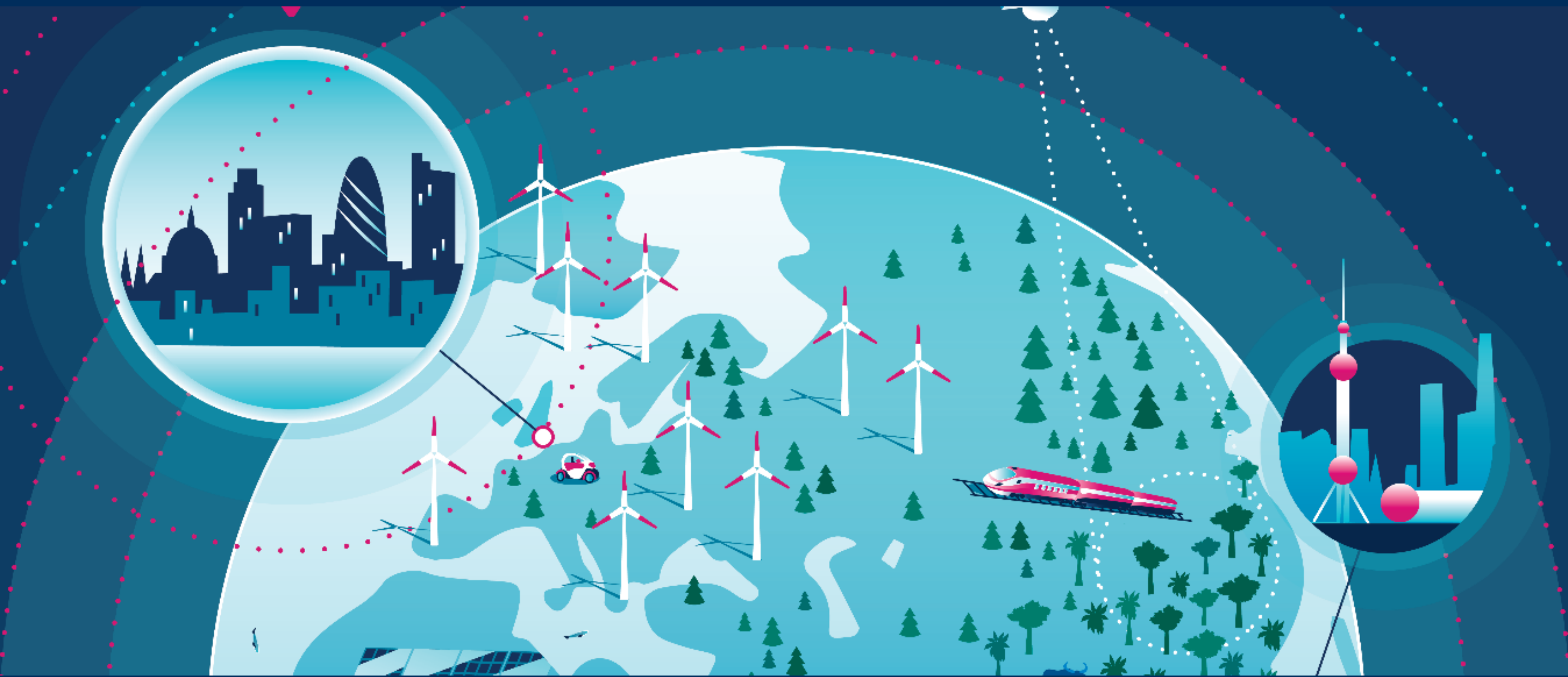
# Enhancement cost assessment modelling for the PR19 Initial Assessment of Plans



Report prepared for Anglian Water

Final report

March 2019

:vivideconomics
putting economics to good use

# Executive Summary

**This report considers Ofwat's use of econometric modelling for enhancement costs in the PR19 Initial Assessment of Plans (IAP) and makes recommendations for how it can be improved during the remainder of the review.** It aims to support the development of enhancement cost assessment over the remainder of PR19, reviewing all aspects of the IAP's use of econometrics and identifying practical ways in which it can be improved upon.

**Enhancement cost assessment is an extremely important component of PR19, but the use of econometrics for this is an inherently challenging undertaking.** With £15.6bn of enhancement totex planned for AMP7, it is critical that Ofwat takes a well-founded view on efficient costs to maintain affordable bills and secure service improvements. However, econometric models are often of limited value in informing such a view of enhancement costs, both because spending tends to be lumpy and bespoke, and because comparable data on costs and drivers is often not available across the sector. This means that, in order to scrutinise expenditure, in some cases it is better to use other approaches. Even in cases where modelling can be a basis for cost assessment, the literature on the most suitable specifications to use is less mature than for base cost modelling, where Ofwat's consultation on cost assessment generated a large body of evidence.

**An overarching limitation of the IAP is the lack of a documented, systematic methodology for model development and selection. This has resulted in some unjustified decisions in material areas.** A systematic process for modelling follows a consistent approach across all cost lines: defining hypotheses on the causes of efficient costs, quality assuring and cleaning data, testing appropriate model specifications, then determining the most suitable method for cost assessment. The evidence provided in the IAP suggests that, unlike in base cost modelling, such a process was not followed in full in the selection of enhancement cost models. As a consequence it is is difficult to have confidence that the best approach has been selected across all lines of expenditure -- and, indeed, it is clear from this review that the models adopted in some specific lines are deficient in ways that have material impacts. Reasons for this include:

- **Lack of fit with engineering or economic logic.** For example, growth models, which cover £4bn in capex, do not account for the drivers of upstream spending that are recovered through customer bills, while the functional forms used in sanitary parameter models, which cover £0.3bn in capex, are inconsistent with any plausible engineering rationale.
- **Lack of statistical fit.** Many of the models produce implausibly wide ranges of efficiency estimates, suggesting they omit important relevant determinants of efficient costs. For example:
  - ◊ **P-removal** models which allocate £2.3bn in capex, estimate the leading company (South West Water) to be more than four times more efficient than the laggard (Thames Water);
  - ◊ **Leakage** models which allocate £0.3bn in capex, take the median of unit costs that vary by a factor of nearly 15, from £0.33m/Ml/d (Portsmouth Water) to £4.85m/Ml/d (Sutton and East Surrey Water);
  - ◊ **Storm tanks** models which allocate £0.5bn in capex, estimate the most efficient company (Severn Trent England) to be 3.5 times more efficient than the least efficient company (Southern Water).
- **Lack of stability.** This is a critical risk in models with small samples where some lines are dominated by a few companies, but does not appear to have been fully assessed. Two extreme examples of this are wastewater growth, where the two models produce estimates of AMP7 industry costs that differ by £1.3bn, and first time sewerage, where two companies (Anglian Water and Severn Trent England) account for 72% of expenditure.

- Inconsistency of approach. This is reflected in mistakes in the implementation of some the models and decisions that are difficult to reconcile with each other, such as the choices of functional form for water and waste growth models: switching from a unit cost approach to random effects (RE) in water growth would reallocate £35m across the sector.

A key step in improving the models is to restructure the sample and include operating costs in dependent variables. This review shows that, while the IAP expended signifcant effort to ensure company cost data was allocated to expenditure lines in a consistent way, two structural decisions on the organisation of data for modelling lack justification and have substantial impacts on modelling outputs.

- Enhancement opex: for most lines, the IAP's models estimate only efficient capital spending, even though operating expenditure is a legitimate component of efficient enhancement costs in many areas, accounting for 10% of overall spending in water enhancement and 7% in waste. By treating enhancement opex as a component of base costs, the IAP will skew coefficients in enhancement models and understate the efficiency of some companies in base cost models

- Hafren Dyfrdwy: in all models that use forecast data, the IAP treats Hafren Dyfrdwy (HDD) and Severn Trent England (SVE) as separate observations of company efficiency. This is not only inconsistent with the treatment of other companies under common management operating in multiple regions, but also introduces an outlying wholesale wastewater company into the sample, which is a major source of instability in many models. For example, in wastewater growth, the decision to separate HDD from SVE leads to the reallocation of £474 million in allowed spending between companies.

Most of the effective efficiency challenge in the IAP arises due to the choices of model specifications rather than explicit adjustments for efficiency. This is symptomatic of underlying weaknesses in the models. Of the 24% of business plan costs in modelled lines that the IAP removes, 66%, or £1.7bn, results from modelled allowances *before any efficiency challenge* being lower than business plan forecasts. This review does not consider it plausible that such a large shortfall could result from inflated business plan costs or a tendency for companies with larger programmes to be less efficient. Rather, the shortfall is more likely to be explained by the models' lack of explanatory power including a downwards bias that results from the way the IAP estimates logged costs, the weight placed on small company data (notably HDD, whose modelled wastewater enhancement costs are more than *five times larger* than its business plan projections), and the absence in models of complexity drivers that can cause costs to increase over time. To address this problem requires a fuller assessment of efficiency score ranges and the effects of outliers in model selection, followed by a consideration of the effects of modelled shortfalls when setting additional efficiency challenges.

Company-specific haircuts used in shallow dives lack justification and do not make use of business plan information. For shallow dive lines, the IAP uses a company-specific efficiency challenge – referred to in some models as 'haircut' – that is mostly derived from performance in base cost models. The evidence used for the haircut is unlikely to be appropriately representative, since relative base cost efficiency reflects investment decisions made in previous AMPs and has a large opex component, whereas enhancement efficiency exclusively concerns forward-looking capex decisions. Moreover, the haircut neglects information in company business plans that can usefully inform cost assessment, which the IAP uses for deep dive lines. There is limited evidence of Ofwat's using companies' written evidence in shallow dives.

A comprehensive assessment of the models themselves supports a range of more specific recommendations. The review subjected all of the IAP's models to a common set of tests in order to assess performance, diagnose any problems and, in many cases, identify improvements that can be implemented with existing data. Table 1 provides a summary of this. It recommends far-reaching changes to the growth models which do not explain variation in costs that are recovered from customers , wastewater quality enhancement models, which lack complexity drivers and in some cases are specified or implemented erroneously, and supply-demand balance models, which conflate willingness to pay information on leakage with estimates of efficient costs.

**There is an opportunity to substantially increase the robustness of enhancement cost assessment during the remainder of the PR19 process.** Ideally, Ofwat would undertake and document a systematic process of model development and selection as it has for base cost models. If this was unfeasible in the time available, this report presents a host of quick wins that would build on and substantially improve existing modelling work. These include adding opex to dependent variables, eliminating HDD from the sample, revisiting model specifications and choices over whether to use modelling or deep dives, replacing the company-specific efficiency challenge, and rebalancing the efficiency challenge away from modelled shortfalls towards more explicit measures.

Table 1 Line-by-line review of enhancement modelling

| | Expenditure line | Model assessment findings | Recommended improvements |
|---|---|---|---|
| Growth | Growth (overarching) | Assessment of costs gross rather than net of grants and contribution creates downside risk for customers | If data on relevant causal factors can be obtained, attempt estimating net costs using appropriate drivers<br>Deep dive clearly preferable to current gross cost modelling |
| | Growth (wastewater) | Unstable RE models, divergence between historical and forecast. Inclusion of sewer flooding, which is funded by ODI | |
| | Growth (water) | More stable than wastewater model. Unreliable company data yields implausible unit cost ranges | |
| | First time sewerage | Two companies dominate expenditure, causing instability | Use a deep dive |
| Waste quality | P-removal | Some errors in implementation, implausible ranges of efficiency scores, absence of complexity driver | Correct errors, triangulate between models, including data on P<1.1mg consents |
| | Chemical removal | Low fit, which disappears when SWB removed. Type of chemical obligations not captured. | Attempt totex modelling, with chemical obligations type. Deep dive if unsuccessful |
| | Event duration monitoring | Reasonable specification but undermined by inconsistent company data | Attempt to improve data. Consider shallow dive if unsuccessful |
| | Flow monitoring | Concerns on data comparability | Attempt to improve data |
| | Flow to full treatment | Variable model performance, with log specifications weaker than linear | Triangulate among the linear models |
| | Sanitary parameters | 'Power' and 'Exponential' specifications inconsistent with engineering logic | Test more transparent specifications, including log model suggested |
| | Spill frequency | Good fit, but instability related to SRN | Diagnose issues in SRN data, consider dropping from sample |
| | Storm tanks | Good overall fit, but range of efficiency scores implausibly wide | Diagnose efficiency score variation to understand omitted variables, reflect in approach to efficiency challenge |
| Water quality | Meeting lead standards | Merger of orthophosphate dosing and replacement of lead pipes does not reflect distinct regulatory drivers. Model highly unstable and produces implausible efficiency score range | Use treated water distribution model for replacement of communication pipes, use shallow dive for orthophosphate |
| Supply demand | Metering | Reasonable statistical fit and stability. Implausible range of efficiency scores | Potential improvement possible from including meter penetration |
| | Leakage | Not valid to use WTP figures to estimate costs. WTP figures are incorrectly taken net of sharing rates.<br>Unit costs highly variable, do not account for increasing marginal costs | Do not use WTP figures, attempt modelling that uses leakage as explanatory variable |
| | 2020-25 schemes | Very wide variation in unit costs, reflecting diversity of schemes | Use a deep or shallow dive |

Source: Vivid Economics analysis

# Contents



# List of tables



# List of figures

# List of boxes

# 1 Introduction

## 1.1 Terms of reference and methodology

Vivid Economics was engaged by Anglian Water to review the development and use of models for enhancement cost assessment by the PR19 Initial Assessment of Plans (IAP). The review recognises the inherent difficulties in modelling efficient enhancement costs and focuses on areas where there are opportunities to improve modelling before the PR19 final determinations.

The project team undertook a full review of the documentation of the IAP's modelling approach and carried out systematic testing of the models used in the IAP. In some areas where potential improvements to the IAP's approach could be readily tested, the team developed alternative models and subjected them to the same tests as conducted on the IAP models. The review relied principally on econometric reasoning and analysis, but was also informed by discussions to establish engineering drivers on the causes of efficient costs, which the team drew from sources including submissions to Ofwat's 2018 consultation on cost assessment modelling as well as expert views provided by Anglian Water's team. The review did not consider the IAP's reallocation of business plan spending between accounting lines, base cost modelling, or elements of 'deep' or 'shallow' dives that place no reliance at all on econometric evidence.

All code used to generate outputs in this report was quality assured by Dr Selma Walther of the University of Sussex.

## 1.2 Structure of this report

The main body of this report presents the main findings of the review. It is structured as follows:

- Section 2 considers overarching features of the IAP's approach to model development and its use of modelling evidence in setting allowances;
- Section 3 presents a more detailed assessment of individual models used in the IAP;
- Section 4 concludes with the recommendations from this review.

The Statistical Annex lays out the statistical evidence produced as part of this review and drawn on in the main report.

The project team has prepared an 'audit log' covering the modelling spreadsheets provided in the IAP, to be shared separately with Ofwat.

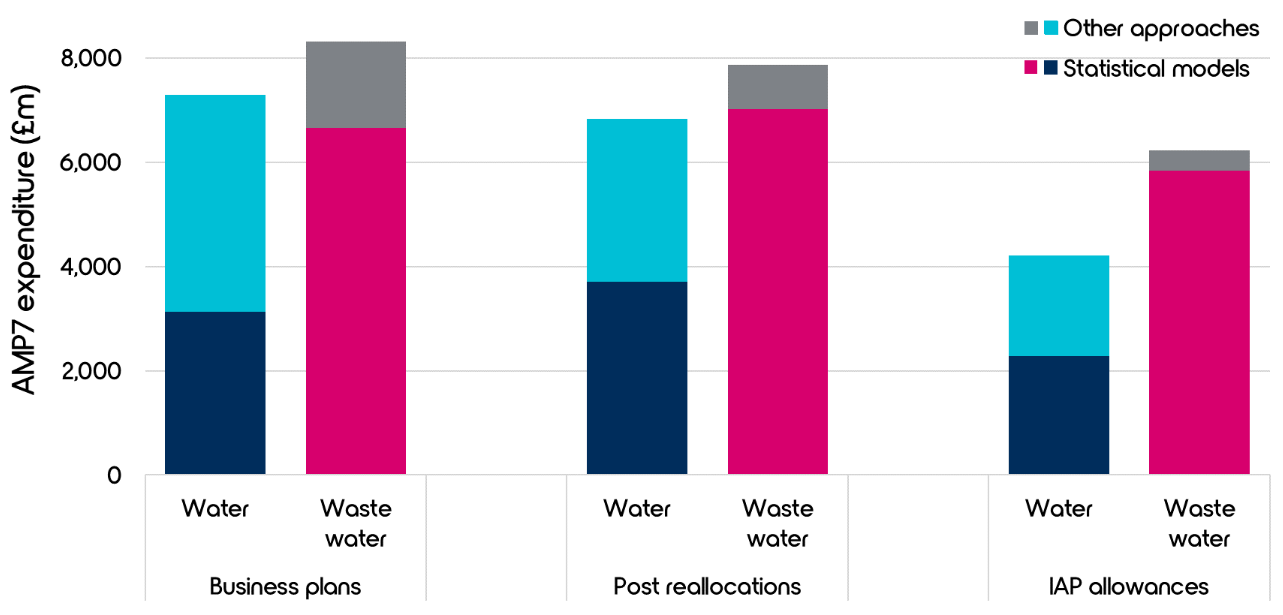
# 2 Approch to enhancement cost assessment

**This section considers overarching aspects of the IAP's use of models in enhancement cost assessment.** Section 2.1 begins with an assessment of how enhancement cost models were developed and selected for use in the IAP, before Section 2.2 reviews a number of issues related to data that cut across the lines of enhancement spending that were modelled. Section 2.3 then appraises the IAP's use of modelling evidence to apply efficiency challenges to modelled allowances.

## 2.1 Model development and selection

**Statistical models play a major role in enhancement cost assessment in the IAP, but models in this area are inherently difficult to develop and do not always provide the most robust evidence on efficient costs.** In principle, models can be a transparent and equitable means of setting allowances and Ofwat places a significant weight on them in assessing enhancement spending in the IAP, using them to allocate £8.1bn out of £10.4bn of enhancement capex (Figure 1). However, compared to base costs, enhancement spending in any particular area tends to be lumpy over time, concentrated among a small number of projects and companies, tailored to project-specific circumstances, and directed towards novel technologies. All of these factors make it more difficult to model future efficient levels of enhancement costs than base costs, particularly in the context of limited industry-wide datasets on spending and drivers. These barriers to modelling, combined with the availability of alternative evidence on enhancement cost efficiency, such as market testing of capital programmes, mean that modelling is often not the most robust form of cost assessment. Ofwat's use of alternative 'deep' or 'shallow dives' in the IAP serves as a testament to this.

Figure 1 Statistical models are used to determine the majority of allowed spend in water and wastewater




Note: Other approaches include the 'shallow dive', 'deep dive' processes and occasionally, immaterial areas where expenditure is allowed in full

Source: Vivid Economics analysis of IAP models

**A systematic approach to model development can reveal the most suitable form of cost assessment in any area.** By proceeding systematically (as recommended in Arup and Vivid Economics, 2018, for example), modellers can identify a full set of candidate models and understand their relative strengths, before selecting the most robust approach to generating allowances – be that through models or other forms of cost assessment. Such an approach would include the following four critical steps:

- **Identify causal drivers and potential explanatory variables**. This sets out the critical engineering or economic drivers that would be expected to affect efficient AMP7 costs in any enhancement line, variables that could be used to measure these drivers, and expectations on the shape and strength of the relationship between costs and the drivers. This understanding allows cost lines to be identified that are suitable for modelling and grouped where they are subject to similar drivers and companies face trade-offs between them, thus avoiding models that misdiagnose trade-offs between lines as efficiency or inefficiency in individual areas.

- **Understand limitations in data and apply mitigations**. Known data issues relate to cost allocation between enhancement lines in company Annual Performance Reports and measurement error on explanatory factors, reflected in confidence grades. Mitigations can be to the data lines themselves, for example through reallocation of reported costs, and by grouping lines for modelling where there is varying practice in cost allocation.

- **Define and test candidate models**. Candidate model specifications can be formulated accounting for narratives and data limitations. These can then be assessed using a common set of tests in order to arrive at a consistent view of whether modelling or another form of cost assessment is most appropriate. Critical aspects of models' performance include:

  - ◊ **Narrative fit:** a basic requirement for models is that the estimated relationship between costs and explanatory variables is consistent with plausible explanations of the causes of efficient costs.

  - ◊ **Statistical fit:** in order to command confidence that models capture the principal drivers of efficient costs, they should explain a reasonable proportion of the observed variation in company costs, as highlighted by $R^2$ and the spread of company efficiency scores. *Excessive* statistical fit, evidenced by very high $R^2$ or narrow spreads of company efficiency scores, can suggest that models explain some of the variation in costs attributable to relative efficiency.

  - ◊ **Stability:** a key risk in enhancement modelling is instability: relationships estimated using the data are not robust over time or across companies and therefore not applicable across the sector in AMP7. Tests can assess the robustness of model coefficients or efficiency scores to dropping potentially unrepresentative individual companies or observations.

  - ◊ **Diversity:** enhancement models are unavoidably subject to error, which can be reduced through the triangulation of multiple models whose errors might reasonably be expected to offset each other, for example due to the inclusion of different sets of explanatory variables. However, simply supplementing stronger models with weaker and similar ones will not reduce risks in this way.

- **Select cost assessment approach.** Model testing can inform a diagnosis of modelling performance, which may then be improved by actions such as removing outlying observations or triangulation of many models. Having arrived at best available approach to modelling and conducted appropriate quality assurance or peer review, a decision can be made on whether to implement this or apply another approach.

**The approach to enhancement cost modelling followed in the IAP does not proceed systematically or follow all of these essential steps.** Ofwat has released significantly less information on its approach to enhancement model development than it has for base cost modelling, but it is evident that not all of the steps set out above have been followed and, to the extent that they have been, this is inconsistent between modelling lines. A review of the available evidence at each of the four steps of modelling development reveals:

- **Some critical cost drivers are not accounted for where explanatory variables are available**. For example, the effect of the stringency of consents on the cost of P-removal enhancement is well documented (see Arup and Vivid Economics, 2018), while data on drivers is available and performs well in models similar to those used in the IAP (see Section 3.2.1). Along with other similar examples in growth and leakage set out in Section 3, this suggests that enhancement models have not been built from an account of key causal narratives. This is in contrast to the base cost models, where such an approach is documented in the IAP (see Econometric approach appendix).

- **Significant efforts have been made to improve cost data, but there remain areas of concern.** The IAP makes substantial reallocations of cost data between lines before modelling is carried out, but there remain areas where cost data used in models does not appear consistent between companies, for example in growth spending (see Section 3.1). It is unclear why no models are estimated using combined forecast and historical data as this would be expected to improve statistical power and, as Section 2.2 explains, there are also some important structural reasons why data used in models is poorly suited to estimation of efficient costs.
- **Model testing does not cover stability or fully assess drivers of cost**. The IAP does not report a full set of model tests and the results presented in the Appendix suggest that not all aspects have been fully considered. At a fundamental level, some models fail simple tests for narrative cogency – for example, the exponential sanitary determinands model implausibly explains log costs with linear drivers. Fit is problematic for some models (for example P-removal, where there is an implausible range of efficiency scores), while stability is a weakness of many of the models, notably wastewater growth and first time sewerage, where two companies account for 72% of expenditure. Finally, in most cases triangulation does not add to diversity, but rather simply combines one model with a very similar, but weaker one (for example in P-Removal).
- **Choices of cost assessment approach are not consistent**. As noted above, there is no evidence in the IAP of a common, overarching review of the models and, perhaps as a consequence of uneven model testing, choices of cost assessment models in the IAP are difficult to reconcile with each other. For example, analysis in Section 3.1 shows that unit cost modelling performs better than random effects modelling for waste growth while the opposite holds for water growth; however, unit cost models are used for water growth and random effects models for waste growth. The IAP does not document decisions over whether to use modelling or other forms of cost assessment, but analysis presented in Section 3 suggests deep dives might be more appropriate than models for wastewater growth and Supply-Demand Balance 2020-25 schemes.

**The lack of a systematic approach to model development reduces confidence in the IAP's approach to enhancement cost assessment.** This review did not carry out a process of model development like that described above. Ideally, such an exercise would be carried out and transparently documented before the PR19 final determinations. This would provide confidence that the most suitable approaches had been adopted.

**This review identifies significant potential to build on the work the IAP's approach, which have substantial impacts on company allowances**. Even without undertaking a full programme of model development, the testing carried out in this review identifies an array of potential improvements, using the IAP models as a starting point. Sections 2.2, 2.3 and 3 all contain constructive suggestions on how cost assessment can be developed through changes to model data, the efficiency challenge, the division between modelling and other approaches, and the models themselves. Improvements to the modelling approach will have substantive effects on the IAP's assessment of efficient costs for companies: to take an example from a single modelling line, P removal allowances across the industry could change by £182 million.

## 2.2 Use of data in models

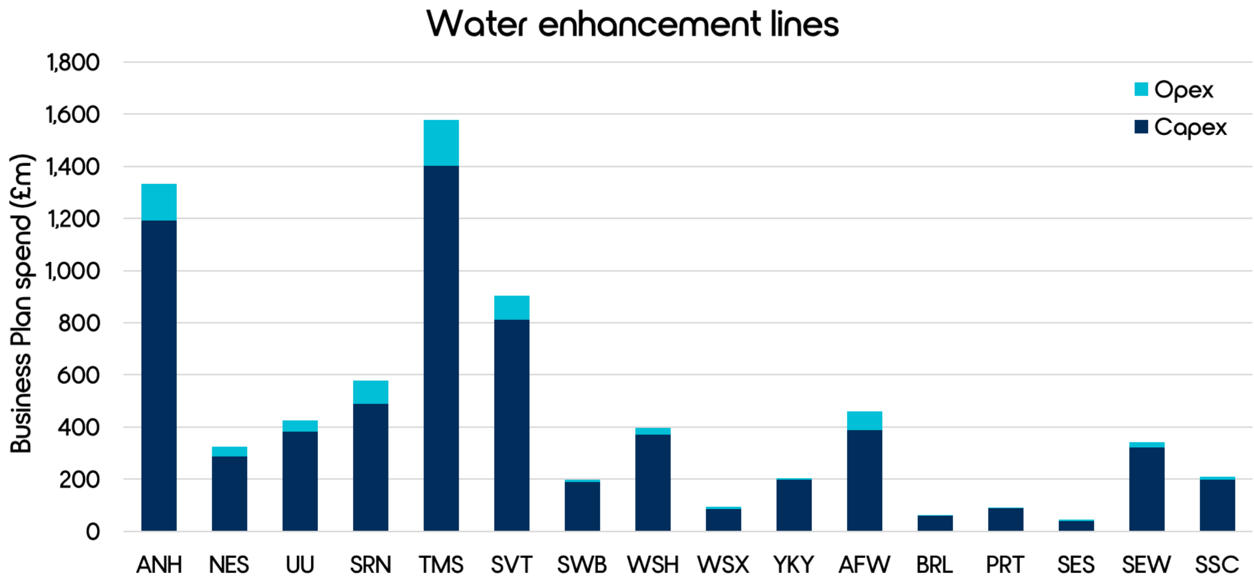
**As Section 2.1 emphasises, understanding and mitigating problems with data is an essential part of model development.** This can correct errors or inconsistencies in data where possible and ensure data on costs and explanatory variables is organised in a suitable way to estimate efficient costs.

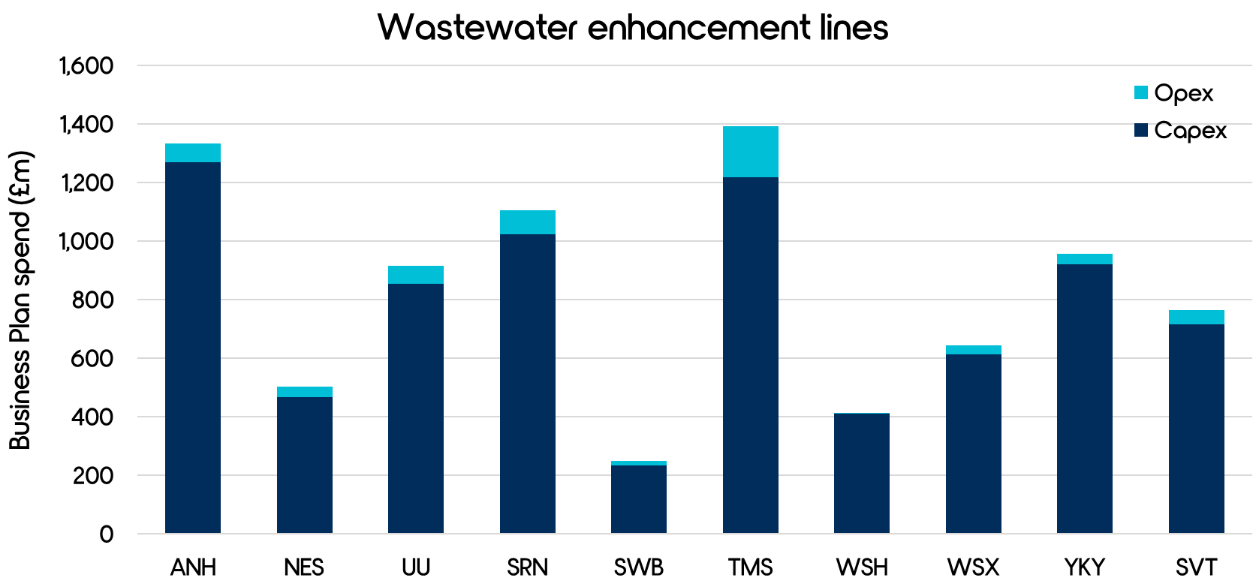
**The IAP has carried out extensive work to make company accounting data consistent and merged expenditure lines that have common drivers.** Vivid has not reviewed the quality of the work to reallocate spending between accounting lines, but evidently considerable effort has gone into this. In some cases, the IAP also sensibly merges some expenditure lines for the purpose of modelling where there are common drivers and trade-offs between spending across the areas: an example of this is in the joint modelling of spending on P-removal at activated sludge and filter bed works.

**However, the way the IAP structures data for modelling is materially detrimental to the quality of the models in a variety of respects.** Two particularly important aspects, which the remainder of this section expands upon, are in modelling enhancement capex rather than totex and in including Hafren Dyfrdwy (HDD) as a data point in models that use forecast data. Another cross-cutting issue with the IAP's approach, which Section 3 considers at the level of individual models, is the apparent lack of exploration of models that estimate relationships using combined historical and forecast datasets, a method that would be expected to improve statistical power compared to the disjoint approach in the IAP. Cumulative enhancement capex data has also not been included in the IAP – while there are concerns around data quality and comparability, cumulative capex data could have improved models by better aligning lumpy capex with volume drivers. Further issues concern the consistency of cost modelling with cost adjustment claims, which does not appear to have been explored in the IAP, and with ODIs, where there are problems related to sewer flooding, SuDS, and leakage (see Section 3).

**By accounting for most enhancement opex in base costs, the IAP skews estimated relationships in enhancement models and understates estimated base cost efficiency.** Enhancement opex is material, representing 10% of enhancement totex in water, and 7% in waste, and may feature in optimal solutions to enhancement drivers. As Figure 2 highlights, the proportion of enhancement opex varies significantly between companies (2% (Portsmouth Water) – 16% (Affinity Water)), so by not accounting for it in enhancement models, the IAP will skew estimated relationships. Furthermore, by accounting for enhancement opex instead in base cost models, which do not include specific enhancement drivers, the IAP will make companies with significant programmes of enhancement opex appear less efficient.

Figure 2 Enhancement opex is material and differs between companies



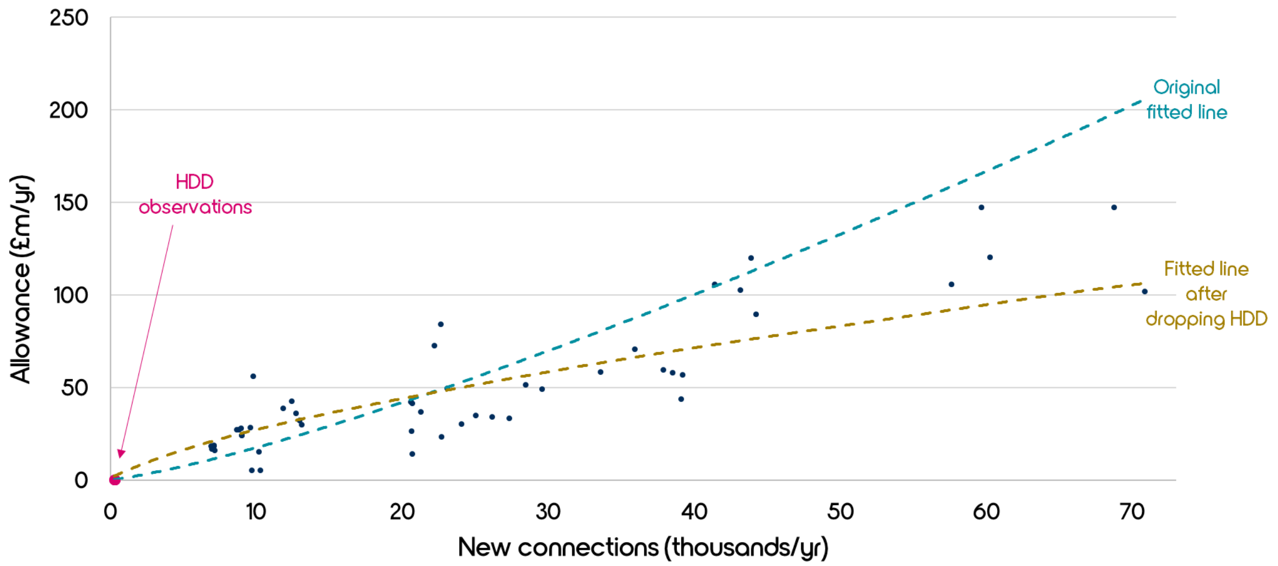




Source: Vivid Economics analysis of PR19 Data Tables

**By treating Severn Trent England (SVE) and Hafren Dyfrdwy (HDD) as distinct entities, the models place excessive weight on the efficiency of one management model.** Separating the two companies in modelling is inconsistent with the IAP's treatment of other companies that operate across a variety of geographical regions – and has the effect of treating the common management of the two companies as distinct observations of efficiency. This is particularly problematic where cost allocation decisions between the two entities can influence their estimated relative efficiency. Given the lack of historical data on HDD's costs, it is difficult to assess how consistent cost allocation is between the two licensees, especially for wastewater.

**The inclusion of HDD as a separate data point has a highly material impact on modelled allowances and reduces model robustness.** Figure 3 shows the example of wastewater growth, where the inclusion of HDD as a data point causes a change in modelled allowances worth £939 million. More generally, HDD is a clear outlier in all wastewater models, which can mean it has undue influence on modelled relationships.

Figure 3 The wastewater growth model fitted line is highly sensitive to the inclusion of HDD in the sample




Source: Vivid Economics analysis of Ofwat IAP models

## 2.3 Efficiency challenge

**Efficiency challenges are a valuable but inherently difficult and risky component of cost assessment.** Since some companies are inevitably more efficient than others in delivering enhancement programmes and the sector at large is expected to become more productive over time, it is legitimate for Ofwat to consider challenging companies to deliver enhancement expenditure programmes in AMP7 at a lower cost than a comparable historical industry average. However, since models of enhancement costs cannot fully account for the drivers of efficient costs, there is no straightforward way of inferring relative efficiency from model outputs, and there is a risk that the models themselves already encode implicit challenges for at least some companies if they do not fully account for efficient costs.

**This section considers the judgements Ofwat appears to have made in applying efficiency challenges**, taking into account the quality of modelling and other evidence on plausible ranges of efficiency. It begins by reviewing efficiency challenges applied to modelled cost lines in Section 2.3.1, before considering challenges applied to other lines in Section 2.3.2.

### 2.3.1 Modelled enhancement costs

**Around two thirds of the challenge to company business plan costs in modelled enhancement lines comes from the models themselves, rather than explicit measures to account for efficiency.** 24% of business plan spending across these lines is disallowed in the IAP, which, as shown in Table 2, can be decomposed into three challenges:

- A 'modelling shortfall', equal to the difference between business plan costs (after reallocations between lines made by the IAP) and modelled allowances. This makes up 85% of the total challenge in water, 44% in waste and 66% overall.
- An efficiency challenge applied uniformly to modelled allowances for some lines of costs: for example an upper quartile challenge of 14% for flow to full treatment costs. This is 0% of the total challenge in water, 20% in waste and 9% overall.

- A claw back applied to all modelled allowances. This allows companies the lower of the modelled allowance (after the explicit efficiency challenge) and their business plan claim. This makes up 15% of the challenge in water, 34% in waste and 24% overall.

Table 2 Decomposition of challenge to business plan costs across modelled enhancement lines (£m)

| | Modelling shortfall | Efficiency challenge | Claw back | Total | % of business plan claim |
|---|---|---|---|---|---|
| Water | 1,209 | 0 | 217 | **1,426** | **38%** |
| Wastewater | 525 | 262 | 399 | **1,185** | **17%** |
| **Total** | **1,734** | **262** | **615** | **2,611** | **24%** |

Note: Business plans assessed after IAP reallocation
Source: Vivid Economics analysis

**Modelling shortfalls can sometimes be rationalised, but the magnitude of the shortfall in the IAP models is symptomatic of more fundamental problems with the models.** Total modelled allowances across the sector are expected to be approximately equal to total planned spending, if planned spending is at a level of efficiency similar to that observed in the sample data. Mathematically, it is highly unlikely that the two figures will match exactly – and a shortfall could arise if company business plans were less efficient than the sample data. However, the very large deviation between modelled and planned spending seen in the IAP is more likely to be explained by shortcomings in the IAP enhancement models identified in this report. These include:

- **treatment of logged costs**. Modelled estimates of *logged* costs cannot simply be transformed by the exponential function to arrive at an estimate of costs, as in logged models the error term, expressed in unlogged terms, has a skew distribution with a positive expected value. Of the IAP models, only the flow to full treatment model accounts for this fact in any way, and does so in a theoretically incorrect manner. *All other logged models underestimate industry costs*. As Figure 4 below illustrates, the extent of such an underestimate can be significant and will be larger for companies with larger programmes.

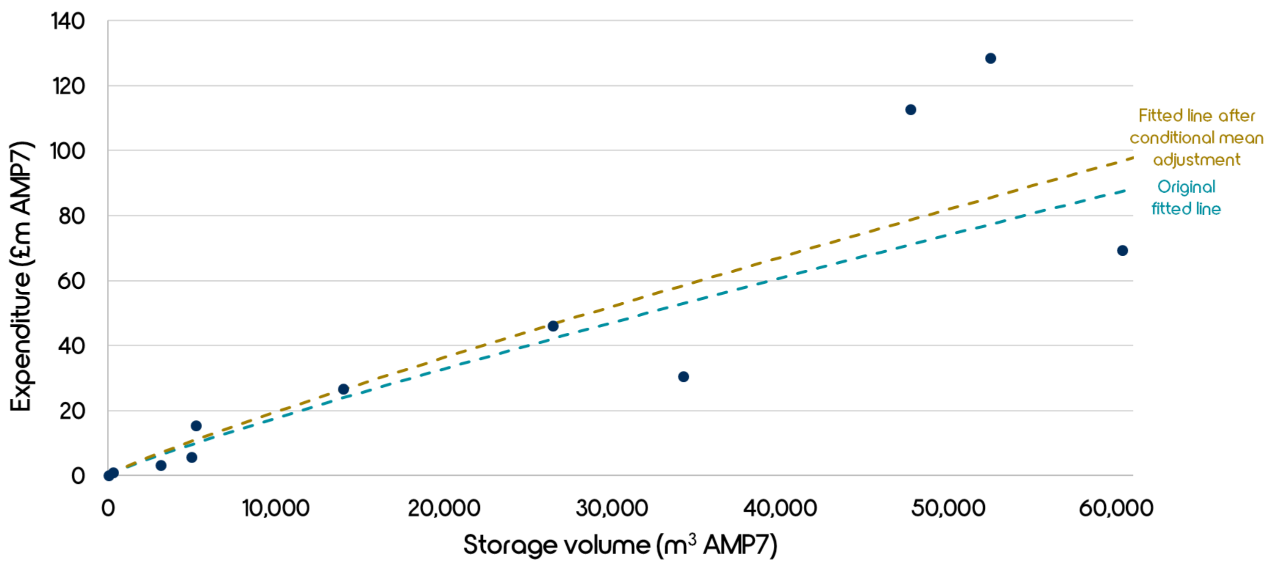
- **use of models with weak explanatory power** (see Section 3). Where log OLS or RE specifications are used, such models will generate allowances with large magnitude errors for companies with large programmes. If these are systematically adverse to the higher spending companies, perhaps as a result of the point immediately following, a high modelling shortfall would be expected.

- **influence of unrepresentative data points** (see Section 2.2). Notably, Hafren Dyfrdwy, by far the smallest wastewater company, has a modelled allowance in waste enhancement lines more than five times greater than its business plan costs. As Figure 3 and tables in the Appendix highlight, this materially affects the allowances of other companies.

- **omitted variables** that will tend to increase future spending and that can make historical models unrepresentative of future costs. For example, sewage treatment enhancement models lack quality drivers, which are expected to lead to increased unit costs as more advanced technologies are deployed to meet tightening consents (see Section 3.2).

**The IAP's sparing use of efficiency challenges and application of clawbacks has a clear rationale in this context, but could have been implemented in a less risky way.** Given the significant challenge many companies face as a consequence of the choice of models, the IAP's application of relatively modest efficiency challenges to modelled allowances is reasonable. Furthermore, since there are examples of companies who are overcompensated by models, it is also reasonable to use clawbacks to protect customers, though the lack of any further challenge to business plan costs may set unhelpful incentives for affected companies over the remainder of the price review. However, by implementing these mechanisms on a line-by-line basis, the IAP increases risks: had it considered modelling evidence of efficiency across multiple lines, some errors would have tended to cancel each other out and the overall challenge would have been more robust.

Figure 4 A failure to adjust estimated costs in logged models can yield material underestimates of cost




Notes: Results shown are for storm tanks model 1, which regresses log capex on log storage volume
Adjustment shown is based on the conditional mean estimator, and accounts for the lognormal distribution of model predictions, in the case when errors are normally distributed

Source: Vivid Economics analysis of Ofwat IAP models

### 2.3.2 Unmodelled enhancement costs

**The IAP uses modelling evidence on relative efficiency to set company specific efficiency challenges for unmodelled expenditure.** It does this for both deep and shallow dives:

- For shallow dives, it applies a company-specific challenge referred to as a 'haircut' in some models to business plan costs. For most companies, this haircut is a percentage calculated as follows. For water enhancement, the challenge is the smaller of 15% and the estimated *inefficiency* of the company's business plan in wholesale water base costs – calculated as the ratio of business plan costs and allowed costs minus one. For waste, the challenge is the smaller of 10% and a composite measure of business plan inefficiency, made up of 70% estimated inefficiency of wholesale base costs and 30% of the estimated inefficiency of enhancement costs for P-removal. A steeper efficiency challenge appears to have been adopted for significant scrutiny companies.

- Deep dive assessments consider evidence provided by companies on the efficiency of spending and apply an adjustment where this found to be unconvincing. In many cases, this includes the shallow dive haircut, plus a further 20% of the claim where evidence on 'need' is deemed insufficient.

**Company-specific efficiency factors are not well founded.** In using evidence on base cost efficiency to assess the efficiency of enhancement spending, the IAP's approach to haircuts diverges starkly from the efficiency challenges applied to modelled cost lines, where, in setting any given challenge, there was no explicit consideration of modelling evidence of relative efficiency elsewhere. While Section 2.3.1 explains that the assessment of relative efficiency across a number of enhancement lines may helpfully serve to reduce risks in the efficiency challenge, the haircuts suffer from the opposite problem: they are based on evidence from spending in other areas that is unlikely to be informative. For whereas relative base cost efficiency reflects investment decisions made in previous AMPs and has a large opex component, relative enhancement efficiency exclusively concerns forward-looking capex decisions. This, and the apparently arbitrary caps and triangulation between efficiency scores in base and P-removal models (see Section 3.2), mean the haircuts lack justification.

**The use of these exacerbates problems related to enhancement opex.** Companies with high enhancement opex are doubly disadvantaged by the use of haircuts: first, since the IAP applies the haircut only to enhancement capex, they start with a lower baseline than companies that choose capex solutions; and, second, since they *appear* unrealistically inefficient in the base cost models (see Section 2.2), an exaggeratedly large haircut is applied to this baseline. The second of these effects applies even where enhancement opex is included in specific expenditure lines.

**More detailed assessment of company submissions could improve upon the use of company specific efficiency factors.** Companies have provided detailed information on the need for and efficiency of their programmes, which could be assessed in greater depth to inform a view on efficiency. This information was taken into account by the IAP in the deep dives.

# 3 Enhancement cost models

**This section sets out key findings on each enhancement line modelled by the IAP.** Enhancement lines that are subject to deep or shallow dives and do not include cost modelling are not discussed here.[1] The remainder of the section groups together modelled lines under 'growth' (Section 3.1), 'wastewater quality' (Section 3.2), 'water quality' (Section 3.3), and 'supply demand balance' (Section 3.4). The Statistical Annex sets out in full the statistical evidence that this section draws on.

## 3.1 Growth

### 3.1.1 Modelling approach

**The IAP's growth models estimate capex gross of grants and contribution, which differs very substantially from the sum efficient companies require to recover through customers.** Since a significant proportion of spending related to growth can be recouped through developer charges, in order to set revenue caps for customer bills, Ofwat adjusts allowances generated by the models by an estimated efficient recovery rate, which is assumed to be uniform across companies. The authors of this report have not seen materials on how these recovery rates were derived.

**By estimating gross rather than net spending on growth, the IAP models are likely to increase customer bills.** New connections, the explanatory variable used in the gross spending models, appears a sensible driver of these costs, but it is unlikely to explain much of the variation in net spending, which depends on factors such as the balance of on- or off-site spending required to accommodate growth and opportunities to install Sustainable Drainage Schemes (SuDS), which cannot be funded by developers. Since these factors vary between companies, the use of the IAP gross cost models and a uniform recovery rate creates a significant risk that companies are over-or under-funded. As this review understands company charges to developers will not be allowed to exceed either outturn costs or allowed costs, over-funded companies will be allowed to recover excessive revenue from customers, while customers of under-funded companies will pay for a significant share of the resulting shortfall that could otherwise have been recouped from developers.

**A more equitable approach would estimate efficient net spending directly and rely on other mechanisms to set developer contributions.** Drivers of net spending in models could include factors related to offsite assets, such as available headroom, or attributes of new development, such as the geographical concentration of growth. If these were not viable, deep dives would be more suitable than the current approach. Residual spending that is passed through to developers can then be controlled by competition, since much of the work is contestable, and regulation.

### 3.1.2 Wastewater growth

**Efficient wastewater growth allowances are estimated using two log-log panel data (RE) specifications: a 'historical' model (2012/13 – 17/18) and a 'forecast' (business plan data) model (2020/21 – 24/25).** The cost variable for both models is the sum of new developments and growth, growth at sewage treatment works (excluding sludge treatment) and reductions in flooding risk expenditure, taken as gross of grants & contributions. The cost driver is total new connections (household and non-household). Both variables are smoothed using a 3-year rolling average and logged; as with water growth, no apparent use has been made

[1] For wholesale water, these lines are addressing low pressure, drinking water protected areas schemes, making ecological improvements at abstractions, eel regulations, improvements to river flows, invasive non-native species measures, environmental investigations and options appraisals, raw water deterioration, resilience, SEMD and non-SEMD, taste odour and colour improvements, water framework directive measures, selected Supply-Demand balance subcomponents, and freeform lines. For wholesale wastewater, these lines are chemical investigations, conservation drivers, discharge relocation schemes, eel regulations, groundwater schemes, WINEP/NEP investigations, monitoring flows at CSOs, N removal, odour, P removal technology investigations, resilience, SEMD and non-SEMD, sludge quality and growth, transferred private sewers and pumping stations, UV disinfection, and freeform lines.

of company data in Business Plan Data Table WWS2a which aligns the drivers of schemes with spending associated with them over time. Model-level allowances are triangulated 50:50 to estimate company modelled allowances. No explicit additional efficiency challenge is applied to these allowances.

**While combining new developments and growth, and growth at sewage treatment works (STWs) in a model is prudent, the inclusion of sewer flooding is questionable.** Trade-offs between new developments and growth, and growth at STWs activity mean individual models are likely to have weak narrative fit and statistical power. By contrast, sewer flooding expenditure has few trade-offs with the two other lines, and is not well explained by volume of new connections – the cost driver affects only the share of sewer flooding expenditure associated with mitigating service quality deteriorations from new connections – and in any case, is funded separately through an ODI.

**Although wastewater growth model coefficients are significant, the overall models are highly unstable and generate implausibly wide ranges of estimated costs.** The ‘historical’ model slope coefficient implies a 4.8% rise in company costs when new connections rise by 10%, while the ‘forecast’ model slope coefficient suggests a 12.6% rise. While changes in the constant term partially offset this, it is implausible that growth was a high fixed cost – low unit cost activity historically, but will be a high unit cost – low fixed cost activity going forward. This instability is evident in allowances from the individual models: the ‘historical’ model forecasts industry spend of £1.7bn across AMP7, while the ‘forecast’ model estimates industry spend at £3.0bn. The ‘forecast’ model is particularly sensitive to excluding individual companies, with the omission of HDD changing the average company efficiency score by 42 percentage points, and reallocating £0.9bn across the industry (Statistical Annex, Figure 18).

**The use of RE specifications is a source of potential risk for companies and customers**. RE makes restrictive assumptions about the underlying data, and uses an estimation technique that may not perform well in small samples. Equivalent OLS specifications appear to have more stable coefficients across ‘historical’ and ‘forecast’ models (Figure 5), and are more robust to excluding companies (Figure 20).

Figure 5 ‘Combined’ data OLS model performs well, while removal of sewer flooding causes some loss of fit

| | IAP Historical | IAP Forecast | ‘Combined’ data | ‘Combined’ data OLS | ‘Combined’ data OLS w/o sewer flooding |
|---|---|---|---|---|---|
| New connections | 0.48 (0.03) | 1.26 (0.00) | 1.10 (0.00) | 1.17 (0.00) | 1.14 (0.00) |
| Constant | 2.00 (0.00) | -0.03 (0.94) | 0.32 (0.56) | 0.18 (0.71) | 0.28 (0.34) |
| Estimation technique | RE | RE | RE | OLS | OLS |
| N | 49 | 55 | 124 | 124 | 124 |
| $R^2$ | 0.62 | 0.91 | 0.87 | 0.87 | 0.80 |

Note: RE estimation technique is random effects; OLS estimation technique is Ordinary Least Squares
‘Combined’ dataset refers to the use of all observations from the IAP dataset (2011/12 – 24/25)
Cost driver and cost variable are logged and smoothed (3yr average)
P-values in parentheses

Source: Vivid Economics

**Wastewater growth could be better assessed using the ‘deep dive’ process due to the instability of all tested cost benchmarking approaches**. An alternative specification based on OLS, a combined historical-forecast dataset (2011/12 – 24/25), and excluding sewer flooding expenditure was tested to address concerns noted above. The model has better statistical fit than the ‘historical’ and ‘forecast’ specifications (Figure 5). However, results remain highly sensitive to the inclusion of Hafren Dyfrdwy, with its exclusion leading to a capex reallocation of £0.2bn (Figure 20). Given the importance of wastewater growth (30% of industry BP wastewater enhancement capex), a ‘deep dive’ is proposed as an alternative to cost benchmarking.

### 3.1.3 Water growth

**The industry median unit cost is used to estimate efficient allowances for water new developments and growth, using both 'historical' (2011/12 – 17/18) and 'forecast' (2020/21 – 24/25) data.** The cost variable for both models is new connections and development costs, taken as gross of grants & contributions, and the cost driver is total new connections. Unit cost is calculated as total expenditure divided by total volume for the 'historical' and 'forecast' periods respectively. Industry median unit costs are calculated for the two datasets, and used to predict company costs, with company modelled allowances triangulated using 50:50 weights. As noted in Section 3.1.2 for waste, company data in Business Plan Data Table WS2A is not used. No explicit additional efficiency challenge is applied to these allowances.

**The comparability of company data on costs per new connection in water growth models is a concern.** Off-site network reinforcement needs are not equal across companies, and are driven by a number of factors including existing network layout, headroom, demographic patterns, resilience and supply-demand balance. These factors not only drive off-site activity levels, but also present cost allocation issues – companies may allocate network reinforcement activity which overlaps growth, resilience and SDB to each line in different proportions. Companies undertake different levels of on-site 'contestable' work (company's share of work that could be undertaken by the company or by developers). Unit cost models which are gross of grants & contributions will lead to companies with high shares of 'contestable' appearing relatively inefficient.

**Water growth unit cost models are relatively stable results across datasets and when dropping companies, however, the range of unit costs across the industry is implausible.** The 'historical' median unit cost is £890 per connection, with an industry allowance of £1.1bn; the 'forecast' median unit cost is £1,140 per connection, with an industry allowance of £1.5bn. The median unit cost approach is robust to excluding individual companies – the average company efficiency score changes by at most 5% when an individual company is excluded, with £48mn reallocated across the industry (Statistical Annex, Figure 23). However, the range of company-level unit costs is £500 (Sutton and East Surrey Water) – £2,100 (Severn Trent Water) in the 'historical' data, and £400 (Yorkshire Water) – £3,100 (South East Water) in the 'forecast' data. Across the two datasets, South East Water's unit cost changes from £800 ('historical') to £3,100 ('forecast'). The range of unit costs across the industry and changes across AMPs suggests that omitted variable bias is a considerable problem and differences between company modelled and claimed allowances cannot be explained by efficiency alone.

**No justification is provided for selecting the median unit cost approach over RE specifications for water growth, which appears to be inconsistent with the model selection for wastewater growth.** Water RE models were tested using log smoothed growth expenditure as the dependent variable, and log smoothed new connections as the cost driver, with smoothing based on a 3 year rolling average of both variables. 'Historical' model results implied a 10.7% rise in cost when new connections rise by 10%, while 'forecast' model results implied a 9.3% rise in costs. Coefficient estimates and model fit were much more stable across the datasets, and when dropping companies than those in wastewater RE models (Statistical Annex, Figure 24).

**The critical step to further improve modelling is to obtain more reliable company data.** This review tested alternative specifications based on OLS and a 'combined' dataset (2011/12 – 24/25) and found they perform similarly to the RE model as set out in Figure 6. However, until the data issues described earlier has been addressed, it is difficult to assess the potential for modelling to improve, particularly given the narrative concerns around off-site work. As with waste growth, a 'deep dive' process could reduce risks for customers, developers and companies.

Figure 6 'Combined' data OLS model performance is in line with IAP models

| | IAP Historical | IAP Forecast | 'Combined' data | 'Combined data OLS |
|---|---|---|---|---|
| New connections | 1.07 (0.00) | 0.93 (0.00) | 1.06 (0.00) | 1.04 (0.00) |
| Constant | -0.24 (0.34) | 0.21 (0.43) | 0.08 (0.72) | 0.12 (0.63) |
| Estimation technique | RE | RE | RE | OLS |
| N | 86 | 85 | 205 | 205 |
| $R^2$ | 0.82 | 0.81 | 0.82 | 0.82 |

Note: RE estimation technique is random effects; OLS estimation technique is Ordinary Least Squares
'Combined' dataset refers to the use of all observations from the IAP dataset (2011/12 – 24/25)
Cost driver and cost variable are logged and smoothed (3yr average)
P-values in parentheses

Source: Vivid Economics

### 3.1.4 First time sewerage

**First time sewerage allowances are set using two linear pooled OLS (POLS) models: a 'historical' model (2012/13 – 17/18) and a 'forecast' (business plan data) model (2020/21 – 24/25).** The cost variable in both models is first time sewerage costs, and the cost drivers are connectable properties served by s101A schemes, and the same variable squared. Both variables are smoothed using a 3-year rolling average. Model-level allowances are triangulated 50:50 to estimate company modelled allowances. No explicit additional efficiency challenge is applied in this enhancement area.

**The IAP models' use of properties$^2$ can capture narratives around economies of scale, but other geographical factors are omitted.** Other potentially relevant exogenous factors not included in the model include the sparsity of properties served by s101A schemes, a factor considered relevant in the IAP's base cost modelling, and increasing unit costs as the share of unconnected properties falls.

**Models have significant coefficients and reasonable explanatory power, but are unstable when omitting individual companies.** Statistical fit is fairly high in both models, with an $R^2$ of 0.83 in the 'historical' model, and 0.90 in the 'forecast' model. Properties and properties$^2$ coefficients are significant or almost significant in both specifications. The negative coefficient on properties$^2$ shows there is statistical evidence for economies of scale in properties served by s101A schemes. The cost of connecting an additional s101A property is £34,000 on average in the 'historical' model, and £41,500 on average in the 'forecast' model. These elasticities vary very modestly across the industry due to the inclusion of the properties$^2$ term, showing that the properties$^2$ term has a modest impact on model allowances. However, when Anglian is removed from the sample, there is a dramatic loss of statistical fit, with $R^2$ falling to 0.53, and a negative allowance predicted for Anglian. Efficiency scores in the 'historical' model are highly sensitive to the inclusion of Anglian as shown in Figure 29.

**This instability stems from the dominance of spending by two companies, which suggests this line is unsuitable for modelling.** Over AMP7, Anglian Water (47%) and Severn Trent England (25%) have 72% of forecast properties served by s101A schemes (Figure 26), and 55% of business plan expenditure between them (Figure 27). In the 'historical' dataset (2011/12 – 17/18), the distribution of activity is even more uneven – Anglian Water has 77% of the industry's properties served by s101A schemes, and 63% of expenditure. Cost models where one or two companies have such disproportionate weight are unlikely to be a suitable basis for cost benchmarking.

## 3.2 Waste service: quality

### 3.2.1 P removal

**Company allowances are calculated using two 'forecast' data (2020/21 – 24/25) linear models – a capex model, and a totex model.** Expenditure on P removal at activated sludge STWs, and filter bed STWs is summed for the cost variable, with total population equivalent (PE) served by STWs with tightened or new consents, and number of schemes used as cost drivers. Company-level variables are summed over the 'forecast' data period, with the models run on the resulting cross-sectional data. An efficiency challenge of 6%, based on upper quartile efficiency scores, is applied to estimates from both models. Allowances from the totex model are adjusted down for industry average 'implied opex', which is calculated by differencing predicted values from the totex and capex models. Allowances from these models are triangulated 50:50 to estimate company modelled allowances.

**The models' aggregation of activity and expenditure at activated sludge STWs and filter bed STWs is justified, as is the use of a totex model and the inclusion of an economies of scale driver.** The IAP model's approach to aggregate expenditure and activity for the two technologies is preferable to separate modelling as choices between spending in the two areas are, at least partially, under company control. Totex models are more appropriate than capex only models given operational trade-offs between capex and opex-intensive solutions for P removal. However, totex model cost drivers (sites and PE) should include both capex and opex-based schemes, to capture operational trade-offs. Number of sites captures scale economies in the two models.

**Complexity and quality are key omitted variables in the IAP P removal models.** While economies of scale are reflected through the number of sites, no account is made for the stringency of consents (quality). Consent standards are an important driver of base costs as shown by the use of a tight ammonia consents (<3mg/l) variable in the IAP sewage treatment and bioresources plus models. Consents are expected to have a similar effect on costs for wastewater quality enhancement programmes such as P removal, and relevant driver data is available from the PR19 Water Industry National Environment Programme (WINEP) dataset.

**There are some errors in the IAP's implementation of the models and the decision to select linear rather than log models is not substantiated.** 'Implied opex' is calculated as the difference between allowances from the totex and capex models – this is not valid from an econometric perspective as modelling noise will be conflated with opex. The correct approach would be to calculate the industry's opex share of totex based on company business plans, and adjust down totex model allowances using this value. In addition, efficiency scores for the log capex and totex specifications have been calculated incorrectly as shown in the Appendix. No justification is provided for the decision to select a linear, rather than a log-log functional form.

**Both models perform reasonably well on tests of statistical fit and stability, although the wide range of efficiency scores is of concern.** Statistical fit is high with $R^2$ of 0.93, and population equivalent (PE) and number of schemes variable coefficients significant at the 1% level in both models. Coefficients in the capex model imply that raising PE by 10% raises capex by 0.75%, and raising sites by 10% raises costs by 12.9%, reflecting economies of scale at sites. Models are robust to dropping companies with the average company efficiency score changing by at most 6 percentage points when an individual company is omitted (South West Water). Model efficiency score ranges are not plausible: 0.32 (South West Water) – 1.37 (Thames Water) in the capex model and 0.32 (South West Water) – 1.33 (Thames Water) in the totex model. The range of scores suggests differences are unlikely to be primarily driven by relative efficiency.

**The addition of a consents variable, % of STWs with P consents >1.1mg/l, improves model performance, and narrows efficiency scores to a more plausible range.** The variable was constructed using industry data from the PR19 WINEP dataset, and performs well in models, with the negative, and significant, coefficients consistent with the narrative that laxer constraints lower costs as shown in Figure 7. The specification with PE, number of sites, and % STWs with P consents >1.1mg/l performs well, but may be overfitted due to the small sample size (10) and the inclusion of 4 variables (including constant). The variable continues to perform well when number of sites is excluded, and remains significant across log model specifications. The inclusion of consents significantly narrows the efficiency score range across the industry as shown in Figure 39.

Figure 7 P removal models are improved by the addition of % STWs with P consents >1.1mg variable

| | IAP models | | | | Add consents variable | | | | Add consents variable, remove number of sites | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) |
| PE | 0.07 (0.00) | | 0.08 (0.00) | | 0.08 (0.00) | | 0.09 (0.00) | | 0.10 (0.00) | | 0.11 (0.00) | |
| Log PE | | 0.52 (0.00) | | 0.52 (0.00) | | 0.50 (0.00) | | 0.50 (0.00) | | 0.57 (0.00) | | 0.59 (0.00) |
| No. of sites | 1.29 (0.01) | | 1.48 (0.00) | | 1.04 (0.01) | | 1.24 (0.01) | | | | | |
| Log no. of sites | | 0.19 (0.34) | | 0.24 (0.25) | | 0.21 (0.13) | | 0.25 (0.09) | | | | |
| % STWs P consents >1.1 mg/l | | | | | -238 (0.08) | -1.74 (0.01) | 236 (0.10) | -1.72 (0.02) | -385 (0.05) | -1.70 (0.02) | -412 (0.07) | -1.67 (0.04) |
| Constant | 43.1 (0.14) | 0.98 (0.12) | 39.6 (0.19) | 0.87 (0.16) | 110.1 (0.03) | 1.54 (0.01) | 106.3 (0.04) | 1.42 (0.01) | 192.2 (0.01) | 1.85 (0.00) | 204 (0.01) | 1.80 (0.00) |
| Estimation technique | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS |
| N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| $R^2$ | 0.93 | 0.91 | 0.93 | 0.92 | 0.96 | 0.97 | 0.96 | 0.97 | 0.88 | 0.95 | 0.86 | 0.95 |

Note: OLS estimation technique is Ordinary Least Squares
P-values in parentheses

Source: Vivid Economics

**This suggests that modelling can be improved by triangulating across a suite of P removal models, capturing economies of scale (number of sites) and consents (% STWs with P consents >1.1mg/l).**

### 3.2.2 Chemical removal

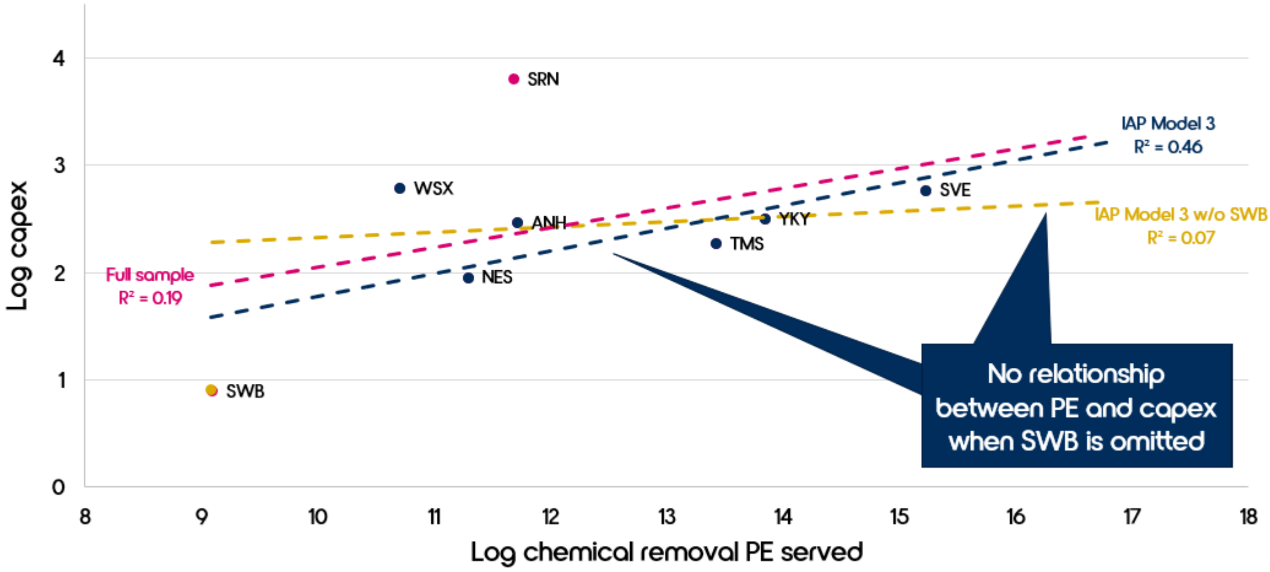
**Efficient allowances for chemical removal are set using four log-log models, all based on 'forecast' data (2020/21 – 24/25).** The four models are:

1. Log capex upon log PE, excluding Southern Water and Severn Trent England
2. Log capex upon log PE and log $PE^2$, excluding Southern Water and Severn Trent England
3. Log capex upon log PE, excluding Southern Water
4. Log capex upon log PE and log $PE^2$, excluding Southern Water

**The cost variable in all four models is capex on chemical removal schemes, while the driver is population equivalent served by works with chemical removal technologies in place.** The models are all cross-sectional, with all variables aggregated and logged. Allowances from the models are triangulated with equal weights, and no explicit additional efficiency challenge is applied.

**Model performance is very poor, with low statistical fit across specifications, and a number of insignificant coefficients.** Models (1) and (3) have much lower fit than models (2) and (4), with $R^2$ of 0.39 and 0.46 in the former models, and 0.70 and 0.62 in the latter two. In addition, the decision to remove Southern Water and Severn Trent England does not seem principled – removing just one more company from the sample, SWB, causes the observed relationship in models (1) and (3) to almost entirely disappear as shown in Figure 8.

Figure 8 Comparison of IAP Model 3 fitted regression line before and after South West Water is omitted




Note: IAP model 3 already excludes SRN
Source: Vivid Economics

**The weak statistical relationship in all 4 IAP models is unsurprising given the wide range of chemicals covered by company obligations.** Companies have substance-specific obligations, covering over ten different chemicals as outlined in WINEP. Technology requirements differ by substance type, and as a result, unit costs (capex/PE) are expected to differ across companies. As obligations are outside of management control, models which fail to control for this will misstate efficient costs.

**Engineering narratives and statistical performance suggest the IAP chemical removal models are not fit for purpose.** A 'shallow' or 'deep dive' would be more appropriate unless a suitable quality driver for the type of chemical improvement obligations can be developed.

### 3.2.3 Event duration monitoring

**Allowances for event duration monitoring (EDM) are estimated using the industry median unit cost, based on 'forecast' (2020/21 – 24/25) data.** The cost variable is capex on event duration monitoring at intermittent discharge sites, and the volume driver is number of intermittent discharge sites with event duration monitoring. The median unit cost is combined with company forecasts on the number of schemes required to estimate company allowances. No explicit additional efficiency challenge is applied in this enhancement line.

**The median industry unit cost is £12,500 per site. Though the use of the median mitigates the effects of significant outliers, the latter points towards significant problems with the data.** While the unit cost approach seems reasonable according to engineering narratives, data quality issues appear to be a significant problem in the IAP unit cost model. There is a large range of variation around the median value used to set allowances: Hafren Dyfrdwy's unit cost is calculated at £140 per site, while Southern Water's unit cost is £85,000. An industry unweighted mean would result in a unit cost estimate of £19,000 per site, while a weighted average leads to an estimate of £15,400 per site, both significantly more than the median value used. Some variation across companies may be due to inconsistencies around the inclusion of Environment Agency permit fees (c.£6,500 per site) within PR19 data tables.

**Improvement to industry data can improve the credibility of modelling in this area.** If data quality is not improved, a 'shallow dive' may be a more appropriate way of setting allowances for EDM.

### 3.2.4 Flow monitoring

**Flow monitoring costs are benchmarked using the industry median unit cost based on 'forecast' (2020/21 – 24/25) data.** The cost variable is capex on flow monitoring at STWs, and the volume driver is the number of monitors for flow monitoring at STWs. Company allowances are estimated using company forecasts on the number of schemes, and the industry median unit cost. No explicit additional efficiency challenge is applied in this enhancement line.

**The industry median unit cost is £50,000.** However, as in the EDM model, company-level variation around this figure is large: Hafren Dyfrdwy's unit cost is £8,750 per monitor, while Wessex Water's unit cost is £231,000. Ofwat's IAP model does note potential issues around Wessex's data on the number of flow monitor schemes, which could affect its position as an outlier, but does not flag other outlying points such as Thames Water at £186,000. The industry unweighted mean is £73,000, and the industry weighted mean is £63,000. If these statistics were used instead of the median, an additional £39m or £23m would be allocated across the industry respectively.

**As with EDM, while the median unit cost does mitigate the effect of outliers, the implausibly wide variation in the data from multiple companies means that improved data would increase the credibility of the modelling.**

### 3.2.5 Flow to full treatment schemes

**Six models are used to triangulate allowances in flow to full treatment (FFT):**

1. Capex upon schemes (linear)
2. Capex upon FFT shortfall (linear)
3. Capex upon schemes and FFT shortfall (linear)
4. Log capex upon log schemes
5. Log capex upon log FFT shortfall
6. Log capex upon log schemes and log FFT shortfall

**The cost variable in all six models is capex on schemes to increase FFT.** Cost drivers are the number of such schemes in company business plans, and the shortfall in FFT each company faces. All six models are cross-sectional and allowances are triangulated with equal weights for each of the models. Log model allowances are 'standardised' so the same capex is allocated across the industry in log models as is under linear models. An upper quartile efficiency challenge of 14% is applied to company fitted allowances to estimate final allowances.

**Statistical performance is variable across models, with log models showing lower model fit.** The log shortfall model is particularly weak in this regard, with an $R^2$ of 0.49. By contrast, linear models (2) and (3) both have $R^2$ of 0.89. Results are sensitive to dropping companies, with the omission of Yorkshire Water changing average triangulated efficiency scores by 9 percentage points. Models with shortfall to FFT and number of schemes suffer from multicollinearity – the correlation between the two variables is 0.91.

**The triangulation approach does not serve to reduce risks and modelling could be improved by dropping some of the specifications.** Only models of approximately equal quality should be triangulated – in this case, the linear models seem to outperform log models, and the combined shortfalls and schemes model produces an insignificant schemes coefficient. Out of the 3 linear models, the shortfall model produces the most stable efficiency scores.

### 3.2.6 Sanitary parameters

**Allowances for sanitary parameters are set using two models: an 'Exponential' model, and a 'Power' model.** The cost variable in both models is capex per population equivalent (PE), where expenditure is on reduction of sanitary parameters, and PE covers the population equivalent served by STWs that have new or tightened consents for one or more sanitary parameters. The cost driver in both models is PE per site, where number of sites is the number of STWs which have new or tightened sanitary parameter consents. The models are cross-sectional, and allowances are triangulated 50:50. No explicit additional efficiency challenge is applied in this enhancement area.

**Both models have functional forms that lack justification in economic or engineering narratives and the statistical fit and stability of the models is weak.** The coefficients are hard to interpret in the IAP model's reported functional forms. However, as shown in , the models can be rewritten with logged capex as the dependent variable. The rewritten specifications show that the 'Power' and 'Exponential' models have unexpected functional forms that are hard to justify: the 'Exponential' coefficient implicitly imposes the constraint that the coefficient on sites is equal to minus the coefficient on PE plus 1, while the 'Power' model makes log capex a linear function of PE per site and log PE. Furthermore, neither model has strong statistical fit, with an $R^2$ of 0.64 in the 'Power' model, and 0.71 in the 'Exponential' model. The slope coefficient in the 'Power' model is interpreted as a one unit change in PE / sites leading to a 0.055% increase in capex. The slope coefficient in the 'Exponential' model means that 10% more PE raises costs by 2.2%, while a 10% increase in sites lowers costs by 7.8%. Both models are also highly unstable, with average efficiency score changes of 13 percentage points in the 'Power' model and 12 percentage points in the 'Exponential' model when Severn Trent England and United Utilities are excluded respectively.

Box 1 Functional forms for Ofwat IAP 'Power' and 'Exponential' sanitary parameters models

**'Power' model**

- Original specification:

$$\frac{Capex}{PE} = \alpha e^{\beta \frac{PE}{Sites}}$$

- Rewritten specification:

$$log(Capex) = \tilde{\alpha} + \beta \frac{PE}{Sites} + log(PE), where\ \tilde{\alpha} = log(\alpha)$$

**'Exponential' model**

- Original specification:

$$\frac{Capex}{PE} = \alpha \left(\frac{PE}{Sites}\right)^{\beta}$$

- Rewritten specification:

$$log(Capex) = \tilde{\alpha} + (\beta + 1)log(PE) - \beta log(Sites)\ where\ \tilde{\alpha} = log(\alpha)$$

A simple improvement on the 'Exponential' model is an 'unconstrained' version of the model as set out below:

$$log(Capex) = \alpha + \beta log(PE) + \gamma log(Sites) + \varepsilon$$

**Coefficients for this model are set out in Figure 9 below.** This model has stronger economic rationale, as there is no arbitrary constraint on the relationship between capex and PE versus capex and sites. However, the model still has an implausibly wide range of efficiency scores, as shown in Figure 44, albeit slightly narrower than the original models.

Figure 9 The 'Unconstrained Log-log' model performs similarly to the 'Power' model

| | 'Exponential' | 'Power' | 'Unconstrained Log-log' |
|---|---|---|---|
| PE/site | -0.06 (0.01) | -0.78 (0.00) | |
| Log PE | | | 0.17 (0.35) |
| Log sites | | | 0.57 (0.04) |
| Constant | 0.42 (0.01) | 1.11 (0.82) | 0.86 (0.26) |
| Estimation technique | OLS | OLS | OLS |
| N | 10 | 10 | 10 |
| $R^2$ | 0.64 | 0.71 | 0.68 |

Note: OLS estimation technique is Ordinary Least Squares
Cost variable is capex / PE in the 'Exponential' and 'Power' models, and log capex in the 'Unconstrainted Log-log' model
P-values in parentheses

Source: Vivid Economics

### 3.2.7 Spill frequency

**Spill frequency allowances are set using a single OLS log-log model, which explains capex on storage schemes to reduce spill frequency at CSOs (cost variable) using volume of additional storage provided.** The model is cross-sectional based on 'forecast' data (2020/21 – 24/25). An efficiency challenge of 10%, based on model quality, is applied to model allowances to reach company final allowances.

**The log-log model has a high degree of statistical fit, with an $R^2$ of 0.96, and significant slope and constant terms. However, stability is a cause for concern.** Coefficients are consistent with a 10% increase in storage volume raising costs by 7.6%. However, model stability is poor – efficiency scores change by 16.5 percentage points on average when Southern Water, the smallest volume company in this category, is excluded from the model (Figure 46).

**The application of a 10% efficiency challenge based on model quality appears arbitrary.** No other lines are given the same challenge based on model quality and the majority of models include no challenge.

### 3.2.8 Storm tanks

**Storm tank allowances are set using two log-log 'forecast' data models (2020/21 – 24/25) – a specification with just storage as the driver, and another with storage and number of schemes as drivers.** The cost variable in both models is capex on storage schemes at STWs to increase storm tank capacity. Model 1 has only the volume of storage provided at CSOs and storm tanks as the driver, while model 2 has volume, and the number of STWs at which new or additional storage is to be provided. The models are cross-sectional, with all variables aggregated and logged. Allowances from the two models are triangulated 25:75, in favour of the model containing schemes, based on an assessment of model quality. An efficiency challenge of 5% is applied to triangulated allowances.

**The models do not fit all company costs and there are concerns around model stability.** Model fit of the overall sample is high in both models with $R^2$ of around 0.97, with storage is always highly significant (<1% level), and while number of schemes is insignificant, its p-value is fairly low at 0.26. Combined with the strong narrative for including number of schemes (economies of scale), the inclusion of schemes in model 2 is appropriate. Coefficients in model 2 suggest a 10% increase in volume of storage raises company costs by 7.8%, while a 10% increase in number of sites raises costs by 2.3%. However, the wide range of efficiency scores suggests the models do not account for all drivers of efficiency costs, with a range of 0.47 (Severn Trent England) – 1.73 (Southern Water). Stability is a concern, with the average efficiency score changing by 9 percentage points when Hafren Dyfrdwy or Severn Trent England are excluded from the models (Figure 48).

## 3.3 Water service: quality

### 3.3.1 Meeting lead standards

**Efficient costs for meeting lead standards are estimated using two linear random effects (RE) models – a 'historical' data model (2011/12 – 17/18), and a 'forecast' data model (2020/21 – 24/25).** The IAP model's cost variable is smoothed expenditure on meeting lead standards, and the drivers are smoothed number of lead communication pipes, and the smoothed number of lead communication pipes replaced. Smoothing is based on taking a 3-year rolling average of each variable. Allowances from the two models are triangulated 50:50 to estimate company modelled allowances. No explicit additional efficiency challenge is applied in this enhancement area.

**From an engineering perspective, merging water treatment and treated water distribution expenditure for meeting lead standards does not appear justified.** The two cost lines have different drivers, with no real operational trade-offs between the two subservices. Water treatment expenditure consists of spending on the construction of new orthophosphate dosing plants, while treated water distribution expenditure is allocated to the replacement of communication pipes. Of these two activities, lead communication pipe replacement is typically more expensive than orthophosphate dosing to meet lead standards, but is applied by companies in response to regulatory requirements imposed by the Drinking Water Inspectorate (DWI). As IAP models contain no relevant drivers for new orthophosphate dosing plant requirements, this component of lead standards costs is not explained well by the model.

**The models do not fit all of the sample data well and suffer from instability.** Statistical fit is reasonably high, with $R^2$ of 0.84 in the 'historical' model, and 0.78 in the 'forecast' model. Both drivers are significant, with coefficients fairly stable across the two specifications. Coefficient values are interpreted as the cost of replacing an additional lead communication pipe being £344 in the 'historical' model, and £452 in the 'forecast' model. Efficiency score variation is large: 0.00 (United Utilities & South East Water) – 2.65 (Hafren Dyfrdwy). Stability when dropping companies is poor, with the exclusion of Yorkshire Water changing the average efficiency score by 33 percentage points, and producing a negative allowance for Hafren Dyfrdwy.

**The addition of opex to the IAP lead standards models improves model performance.** Northumbrian Water, United Utilities and Yorkshire Water have reported all, or a majority of lead standards expenditure under opex rather than capex. This may reflect water supply pipe replacements, in line with DWI recommendations, which cannot be capitalised as part of some company's capitalisation policies. Figure 51 shows capex and opex across the industry.

**Alternative specifications based for TWD subservice-level totex can improve performance.** Results for these specifications are set out in Figure 10, with the models having slightly better statistical fit than IAP 'historical' and 'forecast' data models. WTW costs are modest across the entire industry, with just three companies reporting costs in this area: Anglian Water, South West Water, and Severn Trent England. WTW costs cannot be modelled effectively given the small set of companies – costs could instead be assessed using the 'shallow dive' process.

Figure 10 Meeting lead standards treated water distribution totex models perform well

| | IAP Historical | IAP Forecast | Forecast with totex | Forecast with TWD totex | Forecast with TWD totex and w/o pipe stock |
|---|---|---|---|---|---|
| Communication pipes | 2.40*10^-6 (0.02) | 3.22*10^-6 (0.07) | 3.31*10^-6 (0.09) | 2.86*10^-6 (0.10) | |
| Communication pipes replaced | 3.44*10^-4 (0.00) | 4.5*10^-4 (0.00) | 5.9*10^-4 (0.00) | 6.0*10^-4 (0.00) | 6.9*10^-4 (0.00) |
| Constant | 0.09 (0.68) | 0.19 (0.57) | 0.28 (0.39) | 0.07 (0.81) | 0.95 (0.02) |
| Estimation technique | RE | RE | RE | RE | RE |
| N | 64 | 71 | 71 | 71 | 75 |
| R2 | 0.84 | 0.78 | 0.82 | 0.80 | 0.81 |

Note: RE estimation technique is random effects; OLS estimation technique is Ordinary Least Squares
Cost driver and cost variable are logged and smoothed (3yr average) in all specifications
Totex models have capex + opex in the cost line; TWD totex is treated water distribution totex only
P-values in parentheses

Source: Vivid Economics

## 3.4 Supply-Demand Balance

### 3.4.1 Metering

**Allowances for metering are set using two 'historical' data models (2011/12 – 17/18) – a linear (levels) model, and a log-log model.** The cost variable in both models is metering costs for meters requested by optants, introduced by companies (selective) and for businesses. The cost driver is the total number of optant and selective meters installed. The model is cross-sectional, with both costs and the cost driver aggregated over 'historical' data years. Thames Water and Southern Water are removed from the sample before the model is run, as the model fit drops considerably when they are included. Allowances from the two models are triangulated 33:67, with the weighting in favour of the log-log model, due to its higher $R^2$ and greater data distribution. No explicit additional efficiency challenge is applied in this enhancement area.

**The relationship estimated by the model has a clear engineering narratives, though meter penetration may also affect this.** Higher meter penetration rates may be expected to raise unit costs if the connections which remain without meters are often more costly to fit.

**The models cannot explain company costs across the sector, though the models' overall statistical fit and stability is reasonable.** $R^2$ is 0.88 for the levels model, and 0.96 for the log-log model. Meters installed is highly significant in both models. Coefficients imply that the cost of an additional meter is £234 in the linear model, while in the log-log model, a 10% increase in the number of meters installed, raises capex by 8.8%. The range of efficiency scores is implausibly wide: 0.41 (Hafren Dyfrdwy) – 3.73 (Thames Water). Efficiency score stability when excluding companies is fair, with the average efficiency score changing by around 7 percentage points when South East Water is omitted.

**Models that account for meter penetration offer some improvement.** Specifications using meter penetration are set out in Figure 11 below. These models include the 'combined' sample of 2011/12 – 24/25 – the use of a larger set of observations for the cross-sectional econometric models seems to improve results. 'Forecast' (2020/21 – 24/25) data models also seem to outperform the IAP 'historical' models based on statistical fit.

Figure 11 Metering model coefficients when adding % meter penetration rate, and testing the 'combined' dataset

| | 'Historical' dataset | | | | 'Combined' dataset | |
|---|---|---|---|---|---|---|
| | IAP Levels | IAP Log-log | Levels with % meter penetration rate | Log-log with % meter penetration rate | Levels with % meter penetration rate | Log-log with % meter penetration rate |
| Meters installed | 0.23 (0.00) | | 0.23 (0.00) | | 0.24 (0.00) | |
| Log meters installed | | 0.88 (0.00) | | 0.87 (0.00) | | 0.89 (0.00) |
| % meter penetration rate | | | 8.01 (0.71) | 0.46 (0.23) | 7.65 (0.77) | 0.17 (0.62) |
| Constant | 1.80 (0.69) | -0.80 (0.00) | -2.03 (0.86) | 0.99 (0.00) | -3.21 (0.84) | -0.88 (0.01) |
| Estimation technique | OLS | OLS | OLS | OLS | OLS | OLS |
| N | 15 | 15 | 15 | 15 | 14 | 14 |
| $R^2$ | 0.88 | 0.96 | 0.88 | 0.97 | 0.94 | 0.98 |

Note: OLS estimation technique is Ordinary Least Squares
'Historical' dataset is the summation of volume and spend over 2011/12 – 17/18; 'combined' dataset is the same over 2011/12 – 24/25
% meter penetration is share of properties served which are metered
P-values in parentheses

Source: Vivid Economics

### 3.4.2 Leakage SDB

**Allowed unit costs for leakage are estimated by the triangulation of three industry statistics.** The leakage Supply-Demand Balance (SDB) median 'reported unit cost' is estimated as the median across the industry of expenditure on leakage reduction per Ml/d of leakage reduction, using business plan data. The other two industry median statistics are the median leakage underperformance and outperformance outcome delivery incentive (ODI) rates. The average of these three industry median estimates is taken as the industry-wide estimate of leakage SDB unit cost (£/Ml/d). Allowances are set by multiplying the maximum of leakage benefits beyond 15%, or beyond UQ performance in 2024/25 with the minimum of the industry unit cost estimate of £1.60m/Ml/d, and company unit cost.

**The use of ODI rates to set allowances is inappropriate.** ODI out- and underperformance rates for leakage depend on the social benefits of leakage reduction (customer willingness to pay for leakage reductions) – which, given widespread reduction beyond the economic level of leakage across the sector by 2024/25, would be expected to exceed the costs to companies. These rates therefore understate the efficient costs of leakage reduction so should not be used in benchmarking. This has a material impact on companies as the median 'reported unit cost' is £2.07m/Ml/d, whereas the ODI out- and underperformance rates are £1.64m/Ml/d and £1.10m/Ml/d respectively.

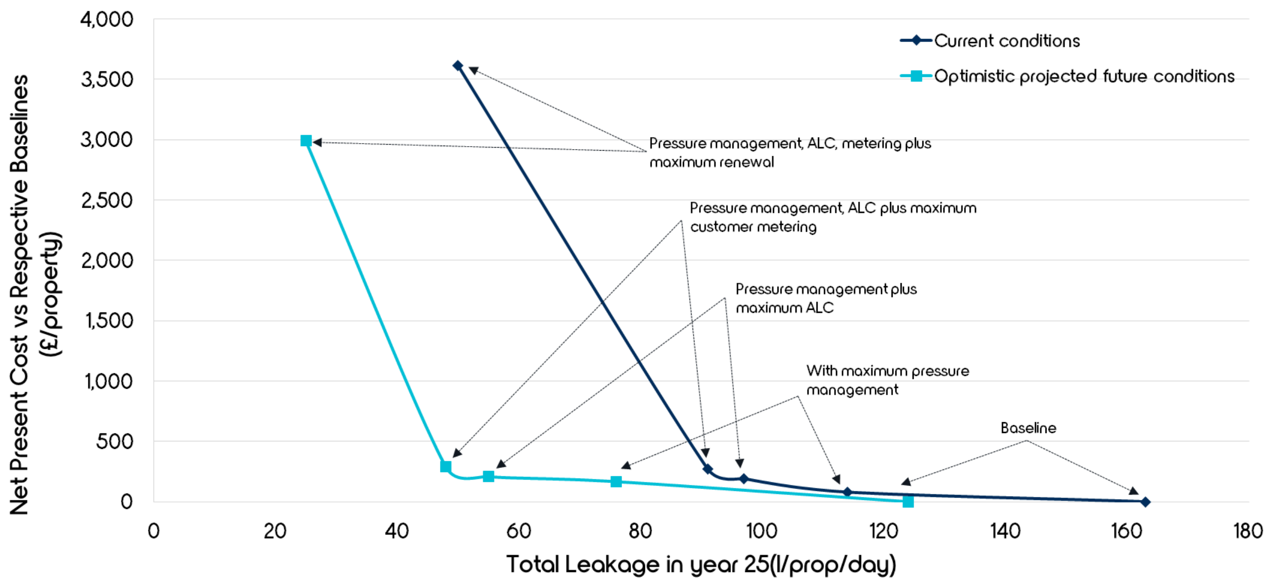
**Even if the use of ODI rates were appropriate, the rates used to triangulate industry median unit cost are not an accurate reflection of social benefits.** ODI rates reported by companies are net of the 'sharing factor' of around 50% – which means they do not reflect the full societal benefits from out- or underperformance. Furthermore, reported ODI rates have been multiplied by five to estimate social benefits across AMP7 – this is an unsuitable way to calculate the net present value of reduced leakage over the AMP.

**The inclusion of Essex and Suffolk Water as a separate data point is anomalous.** The inclusion of the company, which forms part of Northumbrian Water, is anomalous and overweights the group-wide management model in efficiency assessments. The same point is explained in more detail for the case of Hafren Dyfrdwy in Section 2.2.

**Baseline levels of leakage are an important determinant of leakage costs, which is not accounted for in the models.** UK Water Industry Research (UKWIR) found leakage unit costs to be decreasing with leakage levels in £/property terms as shown in Figure 12. This relationship appears to be acknowledged elsewhere in the IAP, for example where a cost adjustment claim for base costs from Anglian is partially allowed on a similar basis. The range of 'reported unit cost' across the industry is consistent with the omission of important drivers: the lowest cost company, PRT, has a unit cost of £0.33m/Ml/d, while the highest cost company, Sutton and East Surrey Water, has a unit cost of £4.85m/Ml/d. It is not plausible that efficiency differences alone explain this 15-fold variation in unit costs.

**The omission of ODI rates from the calculation of industry median unit costs and accounts for baseline levels of leakage could improve the IAP approach.**

Figure 12 Analysis for UKWIR found leakage unit cost to be increasing as total leakage levels are driven down




Source: Reproduced from Long Term Leakage Goals, UKWIR

### 3.4.3 2020 – 25 SDB

**2020 – 25 SDB enhancement allowances are set using the industry median unit cost across both supply and demand-side schemes.** The median unit cost is calculated across the industry as total expenditure (totex) on new supply and demand-side schemes per Ml/d of SDB benefits delivered. Forecast data is used, as in the leakage SDB model. Efficient allowances are calculated by multiplying forecast benefit (Ml/d) with the minimum of the industry median and company unit costs. The industry median unit cost is £1.39m/Ml/d, with an industry range of 0.17 (Yorkshire Water) – 3.71 (Dwr Cymru). Amongst the two companies (Anglian Water, Southern Water) which report benefits and expenditure for more than one supply-side scheme, there is a wide span of unit costs: 1.31 – 4.22 (Anglian Water), and 0.50 – 1.05£m/Ml/d (Southern Water) respectively.

**Complexity and scheme-level economies of scale are not currently captured, and may explain large variation in company-level unit costs.** Supply-side interventions range in complexity from the development of new groundwater sources, to final effluent treatment re-use depending on resource availability and operational constraints. Scheme-level economies of scale may also exist, with fixed costs spread over larger benefit volumes for companies with large schemes. The aggregation of supply and demand-side 2020 – 25 schemes into a single category is questionable given that demand-side schemes have historically been more expensive.

### 3.4.4 Other SDB models

Other Supply-Demand Balance subcomponents (long-term enhancement, strategic regional solution, internal interconnections, and investigations and future planning), are not subject to industry-wide cost benchmarking, and are therefore not assessed in this report.

# 4 Conclusions

**This report highlights opportunities to substantially increase the robustness of enhancement cost assessment during the remainder of the PR19 process.** Ideally, Ofwat would undertake and document a systematic process of model development and selection along the lines laid out in Section 2.1 as it has for base cost models. If this was unfeasible in the time available, this report presents a host of quick wins that would build on and substantially improve existing modelling work. These include a set of overarching changes, including adding opex to dependent variables, eliminating Hafren Dyfrdwy from the sample, replacing the company-specific 'haircut', and rebalancing the efficiency challenge away from modelled shortfalls towards more explicit measures. It also includes a set of more targeted changes to the models already developed in the IAP, including decisions over whether to use modelling or other approaches. Table 3 below summarises these.

Table 3 Line-by-line review of enhancement modelling

| | Expenditure line | Model assessment findings | Recommended improvements |
|---|---|---|---|
| Growth | Growth (overarching) | Assessment of costs gross rather than net of grants and contribution creates downside risk for customers | If data on relevant causal factors can be obtained, attempt estimating net costs using appropriate drivers<br>Deep dive clearly preferable to current gross cost modelling |
| | Growth (wastewater) | Unstable RE models, divergence between historical and forecast. Inclusion of sewer flooding, which is funded by ODI | |
| | Growth (water) | More stable than wastewater model. Unreliable company data yields implausible unit cost ranges | |
| | First time sewerage | Two companies dominate expenditure, causing instability | Use a deep dive |
| Waste quality | P-removal | Some errors in implementation, implausible ranges of efficiency scores, absence of complexity driver | Correct errors, triangulate between models, including data on P<1.1mg consents |
| | Chemical removal | Low fit, which disappears when SWB removed. Type of chemical obligations not captured. | Attempt totex modelling, with chemical obligations type. Deep dive if unsuccessful |
| | Event duration monitoring | Reasonable specification but undermined by inconsistent company data | Attempt to improve data. Consider shallow dive if unsuccessful |
| | Flow monitoring | Concerns on data comparability | Attempt to improve data |
| | Flow to full treatment | Variable model performance, with log specifications weaker than linear | Triangulate among the linear models |
| | Sanitary parameters | 'Power' and 'Exponential' specifications inconsistent with engineering logic | Test more transparent specifications, including log model suggested |
| | Spill frequency | Good fit, but instability related to SRN | Diagnose issues in SRN data, consider dropping from sample |
| | Storm tanks | Good overall fit, range of efficiency scores implausibly wide | Diagnose efficiency score variation to understand omitted variables, reflect in approach to efficiency challenge |
| Water quality | Meeting lead standards | Merger of orthophosphate dosing and replacement of lead pipes does not reflect distinct regulatory drivers. Model highly unstable and produces implausible efficiency score range | Use treated water distribution model for replacement of communication pipes, use shallow dive for orthophosphate |
| Supply demand | Metering | Reasonable statistical fit and stability. Implausible range of efficiency scores | Potential improvement possible from including meter penetration |
| | Leakage | Not valid to use WTP figures to estimate costs WTP figures are incorrectly taken net of sharing rates.<br>Unit costs highly variable, do not account for increasing marginal costs | Do not use WTP figures, attempt modelling that uses leakage as explanatory variable |
| | 2020-25 schemes | Very wide variation in unit costs, reflecting diversity of schemes | Use a deep or shallow dive |

Source: Vivid Economics analysis

# 5 Statistical Annex

## 5.1 Introduction

**This section sets out supporting statistical evidence for the report.** The results in this section are the basis of the statistical assessments made in Section 3, and are used as supporting evidence throughout the report. Evidence comprises three sets of tests set out below:

**Model statistical fit:** Across all models, the estimated relationship between costs and explanatory variables should be aligned with engineering narratives and, ideally, be statistically significant. Models should be able to explain a reasonable proportion of the observed variation in company costs if we are to be confident that they capture the main casual drivers of efficient costs. The key exhibits in this section are Ofwat model regression tables, with the main evidence being:

- Significance of coefficients – signs of coefficients aligned with engineering narratives, any evidence of collinearity between cost drivers in multivariate models
- Overall model fit –$R^2$ and company efficiency score range

**Model stability:** Given the often uneven trends and distribution of enhancement expenditure across time and companies, there is a real risk that relationships estimated from statistical models are not robust, and therefore unsuitable for cost benchmarking purposes. The key exhibits in this area are coefficients across models where multiple specifications are used, and changes in company efficiency scores:

- Coefficients across triangulated specifications – how consistent are coefficients across models, for instance, 'forecast' vs. 'historical' model results
- Impact of dropping individual companies – assessing the changes in company efficiency scores when a single company is removed from the model dataset

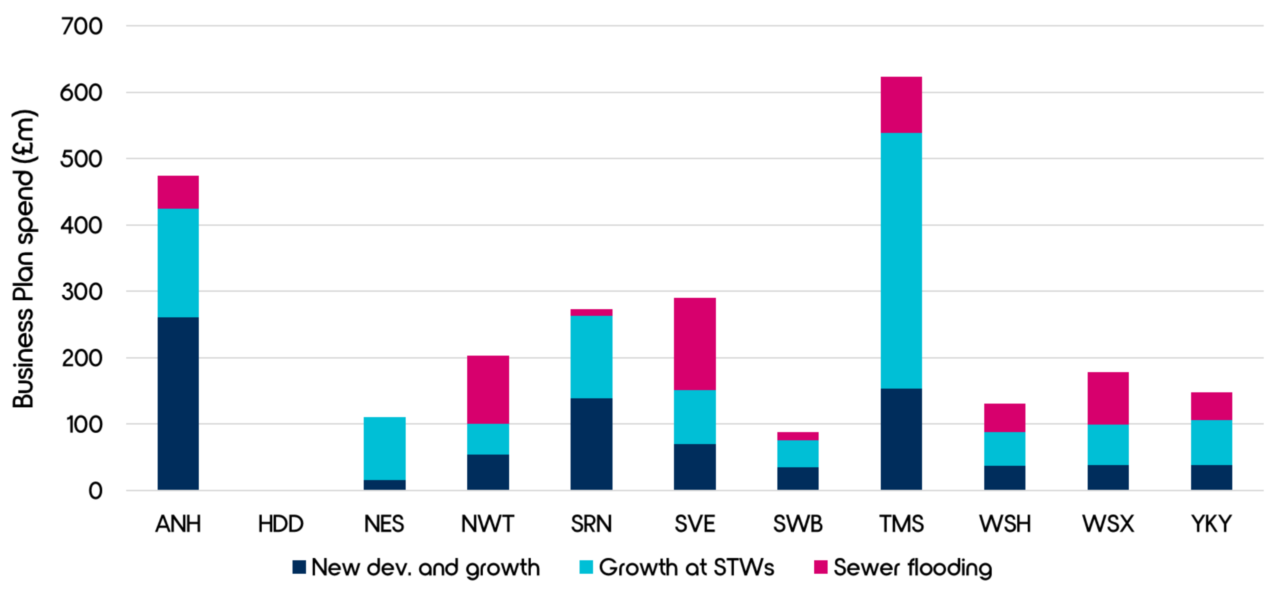
**Alternative specifications:** Statistical models will always be subject to a degree of error – even if the key drivers have been identified, different model fits will lead to different results. Testing alternative specifications can help us be more confident in our results and identify potential improvements in fit. Key exhibits in this section are regression tables presenting plausible, alternative model specifications and comparing the statistical parameters of those models with the originals. Alternative specifications tested included:

- Ordinary Least Squares (OLS) vs. Random Effects (RE)
- Linear vs. log models
- 'Historical' (2011/12 – 17/18) vs. 'forecast' (2020/21 – 24/25) vs. 'combined' (2011/12 – 24/25) datasets
- Changes to included cost lines, drivers and functional forms

## 5.2 Growth

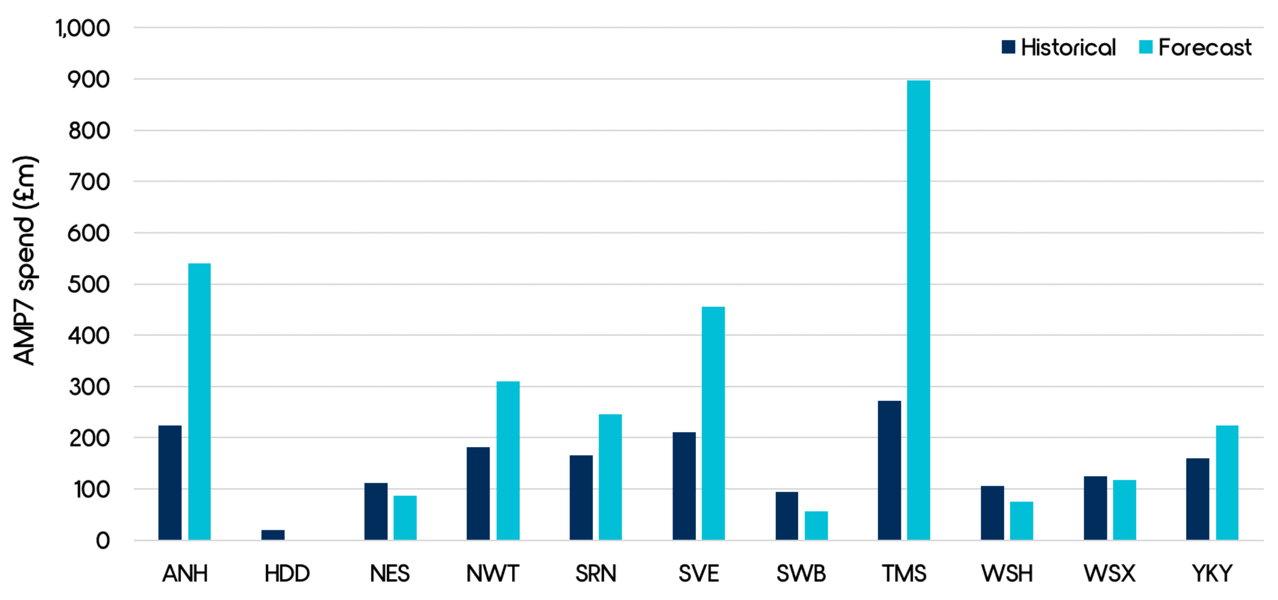
### 5.2.1 Wastewater growth

Figure 13 Wastewater growth business plan spend by enhancement expenditure line



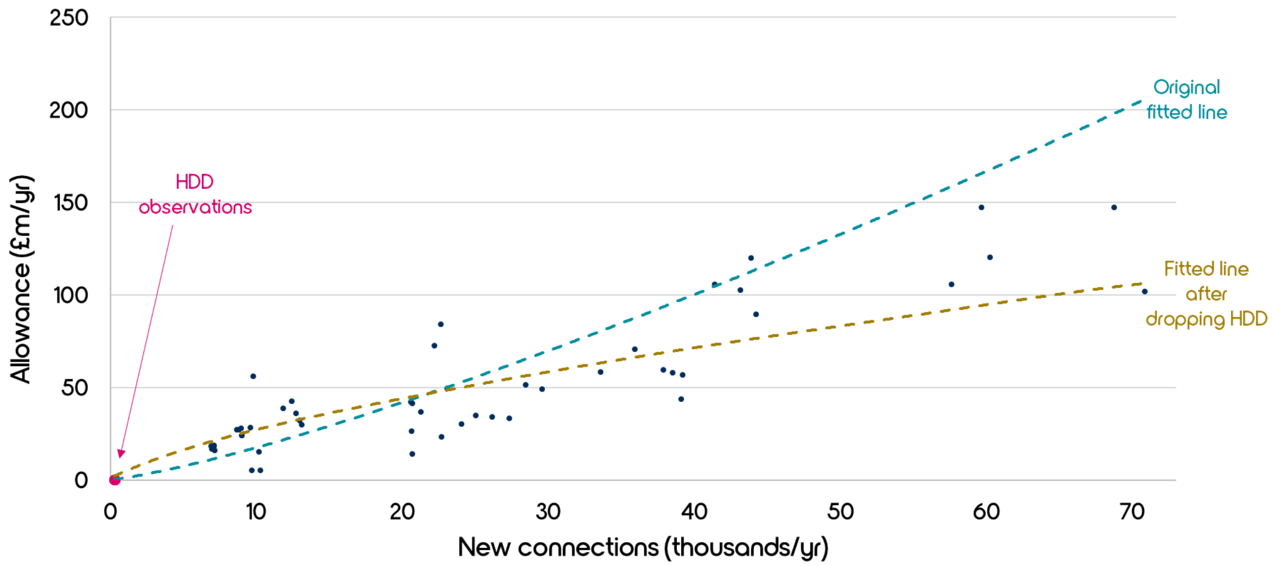

Source: Ofwat IAP wastewater growth model

Figure 14 Company allowances from Ofwat 'historical' and 'forecast' wastewater growth models




Source: Ofwat IAP wastewater growth model

Figure 15 Comparison of fitted line in Ofwat's 'forecast' waste growth model with and without Hafren Dyfrdwy




Note: Line is curved as a result of log-log model specification
Source: Vivid Economics

Figure 16 Coefficients in Ofwat IAP wastewater growth models

| | Historical | Forecast |
|---|---|---|
| New connections | 0.48 (0.03) | 1.26 (0.00) |
| Constant | 2.00 (0.00) | -0.03 (0.94) |
| Estimation technique | RE | RE |
| N | 49 | 55 |
| $R^2$ | 0.62 | 0.91 |

Note: RE estimation technique is random effects
Cost driver and cost variable are logged and smoothed (3yr average)
P-values in parentheses
Source: Vivid Economics replication of Ofwat IAP wastewater growth Random Effects (RE) models

Figure 17 Efficiency scores when dropping companies from IAP 'historical' waste growth model

| Company | Original model | Models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.91 | 1.94 | 1.40 | 1.79 | 1.96 | 2.03 | 2.09 | 2.72 | 1.86 | 1.83 | 1.83 |
| HDD | 0.03 | 0.03 | 0.09 | 0.04 | 0.02 | 0.02 | 0.01 | 0.01 | 0.03 | 0.03 | 0.03 |
| NES | 1.76 | 1.79 | 2.02 | 1.85 | 1.73 | 1.76 | 1.52 | 1.82 | 1.73 | 1.68 | 1.78 |
| NWT | 0.89 | 0.90 | 0.74 | 0.86 | 0.90 | 0.93 | 0.90 | 1.14 | 0.86 | 0.85 | 0.86 |
| SRN | 1.61 | 1.63 | 1.43 | 1.59 | 1.62 | 1.67 | 1.59 | 1.99 | 1.57 | 1.54 | 1.58 |
| SVE | 1.40 | 1.42 | 1.07 | 1.33 | 1.43 | 1.48 | 1.49 | 1.93 | 1.36 | 1.34 | 1.35 |
| SWB | 0.99 | 1.01 | 1.26 | 1.07 | 0.96 | 0.97 | 0.81 | 0.95 | 0.98 | 0.95 | 1.01 |
| TMS | 2.23 | 2.26 | 1.45 | 2.03 | 2.32 | 2.42 | 2.61 | 3.47 | 2.16 | 2.14 | 2.11 |
| WSH | 1.52 | 1.55 | 1.80 | 1.62 | 1.49 | 1.51 | 1.29 | 1.54 | 1.50 | 1.46 | 1.55 |
| WSX | 1.37 | 1.40 | 1.46 | 1.42 | 1.36 | 1.38 | 1.23 | 1.49 | 1.35 | 1.31 | 1.37 |
| YKY | 1.10 | 1.12 | 1.00 | 1.09 | 1.11 | 1.14 | 1.08 | 1.35 | 1.08 | 1.06 | 1.09 |
| Average change | | 0.02 | 0.27 | 0.07 | 0.03 | 0.05 | 0.14 | 0.34 | 0.03 | 0.06 | 0.04 |

Note: SVE and HDD are not included in the 'historical' data model; capex used in efficiency score calculations is smoothed capex rather than company reported capex in BP years (2020/21 – 24/25)
Average change is the average absolute change in efficiency scores compared to the 'Original model'

Source: Vivid Economics

Figure 18 Efficiency scores when dropping companies from IAP 'forecast' waste growth model

| Company | Original model | Models excluding company observations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | HDD | NES | NWT | SRN | SVE | SWB | TMS | WSH | WSX | YKY |
| ANH | 0.79 | 0.78 | 1.15 | 0.85 | 0.74 | 0.82 | 0.74 | 0.82 | 0.71 | 0.85 | 0.81 | 0.74 |
| HDD | 0.56 | 0.56 | 0.05 | 0.61 | 0.53 | 0.55 | 0.60 | 0.63 | 0.66 | 0.64 | 0.60 | 0.62 |
| NES | 2.26 | 2.23 | 1.46 | 2.44 | 2.12 | 2.29 | 2.19 | 2.39 | 2.18 | 2.46 | 2.35 | 2.22 |
| NWT | 0.52 | 0.51 | 0.59 | 0.56 | 0.49 | 0.53 | 0.49 | 0.54 | 0.47 | 0.56 | 0.53 | 0.49 |
| SRN | 1.09 | 1.07 | 1.11 | 1.17 | 1.02 | 1.11 | 1.03 | 1.13 | 1.00 | 1.17 | 1.12 | 1.04 |
| SVE | 0.64 | 0.63 | 0.87 | 0.69 | 0.60 | 0.66 | 0.60 | 0.67 | 0.58 | 0.69 | 0.66 | 0.61 |
| SWB | 1.66 | 1.64 | 0.88 | 1.80 | 1.56 | 1.67 | 1.62 | 1.77 | 1.63 | 1.81 | 1.73 | 1.65 |
| TMS | 0.68 | 0.67 | 1.24 | 0.73 | 0.63 | 0.70 | 0.62 | 0.69 | 0.59 | 0.72 | 0.69 | 0.62 |
| WSH | 2.13 | 2.10 | 1.29 | 2.30 | 2.00 | 2.15 | 2.07 | 2.26 | 2.07 | 2.32 | 2.21 | 2.10 |
| WSX | 1.46 | 1.44 | 1.08 | 1.58 | 1.37 | 1.48 | 1.40 | 1.54 | 1.39 | 1.58 | 1.51 | 1.42 |
| YKY | 0.79 | 0.78 | 0.78 | 0.85 | 0.74 | 0.81 | 0.75 | 0.82 | 0.73 | 0.85 | 0.81 | 0.76 |
| Average change | | 0.02 | 0.42 | 0.09 | 0.07 | 0.02 | 0.05 | 0.06 | 0.07 | 0.09 | 0.04 | 0.04 |

Note: SVT is not included in the 'forecast' data model; capex used in efficiency score calculations is smoothed capex rather than company reported capex in BP years (2020/21 – 24/25)
Average change is the average absolute change in efficiency scores compared to the 'Original model'

Source: Vivid Economics

Figure 19 Coefficients in 'combined' data model with OLS and without sewer flooding in the dependent variable

| | IAP Historical | IAP Forecast | 'Combined' data | 'Combined' data OLS | 'Combined' data OLS w/o sewer flooding |
|---|---|---|---|---|---|
| New connections | 0.48 (0.03) | 1.26 (0.00) | 1.10 (0.00) | 1.17 (0.00) | 1.14 (0.00) |
| Constant | 2.00 (0.00) | -0.03 (0.94) | 0.32 (0.56) | 0.18 (0.71) | 0.28 (0.34) |
| Estimation technique | RE | RE | RE | OLS | OLS |
| N | 49 | 55 | 124 | 124 | 124 |
| $R^2$ | 0.62 | 0.91 | 0.87 | 0.87 | 0.80 |

Note: RE estimation technique is random effects; OLS estimation technique is Ordinary Least Squares
'Combined' dataset refers to the use of all observations from the IAP dataset (2011/12 – 24/25)
Cost driver and cost variable are logged and smoothed (3yr average)
P-values in parentheses
Source: Vivid Economics

Figure 20 Efficiency scores when dropping companies from 'combined' data model with OLS without sewer flooding

| Company | 'New model' | Model excluding company observations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | HDD | NES | NWT | SRN | SVE | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.42 | 1.51 | 1.74 | 1.45 | 1.33 | 1.50 | 1.36 | 1.35 | 1.46 | 1.35 | 1.47 | 1.43 | 1.36 |
| HDD | 0.62 | 0.59 | 0.24 | 0.62 | 0.65 | 0.62 | 0.65 | 0.64 | 0.75 | 0.67 | 0.71 | 0.60 | 0.61 |
| NES | 3.60 | 3.70 | 3.14 | 3.65 | 3.49 | 3.72 | 3.54 | 3.50 | 3.86 | 3.55 | 3.82 | 3.57 | 3.47 |
| NWT | 0.48 | 0.51 | 0.53 | 0.49 | 0.46 | 0.51 | 0.47 | 0.46 | 0.50 | 0.46 | 0.50 | 0.48 | 0.46 |
| SRN | 1.94 | 2.03 | 2.04 | 1.96 | 1.84 | 2.02 | 1.87 | 1.86 | 2.02 | 1.87 | 2.02 | 1.93 | 1.86 |
| SVE | 0.64 | 0.68 | 0.75 | 0.65 | 0.60 | 0.67 | 0.61 | 0.61 | 0.66 | 0.61 | 0.66 | 0.64 | 0.61 |
| SWB | 2.38 | 2.43 | 1.91 | 2.41 | 2.32 | 2.45 | 2.36 | 2.33 | 2.58 | 2.37 | 2.54 | 2.36 | 2.30 |
| TMS | 1.17 | 1.26 | 1.57 | 1.19 | 1.09 | 1.24 | 1.11 | 1.11 | 1.18 | 1.10 | 1.20 | 1.18 | 1.12 |
| WSH | 2.62 | 2.68 | 2.22 | 2.65 | 2.54 | 2.70 | 2.58 | 2.55 | 2.82 | 2.59 | 2.78 | 2.60 | 2.52 |
| WSX | 1.49 | 1.54 | 1.37 | 1.51 | 1.43 | 1.54 | 1.45 | 1.44 | 1.58 | 1.46 | 1.57 | 1.48 | 1.43 |
| YKY | 1.09 | 1.14 | 1.13 | 1.11 | 1.04 | 1.14 | 1.06 | 1.05 | 1.14 | 1.05 | 1.14 | 1.09 | 1.05 |
| Average change | | 0.06 | 0.26 | 0.02 | 0.07 | 0.06 | 0.04 | 0.05 | 0.10 | 0.04 | 0.09 | 0.01 | 0.06 |

Note: 'Combined' data model includes SVT observations prior to 2018/19, and HDD / SVE observations after
Average change is the average absolute change in efficiency scores compared to the 'New model'
Source: Vivid Economics

### 5.2.2 Water growth

Figure 21 Model coefficients for Ofwat IAP water growth random effects (RE) models

| | Historical | Forecast |
|---|---|---|
| New connections | 1.07 (0.00) | 0.93 (0.00) |
| Constant | -0.24 (0.34) | 0.21 (0.43) |
| Estimation technique | RE | RE |
| N | 86 | 85 |
| $R^2$ | 0.82 | 0.81 |

Note: RE estimation technique is random effects
Cost driver and cost variable are logged and smoothed (3yr average)
Minor differences between replication attempt coefficients and Ofwat IAP water growth model spreadsheet
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP water growth Random Effects (RE) models

Figure 22 Model coefficients from 'combined' data OLS water growth model

| | IAP Historical | IAP Forecast | 'Combined' data | 'Combined data OLS |
|---|---|---|---|---|
| New connections | 1.07 (0.00) | 0.93 (0.00) | 1.06 (0.00) | 1.04 (0.00) |
| Constant | -0.24 (0.34) | 0.21 (0.43) | 0.08 (0.72) | 0.12 (0.63) |
| Estimation technique | RE | RE | RE | OLS |
| N | 86 | 85 | 205 | 205 |
| $R^2$ | 0.82 | 0.81 | 0.82 | 0.82 |

Note: RE estimation technique is random effects; OLS estimation technique is Ordinary Least Squares
'Combined' dataset refers to the use of all observations from the IAP dataset (2011/12 – 24/25)
Cost driver and cost variable are logged and smoothed (3yr average)
P-values in parentheses

Source: Vivid Economics

Figure 23 Efficiency scores when dropping companies from Ofwat IAP water growth median unit cost models

| Company | Original models | Model excluding company observations | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AFW | ANH | BRL | BWH | DVW | HDD | NES | NWT | PRT | SES | SEW | SRN | SSC | SVE | SVT | SWB | SWT | TMS | WSH | WSX | YKY |
| AFW | 0.65 | 0.65 | 0.68 | 0.65 | 0.66 | 0.66 | 0.67 | 0.65 | 0.65 | 0.65 | 0.65 | 0.67 | 0.67 | 0.68 | 0.67 | 0.66 | 0.68 | 0.65 | 0.68 | 0.67 | 0.65 | 0.65 |
| ANH | 1.49 | 1.48 | 1.55 | 1.50 | 1.50 | 1.50 | 1.54 | 1.49 | 1.50 | 1.48 | 1.48 | 1.53 | 1.53 | 1.55 | 1.54 | 1.50 | 1.55 | 1.49 | 1.55 | 1.52 | 1.48 | 1.48 |
| BRL | 0.91 | 0.90 | 0.94 | 0.91 | 0.91 | 0.91 | 0.94 | 0.90 | 0.91 | 0.90 | 0.90 | 0.93 | 0.93 | 0.94 | 0.94 | 0.91 | 0.94 | 0.90 | 0.94 | 0.92 | 0.90 | 0.90 |
| HDD | 2.21 | 2.19 | 2.30 | 2.22 | 2.23 | 2.23 | 2.28 | 2.21 | 2.22 | 2.19 | 2.19 | 2.27 | 2.27 | 2.30 | 2.28 | 2.23 | 2.30 | 2.20 | 2.30 | 2.26 | 2.19 | 2.19 |
| NES | 0.97 | 0.96 | 1.00 | 0.97 | 0.97 | 0.97 | 1.00 | 0.96 | 0.97 | 0.96 | 0.96 | 0.99 | 0.99 | 1.00 | 1.00 | 0.97 | 1.00 | 0.96 | 1.00 | 0.98 | 0.96 | 0.96 |
| NWT | 1.00 | 0.99 | 1.04 | 1.01 | 1.01 | 1.01 | 1.04 | 1.00 | 1.01 | 0.99 | 0.99 | 1.03 | 1.03 | 1.04 | 1.04 | 1.01 | 1.04 | 1.00 | 1.04 | 1.02 | 0.99 | 0.99 |
| PRT | 0.53 | 0.52 | 0.55 | 0.53 | 0.53 | 0.53 | 0.55 | 0.53 | 0.53 | 0.52 | 0.52 | 0.54 | 0.54 | 0.55 | 0.55 | 0.53 | 0.55 | 0.53 | 0.55 | 0.54 | 0.52 | 0.52 |
| SES | 0.74 | 0.73 | 0.76 | 0.74 | 0.74 | 0.74 | 0.76 | 0.73 | 0.74 | 0.73 | 0.73 | 0.75 | 0.75 | 0.76 | 0.76 | 0.74 | 0.76 | 0.73 | 0.76 | 0.75 | 0.73 | 0.73 |
| SEW | 3.02 | 2.99 | 3.14 | 3.03 | 3.05 | 3.05 | 3.12 | 3.01 | 3.03 | 2.99 | 2.99 | 3.10 | 3.10 | 3.14 | 3.12 | 3.05 | 3.14 | 3.01 | 3.14 | 3.08 | 2.99 | 2.99 |
| SRN | 1.56 | 1.54 | 1.62 | 1.56 | 1.57 | 1.57 | 1.61 | 1.55 | 1.56 | 1.54 | 1.54 | 1.60 | 1.60 | 1.62 | 1.61 | 1.57 | 1.62 | 1.55 | 1.62 | 1.59 | 1.54 | 1.54 |
| SSC | 1.81 | 1.80 | 1.88 | 1.82 | 1.83 | 1.83 | 1.87 | 1.81 | 1.82 | 1.80 | 1.80 | 1.86 | 1.86 | 1.88 | 1.87 | 1.83 | 1.88 | 1.81 | 1.88 | 1.85 | 1.80 | 1.80 |
| SVE | 1.96 | 1.94 | 2.04 | 1.97 | 1.98 | 1.98 | 2.02 | 1.96 | 1.97 | 1.94 | 1.94 | 2.01 | 2.01 | 2.04 | 2.02 | 1.98 | 2.04 | 1.95 | 2.04 | 2.00 | 1.94 | 1.94 |
| SWB | 1.18 | 1.17 | 1.23 | 1.18 | 1.19 | 1.19 | 1.22 | 1.18 | 1.18 | 1.17 | 1.17 | 1.21 | 1.21 | 1.23 | 1.22 | 1.19 | 1.23 | 1.18 | 1.23 | 1.21 | 1.17 | 1.17 |
| TMS | 1.15 | 1.13 | 1.19 | 1.15 | 1.16 | 1.16 | 1.18 | 1.14 | 1.15 | 1.13 | 1.13 | 1.18 | 1.18 | 1.19 | 1.18 | 1.16 | 1.19 | 1.14 | 1.19 | 1.17 | 1.13 | 1.13 |
| WSH | 1.12 | 1.11 | 1.17 | 1.13 | 1.13 | 1.13 | 1.16 | 1.12 | 1.13 | 1.11 | 1.11 | 1.15 | 1.15 | 1.17 | 1.16 | 1.13 | 1.17 | 1.12 | 1.17 | 1.15 | 1.11 | 1.11 |
| WSX | 0.50 | 0.50 | 0.52 | 0.50 | 0.50 | 0.50 | 0.52 | 0.50 | 0.50 | 0.50 | 0.50 | 0.51 | 0.51 | 0.52 | 0.52 | 0.50 | 0.52 | 0.50 | 0.52 | 0.51 | 0.50 | 0.50 |
| YKY | 0.37 | 0.36 | 0.38 | 0.37 | 0.37 | 0.37 | 0.38 | 0.37 | 0.37 | 0.36 | 0.36 | 0.38 | 0.38 | 0.38 | 0.38 | 0.37 | 0.38 | 0.37 | 0.38 | 0.38 | 0.36 | 0.36 |
| Average change | | 0.01 | 0.05 | 0.00 | 0.01 | 0.01 | 0.04 | 0.00 | 0.00 | 0.01 | 0.01 | 0.03 | 0.03 | 0.05 | 0.04 | 0.01 | 0.05 | 0.01 | 0.05 | 0.02 | 0.01 | 0.01 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original models'
Source: Vivid Economics

Figure 24 Efficiency scores when dropping companies from Ofwat IAP water growth random effects (RE) models

| Company | Original models | Model excluding company observations | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AFW | ANH | BRL | BWH | DVW | HDD | NES | NWT | PRT | SES | SEW | SRN | SSC | SVE | SVT | SWB | SWT | TMS | WSH | WSX | YKY |
| AFW | 0.67 | 0.61 | 0.71 | 0.68 | 0.67 | 0.67 | 0.67 | 0.67 | 0.69 | 0.65 | 0.67 | 0.70 | 0.68 | 0.69 | 0.69 | 0.70 | 0.67 | 0.67 | 0.68 | 0.67 | 0.65 | 0.64 |
| ANH | 1.54 | 1.31 | 1.66 | 1.55 | 1.49 | 1.53 | 1.48 | 1.53 | 1.58 | 1.51 | 1.57 | 1.63 | 1.57 | 1.55 | 1.59 | 1.64 | 1.54 | 1.53 | 1.55 | 1.55 | 1.50 | 1.44 |
| BRL | 0.93 | 0.90 | 0.94 | 0.94 | 0.94 | 0.94 | 0.98 | 0.92 | 0.94 | 0.89 | 0.89 | 0.95 | 0.93 | 0.97 | 0.95 | 0.95 | 0.93 | 0.92 | 0.94 | 0.92 | 0.90 | 0.89 |
| HDD | 2.20 | 2.42 | 2.00 | 2.18 | 2.33 | 2.24 | 2.65 | 2.19 | 2.24 | 2.02 | 1.91 | 2.16 | 2.12 | 2.42 | 2.20 | 2.14 | 2.21 | 2.18 | 2.27 | 2.14 | 2.14 | 2.15 |
| NES | 0.99 | 0.90 | 1.05 | 1.00 | 0.98 | 0.99 | 0.99 | 0.99 | 1.01 | 0.97 | 0.99 | 1.04 | 1.01 | 1.01 | 1.02 | 1.04 | 1.00 | 0.99 | 1.00 | 1.00 | 0.97 | 0.94 |
| NWT | 1.03 | 0.90 | 1.11 | 1.04 | 1.01 | 1.03 | 1.01 | 1.03 | 1.06 | 1.01 | 1.05 | 1.09 | 1.05 | 1.04 | 1.07 | 1.09 | 1.04 | 1.03 | 1.04 | 1.04 | 1.01 | 0.97 |
| PRT | 0.54 | 0.56 | 0.52 | 0.54 | 0.56 | 0.54 | 0.60 | 0.53 | 0.54 | 0.50 | 0.49 | 0.54 | 0.53 | 0.57 | 0.54 | 0.54 | 0.54 | 0.53 | 0.55 | 0.53 | 0.52 | 0.52 |
| SES | 0.75 | 0.76 | 0.73 | 0.75 | 0.77 | 0.76 | 0.82 | 0.74 | 0.76 | 0.70 | 0.69 | 0.75 | 0.74 | 0.79 | 0.76 | 0.75 | 0.75 | 0.74 | 0.76 | 0.74 | 0.73 | 0.72 |
| SEW | 3.11 | 2.94 | 3.19 | 3.13 | 3.12 | 3.12 | 3.19 | 3.08 | 3.16 | 2.99 | 3.01 | 3.20 | 3.13 | 3.20 | 3.18 | 3.20 | 3.11 | 3.09 | 3.13 | 3.09 | 3.02 | 2.95 |
| SRN | 1.60 | 1.48 | 1.67 | 1.62 | 1.60 | 1.61 | 1.62 | 1.59 | 1.63 | 1.55 | 1.58 | 1.66 | 1.62 | 1.64 | 1.65 | 1.67 | 1.61 | 1.60 | 1.62 | 1.60 | 1.56 | 1.52 |
| SSC | 1.86 | 1.77 | 1.91 | 1.88 | 1.88 | 1.87 | 1.93 | 1.85 | 1.89 | 1.79 | 1.80 | 1.91 | 1.87 | 1.92 | 1.91 | 1.91 | 1.87 | 1.85 | 1.88 | 1.85 | 1.81 | 1.77 |
| SVE | 2.02 | 1.78 | 2.16 | 2.04 | 1.98 | 2.01 | 1.98 | 2.01 | 2.07 | 1.98 | 2.04 | 2.13 | 2.06 | 2.05 | 2.08 | 2.14 | 2.03 | 2.01 | 2.04 | 2.03 | 1.97 | 1.90 |
| SWB | 1.21 | 1.15 | 1.25 | 1.22 | 1.22 | 1.22 | 1.25 | 1.21 | 1.24 | 1.17 | 1.18 | 1.25 | 1.22 | 1.25 | 1.24 | 1.25 | 1.22 | 1.21 | 1.23 | 1.21 | 1.18 | 1.15 |
| TMS | 1.18 | 1.00 | 1.28 | 1.19 | 1.14 | 1.17 | 1.13 | 1.17 | 1.21 | 1.16 | 1.21 | 1.26 | 1.21 | 1.18 | 1.22 | 1.26 | 1.18 | 1.17 | 1.19 | 1.19 | 1.15 | 1.11 |
| WSH | 1.15 | 1.09 | 1.19 | 1.16 | 1.16 | 1.16 | 1.19 | 1.15 | 1.17 | 1.11 | 1.12 | 1.19 | 1.16 | 1.19 | 1.18 | 1.19 | 1.16 | 1.15 | 1.17 | 1.15 | 1.12 | 1.10 |
| WSX | 0.51 | 0.50 | 0.52 | 0.52 | 0.52 | 0.52 | 0.54 | 0.51 | 0.52 | 0.49 | 0.49 | 0.53 | 0.51 | 0.53 | 0.53 | 0.53 | 0.51 | 0.51 | 0.52 | 0.51 | 0.50 | 0.49 |
| YKY | 0.38 | 0.34 | 0.40 | 0.38 | 0.37 | 0.38 | 0.37 | 0.38 | 0.39 | 0.37 | 0.38 | 0.40 | 0.39 | 0.38 | 0.39 | 0.40 | 0.38 | 0.38 | 0.38 | 0.38 | 0.37 | 0.36 |
| Average change | | 0.11 | 0.06 | 0.01 | 0.03 | 0.01 | 0.06 | 0.01 | 0.02 | 0.05 | 0.05 | 0.05 | 0.02 | 0.04 | 0.03 | 0.05 | 0.00 | 0.01 | 0.01 | 0.01 | 0.03 | 0.06 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original models'
Source: Vivid Economics

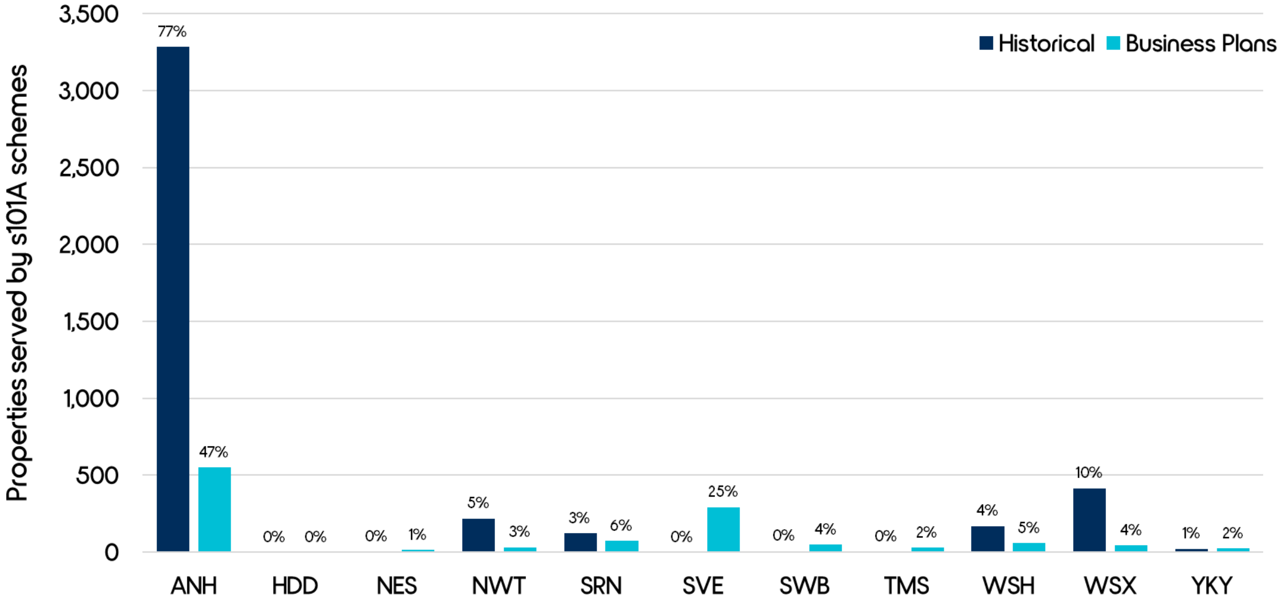
Figure 25 Efficiency scores when dropping companies from 'combined' dataset ordinary least squares (OLS) water growth model

| Company | New model | Model excluding company observations | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AFW | ANH | BRL | BWH | DVW | HDD | NES | NWT | PRT | SES | SEW | SRN | SSC | SVE | SVT | SWB | SWT | TMS | WSH | WSX | YKY |
| AFW | 0.67 | 0.63 | 0.69 | 0.67 | 0.66 | 0.66 | 0.65 | 0.66 | 0.68 | 0.66 | 0.66 | 0.68 | 0.67 | 0.68 | 0.69 | 0.68 | 0.67 | 0.66 | 0.67 | 0.66 | 0.65 | 0.62 |
| ANH | 1.47 | 1.38 | 1.55 | 1.47 | 1.45 | 1.46 | 1.32 | 1.44 | 1.52 | 1.52 | 1.51 | 1.51 | 1.49 | 1.50 | 1.56 | 1.53 | 1.48 | 1.47 | 1.48 | 1.47 | 1.44 | 1.34 |
| BRL | 0.96 | 0.94 | 0.98 | 0.98 | 0.97 | 0.97 | 1.03 | 0.96 | 0.97 | 0.91 | 0.92 | 0.98 | 0.97 | 0.99 | 0.98 | 0.98 | 0.97 | 0.96 | 0.97 | 0.96 | 0.94 | 0.93 |
| HDD | 2.61 | 2.67 | 2.51 | 2.73 | 2.72 | 2.66 | 3.50 | 2.63 | 2.56 | 2.19 | 2.26 | 2.63 | 2.59 | 2.74 | 2.50 | 2.56 | 2.62 | 2.58 | 2.59 | 2.59 | 2.51 | 2.72 |
| NES | 0.98 | 0.93 | 1.02 | 0.99 | 0.98 | 0.98 | 0.94 | 0.96 | 1.00 | 0.98 | 0.98 | 1.00 | 0.99 | 1.00 | 1.02 | 1.01 | 0.98 | 0.97 | 0.98 | 0.98 | 0.96 | 0.91 |
| NWT | 1.00 | 0.94 | 1.05 | 1.00 | 0.99 | 1.00 | 0.92 | 0.98 | 1.03 | 1.02 | 1.02 | 1.03 | 1.01 | 1.02 | 1.05 | 1.04 | 1.01 | 1.00 | 1.01 | 1.00 | 0.98 | 0.92 |
| PRT | 0.59 | 0.59 | 0.58 | 0.61 | 0.60 | 0.60 | 0.70 | 0.59 | 0.59 | 0.53 | 0.54 | 0.60 | 0.59 | 0.61 | 0.58 | 0.59 | 0.59 | 0.58 | 0.59 | 0.59 | 0.57 | 0.59 |
| SES | 0.81 | 0.80 | 0.81 | 0.83 | 0.83 | 0.82 | 0.93 | 0.81 | 0.81 | 0.74 | 0.75 | 0.82 | 0.81 | 0.84 | 0.81 | 0.81 | 0.81 | 0.80 | 0.81 | 0.80 | 0.78 | 0.80 |
| SEW | 3.15 | 3.04 | 3.23 | 3.20 | 3.17 | 3.16 | 3.22 | 3.12 | 3.20 | 3.05 | 3.07 | 3.22 | 3.18 | 3.24 | 3.23 | 3.22 | 3.17 | 3.14 | 3.16 | 3.14 | 3.07 | 2.99 |
| SRN | 1.60 | 1.54 | 1.65 | 1.62 | 1.60 | 1.60 | 1.59 | 1.58 | 1.63 | 1.58 | 1.58 | 1.64 | 1.62 | 1.65 | 1.66 | 1.64 | 1.61 | 1.60 | 1.61 | 1.60 | 1.57 | 1.51 |
| SSC | 1.90 | 1.84 | 1.94 | 1.93 | 1.91 | 1.91 | 1.96 | 1.88 | 1.93 | 1.83 | 1.84 | 1.94 | 1.92 | 1.96 | 1.95 | 1.94 | 1.91 | 1.89 | 1.91 | 1.89 | 1.85 | 1.81 |
| SVE | 1.96 | 1.85 | 2.05 | 1.97 | 1.95 | 1.95 | 1.82 | 1.93 | 2.02 | 1.99 | 1.99 | 2.01 | 1.99 | 2.00 | 2.06 | 2.03 | 1.97 | 1.95 | 1.97 | 1.96 | 1.92 | 1.80 |
| SWB | 1.23 | 1.19 | 1.26 | 1.25 | 1.24 | 1.24 | 1.26 | 1.22 | 1.25 | 1.19 | 1.20 | 1.26 | 1.24 | 1.27 | 1.27 | 1.26 | 1.24 | 1.23 | 1.24 | 1.23 | 1.20 | 1.17 |
| TMS | 1.12 | 1.05 | 1.19 | 1.12 | 1.11 | 1.11 | 1.00 | 1.10 | 1.16 | 1.17 | 1.16 | 1.15 | 1.14 | 1.14 | 1.19 | 1.17 | 1.13 | 1.12 | 1.13 | 1.12 | 1.10 | 1.01 |
| WSH | 1.17 | 1.13 | 1.20 | 1.19 | 1.18 | 1.18 | 1.20 | 1.16 | 1.19 | 1.13 | 1.14 | 1.20 | 1.18 | 1.21 | 1.20 | 1.20 | 1.18 | 1.17 | 1.18 | 1.17 | 1.14 | 1.11 |
| WSX | 0.53 | 0.51 | 0.54 | 0.54 | 0.53 | 0.53 | 0.56 | 0.52 | 0.54 | 0.50 | 0.51 | 0.54 | 0.53 | 0.55 | 0.54 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 | 0.52 | 0.51 |
| YKY | 0.37 | 0.35 | 0.39 | 0.37 | 0.37 | 0.37 | 0.35 | 0.36 | 0.38 | 0.37 | 0.37 | 0.38 | 0.37 | 0.38 | 0.39 | 0.38 | 0.37 | 0.37 | 0.37 | 0.37 | 0.36 | 0.34 |
| Average change | | 0.05 | 0.04 | 0.02 | 0.02 | 0.01 | 0.12 | 0.02 | 0.02 | 0.06 | 0.05 | 0.03 | 0.01 | 0.04 | 0.05 | 0.03 | 0.01 | 0.01 | 0.01 | 0.00 | 0.03 | 0.08 |

Note: Average change is the average absolute change in efficiency scores compared to the 'New model'
Source: Vivid Economics

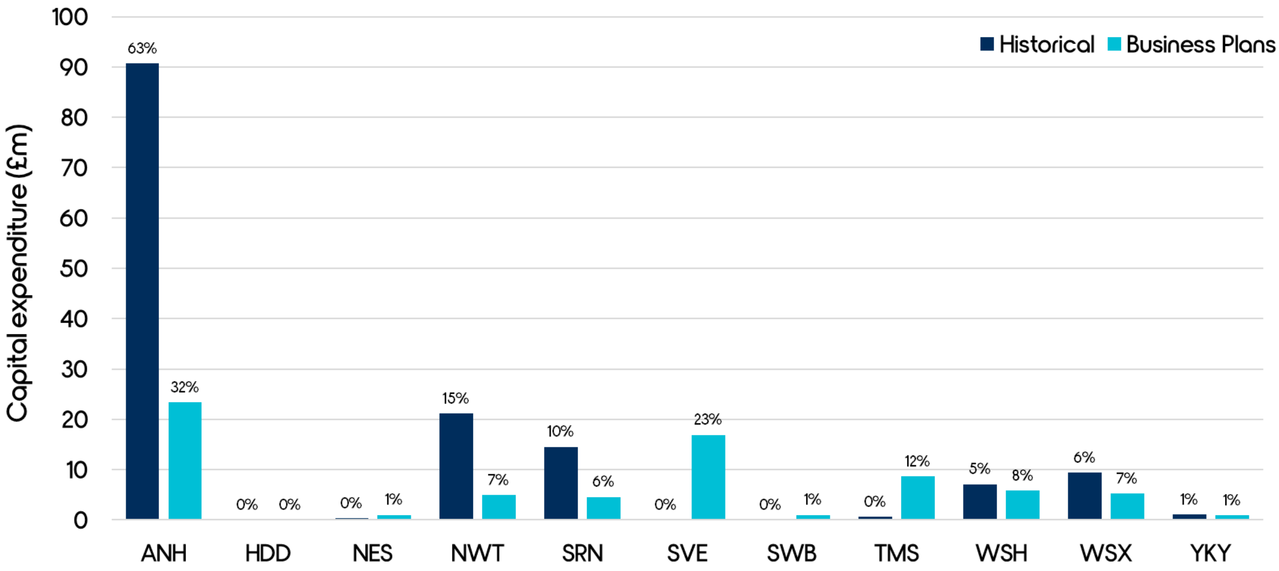
### 5.2.3 First time sewerage

Figure 26 Properties served by s101A schemes in the 'historical' and 'forecast' data across the industry




Note: Percentages denote each company's share of total industry s101A properties in each dataset
Source: Ofwat IAP first time sewerage model

Figure 27 Capital expenditure on connecting s101A scheme properties in the 'historical' and 'forecast' datasets




Note: Percentages denote each company's share of total industry capex in each dataset
Source: Ofwat IAP first time sewerage model

Figure 28 Model coefficients in IAP first time sewerage models

| | Historical | Forecast |
|---|---|---|
| s101A properties | 0.03 (0.00) | 0.04 (0.00) |
| s101A properties2 | -0.00002 (0.02) | -0.00004 (0.00) |
| Constant | 0.40 (0.28) | 0.44 (0.00) |
| Estimation technique | Pooled OLS | Pooled OLS |
| N | 38 | 47 |
| R2 | 0.83 | 0.90 |

Note: Pooled OLS technique is pooled (clustered standard errors) ordinary least squares
Cost driver and cost variable are smoothed (3yr average)
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP first time sewerage Pooled OLS models

Figure 29 Efficiency scores when dropping companies from IAP first time sewerage models

| Company | Original models | Models excluding company observations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVE | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.23 | N/A* | 1.23 | 1.26 | 1.21 | 1.25 | 1.24 | 1.22 | 1.24 | 1.23 | 1.16 | 1.23 |
| NES | 0.37 | 0.52 | 0.34 | 0.51 | 0.35 | 0.38 | 0.41 | 0.31 | 0.41 | 0.38 | 0.39 | 0.32 |
| NWT | 1.51 | 1.74 | 1.42 | 1.91 | 1.44 | 1.53 | 1.62 | 1.30 | 1.61 | 1.52 | 1.53 | 1.34 |
| SRN | 0.96 | 1.03 | 0.92 | 1.11 | 0.91 | 0.97 | 1.00 | 0.86 | 0.99 | 0.96 | 0.94 | 0.89 |
| SVE | 1.35 | 1.34 | 1.34 | 1.41 | 1.30 | 1.38 | 1.37 | 1.32 | 1.36 | 1.35 | 1.27 | 1.33 |
| SWB | 0.25 | 0.26 | 0.24 | 0.30 | 0.24 | 0.26 | 0.27 | 0.22 | 0.26 | 0.25 | 0.25 | 0.23 |
| TMS | 2.75 | 3.34 | 2.57 | 3.53 | 2.61 | 2.78 | 2.97 | 2.34 | 2.95 | 2.77 | 2.79 | 2.43 |
| WSH | 1.38 | 1.40 | 1.31 | 1.64 | 1.31 | 1.40 | 1.45 | 1.23 | 1.44 | 1.38 | 1.36 | 1.26 |
| WSX | 1.39 | 1.48 | 1.32 | 1.70 | 1.33 | 1.41 | 1.48 | 1.22 | 1.47 | 1.40 | 1.39 | 1.26 |
| YKY | 0.32 | 0.39 | 0.30 | 0.41 | 0.30 | 0.32 | 0.34 | 0.27 | 0.34 | 0.32 | 0.32 | 0.28 |
| Average change | | 0.14* | 0.05 | 0.23 | 0.05 | 0.02 | 0.06 | 0.12 | 0.06 | 0.01 | 0.03 | 0.09 |

Note: HDD is omitted from ‘forecast’ models due to 0 volume in all years, and is not in ‘historical’ dataset
Average change is the average absolute change in efficiency scores compared to the ‘Original models’
N/A indicates that Anglian Water receives a negative allowance under the specification in question; average change in efficiency score calculations exclude Anglian Water

Source: Vivid Economics

## 5.3 Waste service: quality

### 5.3.1 Chemical removal

Figure 30 Coefficients in Ofwat IAP chemical removal models

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Log PE | 0.24 (0.19) | 3.48 (0.15) | 0.21 (0.10) | 1.84 (0.21) |
| Log PE² | | -0.14 (0.17) | | -0.07 (0.26) |
| Constant | -0.61 (0.75) | -19.00 (0.17) | -0.34 (0.80) | -10.00 (0.25) |
| Estimation technique | OLS | OLS | OLS | OLS |
| N | 6 | 6 | 7 | 7 |
| $R^2$ | 0.39 | 0.70 | 0.46 | 0.62 |

Note: Models 1 and 2 exclude both SRN and SVE; Models 3 and 4 exclude SRN
P-values in parentheses
Source: Vivid Economics replication of Ofwat IAP chemical removal models

Figure 31. Efficiency scores in IAP chemical removal models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | SVE | SWB | TMS | WSX | YKY |
| ANH | 1.18 | 1.21 | 1.05 | 1.12 | 1.14 | 1.10 | 1.47 | 1.15 |
| NES | 0.79 | 0.81 | 0.70 | 0.75 | 0.65 | 0.75 | 1.00 | 0.77 |
| SRN | 4.52 | 4.65 | 4.04 | 4.32 | 4.36 | 4.25 | 5.66 | 4.42 |
| SVE | 1.17 | 1.21 | 1.21 | 1.31 | 0.82 | 1.00 | 1.05 | 1.10 |
| SWB | 0.63 | 0.66 | 0.62 | 0.66 | 0.08 | 0.65 | 0.84 | 0.65 |
| TMS | 0.76 | 0.79 | 0.73 | 0.79 | 0.91 | 0.68 | 0.82 | 0.76 |
| WSX | 2.16 | 2.23 | 1.95 | 2.08 | 1.32 | 2.10 | 2.81 | 2.12 |
| YKY | 0.94 | 0.97 | 0.93 | 1.00 | 1.07 | 0.83 | 0.96 | 0.93 |
| Average change | | 0.05 | 0.12 | 0.08 | 0.29 | 0.10 | 0.34 | 0.04 |

Note: SRN excluded from all IAP chemical removal models
Average change is the average absolute change in efficiency scores compared to the 'Original models'
Source: Vivid Economics

### 5.3.2 Event duration monitoring

Figure 32 Efficiency scores in IAP Event Duration Monitoring when dropping companies

| Company | Original model | Models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SWB | TMS | WSH | WSX | YKY | HDD |
| ANH | 1.81 | 1.93 | 1.70 | 1.93 | 1.93 | 1.70 | 1.70 | 1.70 | 1.93 | 1.93 | 1.70 |
| NES | 0.45 | 0.48 | 0.42 | 0.48 | 0.48 | 0.42 | 0.42 | 0.42 | 0.48 | 0.48 | 0.42 |
| NWT | 1.06 | 1.13 | 1.00 | 1.13 | 1.13 | 1.00 | 1.00 | 1.00 | 1.13 | 1.13 | 1.00 |
| SRN | 6.77 | 7.21 | 6.38 | 7.21 | 7.21 | 6.38 | 6.38 | 6.38 | 7.21 | 7.21 | 6.38 |
| SWB | 0.88 | 0.93 | 0.83 | 0.93 | 0.93 | 0.83 | 0.83 | 0.83 | 0.93 | 0.93 | 0.83 |
| TMS | 0.94 | 1.00 | 0.89 | 1.00 | 1.00 | 0.89 | 0.89 | 0.89 | 1.00 | 1.00 | 0.89 |
| WSH | 0.35 | 0.38 | 0.33 | 0.38 | 0.38 | 0.33 | 0.33 | 0.33 | 0.38 | 0.38 | 0.33 |
| WSX | 1.85 | 1.97 | 1.74 | 1.97 | 1.97 | 1.74 | 1.74 | 1.74 | 1.97 | 1.97 | 1.74 |
| YKY | 1.28 | 1.37 | 1.21 | 1.37 | 1.37 | 1.21 | 1.21 | 1.21 | 1.37 | 1.37 | 1.21 |
| HDD | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Average change | | 0.10 | 0.09 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 | 0.10 | 0.09 |

Note: SVE reports 0 volume and is excluded from the unit cost model
Average change is the average absolute change in efficiency scores compared to the 'Original model'
Source: Vivid Economics

### 5.3.3 Flow monitoring

Figure 33 Efficiency scores in IAP flow monitoring model when dropping companies

| Company | Original model | Models excluding company observations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | HDD | NES | NWT | SRN | SVE | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.00 | 0.91 | 0.91 | 1.00 | 0.91 | 1.00 | 0.91 | 0.91 | 1.00 | 0.91 | 1.00 | 1.00 |
| HDD | 0.17 | 0.16 | 0.16 | 0.17 | 0.16 | 0.17 | 0.16 | 0.16 | 0.17 | 0.16 | 0.17 | 0.17 |
| NES | 1.34 | 1.22 | 1.22 | 1.35 | 1.22 | 1.35 | 1.22 | 1.22 | 1.35 | 1.22 | 1.35 | 1.35 |
| NWT | 0.97 | 0.89 | 0.89 | 0.98 | 0.89 | 0.98 | 0.89 | 0.89 | 0.98 | 0.89 | 0.98 | 0.98 |
| SRN | 1.20 | 1.09 | 1.09 | 1.20 | 1.09 | 1.20 | 1.09 | 1.09 | 1.20 | 1.09 | 1.20 | 1.20 |
| SVE | 0.38 | 0.35 | 0.35 | 0.38 | 0.35 | 0.38 | 0.35 | 0.35 | 0.38 | 0.35 | 0.38 | 0.38 |
| SWB | 0.33 | 0.30 | 0.30 | 0.33 | 0.30 | 0.33 | 0.30 | 0.30 | 0.33 | 0.30 | 0.33 | 0.33 |
| TMS | 3.71 | 3.39 | 3.38 | 3.72 | 3.38 | 3.72 | 3.38 | 3.38 | 3.72 | 3.38 | 3.72 | 3.72 |
| WSH | 0.99 | 0.91 | 0.90 | 1.00 | 0.90 | 1.00 | 0.90 | 0.90 | 1.00 | 0.90 | 1.00 | 1.00 |
| WSX | 4.60 | 4.21 | 4.19 | 4.62 | 4.19 | 4.62 | 4.19 | 4.19 | 4.62 | 4.19 | 4.62 | 4.62 |
| YKY | 1.32 | 1.20 | 1.20 | 1.32 | 1.20 | 1.32 | 1.20 | 1.20 | 1.32 | 1.20 | 1.32 | 1.32 |
| Average change | | 0.13 | 0.13 | 0.01 | 0.13 | 0.01 | 0.13 | 0.13 | 0.01 | 0.13 | 0.01 | 0.01 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original model'
Source: Vivid Economics

### 5.3.4 Flow to full schemes

Figure 34 Coefficients in Ofwat IAP Flow to full schemes models

| | Linear schemes | Linear shortfall | Linear schemes & shortfall | Log schemes | Log shortfall | Log schemes & shortfall |
|---|---|---|---|---|---|---|
| Schemes | 1.98 (0.00) | | 0.49 (0.48) | | | |
| Shortfall | | 0.04 (0.00) | 0.03 (0.04) | | | |
| Log schemes | | | | 0.79 (0.00) | | 0.66 (0.05) |
| Log shortfall | | | | | 0.52 (0.02) | 0.14 (0.55) |
| Constant | 12.51 (0.25) | 14.19 (0.09) | 12.42 (0.16) | 1.57 (0.02) | 0.37 (0.77) | 1.03 (0.34) |
| Estimation technique | OLS | OLS | OLS | OLS | OLS | OLS |
| N | 10 | 10 | 10 | 10 | 10 | 10 |
| $R^2$ | 0.80 | 0.89 | 0.89 | 0.70 | 0.49 | 0.72 |

Note: P-values in parentheses
Source: Vivid Economics replication of Ofwat IAP Flow to full schemes models

Figure 35. Efficiency scores in Flow to full schemes models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVE | SWB | TMS | WSH | WSX | YKY |
| ANH | 0.88 | 0.83 | 0.88 | 0.89 | 0.95 | 0.85 | 0.87 | 0.89 | 0.87 | 0.90 | 0.88 |
| NES | 1.01 | 0.98 | 1.04 | 0.96 | 1.00 | 0.96 | 1.02 | 1.02 | 1.00 | 1.02 | 1.14 |
| NWT | 0.48 | 0.49 | 0.50 | 0.43 | 0.46 | 0.43 | 0.50 | 0.49 | 0.48 | 0.56 | 0.56 |
| SRN | 1.11 | 1.05 | 1.11 | 1.12 | 1.21 | 1.07 | 1.09 | 1.13 | 1.10 | 1.16 | 1.08 |
| SVE | 0.34 | 0.34 | 0.35 | 0.32 | 0.35 | 0.32 | 0.35 | 0.35 | 0.34 | 0.38 | 0.37 |
| SWB | 0.92 | 0.88 | 0.94 | 0.87 | 0.89 | 0.87 | 0.91 | 0.92 | 0.90 | 0.89 | 1.07 |
| TMS | 1.07 | 1.03 | 1.07 | 1.05 | 1.12 | 1.02 | 1.07 | 1.08 | 1.06 | 1.12 | 1.09 |
| WSH | 0.85 | 0.86 | 0.88 | 0.75 | 0.80 | 0.76 | 0.88 | 0.86 | 0.84 | 0.96 | 1.01 |
| WSX | 1.42 | 1.41 | 1.42 | 1.34 | 1.46 | 1.31 | 1.42 | 1.44 | 1.41 | 1.61 | 1.45 |
| YKY | 1.70 | 1.68 | 1.75 | 1.53 | 1.60 | 1.55 | 1.74 | 1.71 | 1.67 | 1.80 | 2.01 |
| Average change | | 0.03 | 0.02 | 0.06 | 0.05 | 0.06 | 0.01 | 0.01 | 0.01 | 0.07 | 0.09 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'Original models'
Source: Vivid Economics analysis of Ofwat IAP models

### 5.3.5 P removal

Figure 36 Coefficients in Ofwat IAP P removal linear models

| | Model 1 (capex) | Model 2 (totex) |
|---|---|---|
| PE | 0.07 (0.00) | 0.08 (0.00) |
| No. of sites | 1.29 (0.01) | 1.48 (0.00) |
| Constant | 43.08 (0.14) | 39.62 (0.19) |
| Estimation technique | OLS | OLS |
| N | 10 | 10 |
| $R^2$ | 0.93 | 0.93 |

Note: OLS estimation technique is Ordinary Least Squares
Expenditure and cost driver variables are linear in both specifications
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP P removal models

**Note on log-log model efficiency scores in IAP P removal model vs. other models**

In the P removal log-log models, efficiency scores are calculated as:

$$Company\ efficiency\ score\ = \frac{Log(capex)}{Log\ model\ allowance}$$

In other enhancement models and at PR14, efficiency scores for log-log models are calculated as:

$$Company\ efficiency\ score\ = \frac{Capex}{Exp(Log\ model\ allowance)}$$

The resulting efficiency scores are not the same in general, and the error has a material impact on the upper quartile efficiency scores from the two log-log P removal models

1. Model 2 Upper Quartile efficiency score: 0.98 (IAP approach) → 0.87 (correct approach)
2. Model 4 Upper Quartile efficiency score: 0.98 (IAP approach) → 0.87 (correct approach)

Figure 37. Efficiency scores in IAP P removal linear models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 0.98 | 0.95 | 0.98 | 0.98 | 1.06 | 0.95 | 0.99 | 0.97 | 0.99 | 0.98 | 0.98 |
| NES | 1.01 | 1.03 | 1.02 | 1.02 | 1.13 | 0.98 | 0.86 | 1.14 | 1.00 | 1.03 | 0.97 |
| NWT | 0.94 | 0.94 | 0.94 | 0.90 | 0.93 | 0.92 | 0.94 | 0.95 | 0.94 | 0.94 | 1.02 |
| SRN | 1.32 | 1.28 | 1.32 | 1.33 | 1.45 | 1.26 | 1.28 | 1.33 | 1.32 | 1.32 | 1.29 |
| SVT | 0.81 | 0.79 | 0.81 | 0.80 | 0.85 | 0.78 | 0.80 | 0.81 | 0.81 | 0.81 | 0.83 |
| SWB | 0.32 | 0.32 | 0.32 | 0.33 | 0.36 | 0.31 | 0.27 | 0.36 | 0.32 | 0.33 | 0.31 |
| TMS | 1.35 | 1.39 | 1.36 | 1.32 | 1.41 | 1.33 | 1.21 | 1.49 | 1.34 | 1.37 | 1.40 |
| WSH | 0.94 | 0.96 | 0.95 | 0.95 | 1.04 | 0.91 | 0.81 | 1.06 | 0.93 | 0.96 | 0.92 |
| WSX | 1.05 | 1.05 | 1.05 | 1.03 | 1.10 | 1.02 | 0.98 | 1.10 | 1.04 | 1.06 | 1.07 |
| YKY | 1.07 | 1.07 | 1.07 | 1.02 | 1.05 | 1.05 | 1.08 | 1.07 | 1.07 | 1.07 | 1.16 |
| Average change | | 0.02 | 0.00 | 0.02 | 0.06 | 0.03 | 0.06 | 0.05 | 0.01 | 0.01 | 0.04 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'Original models'
Source: Vivid Economics

Figure 38 P removal models are improved by consents variable: % STWs with P removal consents >1.1mg/l

| | IAP models | | | | Add consents variable | | | | Add consents variable, remove number of sites | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) | Model 1 (linear, capex) | Model 2 (log, capex) | Model 3 (linear, totex) | Model 4 (log, totex) |
| PE | 0.07 (0.00) | | 0.08 (0.00) | | 0.08 (0.00) | | 0.09 (0.00) | | 0.10 (0.00) | | 0.11 (0.00) | |
| Log PE | | 0.52 (0.00) | | 0.52 (0.00) | | 0.50 (0.00) | | 0.50 (0.00) | | 0.57 (0.00) | | 0.59 (0.00) |
| No. of sites | 1.29 (0.01) | | 1.48 (0.00) | | 1.04 (0.01) | | 1.24 (0.01) | | | | | |
| Log No. of sites | | 0.19 (0.34) | | 0.24 (0.25) | | 0.21 (0.13) | | 0.25 (0.09) | | | | |
| % STWs P consents >1.1 mg/l | | | | | -238 (0.08) | -1.74 (0.01) | 236 (0.10) | -1.72 (0.02) | -385 (0.05) | -1.70 (0.02) | 412 (0.07) | -1.67 (0.04) |
| Constant | 43.1 (0.14) | 0.98 (0.12) | 39.6 (0.19) | 0.87 (0.16) | 110.1 (0.03) | 1.54 (0.01) | 106.3 (0.04) | 1.42 (0.01) | 192.2 (0.01) | 1.85 (0.00) | 204 (0.01) | 1.80 (0.00) |
| Estimation technique | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS | OLS |
| N | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| $R^2$ | 0.93 | 0.91 | 0.93 | 0.92 | 0.96 | 0.97 | 0.96 | 0.97 | 0.88 | 0.95 | 0.86 | 0.95 |

Note: OLS estimation technique is Ordinary Least Squares
P-values in parentheses
Source: Vivid Economics

Figure 39 P removal models with consents have narrower efficiency score ranges and similar stability

| Company | 'New models' | Log w/consents models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 0.84 | 0.64 | 0.86 | 0.84 | 0.96 | 0.81 | 0.91 | 0.87 | 0.84 | 0.84 | 0.84 |
| NES | 1.23 | 1.19 | 1.36 | 1.23 | 1.29 | 1.24 | 1.11 | 1.14 | 1.21 | 1.23 | 1.18 |
| NWT | 1.00 | 1.00 | 0.97 | 1.00 | 1.00 | 0.96 | 1.01 | 0.97 | 1.00 | 1.00 | 1.12 |
| SRN | 1.26 | 1.05 | 1.32 | 1.26 | 1.39 | 1.22 | 1.24 | 1.28 | 1.25 | 1.26 | 1.25 |
| SVT | 0.85 | 0.78 | 0.84 | 0.85 | 0.88 | 0.81 | 0.80 | 0.88 | 0.85 | 0.85 | 0.90 |
| SWB | 0.90 | 0.96 | 0.95 | 0.90 | 0.91 | 0.87 | 0.60 | 0.98 | 0.89 | 0.90 | 0.90 |
| TMS | 0.87 | 0.90 | 0.93 | 0.87 | 0.86 | 0.89 | 0.98 | 0.68 | 0.86 | 0.87 | 0.89 |
| WSH | 0.96 | 0.95 | 1.04 | 0.96 | 0.98 | 0.96 | 0.89 | 0.85 | 0.94 | 0.96 | 0.94 |
| WSX | 1.00 | 1.03 | 1.02 | 1.00 | 1.00 | 0.98 | 0.94 | 0.95 | 1.00 | 1.00 | 1.07 |
| YKY | 1.21 | 1.21 | 1.16 | 1.21 | 1.19 | 1.15 | 1.21 | 1.17 | 1.21 | 1.21 | 1.36 |
| Average change | | 0.07 | 0.05 | 0.00 | 0.04 | 0.03 | 0.08 | 0.07 | 0.01 | 0.00 | 0.05 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'New models'
Source: Vivid Economics

Figure 40 Efficiency score range remains narrow when sites is removed from P-removal models including consents

| Company | 'New models' | Log w/consents and w/o sites models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVT | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.04 | 1.05 | 1.07 | 1.03 | 1.12 | 1.03 | 1.10 | 0.91 | 1.02 | 1.03 | 1.06 |
| NES | 1.23 | 1.24 | 1.36 | 1.24 | 1.31 | 1.24 | 1.18 | 1.11 | 1.19 | 1.23 | 1.20 |
| NWT | 0.94 | 0.95 | 0.91 | 0.92 | 0.97 | 0.92 | 0.94 | 0.95 | 0.95 | 0.93 | 1.00 |
| SRN | 1.46 | 1.47 | 1.52 | 1.45 | 1.56 | 1.45 | 1.47 | 1.32 | 1.43 | 1.45 | 1.47 |
| SVT | 0.90 | 0.91 | 0.90 | 0.89 | 0.93 | 0.88 | 0.89 | 0.90 | 0.90 | 0.89 | 0.95 |
| SWB | 0.95 | 0.94 | 1.01 | 0.95 | 0.94 | 0.94 | 0.80 | 1.02 | 0.93 | 0.94 | 0.95 |
| TMS | 0.70 | 0.71 | 0.75 | 0.70 | 0.76 | 0.70 | 0.72 | 0.60 | 0.68 | 0.70 | 0.70 |
| WSH | 0.90 | 0.91 | 0.98 | 0.90 | 0.96 | 0.90 | 0.87 | 0.81 | 0.87 | 0.90 | 0.89 |
| WSX | 0.93 | 0.94 | 0.94 | 0.92 | 0.96 | 0.92 | 0.90 | 0.92 | 0.93 | 0.92 | 0.96 |
| YKY | 1.13 | 1.13 | 1.08 | 1.10 | 1.15 | 1.09 | 1.12 | 1.15 | 1.14 | 1.11 | 1.20 |
| Average change | | 0.01 | 0.05 | 0.01 | 0.05 | 0.01 | 0.04 | 0.07 | 0.02 | 0.01 | 0.03 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'New models'
Source: Vivid Economics

### 5.3.6 Sanitary parameters

Figure 41 Coefficients in Ofwat IAP sanitary parameters models

| | 'Exponential' | 'Power' |
|---|---|---|
| PE/site | -0.06 (0.01) | -0.78 (0.00) |
| Constant | 0.42 (0.01) | 1.11 (0.82) |
| Estimation technique | OLS | OLS |
| N | 10 | 10 |
| $R^2$ | 0.64 | 0.71 |

Note: OLS estimation technique is Ordinary Least Squares
Cost variable is capex / PE in the 'Exponential' and 'Power' models
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP sanitary parameters models

Figure 42 Efficiency scores in IAP sanitary parameters models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVE | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.10 | 1.12 | 1.11 | 1.17 | 1.25 | 0.95 | 1.04 | 1.07 | 1.10 | 1.18 | 1.08 |
| NES | 1.15 | 1.16 | 1.17 | 1.03 | 1.20 | 1.05 | 1.08 | 1.29 | 1.17 | 1.20 | 1.12 |
| NWT | 0.72 | 0.71 | 0.75 | 0.56 | 0.70 | 0.70 | 0.68 | 0.91 | 0.72 | 0.74 | 0.71 |
| SRN | 1.87 | 1.91 | 1.89 | 1.90 | 2.09 | 1.64 | 1.76 | 1.88 | 1.91 | 1.99 | 1.82 |
| SVE | 0.40 | 0.41 | 0.40 | 0.39 | 0.44 | 0.35 | 0.37 | 0.41 | 0.41 | 0.42 | 0.39 |
| SWB | 0.58 | 0.59 | 0.59 | 0.55 | 0.63 | 0.52 | 0.54 | 0.62 | 0.60 | 0.61 | 0.56 |
| TMS | 1.59 | 1.59 | 1.65 | 1.31 | 1.59 | 1.53 | 1.51 | 1.95 | 1.60 | 1.65 | 1.57 |
| WSH | 1.11 | 1.13 | 1.11 | 1.23 | 1.29 | 0.95 | 1.06 | 1.04 | 1.08 | 1.19 | 1.09 |
| WSX | 1.61 | 1.64 | 1.63 | 1.56 | 1.75 | 1.43 | 1.51 | 1.67 | 1.66 | 1.70 | 1.56 |
| YKY | 0.80 | 0.82 | 0.82 | 0.77 | 0.87 | 0.72 | 0.75 | 0.85 | 0.83 | 0.85 | 0.78 |
| Average change | | 0.02 | 0.02 | 0.09 | 0.09 | 0.11 | 0.06 | 0.10 | 0.02 | 0.06 | 0.03 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'Original models'

Source: Vivid Economics

Figure 43 Coefficients in 'Unconstrained Log-log' sanitary parameters model

| | 'Exponential' | 'Power' | 'Unconstrained Log-log' |
|---|---|---|---|
| PE/site | -0.06 (0.01) | -0.78 (0.00) | |
| Log PE | | | 0.17 (0.35) |
| Log sites | | | 0.57 (0.04) |
| Constant | 0.42 (0.01) | 1.11 (0.82) | 0.86 (0.26) |
| Estimation technique | OLS | OLS | OLS |
| N | 10 | 10 | 10 |
| $R^2$ | 0.64 | 0.71 | 0.68 |

Note: OLS estimation technique is Ordinary Least Squares
Cost variable is Capex / PE in the 'Exponential' and 'Power' models, and log capex in the 'Unconstrained Log-log' model
P-values in parentheses

Source: Vivid Economics

Figure 44 Efficiency scores in 'Unconstrained Log-log' sanitary parameters model when dropping companies

| Company | 'New model' | Unconstrained log model excluding company observations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVE | SWB | TMS | WSH | WSX | YKY |
| ANH | 1.08 | 1.12 | 1.09 | 1.12 | 1.25 | 0.88 | 1.06 | 1.05 | 0.97 | 1.14 | 1.07 |
| NES | 0.94 | 0.94 | 0.90 | 0.90 | 0.97 | 1.18 | 0.84 | 1.02 | 0.93 | 1.03 | 0.84 |
| NWT | 0.72 | 0.72 | 0.71 | 0.57 | 0.72 | 0.60 | 0.68 | 0.97 | 0.78 | 0.75 | 0.71 |
| SRN | 1.94 | 1.98 | 1.94 | 1.94 | 2.17 | 1.72 | 1.85 | 1.97 | 1.79 | 2.06 | 1.88 |
| SVE | 0.67 | 0.68 | 0.68 | 0.59 | 0.74 | 0.42 | 0.67 | 0.77 | 0.65 | 0.68 | 0.70 |
| SWB | 0.61 | 0.61 | 0.59 | 0.57 | 0.64 | 0.62 | 0.56 | 0.65 | 0.59 | 0.65 | 0.57 |
| TMS | 1.75 | 1.74 | 1.73 | 1.42 | 1.77 | 1.49 | 1.65 | 2.26 | 1.85 | 1.82 | 1.71 |
| WSH | 0.80 | 0.83 | 0.80 | 0.90 | 0.94 | 0.75 | 0.76 | 0.70 | 0.68 | 0.86 | 0.76 |
| WSX | 1.70 | 1.72 | 1.68 | 1.64 | 1.84 | 1.65 | 1.59 | 1.81 | 1.62 | 1.82 | 1.61 |
| YKY | 0.73 | 0.73 | 0.71 | 0.72 | 0.78 | 0.82 | 0.66 | 0.76 | 0.70 | 0.79 | 0.66 |
| Average change | | 0.02 | 0.02 | 0.08 | 0.09 | 0.15 | 0.06 | 0.13 | 0.07 | 0.07 | 0.05 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'New model'

Source: Vivid Economics

### 5.3.7 Spill frequency

Figure 45 Coefficients in Ofwat IAP spill frequency model

| | 'Power' model |
|---|---|
| Volume of new or additional storage | 0.76 (0.00) |
| Constant | 0.02 (0.00) |
| Estimation technique | OLS |
| N | 9 |
| $R^2$ | 0.96 |

Note: OLS estimation technique is Ordinary Least Squares
'Power' model is equivalent to log capex on log volume of new or additional storage – coefficients from this model are shown
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP spill frequency model

Figure 46 Efficiency scores in IAP spill frequency model when dropping companies

| Company | Original model | Models excluding company observations | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SVE | SWB | WSH | WSX | YKY |
| ANH | 1.23 | 1.28 | 1.17 | 1.24 | 0.97 | 1.21 | 1.22 | 1.35 | 1.25 | 1.23 |
| NES | 0.65 | 0.67 | 0.62 | 0.66 | 0.61 | 0.63 | 0.63 | 0.69 | 0.66 | 0.65 |
| NWT | 1.11 | 1.13 | 1.05 | 1.15 | 1.26 | 1.07 | 1.07 | 1.16 | 1.14 | 1.12 |
| SRN | 0.79 | 0.85 | 0.77 | 0.78 | 0.41 | 0.80 | 0.81 | 0.92 | 0.80 | 0.79 |
| SVE | 0.83 | 0.84 | 0.78 | 0.85 | 0.86 | 0.80 | 0.80 | 0.87 | 0.85 | 0.83 |
| SWB | 0.85 | 0.86 | 0.80 | 0.87 | 0.91 | 0.82 | 0.82 | 0.88 | 0.86 | 0.85 |
| WSH | 1.71 | 1.77 | 1.62 | 1.72 | 1.37 | 1.68 | 1.68 | 1.86 | 1.74 | 1.71 |
| WSX | 1.16 | 1.19 | 1.10 | 1.18 | 1.09 | 1.13 | 1.13 | 1.24 | 1.18 | 1.16 |
| YKY | 1.02 | 1.03 | 0.96 | 1.05 | 1.15 | 0.98 | 0.98 | 1.05 | 1.04 | 1.02 |
| Average change | | 0.03 | 0.05 | 0.02 | 0.16 | 0.03 | 0.03 | 0.08 | 0.02 | 0.00 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original model'

Source: Vivid Economics

### 5.3.8 Storm tanks

Figure 47 Coefficients in Ofwat IAP storm tanks models

| | Model 1 | Model 2 |
|---|---|---|
| Storage volume | 0.89 (0.00) | 0.78 (0.00) |
| No. of schemes | | 0.23 (0.26) |
| Constant | -5.30 (0.00) | -5.11 (0.00) |
| Estimation technique | OLS | OLS |
| N | 11 | 11 |
| R2 | 0.96 | 0.97 |

Note: OLS estimation technique is Ordinary Least Squares
Both cost and cost driver variables are logged in both models
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP storm tanks models

Figure 48 Efficiency scores in IAP storm tanks models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | ANH | NES | NWT | SRN | SWB | TMS | WSH | WSX | YKY | SVE | HDD |
| ANH | 1.32 | 1.48 | 1.33 | 1.29 | 1.43 | 1.42 | 1.25 | 1.31 | 1.33 | 1.32 | 1.16 | 1.21 |
| NES | 1.27 | 1.28 | 1.38 | 1.26 | 1.25 | 1.33 | 1.27 | 1.22 | 1.28 | 1.27 | 1.12 | 1.50 |
| NWT | 0.99 | 0.99 | 0.97 | 1.02 | 1.13 | 0.97 | 0.88 | 0.92 | 1.03 | 0.99 | 1.01 | 0.99 |
| SRN | 1.73 | 1.82 | 1.71 | 1.75 | 1.93 | 1.76 | 1.58 | 1.65 | 1.78 | 1.73 | 1.65 | 1.66 |
| SWB | 1.27 | 1.39 | 1.33 | 1.24 | 1.31 | 1.36 | 1.24 | 1.25 | 1.27 | 1.27 | 1.09 | 1.28 |
| TMS | 0.61 | 0.64 | 0.61 | 0.62 | 0.68 | 0.62 | 0.56 | 0.58 | 0.63 | 0.61 | 0.58 | 0.60 |
| WSH | 0.65 | 0.65 | 0.67 | 0.65 | 0.69 | 0.65 | 0.61 | 0.61 | 0.66 | 0.65 | 0.61 | 0.69 |
| WSX | 1.21 | 1.24 | 1.22 | 1.23 | 1.32 | 1.22 | 1.12 | 1.14 | 1.25 | 1.21 | 1.16 | 1.25 |
| YKY | 1.02 | 1.10 | 1.03 | 1.01 | 1.11 | 1.07 | 0.96 | 0.99 | 1.04 | 1.02 | 0.93 | 0.98 |
| SVE | 0.47 | 0.50 | 0.50 | 0.47 | 0.49 | 0.50 | 0.46 | 0.46 | 0.48 | 0.47 | 0.42 | 0.50 |
| HDD | 1.17 | 1.10 | 1.32 | 1.18 | 1.10 | 1.18 | 1.20 | 1.09 | 1.18 | 1.17 | 1.05 | 1.58 |
| Average change | | 0.06 | 0.04 | 0.02 | 0.08 | 0.04 | 0.06 | 0.04 | 0.02 | 0.00 | 0.09 | 0.09 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original models'

Source: Vivid Economics

## 5.4 Water service: quality

### 5.4.1 Meeting lead standards

Figure 49 Coefficients in Ofwat IAP meeting lead standards models

| | Model 1 (Historical) | Model 2 (Forecast) |
|---|---|---|
| Communication pipes | 2.40*10^-6 (0.02) | 3.22*10^-6 (0.07) |
| Communication pipes replaced | 3.44*10^-4 (0.00) | 4.5*10^-4 (0.00) |
| Constant | 0.09 (0.68) | 0.19 (0.57) |
| Estimation technique | RE | RE |
| N | 64 | 71 |
| R2 | 0.84 | 0.78 |

Note: RE estimation technique is random effects
Cost drivers and cost variable are logged and smoothed (3yr average) in all specifications
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP meeting lead standards models

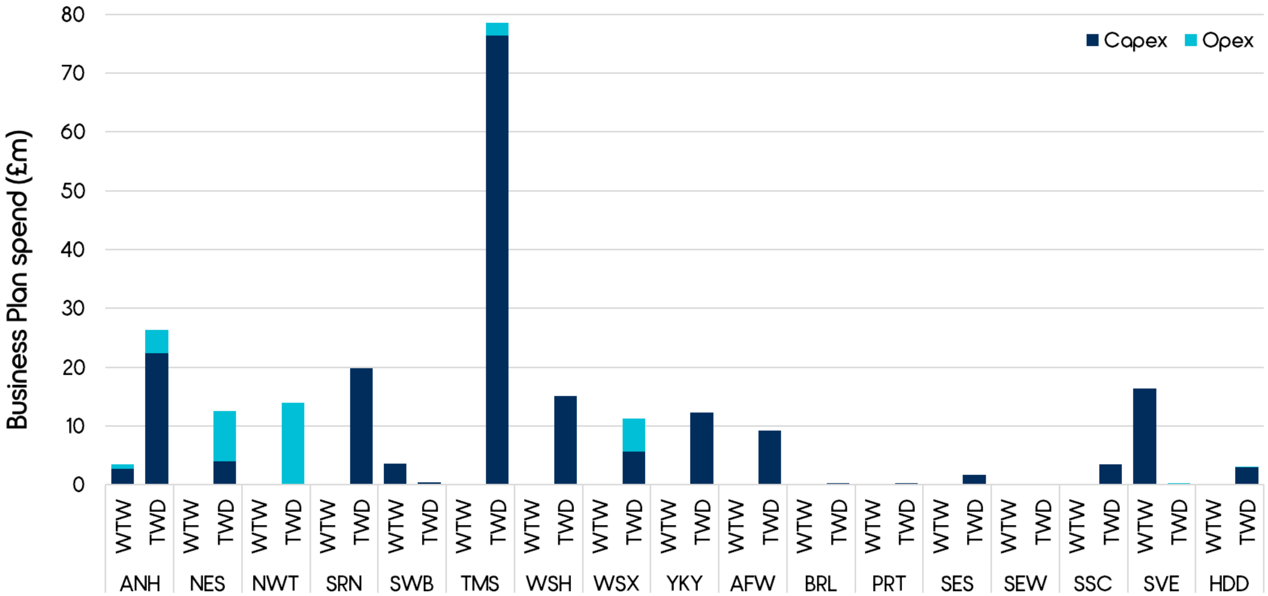
Figure 50 Efficiency scores in Ofwat IAP lead standards models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AFW | ANH | BRL | HDD | NES | NWT | PRT | SES | SEW | SRN | SSC | SVE | SWB | TMS | WSH | WSX | YKY |
| AFW | 1.09 | 1.08 | 1.14 | 1.08 | 1.10 | 1.06 | 1.02 | 1.08 | 1.08 | 1.09 | 1.12 | 1.09 | 1.11 | 1.09 | 1.23 | 1.09 | 1.09 | 1.04 |
| ANH | 2.51 | 2.58 | 2.58 | 2.49 | 2.52 | 2.44 | 2.37 | 2.50 | 2.50 | 2.51 | 2.49 | 2.50 | 2.57 | 2.51 | 3.08 | 2.51 | 2.51 | 2.12 |
| BRL | 0.11 | 0.12 | 0.11 | 0.10 | 0.11 | 0.10 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.12 | 0.11 | 0.11 | 0.12 |
| HDD | 2.65 | 2.77 | 2.47 | 2.26 | 2.82 | 2.07 | 2.84 | 2.36 | 2.40 | 2.65 | 2.43 | 2.53 | 2.71 | 2.64 | 2.03 | 2.65 | 2.61 | 19.19 |
| NES | 0.34 | 0.34 | 0.36 | 0.34 | 0.34 | 0.33 | 0.32 | 0.34 | 0.34 | 0.34 | 0.35 | 0.34 | 0.35 | 0.34 | 0.41 | 0.34 | 0.34 | 0.30 |
| NWT | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| PRT | 0.14 | 0.15 | 0.14 | 0.13 | 0.15 | 0.12 | 0.14 | 0.13 | 0.13 | 0.14 | 0.13 | 0.14 | 0.14 | 0.14 | 0.13 | 0.14 | 0.14 | 0.22 |
| SES | 0.65 | 0.67 | 0.65 | 0.61 | 0.67 | 0.59 | 0.64 | 0.63 | 0.63 | 0.65 | 0.64 | 0.64 | 0.67 | 0.65 | 0.66 | 0.65 | 0.65 | 0.82 |
| SEW | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SRN | 1.47 | 1.35 | 1.59 | 1.45 | 1.48 | 1.44 | 1.34 | 1.46 | 1.46 | 1.47 | 1.60 | 1.47 | 1.47 | 1.47 | 1.52 | 1.47 | 1.47 | 1.61 |
| SSC | 1.03 | 1.06 | 1.03 | 0.98 | 1.05 | 0.94 | 1.00 | 1.00 | 1.00 | 1.03 | 1.00 | 1.01 | 1.05 | 1.03 | 1.10 | 1.03 | 1.02 | 1.11 |
| SVE | 1.48 | 1.53 | 1.51 | 1.47 | 1.48 | 1.44 | 1.40 | 1.47 | 1.47 | 1.48 | 1.45 | 1.47 | 1.51 | 1.47 | 1.87 | 1.48 | 1.47 | 1.20 |
| SWB | 1.44 | 1.43 | 1.45 | 1.35 | 1.47 | 1.29 | 1.40 | 1.37 | 1.38 | 1.44 | 1.43 | 1.41 | 1.46 | 1.44 | 1.40 | 1.44 | 1.43 | 1.90 |
| TMS | 2.03 | 1.95 | 2.16 | 2.03 | 2.03 | 2.01 | 1.86 | 2.03 | 2.03 | 2.03 | 2.13 | 2.03 | 2.05 | 2.03 | 2.34 | 2.03 | 2.03 | 1.83 |
| WSH | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 | 1.46 |
| WSX | 1.15 | 1.08 | 1.22 | 1.11 | 1.17 | 1.08 | 1.07 | 1.12 | 1.13 | 1.15 | 1.23 | 1.14 | 1.16 | 1.16 | 1.12 | 1.15 | 1.15 | 1.49 |
| YKY | 0.57 | 0.60 | 0.59 | 0.58 | 0.57 | 0.57 | 0.54 | 0.58 | 0.58 | 0.57 | 0.57 | 0.57 | 0.59 | 0.57 | 0.75 | 0.57 | 0.57 | 0.44 |
| Average change | | 0.04 | 0.04 | 0.04 | 0.02 | 0.07 | 0.06 | 0.03 | 0.03 | 0.00 | 0.04 | 0.01 | 0.02 | 0.00 | 0.15 | 0.00 | 0.00 | 1.11 |

Notes: Average change is the average absolute change in efficiency scores compared to the 'Original models'
WSH allowance is set through the 'deep dive' process and efficiency score therefore does not vary when dropping companies

Source: Vivid Economics

Figure 51 Meeting lead standards expenditure by subservice area and cost type across the industry




Note: WTW is the Water Treatment subservice; TWD is the treated water distribution subservice
Source: Vivid Economics analysis of PR19 Business Plan data tables

Figure 52 Totex treated water distribution lead standards model coefficients

| | IAP Historical | IAP Forecast | Forecast with totex | Forecast with TWD totex | Forecast with TWD totex and w/o pipe stock |
|---|---|---|---|---|---|
| Communication pipes | 2.40*10^-6 (0.02) | 3.22*10^-6 (0.07) | 3.31*10^-6 (0.09) | 2.86*10^-6 (0.10) | |
| Communication pipes replaced | 3.44*10^-4 (0.00) | 4.5*10^-4 (0.00) | 5.9*10^-4 (0.00) | 6.0*10^-4 (0.00) | 6.9*10^-4 (0.00) |
| Constant | 0.09 (0.68) | 0.19 (0.57) | 0.28 (0.39) | 0.07 (0.81) | 0.95 (0.02) |
| Estimation technique | RE | RE | RE | RE | RE |
| N | 64 | 71 | 71 | 71 | 75 |
| R2 | 0.84 | 0.78 | 0.82 | 0.80 | 0.81 |

Note: RE estimation technique is random effects
Cost drivers and cost variable are logged and smoothed (3yr average) in all specifications
Totex models have capex + opex in the cost line; TWD totex is treated water distribution totex only
P-values in parentheses
Source: Vivid Economics

## 5.5 Supply-Demand Balance

### 5.5.1 Metering

Figure 53 Coefficients in Ofwat IAP metering models

| | Model 1 (levels) | Model 2 (log-log) |
|---|---|---|
| Meters Installed | 0.23 (0.00) | |
| Log Meters Installed | | 0.88 (0.00) |
| Constant | 1.80 (0.69) | -0.80 (0.00) |
| Estimation technique | OLS | OLS |
| N | 15 | 15 |
| R2 | 0.88 | 0.96 |

Note: OLS estimation technique is Ordinary Least Squares
P-values in parentheses

Source: Vivid Economics replication of Ofwat IAP metering models

Figure 54 Efficiency scores in IAP metering models when dropping companies

| Company | Original models | Models excluding company observations | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | AFW | ANH | BRL | DVW | NES | NWT | PRT | SES | SEW | SSC | SVT | SWB | WSH | WSX | YKY |
| ANH | 1.10 | 1.11 | 1.10 | 1.10 | 1.13 | 1.13 | 1.10 | 1.09 | 1.09 | 1.10 | 1.09 | 1.09 | 1.11 | 1.07 | 1.10 | 1.09 |
| NES | 0.89 | 0.89 | 0.86 | 0.89 | 0.88 | 0.92 | 0.88 | 0.89 | 0.89 | 0.94 | 0.89 | 0.87 | 0.89 | 0.87 | 0.89 | 0.87 |
| NWT | 1.14 | 1.15 | 1.11 | 1.14 | 1.13 | 1.17 | 1.13 | 1.14 | 1.13 | 1.20 | 1.13 | 1.11 | 1.15 | 1.12 | 1.14 | 1.12 |
| SRN | 1.24 | 1.25 | 1.27 | 1.24 | 1.30 | 1.28 | 1.24 | 1.22 | 1.23 | 1.18 | 1.22 | 1.24 | 1.27 | 1.20 | 1.25 | 1.23 |
| SWB | 0.43 | 0.43 | 0.43 | 0.43 | 0.44 | 0.44 | 0.43 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.43 | 0.42 | 0.43 | 0.42 |
| TMS | 3.69 | 3.73 | 3.54 | 3.69 | 3.56 | 3.84 | 3.65 | 3.72 | 3.70 | 4.05 | 3.70 | 3.58 | 3.71 | 3.65 | 3.69 | 3.62 |
| WSH | 0.83 | 0.83 | 0.83 | 0.83 | 0.85 | 0.85 | 0.83 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.84 | 0.81 | 0.83 | 0.82 |
| WSX | 0.99 | 0.99 | 1.01 | 0.99 | 1.03 | 1.02 | 0.99 | 0.97 | 0.98 | 0.94 | 0.98 | 0.99 | 1.01 | 0.96 | 1.00 | 0.98 |
| YKY | 0.64 | 0.65 | 0.63 | 0.64 | 0.64 | 0.66 | 0.64 | 0.64 | 0.64 | 0.67 | 0.64 | 0.63 | 0.65 | 0.63 | 0.64 | 0.63 |
| AFW | 1.27 | 1.28 | 1.23 | 1.27 | 1.25 | 1.32 | 1.26 | 1.27 | 1.27 | 1.36 | 1.27 | 1.24 | 1.28 | 1.25 | 1.27 | 1.25 |
| BRL | 0.74 | 0.75 | 0.76 | 0.74 | 0.77 | 0.76 | 0.74 | 0.73 | 0.73 | 0.71 | 0.73 | 0.74 | 0.76 | 0.72 | 0.75 | 0.74 |
| PRT | 0.69 | 0.69 | 0.72 | 0.69 | 0.72 | 0.71 | 0.69 | 0.67 | 0.68 | 0.62 | 0.67 | 0.69 | 0.70 | 0.66 | 0.69 | 0.68 |
| SES | 0.96 | 0.96 | 0.95 | 0.95 | 0.97 | 0.98 | 0.95 | 0.95 | 0.95 | 0.97 | 0.95 | 0.94 | 0.97 | 0.93 | 0.96 | 0.95 |
| SEW | 0.54 | 0.55 | 0.90 | 0.54 | 0.57 | 0.65 | 0.59 | 0.50 | 0.52 | 0.29 | 0.51 | 0.61 | 0.63 | 0.44 | 0.57 | 0.55 |
| SSC | 0.97 | 0.98 | 0.98 | 0.97 | 1.01 | 1.00 | 0.97 | 0.96 | 0.96 | 0.94 | 0.96 | 0.97 | 0.99 | 0.94 | 0.98 | 0.96 |
| SVE | 0.89 | 0.90 | 0.86 | 0.89 | 0.87 | 0.93 | 0.88 | 0.90 | 0.89 | 0.97 | 0.89 | 0.87 | 0.90 | 0.88 | 0.90 | 0.88 |
| HDD | 0.41 | 0.42 | 0.54 | 0.41 | 0.45 | 0.46 | 0.43 | 0.39 | 0.40 | 0.27 | 0.40 | 0.44 | 0.45 | 0.36 | 0.43 | 0.42 |
| Average change | | 0.01 | 0.05 | 0.00 | 0.03 | 0.04 | 0.01 | 0.01 | 0.01 | 0.08 | 0.01 | 0.02 | 0.02 | 0.03 | 0.01 | 0.01 |

Note: Average change is the average absolute change in efficiency scores compared to the 'Original models'
Efficiency scores shown are 'forward-looking', that is, requested capex post-reallocations relative to triangulated model allowances
Thames and Southern Water are excluded from the IAP metering model; HDD and SVE are not present in the 'historical' dataset

Source: Vivid Economics

Figure 55 Coefficients in metering 'historical' and 'combined' dataset models with % meter penetration rate

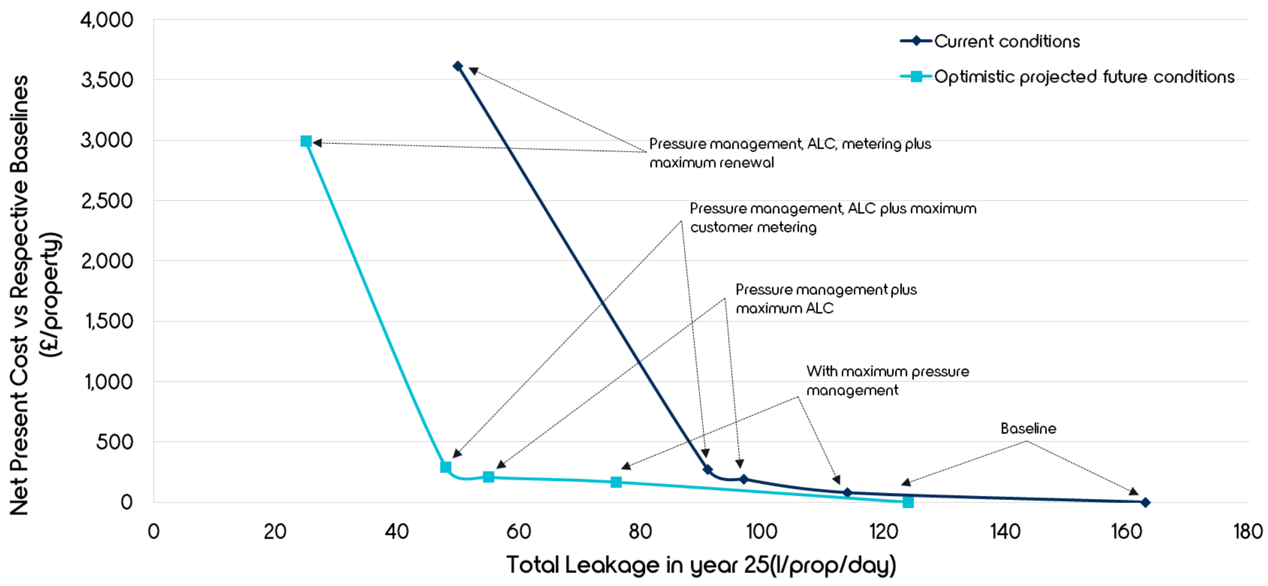
| | 'Historical' dataset | | | | 'Combined' dataset | |
|---|---|---|---|---|---|---|
| | IAP Levels | IAP Log-log | Levels with % meter penetration rate | Log-log with % meter penetration rate | Levels with % meter penetration rate | Log-log with % meter penetration rate |
| Meters installed | 0.23 (0.00) | | 0.23 (0.00) | | 0.24 (0.00) | |
| Log meters installed | | 0.88 (0.00) | | 0.87 (0.00) | | 0.89 (0.00) |
| % meter penetration rate | | | 8.01 (0.71) | 0.46 (0.23) | 7.65 (0.77) | 0.17 (0.62) |
| Constant | 1.80 (0.69) | -0.80 (0.00) | -2.03 (0.86) | 0.99 (0.00) | -3.21 (0.84) | -0.88 (0.01) |
| Estimation technique | OLS | OLS | OLS | OLS | OLS | OLS |
| N | 15 | 15 | 15 | 15 | 14 | 14 |
| $R^2$ | 0.88 | 0.96 | 0.88 | 0.97 | 0.94 | 0.98 |

Note: OLS estimation technique is Ordinary Least Squares
'Historical' dataset is the summation of volume and spend over 2011/12 – 17/18; 'combined' dataset is the same over 2011/12 – 24/25
% meter penetration is share of properties served which are metered
P-values in parentheses

Source: Vivid Economics

### 5.5.2 Leakage SDB

Figure 56 Leakage unit costs increases as absolute levels of leakage fall




Source: Reproduced from UK Water Industry Research, 2009

Figure 57 Industry median leakage unit cost varies from £1.49 – £1.65m/Ml/d when excluding a single company

| Model | AFW | ANH | BRL | ESK | NES | PRT | SES | SEW | SRN | SSC | SVE | SWB | TMS | NWT | WSH | WSX | YKY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Original model | 1.60 | | | | | | | | | | | | | | | | |
| Drop company | 1.64 | 1.49 | 1.65 | 1.61 | 1.61 | 1.65 | 1.49 | 1.49 | 1.55 | 1.65 | 1.65 | 1.64 | 1.49 | 1.65 | 1.49 | 1.49 | 1.50 |

Note: HDD not included in median unit cost calculations as in IAP model
Source: Vivid Economics

### 5.5.3 2020 – 25 SDB

Figure 58 Industry median 2020-25 SDB unit cost varies from £1.33 – £1.45m/Ml/d when excluding a single company

| Model | AFW | ANH | PRT | SES | SEW | SRN | SSC | SVE | TMS | WSH | WSX | YKY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Original model | 1.39 | | | | | | | | | | | |
| Drop company | 1.33 | 1.33 | 1.45 | 1.45 | 1.45 | 1.33 | 1.45 | 1.33 | 1.33 | 1.33 | 1.45 | 1.45 |

Note: BRL, HDD, ESK, NES, SWB and NWT not included in median unit cost calculations as in IAP model
Source: Vivid Economics

Click or tap here to enter text.

# Company profile

Vivid Economics is a leading strategic economics consultancy with global reach. We strive to create lasting value for our clients, both in government and the private sector, and for society at large.

We are a premier consultant in the policy-commerce interface and resource- and environment-intensive sectors, where we advise on the most critical and complex policy and commercial questions facing clients around the world. The success we bring to our clients reflects a strong partnership culture, solid foundation of skills and analytical assets, and close cooperation with a large network of contacts across key organisations.

# Contact us

Vivid Economics Limited
163 Eversholt Street
London NW1 1BU
United Kingdom

T: +44 (0)844 8000 254
enquiries@vivideconomics.com