0673	0.0193	0.321	1.87e-08	0.00149	0.312	0

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 15 YKY's estimated efficiency – Ofwat's models

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	Average*
OLS									
UQ	1.00	1.00	1.00	0.99	1.00	0.97	0.95	0.99	0.99
UQi	0.99	0.98	0.97	0.97	0.99	0.97	0.93	0.95	0.94
UD	0.98	0.96	0.93	0.96	0.94	0.89	0.93	0.94	0.90
RE									
UQ	0.98	0.93	0.95	0.98	0.99	0.93	0.94	0.97	0.98
UQi	0.97	0.91	0.91	0.97	0.96	0.90	0.93	0.95	0.95
UD	0.94	0.90	0.91	0.93	0.93	0.87	0.90	0.93	0.94
SFA									
Four-component model	0.94	0.91	0.91	0.97	1.00	0.95	1.00	0.99	0.97

Note: * Average refers to the triangulated efficiency across all eight models.

Source: Oxera analysis.

Table A 16 YKY's relative position—Ofwat's models

	OWW1	OWW2	OWW3	OWW4	OWW5	OWW6	OWW7	OWW8	Average*
POLS	3	4	3	4	3	7	6	5	4
RE	6	4	6	4	5	7	6	7	6
SFA	6	4	6	4	4	7	6	7	5

Note: * Average refers to the triangulated position across all eight models.

Source: Oxera analysis.

A2.2 Ofwat's growth enhancement models

Ofwat presented seven enhancement models in growth-related activity. Specifically, Ofwat presented three models (OE6 to OE8) for enhancement expenditure related to first-time sewerage and four models (OE9 to OE12) related to enhancement expenditure related to new developments, growth at sewage treatment works, and the reduction of sewer flooding risk for properties. Both costs and cost drivers were smoothed using a three-year moving average.

The results from OLS and RE modelling are presented in Table A 17 and Table A 18 respectively.

Table A 17 Regression results of enhancement models (OLS)

	OE6	OE7	OE8	OE9	OE10	OE11	OE12
	Smooth (first-time sewerage costs)			Smooth (new developments, growth at sewage treatment works, and reducing sewer flooding risk for properties)			
S101a schemes (smooth)	1.438***		0.318**				
Average number of connectable properties per s101a schemes (smooth)	0.0236*						
Connectable properties served by s101a schemes (smooth)		0.0197***	0.0165***				
Resident population (smooth)				0.00476***		0.00277	
Household and non-household properties billed for sewage (smooth)					0.0124***		0.00643
Load per sewage treatment work (smooth)						0.0121	0.0142*
Constant	-0.711	0.666*	0.454	5.197	2.844	7.526	6.379
Observations	59	59	59	40	40	40	40
Estimation technique	POLS	POLS	POLS	POLS	POLS	POLS	POLS
Adjusted R-squared	0.834	0.917	0.922	0.775	0.745	0.807	0.808
RESET test	9.24e-05	0.00167	0.0125	0.0191	0.00195	0.407	0.335
Breusch–Pagan Lagrangian multiplier test	0	1.30e-10	0	7.23e-11	0	3.87e-10	4.44e-10

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 18 Regression results of enhancement models (RE)

	OE6	OE7	OE8	OE9	OE10	OE11	OE12
	Smooth (first time sewerage costs)			Smooth (new developments, growth at sewage treatment works and reducing sewer flooding risk for properties)			
S101a schemes (smooth)	1.225***		0.432***				
Average number of connectable properties per s101a schemes (smooth)	0.00858						
Connectable properties served by s101a schemes (smooth)		0.0165***	0.0121***				
Resident population (smooth)				0.00479***		0.00276*	
Household and non-household properties billed for sewage (smooth)					0.0124***		0.00617*
Load per sewage treatment work (smooth)						0.0123	0.0149**
Constant	0.229	0.964**	0.683*	5.004	2.760	7.405	6.516
Observations	59	59	59	40	40	40	40
Estimation technique	RE	RE	RE	RE	RE	RE	RE
R-squared (overall)	0.824	0.918	0.922	0.780	0.751	0.817	0.818
R-squared (between)	0.853	0.948	0.947	0.803	0.773	0.841	0.842
R-squared (within)	0.626	0.709	0.745	0.0334	0.0114	0.0352	0.0272
Hausman test	0.140	0.000702	0.00529	0.644	0.788	0.912	0.771
Over-identification test	5.34e-06	9.96e-06	0.000108	0.635	0.829	0.837	0.727

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

A2.3 YKY aggregate BOTEX and BOTEX(Growth) models

Regression results from YKY models submitted in the consultation in March 2018 are shown in the table below. The RESET and Breusch–Pagan test have the expected p-value under OLS.

Table A 19 Regression results of YKY BOTEX models (OLS)

	YKYWWW1	YKYWWW2	YKYWWW3	YKYWWW4
	Ln(wholesale wastewater base costs)			
Ln(properties)	0.675***	0.846***	0.810***	0.805***
% load with tight consents (<10mg/l BOD and <1mg/l ammonia)	3.694***	1.388**	1.906**	2.367***
Ln(Pumping station capacity per km sewer)	0.124**	0.257***	0.233***	0.223***
Ln(Number of combined sewer overflows per km sewer)	0.0517			
Proportion of combined sewers to total length of sewers (%)		0.593***	0.520***	0.362**
% of area with more than 4000 people per km^2				-0.401**
% of area with more than 2000 people per km^2	-0.504**		-0.125	
Constant	0.397	-1.389**	-1.061*	-0.982*
Observations	60	60	60	60
Estimation technique	POLS	POLS	POLS	POLS
Adjusted R-squared	0.959	0.969	0.969	0.973
RESET test	0.294	0.281	0.323	0.00454
Breusch–Pagan Lagrangian multiplier test	7.50e-05	0.000402	0.000620	0.0287

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 20 shows the RE regression results for BOTEX.

Table A 20 Regression results of YKY submitted BOTEX models (RE)

	YKYWWW1	YKYWWW2	YKYWWW3	YKYWWW4
	Ln(wholesale wastewater base costs)			
Ln(properties)	0.672***	0.839***	0.834***	0.839***
% load with tight consents (<10mg/l BOD and <1mg/l ammonia)	1.904***	1.610***	1.759***	1.994***
Ln(Pumping station capacity per km sewer)	0.139**	0.275***	0.269***	0.251***
Ln(Number of combined sewer overflows per km sewer)	-0.106*			
Proportion of combined sewers to total length of sewers (%)		0.613***	0.582***	0.432***
% of area with more than 4000 people per km^2				-0.333***
% of area with more than 2000 people per km^2	-0.333**		-0.0749	
Constant	0.00574	-1.376***	-1.306***	-1.276***
Observations	60	60	60	60
Estimation technique	RE	RE	RE	RE
R-squared (overall)	0.945	0.971	0.971	0.974
R-squared (between)	0.960	0.988	0.988	0.992
R-squared (within)	0.340	0.285	0.276	0.262
Hausman test	3.29e-06	0.143	0.0835	0.00776
Over-identification test	5.82e-07	0.0165	0.0290	1.85e-08

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 21 shows the OLS regression results for BOTEX(Growth).

Table A 21 Regression results of YKY BOTEX(Growth) models (OLS)

	YKYWWW5	YKYWWW6	YKYWWW7	YKYWWW8
	Ln(wholesale wastewater base costs)			
Ln(properties)	0.660***	0.797***	0.760***	0.759***
% load with tight consents (<10mg/l BOD and <1mg/l ammonia)	3.558***	1.755**	2.293*	2.648**
Ln(Pumping station capacity per km sewer)	0.0597	0.163***	0.138*	0.133***
Ln(Number of combined sewer overflows per km sewer)	0.0312			
Proportion of combined sewers to total length of sewers (%)		0.441**	0.365*	0.228
% of area with more than 4000 people per km2				-0.368
% of area with more than 2000 people per km2	-0.399		-0.129	
Properties growth (%)	2.026	2.076	2.012	1.315
Constant	0.581	-0.805	-0.465	-0.424
Observations	60	60	60	60
Estimation technique	POLS	POLS	POLS	POLS
Adjusted R-squared	0.958	0.963	0.963	0.966
RESET test	0.395	0.153	0.192	0.0248
Breusch–Pagan Lagrangian multiplier test	2.44e-06	1.15e-07	7.99e-07	8.54e-05

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

Table A 22 shows the RE regression results for BOTEX(Growth).

Table A 22 Regression results of YKY submitted BOTEX(Growth) models (RE)

	YKYWWW5	YKYWWW6	YKYWWW7	YKYWWW8
	Ln(wholesale wastewater base costs)			
Ln(properties)	0.763***	0.846***	0.855***	0.831***
% load with tight consents (<10mg/l BOD and <1mg/l ammonia)	1.389***	1.255***	1.122***	1.616***
Ln(Pumping station capacity per km sewer)	0.104**	0.184***	0.201***	0.162***
Ln(Number of combined sewer overflows per km sewer)	-0.0142			
Proportion of combined sewers to total length of sewers (%)		0.470***	0.512***	0.365**
% of area with more than 4000 people per km2				-0.175
% of area with more than 2000 people per km2	-0.138		0.0783	
Properties growth (%)	3.260***	3.321***	3.374***	3.052**
Constant	-0.326	-1.183**	-1.293**	-1.007**
Observations	60	60	60	60
Estimation technique	RE	RE	RE	RE
R-squared (overall)	0.952	0.965	0.965	0.967
R-squared (between)	0.968	0.980	0.980	0.983
R-squared (within)	0.218	0.225	0.228	0.214
Hausman test	0.176	0.737	0.600	0.246
Over-identification test	5.77e-09	0.000300	4.64e-09	0

Note: *** significant at 0.1%; ** significant at 1%; * significant at 5%.

Source: Oxera analysis.

A2.4 Stochastic Frontier Analysis results

Table A 23 and Table A 24 show YKY's estimated efficiency scores on wholesale wastewater BOTEX and BOTEX(Growth) models with the SFA four-component estimation approach respectively.

The tables also show the p-value for the corresponding likelihood ratio test, which tests for the presence of time-invariant and time-varying inefficiency. Many of Ofwat's and YKY's models are unable to distinguish between noise and inefficiency.

Table A 23 YKY's estimated efficiency on wholesale BOTEX models (SFA four-component)

Model	YKY's estimated efficiency	Likelihood ratio test (time-varying inefficiency)	Likelihood ratio test (time-constant inefficiency)
YKYWWW1	1.00	1.00	1.00
YKYWWW2	1.00	1.00	0.00
YKYWWW3	1.00	1.00	0.30
YKYWWW4	1.00	1.00	1.00
OWWW1	0.94	0.31	0.00
OWWW2	0.91	1.00	0.00
OWWW3	0.91	0.44	0.00
OWWW4	0.97	0.48	0.01
OWWW5	1.00	1.00	0.48
OWWW6	0.95	1.00	0.11
OWWW7	1.00	1.00	0.50
OWWW8	0.99	0.46	0.32

Source: Oxera analysis.

Table A 24 YKY's estimated efficiency on wholesale BOTEX(Growth) models (SFA four-component)

Model	YKY's estimated efficiency	Likelihood ratio test (time-varying inefficiency)	Likelihood ratio test (time-constant inefficiency)
YKYWWW5	1.00	0.26	1.00
YKYWWW6	1.00	0.31	1.00
YKYWWW7	1.00	0.32	1.00
YKYWWW8	1.00	0.27	1.00

Source: Oxera analysis.

YORKSHIREWATER.COM

Yorkshire Water Services Limited, Western House, Halifax Road,
Bradford, BD6 2SZ. Registered in England and Wales No.2366682

YorkshireWater