Reckon LLP

Data analysis | Economic regulation | Competition law

Econometric models for residential retail cost assessment

13 February 2018

Reckon LLP, limited liability partnership registered in England (number OC307897)
31 Southampton Row, London, WC1B 5HJ • www.reckon.co.uk

Table of contents

Section 1: Introduction and summary

Introduction

1. Reckon LLP has been working with United Utilities to develop econometric models that could be used for Ofwat's residential retail cost benchmarking analysis.

2. Our work is intended as a positive contribution to the development of econometric models for Ofwat's upcoming PR19 review of water companies' price limits. We believe that it provides considerable benefits, in terms of the suggested model specifications and the accompanying dataset we have developed on company-level deprivation measures (drawing on data from Equifax). Further model development and refinement may be possible in the future, building on this report.

3. The main body of this report provides an overview of the key elements of water companies' residential retail costs, explains our approach to the development of econometric models for the benchmarking of these costs, and then presents estimation results and other analysis for a range of models. Appendix 1 provides information on the data sources used, Appendix 2 gives an overview of our earlier work to develop measures of deprivation and of arrears risk, and Appendix 3 provides some more technical guidance on the interpretation of model estimation results.

4. This introductory section presents background to the work and a summary of the main outputs and findings. It takes the following points in turn:

   (a) Ofwat's approach to residential retail cost benchmarking.

   (b) Development of econometric models for benchmarking residential retail costs.

   (c) Overview of findings on cross-cutting aspects of model development.

   (d) Overview of findings on the set of cost drivers to incorporate in models.

Ofwat's approach to residential retail cost benchmarking

5. In December 2014, Ofwat published its final determinations to its PR14 periodic review of price controls for water companies in England and Wales, setting new controls covering the period 1 April 2015 to 31 March 2020. For the first time, Ofwat set separate controls for companies' residential retail activities.

6. One of the main ingredients to the calculation of each company's control for residential retail activities was Ofwat's allowance for the "average cost to serve" households with water and wastewater retail services (ACTS). This allowance represented a benchmark level of retail costs.

7. Ofwat's approach for PR14 made some allowances for differences between companies that could affect their retail expenditure requirements. This was partly through the calculation of adjustments to the ACTS benchmarks (e.g. taking account of the higher retail costs for metered customers) and partly through company-specific special cost factor adjustments applied separately (e.g. allowances for the higher bad debt costs from serving customers in a part of the country with higher levels of

economic deprivation). However, Ofwat's calculation of the ACTS benchmarks used averages of relatively simple measures of historical unit cost taken across the companies in the industry.

8. There is an opportunity to make a substantial improvement to Ofwat's retail cost assessment in the future, through the use of econometric modelling approaches, rather than the unit cost measures and industry averages. Ofwat used econometric modelling approaches for PR14 for its assessment of companies' wholesale costs. If anything, retail costs should be more amenable to econometric modelling than wholesale costs, due to the greater comparability in companies' retail activities than their wholesale activities. Drawing on an econometric approach would also align well with Ofwat's stated desire to have a symmetric approach to cost adjustments.[1] Nonetheless, realisation of the opportunities for improvement will depend on the details of implementation. There are risks that, if econometric modelling approaches and model specifications are not well-thought out, the outcomes could be worse than a simpler non-econometric approach.

9. In December 2017, Ofwat published its final methodology for its PR19 review, which will set new price controls from April 2020. Ofwat has confirmed that the retail controls will be set for a five-year period, mirroring the duration of the wholesale control.

10. Ofwat said that it intends to use econometric models to set efficient cost allowances for the residential retail controls.[2] It explained that econometric models have a number of advantages over a unit cost approach, including that the model can simultaneously account for multiple factors that drive differences in costs across companies. Ofwat highlighted factors such as whether customers are billed for a single or dual service and differences in the cost to serve a metered customer versus an unmetered customer, and plans to consider differences between companies in bill sizes and deprivation levels when assessing the levels of bad debt costs.[3]

11. Ofwat also said that, if its econometric models are not sufficiently robust, it will use a non-econometric approach, based on measures of efficient cost to serve (ECTS) derived from a modification to the ATCS approach from PR14. [4]

Development of econometric models for benchmarking residential retail costs

12. In earlier phases of our work for United Utilities, our efforts were directed mainly at identifying datasets relating to economic deprivation, and at investigating their potential use for econometric modelling purposes. United Utilities published our working paper covering that initial work in May 2017.[5] That paper explored the use of variables relating to economic deprivation and arrears risk in econometric

[1] Ofwat (2017) Delivering Water 2020: Our final methodology for the 2019 price review, p149.
[2] Ofwat (2017) Delivering Water 2020: Our final methodology for the 2019 price review, p135
[3] Ofwat (2017) Delivering Water 2020: Our final methodology for the 2019 price review Appendix 11: Securing cost efficiency, p21
[4] Ofwat (2017) Delivering Water 2020: Our final methodology for the 2019 price review Appendix 11: Securing cost efficiency, p21
[5] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities. Available from https://www.unitedutilities.com/globalassets/z_corporate-site/about-us-pdfs/looking-to-the-future/depriviation-and-arrears-risk-in-hh-retail-cost-assessment-100517.pdf

modelling of the bad debt cost element of water companies' residential retail costs. The paper demonstrated grounds for using such variables as part of Ofwat's PR19 retail cost assessment. We presented this work to Ofwat and other water companies at a cost assessment working group meeting in September 2017.

13. Subsequently, we carried out a more comprehensive phase of work to develop and test potential econometric models for use in the benchmarking of water companies' residential retail costs, considering the treatment of all key cost drivers, including cost drivers relating to factors other than deprivation. This report covers this subsequent phase of work.

14. Our approach to model development has been an iterative one, involving judgement, in which we started with initial ideas on how to specify models of retail costs, drawing on a review of candidate cost drivers, and then extended and refined these models in the light of analysis of model estimation results and other indicators. Our objective has been, firstly, the development of econometric models of water companies' (efficient) expenditure requirements on residential retail activities, which could be used instead of Ofwat's non-econometric ACTS approach from PR14, and then the subsequent extension and refinement of such models to improve their likely accuracy. Our overarching aim has been to make progress in the econometric modelling of residential retail costs: to end up in a better place than where we started.

15. The work we have carried out, to develop and test econometric models of companies' residential retail costs, shows that it is possible for Ofwat to use econometric models for PR19, rather than a unit cost approach. An econometric approach allows the benchmarking analysis to take account of the multiple drivers of residential retail costs in a systematic and evidence-based way. The key question is not whether such models are perfect, but how they compare against alternative feasible approaches to cost assessment.

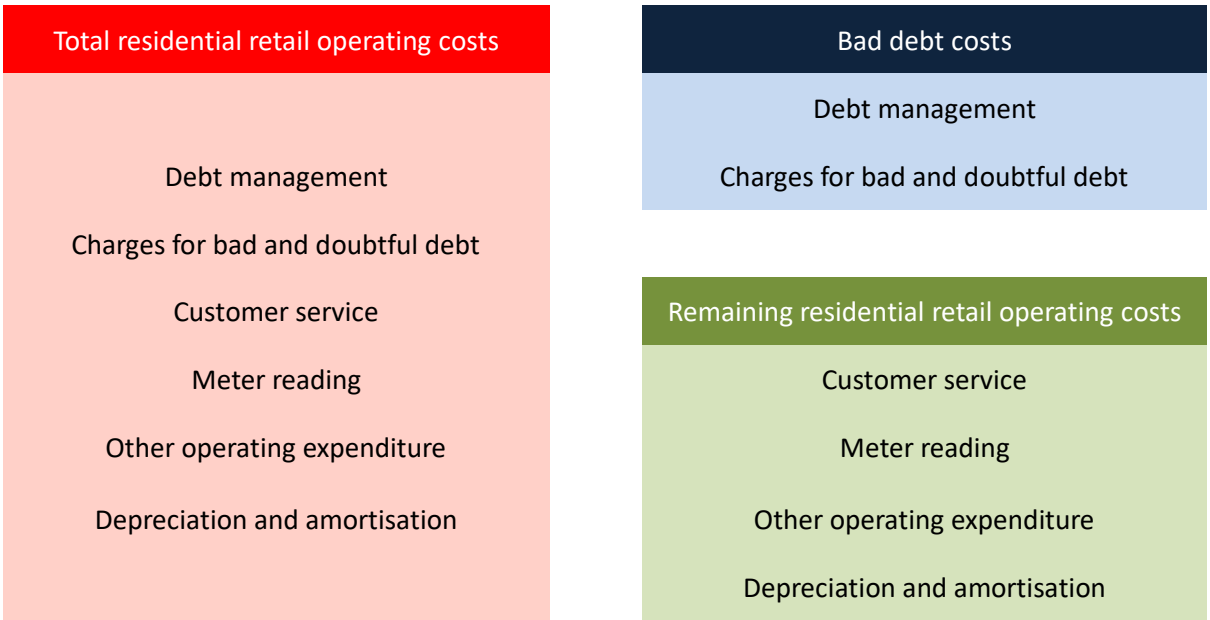
16. We consider that our model development work has achieved significant progress, and provides valuable insight, in terms of the way that a number of residential retail cost drivers could be treated as part of the specification of econometric models to be used for Ofwat's benchmarking purposes.

Overview of findings from model development process: cross-cutting aspects

17. We have developed models covering total residential retail operating costs, and models covering two, more granular, cost categories: one for bad debt costs (including debt management costs) and one for the remainder of retail operating costs. Figure 1 shows the main elements of cost falling within each of the three cost categories that we modelled. These are in line with the benchmarking models that Ofwat proposes to develop, as set out in its December 2017 methodology.[6]

[6] Ofwat (2017) "Delivering Water 2010: Our final methodology for the 2019 price review, Appendix 11: Securing cost efficiency", page 18.

Figure 1 Cost categories modelled



18. Aggregated models have the benefit of being less vulnerable to any inconsistencies across companies in how costs are allocated across cost categories. The benefit of the more granular models is that the functional form and explanatory variables can be more tailored to the specific category of costs, which may allow relationships between costs and cost drivers to come out more accurately in the estimation results. Our modelling results reinforced our view on this benefit of more granular models. It may be possible to produce even more granular models (e.g. models focused on meter reading costs) but we did not explore this for this report.

19. All the models we have considered exclude any allowance for financing costs such as the cost of capital, working capital or profit margins for retail activities. This exclusion reflects Ofwat's approach to water companies' residential retail price controls, which involves a separation between the assessments of companies' requirements for operating costs (cost to serve) and the assessment of financing costs.

20. For most models, we set the dependent variable as a measure of each company's average costs per household supplied (or the natural logarithm of this). These measures represent a development of the average cost to serve concept used by Ofwat for PR14. For bad debt cost models, we also explored models in which the dependent variable was bad debt costs divided by household revenue, but developing this modelling approach further does not seem a priority for future work.

21. We considered models in which the dependent variable is a measure of aggregate retail costs, but did not include these in the model specifications we present. While we would not rule these out, such models can suffer statistically from multi-collinearity and heteroscedasticity issues and can perform worse than unit cost models as a means to predict companies' residential retail costs. We did not identify a strong need for models to allow for economies of scale in retail operations.

22. Some of the other key features of the models we have developed are as follows:

(a) We used a **panel dataset**, representing 18 companies over a four-year sample period from 2013/14 to 2016/17, though data for 2016/17 relates to only 17 companies as a result of the merger of South West Water and Bournemouth Water. This dataset contains an extra year of data compared to that used for our May 2017 working paper.

(b) We used **pooled OLS** applied to the panel dataset. An alternative to consider is a random effects structure estimated using GLS. We did not prioritise the exploration of this approach, as investigation of other aspects of model specification (particularly the incorporation of cost drivers) seemed more likely to yield useful insights. OLS estimation requires fewer assumptions on the structure of the error term to be made and is a more familiar and simpler approach. Another alternative would be to use cross sectional OLS estimation, in which we take averages of data for each company over a four-year period. This might be a reasonable approach to explore. Given the role of Ofwat's econometric benchmarking and cost assessment, and the small sample size, we did not see any merit in the exploration of use of stochastic frontier analysis (SFA) or data envelopment analysis (DEA).

(c) We **deflated cost data** using CPI. We included time dummy variables in model specifications to allow for fluctuations and variations in industry-wide costs over time (e.g. reflecting changes in input prices and improvements in efficiency).

Overview of findings from model development process: cost drivers

23. Our work on the explanatory variables started with the identification of a series of candidate cost drivers and then explored whether, and how, these could be incorporated into econometric models of residential retail costs. For the identification of candidate variables, we drew on the following: (a) the special cost factors Ofwat considered as part of its PR14 final determinations; (b) cost drivers captured in previous work by Ofwat and other parties on econometric models for retail costs in the water industry; (c) discussions with United Utilities, building on its operational knowledge; and (d) our experience and knowledge of water industry benchmarking analysis.

24. We considered a range of explanatory variables, intended to capture a variety of cost drivers. These varied by model, depending in part on the level of granularity and scope of costs covered.

25. Table 1 provides an overview of the retail cost drivers that we identified as candidates for further investigation. In the table, we distinguish between drivers of bad debt costs and drivers of the remaining retail operating costs. We highlight whether each candidate factor is most relevant as a potential driver of bad debt costs or of remaining retail operating costs, or whether it seems to apply to both. We explain the reasoning for the identified relationships in the main body of the report (see Table 4). Any driver that is a potential driver to either cost category would also be potentially relevant to the models of total retail operating costs. The final column in the table indicates whether the cost driver was included in some or all of the econometric

models we present in this report. We set out in the main body of the report the rationale underlying the mapping captured in the table.

Table 1 Overview of identified cost drivers for retail cost models

Candidate factor	Potential driver of bad debt costs	Potential driver of other retail costs	Included in some/all models presented in this report
Number of households supplied	✪	✪	✓
Size of residential retail bills	✪		✓
Proportion of dual service customers (i.e. customers that take both water and wastewater services from the retailer, rather than just one of these)	✪	✪	✓
Measures of economic deprivation and arrears risk within geographic areas served	✪	✪?	✓
Transiency of households within the area served (i.e. rate of occupancy changes)	✪	✪	
Quality of service provided to customers		✪	
Meter penetration rate for customers		✪	✓

26. Besides the cost drivers listed in the table, there may be other relevant factors (e.g. factors that affect the relative ease of taking meter readings at customer premises, such as the geographical density of metered customers and levels of local traffic congestion), but these seemed likely to be less significant than the factors in the table and we did not include work on these in the report on prioritisation grounds.

27. We provide our detailed modelling results in the main body of the report, and the final section of the report draws together the findings from the work in terms of the incorporation of the retail cost drivers in econometric models. Table 2 overleaf provides a high-level summary of that analysis.

28. Finally, our modelling results reinforced our view on the benefits of using models defined at a more granular level than total residential retail costs. We produced separate models for (a) costs relating to bad debt and debt management; and for (b) the remaining retail costs. Our view is that these models are better able to capture (or approximate) relationships between costs and cost drivers than an approach based on models for total residential retail operating costs. We would certainly not rule out the consideration of models that take total residential retail costs together, and these models may, in some circumstances, have some countervailing benefits. But models that take the different categories of residential retail costs separately seem the best starting point for Ofwat's development and estimation of econometric models. If granular models are to be used, the results should be aggregated across models first, before any upper quartile benchmark is calculated.

Table 2 Summary of findings on incorporation of cost drivers in econometric models

Candidate factor	Does our work provide evidence on materiality of cost driver?	Should cost driver be included in econometric models of residential retail costs?	Are there other ways to capture *without* using econometric models of retail costs?
Number of households supplied	Yes	Yes We can specify as denominator in the dependent variable in model	Yes (e.g. assumption of proportional relationship between number of households and total residential retail costs)
Size of household retail bills	Yes	Our work supports use of this cost driver in explanatory variables for econometric models	Could make assumption on relationship between bills and bad debt but risk that assumption lacks evidence
Proportion of dual service customers	Some indication of material positive relationship but not strong evidence	Our work provides some support for this cost driver, but results quite sensitive to model specification	Without Ofwat collecting more granular cost data, it seems difficult to capture this driver without econometric approach
Measures of economic deprivation and arrears risk	Yes	Our work supports use of this cost driver in explanatory variables for econometric models	Difficult to estimate financial impact of varying levels of deprivation on bad debt costs without econometric modelling
Transiency of households within the area served (i.e. rate of occupancy changes)	Cost driver not fully considered in the models due to data limitations. Drawing on the data that were available, no strong evidence of positive relationship.	Not based on the work to date. May be possible to overcome with better measure of transiency.	Difficult to estimate financial impact of varying levels of transiency on retail costs without econometric modelling
Quality of customer service	No (due to modelling issues: does not imply this is not a cost driver in practice)	Not based on work to-date (though may be possible to overcome modelling issues in the future)	There seem to be other ways that companies could be remunerated for the costs of providing higher-quality services
Meter penetration rate for customers	Some indication of material positive relationship but not strong evidence	Our work provides some support for this cost driver, but results quite sensitive to model specification	Meter reading unit cost metric approach seems feasible (either as alternative or complement) though may not capture all meter-related costs

Section 2: Overview of residential retail operating costs

29. This section gives an overview of the main elements that make up residential retail operating costs. The purpose is to set the scene for the modelling analyses presented in subsequent sections. It identifies the activities whose costs are included within the different benchmarking models, and, in so doing, informs on the set of candidate cost drivers for each of those models.

Cost categories associated with provision of residential retail services

30. The Regulatory Accounting Guidelines (RAGs) 4.07 include guidance on the categories of cost for which companies are to break down and report the operating costs relating to the provision of retail services to households. Based on that guidance, Table 3 lists the main categories of costs, and for each, sets out what those costs are intended to cover.

Table 3 Components of residential retail operating costs, as per RAG 4.07

Cost category	What costs refer to
Customer service	• Billing • Payment handling, remittance and cash handling • Charitable trust donations • Vulnerable customer schemes • Non-network and network customer enquiries and complaints • Investigatory visits (where the cause of the visit is not a network issue) Excludes customer services costs incurred in providing services to a third party (e.g. where a water only company bills and collect payments on behalf of a water and sewerage company)
Debt management	All costs relating to the management of debt recovery for household customers Excludes costs incurred relating to the management of debt recovery for a third party (e.g. where a water only company manages debt on behalf of a water and sewerage company)
Doubtful debts	The charge for bad and doubtful debts for household customers Should exclude doubtful debts relating to a third party
Meter reading	Costs associated with meter reading for household customers, including ad hoc read requests, cyclical reading, scheduling, transport, physical reading, reading queries and read processing costs, manager meter data and supervision and management of meter readers. Income from meter reading commission should be netted of these costs Excludes costs associated with meter reading for third parties
Other operating expenditure	Any other operating costs (excluding interest and taxation) incurred serving household customers. Amongst others, includes the costs of: • Provision of offices • Insurance premiums • Net retail expenditure on demand-side water efficiency initiatives • Net retail expenditure on customer side leaks

Cost category	What costs refer to
	• General and support expenditure • Local authority rates • Other business activities
Third party services operating expenditure	The operating costs of providing appointed household retail service to third parties
Depreciation – tangible fixed assets	Depreciation on tangible assets used wholly or principally for the household retail business
Amortisation – intangible fixed assets	Amortisation on intangible assets used wholly or principally for the household retail business

31. As indicated in the table above, the amortisation and depreciation figures are for assets wholly or principally for the residential retail business. Where such assets are also used to provide services for other price control business units, there may be recharges from the residential retail business unit to these other business units. Similarly, the costs identified in the table above do not include amortisation and depreciation for assets that are used to some degree for residential retail activities but which are allocated to other business units (e.g. wholesale wastewater) under the principal use approach, and for which there may be recharges to the residential retail business. For the purposes of this report we have not made adjustments for the net effects of recharges in calculating our relevant measures of residential retail costs, but this may be something to consider in the future, subject to data availability and consistency.

32. Also in relation to the data on amortisation and depreciation, we note that the figures for 2013/14 and 2014/15 relate to current cost depreciation, whilst those for the last two years in our data period, 2015/16 and 2016/17, do not. This reflects a change in the requirements specified in the relevant Regulatory Accounting Guidelines.[7] We have not sought to adjust the data on depreciation that we compiled to address this. Again, this may be something to consider in the future, subject to data becoming available.

Components of residential retail operating costs

33. Across companies, and across 2013/14 to 2016/17, the average residential retail operating cost was £28 per household.[8] Figure 2 shows how this cost is broken down into the main categories listed in Table 3 (we have grouped depreciation on tangible fixed assets and amortisation on intangible fixed assets together and labelled it as depreciation). The chart shows the breakdown for the industry as a whole, and for the water only and the water and wastewater companies separately.

[7] Compare RAG 4.04 from February 2013 with RAG 4.05 from October 2015.
[8] We used the CPI to express values in 2016/17 prices.

Figure 2 Components of residential retail operating costs, 2016/17

34. Across all companies, and across the four-year period, costs relating to debt management and bad debt represented 38 per cent of total residential retail operating costs. Customer services accounted for 29 per cent of, and costs reported under "Other operating expenditure" for 22 per cent. Meter reading costs accounted, on average across all companies, for 5 per cent of total operating costs. The remaining set of costs, around 7 per cent of total, relate to depreciation and amortisation of fixed assets.

35. As shown in Figure 2, there is a marked difference in the cost per household between water only companies (WoCs), and water and wastewater companies (WaSCs). From 2013/14 to 2016/17, the average total operating cost per household of the former was £21 and for the latter it was £33. Figure 2 is also revealing in showing that the key difference in the unit cost between WoCs and WaSCs arises from the differences in the costs relating to debt, in particular from differences in the costs of bad debt. As discussed in the analysis of the econometric models for bad debt costs in Section 4, this difference between WoCs and WaSCs points to the importance of controlling for differences in companies' average bill size when comparing their bad debt costs.

Section 3: Overview of approach to econometric model development

36. This section presents an overview of our approach to the development and assessment of econometric models of retail costs. We take the following points in turn:

    (a) Overall approach to model development.

    (b) Granularity of models.

    (c) Candidate cost drivers.

    (d) Specification of model dynamics and estimation method.

    (e) Criteria used in the development and assessment of models.

Overall approach to model development

37. We do not see the development of econometric models to produce benchmarks of water companies' retail costs as the search for a single underlying true model. At best, the models provide approximations of the relationship between retail costs and the identified cost drivers. The extent of the approximation reflects several factors, such as the relatively small sample size, the lack of variance in explanatory variables over time (as companies' retail supplies and characteristics are similar from one year to the next) and the simplifications inherent in the model specifications (e.g. that a linear relationship exists between the dependent variable and a particular cost driver).

38. Our approach to model development has been an iterative one, involving judgement, in which we started with initial ideas on how to specify models of retail costs, drawing on a review of candidate cost drivers, and then extended and refined these models in the light of analysis of model estimation results and other indicators. Our objective has been the development of more accurate models of water companies' (efficient) expenditure requirements on residential retail activities. The most important feature of our approach is progress: *ending up in a better place than where we started*.

39. We consider that the models presented in this report represent a substantial improvement on the modelling approach that Ofwat used for PR14. It is quite possible that the models we present could be further improved through subsequent work.

40. In our development of models, we gave weight to reviewing the sign and magnitude of estimated coefficients to check their consistency with what we might expect from an economic perspective. We incorporated into that process, an analysis of the sensitivity of the estimated coefficients to variations in the dataset, and considered the results of a series of statistical diagnostic tests. We did not consider that it would be useful, for the purposes of Ofwat's PR19 retail cost assessment, to give emphasis to whether particular models are "robust" or "not robust". The word robust can mean different things to different people. Furthermore, given that Ofwat will need to use benchmarking analysis of some form (whether this is econometric modelling or unit cost comparisons), the question is not so much whether a specific approach is robust,

but whether it is the best available approach or forms part of such an approach. Nonetheless, we have considered the sensitivity of results to model specification and the dataset used as a key part of our model development process.

41. Given that all feasible models are likely to involve a degree of approximation, and to involve limitations, we think it is better to consider a range of econometric models that work in different ways than to focus on one model or one modelling approach. We have therefore sought to explore and present a range of models in this report.

42. In a separate sub-section below, we summarise the statistical results and other factors that we gave weight to as part of our model development process.

Granularity of retail cost models

43. We have developed models for three different categories of retail costs:

    (a) Models encompassing all residential retail operating costs, which include retail operating expenditure and depreciation on retail capital assets.

    (b) Models that focus on bad debt costs only, comprising costs relating to charges for doubtful and bad debt and to debt management costs for residential retail activities.

    (c) Models that focus on the remaining retail operating costs, i.e. residential retail operating costs other than those relating to charges for doubtful debt and to debt management costs.

44. The first type of models captures the sum of the costs under the second and third type. As such, the second and third type are more granular models.

45. In this report we use for convenience the term "bad debt cost models" as shorthand to describe the models in (b) above.

46. For its PR19 cost assessment, Ofwat's plan is to consider the three types of models above for retail costs.[9]

47. The benefit of more granular modelling is that, especially with a small sample size, the econometric modelling may be able to produce a more accurate estimate of the relationship between costs and the relevant cost drivers. This is particularly so if a cost driver acts on only a part of retail costs: stripping out the costs that are not expected to be affected by that cost driver can help in the estimation of the relationship between the cost driver and those elements of costs that it does affect. However, there may also be reasons to consider the aggregated models. For instance, more granular models may suffer from any inconsistencies between companies in the allocation of costs between different categories.

48. It would be possible to develop more granular models for retail costs than those indicated above. We have not sought to consider this within the scope of this report.

[9] Ofwat (2017) "Delivering Water 2010: Our final methodology for the 2019 price review, Appendix 11: Securing cost efficiency", page 18

49. All the models we have considered exclude any allowance for financing costs such as the cost of capital and/or profit margins for retail activities. This exclusion reflects Ofwat's approach to water companies' residential retail price controls, which involves a separation between the assessments of companies' requirements for operating costs (cost to serve) and the assessment of financing costs.

Candidate cost drivers

50. We have considered the potential drivers for the cost of providing water and wastewater retail services to households in England and Wales. These are the factors which could account for differences in costs across companies and which we would wish to take account of in a benchmarking assessment to allow for a more like-for-like comparison between companies.

51. We have drawn on the following:

    (a) The special cost factors Ofwat considered as part of its PR14 final determinations.

    (b) Cost drivers captured in previous work by Ofwat and other parties on econometric models for retail costs in the water industry.

    (c) Discussions with United Utilities, building on its operational knowledge.

    (d) Our experience and knowledge of water industry benchmarking analysis,

52. Table 4 provides an overview of the candidate retail cost drivers that we identified as candidates for further investigation.[10] For each factor, we briefly explain the rationale for its inclusion. We highlight in the table whether each candidate driver is relevant to models of bad debt costs and/or models of remaining retail costs. Any cost driver that is relevant to either bad debt cost models or to the remaining retail cost models would also be relevant to models of total retail operating costs.

Table 4 Overview of identified cost drivers for retail cost models

Candidate factor	Rationale	Potential driver of:	
		Bad debt costs	Remaining operating costs
Number of households supplied with retail services	The number of household customer that are provided with a retail service is a key driver of retail cost, reflecting scale of the retail business.	✓	✓
Measures of economic deprivation and arrears risk within geographic areas	Economic deprivation may be a cause for households to fall behind on payments or not make water bill payments.	✓	?

[10] The table highlights a series of potential relationships between residential retail costs and factors that may differ between companies, but is not intended to be exhaustive of all conceivable relationships. For instance, we consider deprivation to be a cost driver for bad debt costs, but there may also be some relationship between deprivation and other categories of retail costs, such as payment processing (e.g. if customers in more deprived areas make greater use of payment methods that are higher-cost for retailers). We have not sought to investigate all conceivable relationships in this report.

Candidate factor	Rationale	Potential driver of:	
		Bad debt costs	Remaining operating costs
served	Measures of arrears risk in different parts of England and Wales that are based on customer data across a range of sectors are likely to be an indicator of arrears risk for residential retail supplies. Furthermore, to the extent that there are differences in measures of arrears risk between different parts of the country, we would expect these to be primarily (or entirely) driven by differences in factors related to deprivation between these areas		
Size of residential retail bills or wholesale bills	Plausible link to bad debt costs in two ways. The higher the bill the higher the amount of money that is at risk of not being paid in full. And higher bills may themselves bring a greater risk of a household falling into debt.	✓	
Whether customers take both water and wastewater services (dual service), or just one of these services	There are likely to be some retail costs (e.g. costs relating to customer enquiries) that are greater on average per customer (household) when suppling the customer with both water and wastewater rather than a single service. The provision of two services rather than one will also tend to increase the average bill which is identified separately above as a likely driver of bad debt costs.	✓	✓
Transiency of households (i.e. rate of occupancy changes)	Transiency of population is plausibly associated with costs of customer services (e.g. opening and closing of accounts and metering enquiries) as well as of bad debt.	✓	✓
Quality of service provided	The provision of higher quality of customer service may be associated with higher costs.		✓
Meter penetration rate	Costs of meter reading likely to be associated with number of metered customers. In addition, also plausible that there may be differences in cost of serving metered customers and unmetered ones, e.g. if former are more likely to raise queries over bills.		✓
Factors affecting relative ease of taking meter readings at customer premises	There may be factors that affect the costs of meter reading beyond the number of metered customers, such as: dispersion of (metered) customer base across area of appointment; and traffic congestion in area of appointment.		✓

53. The ability to capture each cost driver within econometric models of retail costs will depend on the availability of relevant data on that cost driver. We comment in Table 5 on the data availability for each of the cost drivers from Table 4.

Table 5 Overview of data availability for candidate cost drivers

Candidate factor	Comments on data availability
Number of households supplied with retail services	We treat a household that is billed by a water retailer as a single customer of that retailer. This customer may take one or two services from the retailer. Data is available from companies' regulatory accounts (within the Annual Performance Report) on the number of customers (households) falling into categories representing the services provided to them (e.g. measured water and wastewater; measured water only, etc.). The total number of households supplied with retail services can be calculated by aggregating across these categories. In line with the Regulatory Accounting Guidelines, the number of customers is the average number in the year, calculated at least on a monthly basis. Void properties are excluded.
Measures of economic deprivation and arrears risk within geographic areas served	As explained in Reckon (2017), for the purposes of Ofwat's cost assessment for PR19, there are significant limitations in the data on economic deprivation available from the DCLG and Statistics for Wales. [11] We have worked with United Utilities to produce measures of economic deprivation and arrears risk based on data from Equifax, which relate to: • Measures of the proportion of households within a company's area of appointment that are within those geographic areas identified as the most deprived (e.g. 10% or 20% most deprived) or the highest risk; • A weighted average measure of deprivation or arrears risk across each company's area of appointment.
Size of residential retail bills or wholesale bills	An average retail bill can be calculated by taking data from the regulatory accounts on total revenue from households (which comprises retail revenues and wholesale revenues) and dividing this by the number of households supplied with retail services (see above).
Whether customers take both water and wastewater services (dual service), or just one of these services	The information from companies' regulatory accounts identifies the number of customers (households) serviced by each company that takes a water and wastewater service (i.e. dual service customers); a water service only; or a wastewater service only.
The transiency of households within the area served (i.e. rate of occupancy changes)	We did not identify a good data source on transiency within companies' regulatory accounts. We explored two other sources: • ONS data on migration flows to and from Local Authority Districts; • Equifax data on changes to electoral registry, reported at the postcode level.
Quality of service provided to customers	Companies' regulatory accounts report data on the Ofwat SIM (service incentive mechanism) score which is a measure of customer service performance.
Meter penetration rate for customers	The Annual Performance Report provides data on the number of customer that are metered (or measured customers), distinguishing between different services (e.g. metered water and wastewater versus metered water only). A customer that is metered for both water and wastewater services (both based on the installed water meter) counted as both a metered water

[11] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities.

Candidate factor	Comments on data availability
	customer and as a metered wastewater customer.
Factors affecting relative ease of taking meter readings at customer premises	We felt this was likely to reveal fewer insights compared with the work on other cost drivers and we have not sought to explore data availability for these aspects of meter reading. This is not to say that model incorporating variables for these aspects could not be developed in the future.

54. We provide more detailed information of how we have incorporated data on the cost drivers in our model specifications in the sections further below, which take each of the main types of models in turn.

55. In the specifications of the models that we explored, we considered the role of economies of scale in retail costs. This captures the idea that companies serving a smaller number of households would, all else equal, have a higher level of efficient costs per household. We did not identify this as a key cost driver for the purposes of the benchmarking analysis. While there may be some degree of economies of scale in particular retail activities, there are potential opportunities for smaller companies to mitigate such effects (e.g. joint ventures and outsourcing arrangements). For our econometric model development, we did not focus on this issue and it is not covered in any detail in this report. All the same, we did some modelling which allowed for economies and scale. In particular, we estimated versions of the models where the dependent variable was an aggregate measure of cost (rather than unit cost measure) and where the set of explanatory variables included a measure of customer numbers. The results of that analysis did not indicate that it was unreasonable to assume no economies of scale or that such aggregate models performed better in statistical terms.

Specification of model dynamics and estimation method

56. The main models that we present in this report have the following features:

    (a) The dataset has a panel structure, covering 17 or 18 companies over four years, from 2013/14 to 2016/17. There are 71 observations in total. As explained in Appendix 1, there are changes in the number of companies over the sample period due to the merger between South West Water and Bournemouth Water.

    (b) All cost data are expressed in 2016/17 prices, inflated using the CPI.

    (c) Each model has a series of time dummy variables to allow for industry-level changes in retail costs over time.

    (d) Models are estimated using the ordinary least squares (OLS) method. This means that we have adopted a "pooled OLS" approach to the panel data.

    (e) Where we report estimated standard errors or t-statistic, these are estimated using the "cluster robust" approach to allow for correlations over time in the residuals for each company.

57. There are potential alternative approaches that we could have used, which we briefly discuss below.

58. For its PR14 wholesale cost assessment, Ofwat used random effects models estimated using generalised least squares (GLS) alongside models estimated using OLS. The random effects models tend to give similar results, but the results in some cases can vary significantly. We have used OLS as it is the simpler and more familiar approach, and because it is less demanding regarding the assumptions on the structure of the error terms in the econometric model. We note that the use of GLS in the context of benchmarking water companies has been criticised.[12] We considered that the choice of approach between GLS random effects and OLS was of secondary importance to other aspects of our analysis on model development and we did not explore the issue further for this phase of our work.

59. With respect to controlling for industry-wide differences over time, an alternative to the inclusion of time dummy variables is to include within the set of explanatory variables one that would capture a time trend. We consider that the time trend approach is not as good as the dummy variable approach in this context. We have no need or reason to assume that factors driving changes in industry-level costs (e.g. industry-wide productivity improvements and movements in input prices such as wages relative to CPI) have the same annual effects across our data period.

60. A further alternative would be to structure the model as a cross sectional one, based on data for each company averaged over the four-year period 2013/14 to 2016/17. This was the main approach taken for our working paper in May 2017, which provided an initial exploration of the use of deprivation measures based on Equifax data and focusing on models of bad debt costs only.

61. With respect to the choice of expressing costs in 2016/17 prices using the CPI, we could have taken a different approach. We could have used nominal costs throughout, or have used a different price index, such as the CPIH. In practice, this choice has little impact on the output of the analysis, given other aspects of the specification of the models we explored. For the models of bad debt costs and of total residential retail operating costs that we present in this report (Sections 5 and 7), the choice of price index to use (or the choice of whether to make an adjustment for inflation at all) has no impact on the performance of the models, or on the estimated coefficients, other than on those for the year-specific dummy variable and on the constant. This is because the dependent variable on those models is expressed in logarithms, and the set of explanatory variables includes year-specific dummies. In the case of the models of remaining retail operating costs, where the dependent variable is not expressed in logarithmic terms, the choice does affect the estimated coefficients of all explanatory variables. We have tested the sensitivity of the results of those models to this, re-running the same set of models using data on nominal costs. We found this had only very marginal effects on the size of the estimated coefficients.

[12] Vivid economics and Arup (2017) "Understanding the exogenous drives of wholesale wastewater costs in England and Wales", page 38.

Criteria used in the development and assessment of models

62. We set out in Table 6 the key criteria that we have taken into account as part of the development of models presented in this report and which we discuss further in our assessment of results. We used judgement in the application of these criteria.

63. Later in this report we present key modelling results for a set of short-listed models which emerged following the model development phase, taking account of the criteria in Table 6. These models do not represent the full range of models considered along the way: we explored a greater number of candidate model specifications and modelling approaches.

Table 6 Criteria used in model development and assessment

Consideration	Comments
Economic and business rationale for explanatory variables and functional form of the model	We have given weight to the economic and business rationale for alternative model specifications, including the choice of candidate cost drivers and the way that these are incorporated as explanatory variables within specific models.
Sign and magnitude of estimated coefficients	We have reviewed the sign and magnitude of estimated coefficients to check their consistency with what we would expect (insofar as we have prior expectations) given the identified economic and business rationale for the cost driver.
t-statistics (and estimated variance) for estimated coefficients	The t-statistic for an estimated coefficient is calculated by dividing the value of the estimated coefficient by the standard deviation of the estimated coefficient. The t-statistic provides a measure of the variance of the estimated relationship between a candidate cost driver and costs, which is normalised to allow for meaningful comparison across different explanatory variables and models. A smaller t-statistic implies less precision in the estimation of the relationship between the candidate cost driver and retail costs. The standard deviation of the coefficient (and hence t-statistics) is itself is an estimate subject to uncertainty and estimation error so caution is needed in interpretation.
Sensitivity of results to the dataset used	As part of our model development and assessment, we incorporated analysis of the sensitivity of estimated coefficients to minor variations in the dataset used. For the short-listed models, we carried out an exercise of re-estimating the model using variants on the dataset. Specifically: • Versions of the dataset which feature data for all but one of the companies (across all years); • Versions of the dataset which feature data for three out of four years (and all companies). Where the estimated coefficients for an explanatory variable is highly sensitive to variations in the dataset this may indicate that the econometric modelling is struggling to capture underlying relationships between that variable and costs. This analysis provides information on the precision of the estimated coefficients using a different perspective to the t-statistic. Particularly in a small sample, the t statistic for the coefficients on an explanatory variable may not provide a good guide to the robustness of the estimated relationship between that variable and the dependent variable, and further insight can be gained by considering how sensitive the estimated coefficient is to minor changes in the dataset. If the estimated coefficient is highly sensitive to such changes then we would question whether it is providing a useful

Consideration	Comments
	approximation to the underlying causal factors, at least if we do not have other reasons to explain such sensitivity.
Estimated variance of predicted values	Ultimately, the outcome of the econometric models is to produce cost benchmarks for the purpose of cost assessment, based on the predicted values from the estimation of the econometric models. The statistical software we use can calculate an estimate of the variance of the predicted values produced by the model. This provides information on the statistical precision of these predicted values, having imposed the model on the dataset. It will reflect in part the estimated variance for each of the coefficients for the explanatory variables in the model. We consider this a relevant consideration for model development and review.
R-squared	The R-squared is a measure of the goodness of fit and provides information on the extent to which the variance in the dependent variable can be attributed (in the estimated model) to variance in the explanatory variables. One issue is that R-squared can always be improved by adding in further explanatory variables, but such variables may lead to the results being driven by spurious correlations rather than underlying relationships between costs and cost drivers. While this is sometimes tackled by using measures of the "adjusted R-squared", the way that such measures penalise the inclusion of additional explanatory variables is somewhat arbitrary. As such, we focused on the normal R-squared measure while taking into account this issue. That said, we consider that the R-squared is a relevant measure as part of a wider review of alternative measures. R-squared is not comparable on a like-for-like basis between models with different dependent variables.
Diagnostic tests	We have carried out some diagnostic tests as part of model development and assessment, drawing on the tests used by Ofwat during PR14. We describe the tests further below. As a general point, we do not consider that these tests should be treated as a pass/fail criteria for whether models are acceptable for use in Ofwat's cost assessment modelling: ultimately that comes down to a comparison of the alternative feasible approaches, taking account of the benefits of certain models as well as any identified. Nonetheless, they may reveal aspects of model results to investigate further and opportunities for refinement of model specifications.

64. While we identify in Table 6 that the t-statistic and estimated variance are informative, we did not use any t-tests mechanistically to decide whether to include or exclude variables from the model specification. With a small sample size, this approach would give too much weight to statistical results. It would overlook the sensitivity of t-test results to the details of model specification. And it would treat the non-rejection of a null hypothesis the same as confirmation of the null hypothesis.

65. As highlighted in Table 6, we carried out a series of diagnostic tests on the models. These were:

    (a) **Ramsey RESET test**. This is a test on aspects of the specification of the model. It tests the joint significance of the square, cube and fourth power of the

predicted value, were these to be added to the set of explanatory variables of the model.[13].

(b) **Linktest**. This too is a test on aspects of the specification of the model. It tests the significance of the square of the predicted values in a regression of the dependent variable against the predicted values and the square of the predicted values.

(c) **Shapiro-Wilk test of the residuals**. We applied the Shapiro-Wilk test to the residuals from the regressions of each model to test the hypothesis that these were normally distributed.

66. In each of the tests that we ran, the null hypothesis is that that assumption underpinning the ordinary least squares estimation and inference holds. For example, in the case of the Ramsey RESET test, the null hypothesis is that the estimated coefficient on the powers of the predicted values are 0, and in the case of the Shapiro-Wilk test of the residuals, the null hypothesis is that these are normally distributed

67. In respect of the first two tests above, it is important to recognise that these are not general tests of whether or not models are specified well. The tests are quite narrow and relate primarily to the explanatory power of the squares (and cube, and fourth power) of explanatory variables and of the interactions between those variables. The results from these tests may lead to improvements and refinements in model specification.

68. We have used the cluster robust approach (as implemented in the software Stata) to the estimated standard errors, and, as such, have not reported on a test of heteroscedasticity, such as the Breusch-Pagan test. We do not consider this a significant limitation in our work. We have specified models in a way that is designed to limit the risks of heteroscedasticity emerging, by specifying the dependent variable as a measure of retail costs per household rather than aggregate costs.

[13] This describes the default Ramsey RESET test as implemented in the statistics package STATA. The test can be specified to include higher powers of the fitted values, or to consider powers of some or all of the explanatory variables.

Section 4: Models of residential retail bad debt costs

69. This section presents our analysis of econometric models to benchmark measures of water companies' bad debt costs relating to the provision of residential retail services.

70. In this section, and elsewhere in the report, we use the term "bad debt costs" as shorthand to refer to the sum of companies' costs relating to debt management and to the "charges for bad doubtful debt for household customers".[14]

71. The section is structured as follows:

    (a) We give an overview of how bad debt costs vary across water companies.

    (b) We discuss the specification of the models explored.

    (c) We present the main results from our estimation of the models.

    (d) We present analysis of the sensitivity of those results to variations in the dataset.

    (e) We present the outcome of statistical diagnostic tests on the models.

    (f) We compare companies' actual costs with those predicted by the models.

    (g) We discuss the findings.

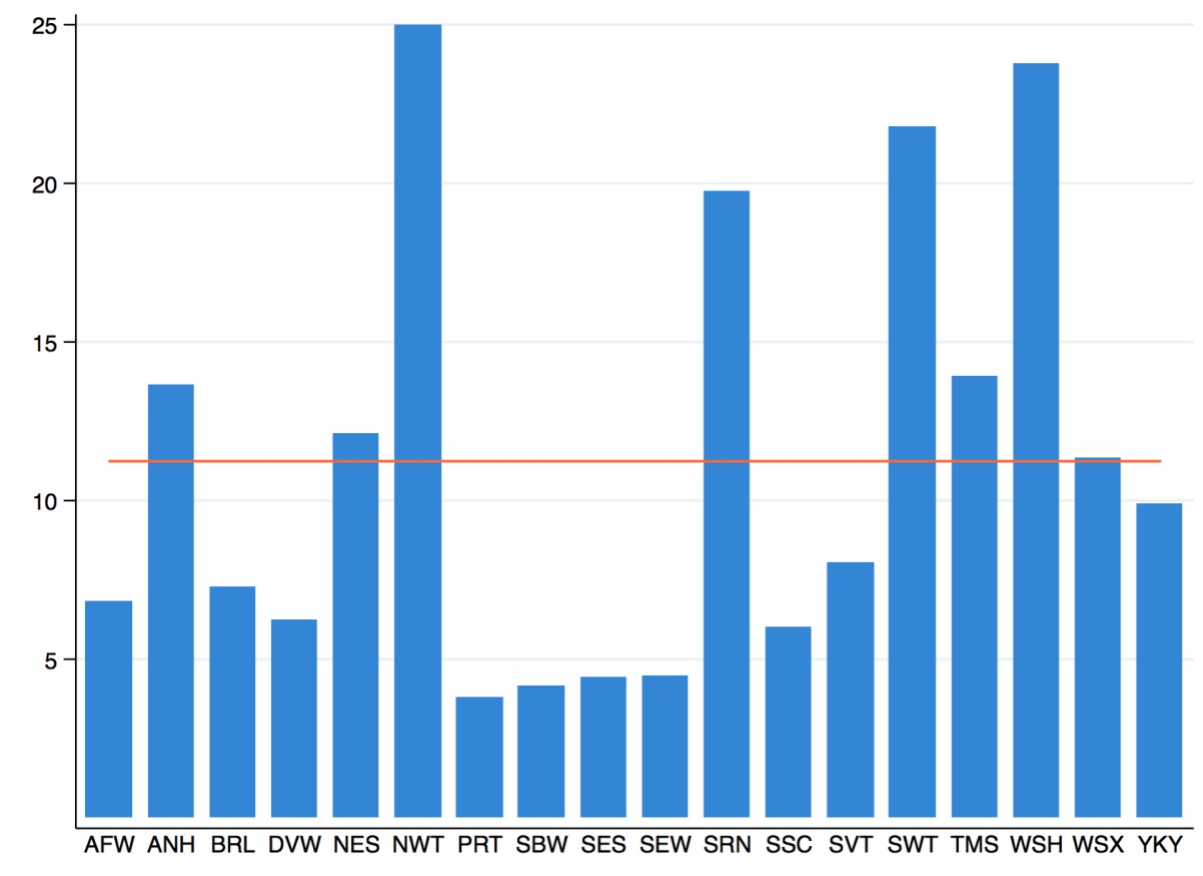
Overview of household bad debt costs

72. Across the four-year period from 2013/14 to 2016/17 and across all companies, the average annual bad debt cost per household was £11[15]. Figure 3 shows the variation of bad debt costs per household across companies. The horizontal orange line in the figure marks the industry average.

[14] Ofwat (2017) "Regulatory Accounting Guidelines 4.07", available from https://www.ofwat.gov.uk/publication/rag-4-07-guideline-table-definitions-annual-performance-report/. See definition of lines 2 and 3 of pro-forma Table 2C.

[15] Values expressed in 2016/17 prices. The values for South West Water and for Bournemouth Water are based on the average from 2013/14 to 2015/16.

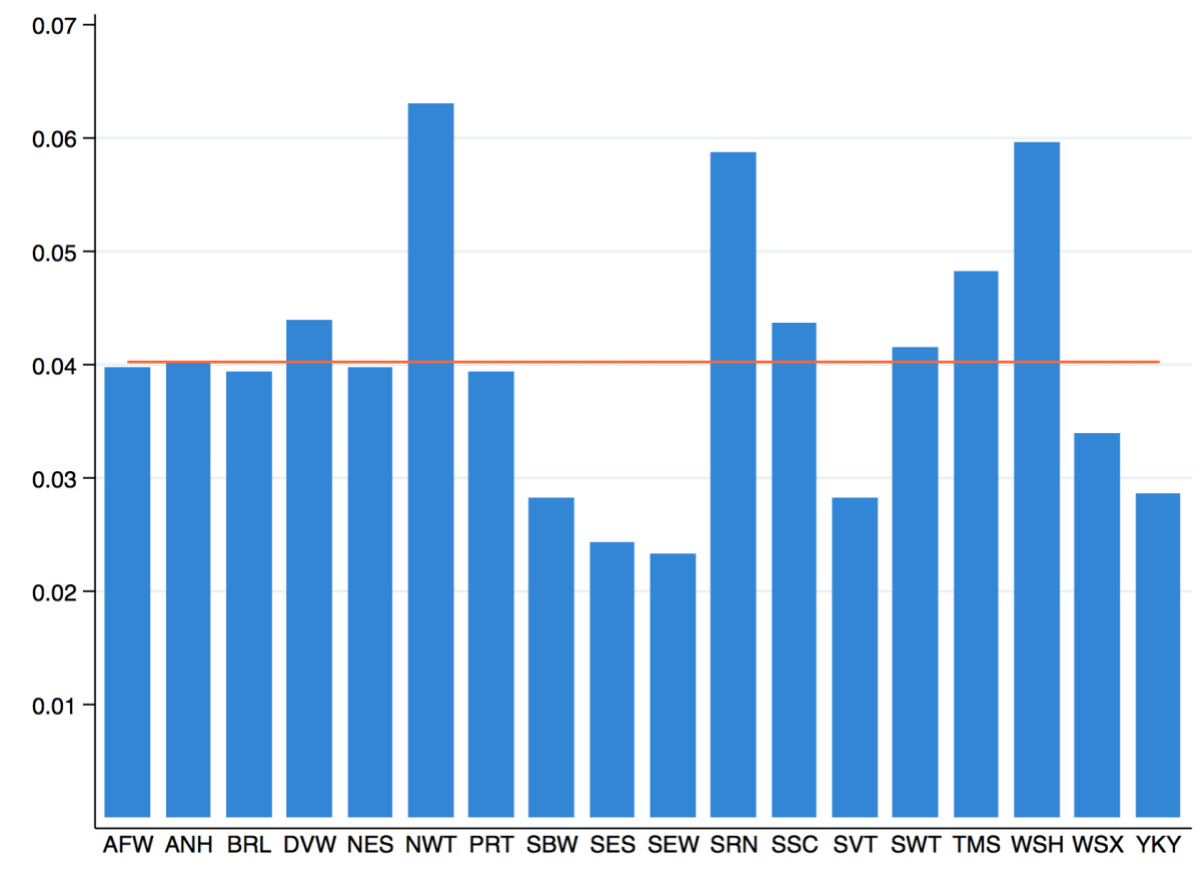
Figure 3 Average annual bad debt cost per household (2013/14 to 2016/17)



73. As shown in Figure 3, bad debt costs per household are higher for water and wastewater companies than for water-only companies. As discussed further below, this is explained, in part, by the finding that the size of a company's average bill (revenue per household) is strongly correlated with bad debt costs per household. The average bill is higher for those companies that provide both water and wastewater services.

74. Figure 4 charts the ratio of bad debt costs to household revenue. On average, across the industry, this ratio was 0.04; this is shown in the figure by the horizontal orange line. The measure being charted in that figure can be seen to control for the differences across companies in average bills. We can see that this measure is more similar across companies than the measure of bad debt costs per household in Figure 3.

Figure 4 Ratio of bad debt cost to household revenue (2013/14 to 2016/17)



Model specification

75. We considered models with the dependent variable based on two different measures of bad debt costs:

    (a) **Bad debt cost per household (unit cost)**. This formulation has a natural interpretation in a benchmarking model. It can be viewed as a measure of cost per unit of output, where, in this instance, a unit of output is serving one household. At PR14, Ofwat estimated an average cost to serve across the industry, based on averaging the unit cost to serve across companies. Developing an econometric model of cost per household is in that same vein.

    (b) **Ratio of bad debt cost to household retail and wholesale revenue.** We also explored models where the dependent variable is defined as the ratio of bad debt costs to household retail and wholesale revenue.

76. To construct the measure of cost of bad debt per household – the first of the above formulations – we gave the same weight to water-only and to wastewater-only households as we did to households who received both services. Ofwat took a different approach to this at PR14 in its calculation of unit cost. In its analysis, Ofwat constructed a measure of "unique customers (adjusted for economies of scope)" as the

sum of the number of water only and wastewater only customers plus 1.3 times the number of dual service customers. The 1.3 factor was intended to adjust for the economies of scope associated with providing both services.[16] We did not think that this approach would make sense in the context of econometric models for bad debt costs (or indeed for the other models we consider in subsequent sections, where it seemed better to include an explanatory variable for dual service rather than imposing an assumption such as 1.3).

77. The level of economic deprivation and arrears risk in the areas served by each company was one of the potential drivers of bad debt costs we set out to explore in our modelling. We drew on a range of alternative measures of this. These fell into two categories:

(a) Measures of average level of deprivation. These are measures constructed as the weighted average of the deprivation level across the lower-layer super output areas (LSOAs) that fall within each company's water supply area or sewerage service area, or within both. We used as weights household numbers and the average bill of the service provided by the company to each of the LSOAs.

(b) Measures of "extreme" deprivation. These are measures of a weighted proportion of households in LSOAs that are within a company's water supply area or sewerage service area, or within both, which are within the top 10 per cent (or, alternatively, top 20 per cent) of the most deprived LSOAs across England and Wales according to a given deprivation measure. We use the average bill of the services provided to each LSOA as weights in the calculation. Because of this, the measure can be regarded as a proxy for the proportion of a company's revenue that is from LSOAs that are within the top 10 per cent, or 20 per cent, of the most deprived LSOAs across Britain.

78. We developed these measures in the course of earlier work. A summary of that work is set out in Appendix 2 of this paper, and a fuller description of it is set out in the working paper we produced in May 2017.[17] Table 7 lists the set of deprivation measures we used in the models reported on in this paper.

Table 7 Deprivation measures in models of bad debt costs

Deprivation measure	Description
Average IMD (predicted) score	This is the average of the IMD (predicted) score across LSOAs, where the LSOAs scores are weighted by the product of household numbers and average bill of service(s) provided by company to LSOA. The "IMD (predicted) score" is the IMD score predicted for the LSOA in the relevant year on the basis of the econometric analysis we developed in earlier phase of the work. Higher levels of the IMD (predicted) score indicate higher levels of deprivation. The measure is calculated for each company, and for each year in period

[16] Ofwat (2014) "Final price control determination notice: policy chapter A5 – household retail costs and revenues", page 35.

[17] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities.

Deprivation measure	Description
	2013/14 to 2016/17.
Average RGC102 score	This is the average across LSOAs of the value of the Equifax variable labelled as "RGC102 – Postcode Risk Navigator Full – Credit Risk score derived from all Insight data". The average is weighted by the product of household numbers and average bill of service(s) provided by the company to households in each LSOA. The RGC102 score is calculated in a way such that higher values are associated with lower levels of deprivation. The measure is calculated for each company, and for each year in period 2013/14 to 2016/17.
Proportion of revenue from 20 per cent most deprived LSOAs, as measured by IMD (predicted) score	The measure is calculated as the proportion of households in the LSOAs served by a company that are within the 20 per cent most deprived LSOAs in England and Wales, according to the their IMD (predicted) score. The number of households in an LSOA are weighted by the average bill of the services provided by the company to that LSOAs, so that the measure can, as shorthand, be interpreted as a proxy for the proportion of a company's revenue that is from LSOAs that are within the 20 per cent most deprived. The measure is calculated for each company, and for each year in period 2013/14 to 2016/17.
Proportion of revenue from 10 per cent most deprived LSOAs, as measured by IMD (predicted) score	As above, but based on revenue from the LSOAs that are within the 10 per cent most deprived LSOAs in England and Wales.
Proportion of revenue from 10 per cent most deprived LSOAs, as measured by RGC102 score	As above, but using the Equifax variable RGC102 to identify the 10 per cent most deprived LSOAs.
Proportion of revenue from 20 per cent most deprived LSOAs, as measured by RGC102 score	As above, but using the Equifax variable RGC102 to identify the 20 per cent most deprived LSOAs.

79. As indicated in the table, we have used explanatory variables based on proxies for the IMD deprivation score for a water company's area of appointment and explanatory variables based on measures of arrears risk amongst the households within that area. While the measure of arrears risk is not a direct measure of deprivation, we consider that the primary drivers (and perhaps only significant drivers) of differences in the arrears risks between different parts of England and Wales will be differences in factors that are linked to economic and social deprivation. And we found that the measures of arrears risk that we have used were highly correlated with the published IMD deprivation measures. We consider it reasonable to treat the arrears risk measures as providing proxies for the differences in deprivation across England and Wales. In any event, measures of arrears risk amongst households, based on data across multiple sectors and services, seems a relevant explanatory variable to consider for models of bad debt costs.

80. In an initial phase of model development, we also examined the potential role for including measures of population transiency in our models. We considered two alternative measures of this:

(a) The variable reported within the Equifax dataset labelled as "EPCF27 – Electoral Roll Postcode Event, average number of occupancy changes per household".

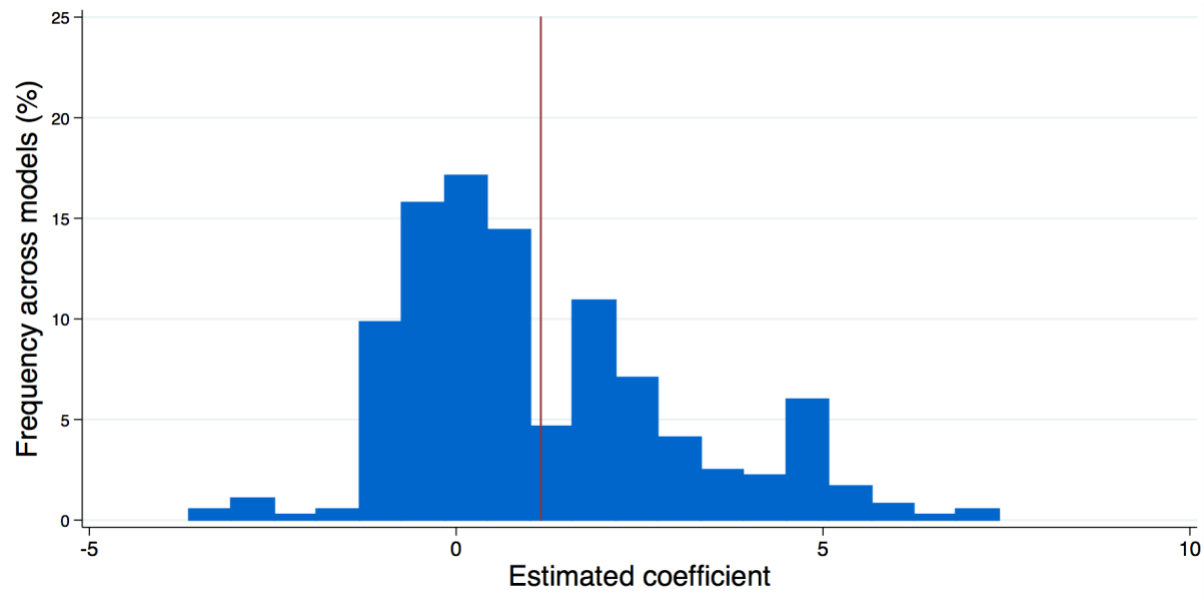
(b) A measure we constructed on the basis of ONS data on population flows into and out of Local Authority Districts (LADs). For each LAD, we computed the measure as the ratio of the sum of the number of people that migrated into the LAD and the number of people that migrated out of the LAD, to the mid-year population in the LAD. The measure takes account of internal migration (i.e. within UK, from one LAD to another) and external migration (i.e. involving moving to or from another country). We constructed a company-wide measure of population transiency by mapping LADs to the areas served by each company.

81. Equifax provided us with details on how it constructs its measure. This revealed that the measure did not adequately capture occupancy changes as would be relevant for our purpose, nor could it be seen as a reasonable proxy for it.

82. We also had concerns about the use of the transiency measure constructed on the basis of the ONS data. In part, this stems from the concern that such a measure does not capture what is intended, which is the turnover of customer accounts. For example, the ONS data on migration flows will not reflect household moves within an LAD, which could account for a significant share of population flows. We drew on data from United Utilities for 2016 to compute the correlation between the number of new accounts in each of the LADs it serves and the net migration flow estimated on the basis of the ONS data. We found this to be 0.70, indicating a positive linear association between the two, though not as strong as might be desirable.

83. All the same, we explored the role of this measure of transiency in explaining variations in companies' costs. We included the measure within the set of explanatory variables of each of the models of bad debt costs and of total operating costs which we discuss in detail in Section 5 and 7. We examined the extent to which the estimated coefficient of that transiency measure was sensitive to the choice of other explanatory variables included (e.g. on the choice of measure used to capture deprivation) and was sensitive to the use of different versions of the dataset (e.g. versions which feature data for all but one company, and versions which feature data for all but on year). We found that the distribution of the estimated coefficient on the ONS transiency measure is wide, covering both positive and negative values. This is illustrated in Figure 5, which shows a histogram of the estimated coefficients for ONS transiency across the models of total operating costs we explored and across the variations in the dataset. A similar pattern is observed for models of bad debt costs.

Figure 5 **Histograms of estimated coefficient on ONS transiency measure in models of total operating costs**



84. We felt that further work on the identification or development of other data sources on transiency was likely to reveal fewer insights compared with the work on other aspects of our analysis and we did not pursue that further. In the light of our reservations about the quality of the data capturing population transiency, and in the light of the results outlined above, we chose not to include such a variable in the set of models we explored in more detail and which we present below.

85. Tables 8 and 9 set out the specification of the models of bad debt costs that we explored for this phase of work.

Table 8 **Specification of models of bad debt costs**

Ref	Model specification
BD1	Dependent variable • Natural logarithm of bad debt costs per unique customer Explanatory variables • A deprivation measure (from those in Table 9 below) • Natural logarithm of revenue per unique customer
BD2	Dependent variable • Ratio of bad debt costs to domestic revenue Explanatory variable • A deprivation measure (from those in Table 9 below)

Table 9 Deprivation measures in models of bad debt costs

Suffix	Deprivation measure
_d1	Proportion of revenue from households served in LSOAs within top-20 per cent, as measured by RGC102
_d2	Proportion of revenue from households served in LSOAs within top-10 per cent, as measured by RGC102
_d3	Proportion of revenue from households served in LSOAs within top-20 per cent, as measured by IMD predicted
_d4	Proportion of revenue from households served in LSOAs within top-10 per cent, as measured by IMD predicted
_d5	Average IMD (predicted) score
_d6	Average RGC102 score

86. The selection of the deprivation measures we considered builds on our earlier work which focused on capturing deprivation in econometric models to benchmark companies' residential retail costs. Our May 2017 working paper sets out that work in detail, and Appendix 2 of this report gives a summary of that.

Estimation results

87. Tables 10 and 11 report the results of the estimated models. Table 10 relates to those models where the dependent variable is the bad debt cost per households, and Table 11 relates to those models where the dependent variable is the ratio of bad debt costs to household revenue. The tables report the estimated coefficient and, in brackets, the t-statistic for each explanatory variable. We have not reported the estimated coefficients for the time-dummy variables (or for the constant).

Table 10 Bad debt cost models: Models BD1

Model ref.	**BD1_d1**	**BD1_d2**	**BD1_d3**	**BD1_d4**	**BD1_d5**	**BD1_d6**
Dependent variable	Logarithm of bad debt cost per household					
Explanatory variables						
Logarithm of revenue per household	1.151	1.16	1.142	1.14	1.115	1.123
	(11.691)	(11.072)	(12.427)	(11.574)	(12.98)	(12.64)
Deprivation measure	1.226	1.463	1.204	1.589	3.001	-0.035
	(1.556)	(1.107)	(1.854)	(1.489)	(2.456)	(-2.265)
R-squared	0.772	0.761	0.778	0.769	0.792	0.792
Observations	71	71	71	71	71	71

Table 11 Bad debt cost models: Models BD2

Model ref.	BD2_d1	BD2_d2	BD2_d3	BD2_d4	BD2_d5	BD2_d6
Dependent variable	Ratio of bad debt costs to household retail revenue					
Explanatory variables						
Deprivation measure	0.054	0.072	0.054	0.077	0.128	-0.001
	(1.657)	(1.292)	(1.979)	(1.703)	(2.555)	(-2.282)
R-squared	0.168	0.130	0.204	0.181	0.258	0.238
Observations	71	71	71	71	71	71

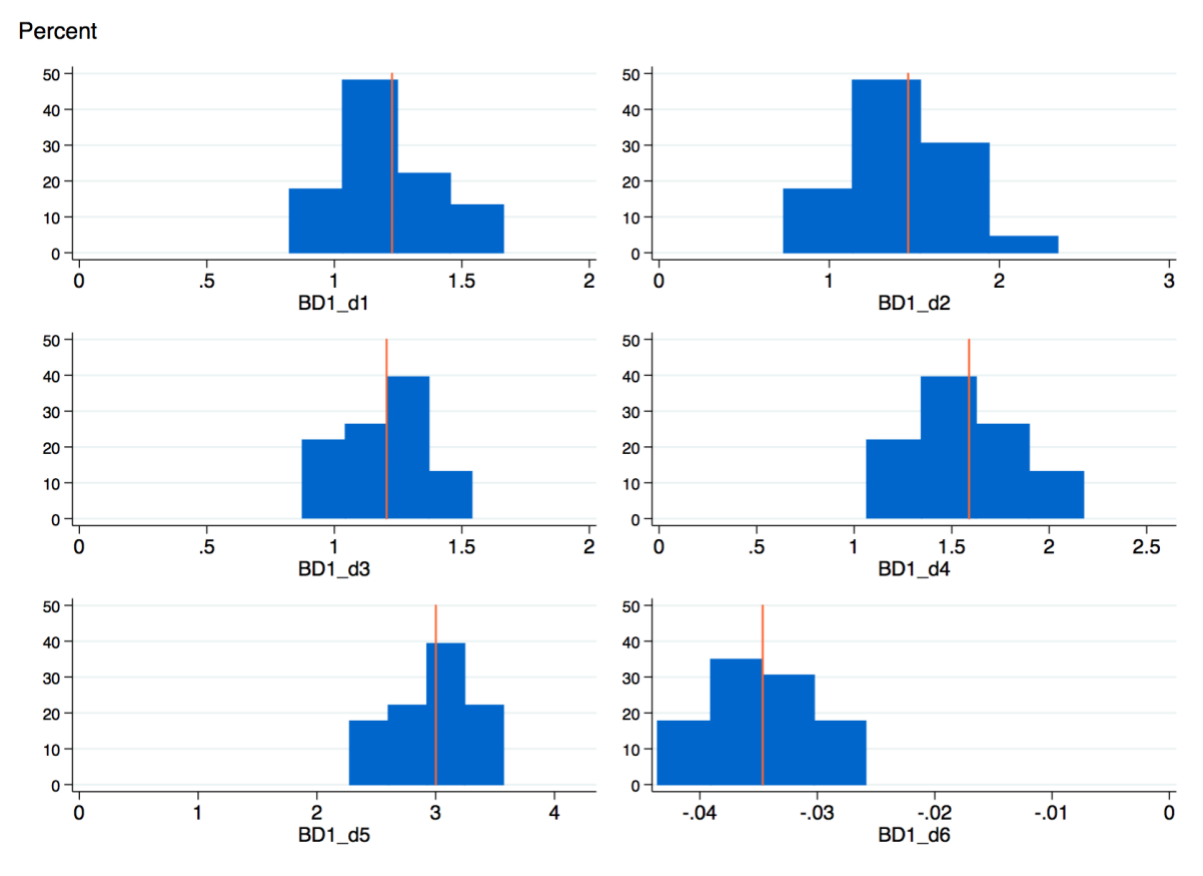
Sensitivity to changes in dataset

88. We examined the sensitivity of the results to variations in the sample period over which the models are estimated. We re-estimated each of the models after dropping, in turn, the observations for each of the four years between 2013/14 to 2016/17. We also examined the sensitivity of the results to dropping, in turn, observations for each of the 18 companies and re-estimating the model. This meant that we re-estimated each of the model for 22 different variations to the dataset. The purpose of the exercise was to examine the extent to which the results obtained from the estimation of the model on the "complete" dataset – covering all companies, and all years in the period 2013/14 to 2016/17 – held when the model was run on a modified dataset. This exercise can provide guidance on the risks that the estimated coefficients are the product of chance, rather than reflective of underlying cost relationships, which is a particular concern in the context of model estimation with a small sample size.

89. In carrying out this exercise, we were particularly interested in understanding whether the positive association between the measures of deprivation and companies' bad debt costs was found consistently across variations in the dataset.

90. Figure 6 shows a set of histograms of the estimated coefficient of the deprivation measures for each of the BD1 models – i.e. those models where the dependent variable is the logarithm of debt cost per households – across the 22 different runs of the models. The vertical line in each of the histograms marks the value of the estimated coefficient when the model is estimated on the full dataset.

Figure 6 Histogram of estimated coefficients on deprivation measures for BD1-type models



91. The set of histograms in Figure 6 shows that there is some variation in the estimated coefficient of the deprivation measures across the different runs of the same model. That is as expected. But the variation is such that the set of histograms shows, consistently across the variations to the datasets, a positive association between deprivation and bad debt costs per household. (The estimated coefficient on the measure RGC102 is negative. However, the construction of that Equifax measure is such that higher values are associated with lower arrears risk/socio-economic deprivation, and so the negative sign is as expected.)

92. The set of histograms of the estimated coefficients on the deprivation measures for the models where the dependent variable is the ratio of bad debt cost to revenue shows a similar pattern to those in Figure 6. The estimated coefficients on the measures of deprivation show a consistently positive association between deprivation and bad debt costs.

93. As shown in the results reported in Table 10, across the set of models where the dependent variable is the logarithm of bad debt cost per household (models BD1), the estimated coefficient on the explanatory variable relating to the logarithm of household revenue is around 1.1. This is the case independently of the choice of deprivation measure used. We found that the estimated coefficients for that variable

also remained close to 1.1 when we ran the models on different variations of the dataset.

Diagnostic tests

94. Tables 12 and 13 show the outcome of the diagnostic tests we carried out. These tests help in the model development stage, potentially identifying modelling issues that would need closer attention. The tables report the p-value of each test. We have colour coded the cells in the table using a 5 per cent significance level as the threshold: the light-green shade across all cells of the two tables indicate that, for each of the tests done, the tests do not reject at the 5 per cent significance level the model structures we have assumed in favour of alternatives postulated by the tests or, in the case of the Shapiro-Wilk test, in favour of non-normally distributed errors.

Table 12 Summary of diagnostic tests: Models BD1

Model Ref.	BD1_d1	BD1_d2	BD1_d3	BD1_d4	BD1_d5	BD1_d6
Ramsey RESET test for model specification	0.622	0.560	0.428	0.561	0.393	0.441
Linktest model specification test	0.238	0.185	0.292	0.244	0.608	0.591
Shapiro-Wilk test for normality of residuals	0.164	0.095	0.130	0.125	0.161	0.101

Table 13 Summary of diagnostic tests: Models BD2

Model Ref.	BD2_d1	BD2_d2	BD2_d3	BD2_d4	BD2_d5	BD2_d6
Ramsey RESET test for model specification	0.597	0.136	0.402	0.23	0.218	0.776
Linktest model specification test	0.572	0.715	0.724	0.329	0.782	0.476
Shapiro-Wilk test for normality of residuals	0.239	0.399	0.316	0.313	0.197	0.305

Comparison of actual and predicted costs

95. We have calculated the costs that the models predict for each company and contrasted these with companies' actual costs. Figures 7 and 8 show this comparison for one of the models, namely model BD1_d1. (The comparison of the results from the other models show a similar pattern). Figure 7 compares, for each company, the actual and the predicted bad debt costs per household, averaged over the four-year period from 2013/14 to 2016/17. We have anonymised the companies. For ease of reading the graph we have numbered companies 1 to 18 in ascending order of predicted bad debt cost per household. In Figure 8, we chart the ratio of companies' predicted to actual costs, providing a different perspective on the comparison to that offered by Figure 7.

Figure 7 Actual and predicted bad debt costs per household, model BD1_d1 (2013/14 to 2016/17)

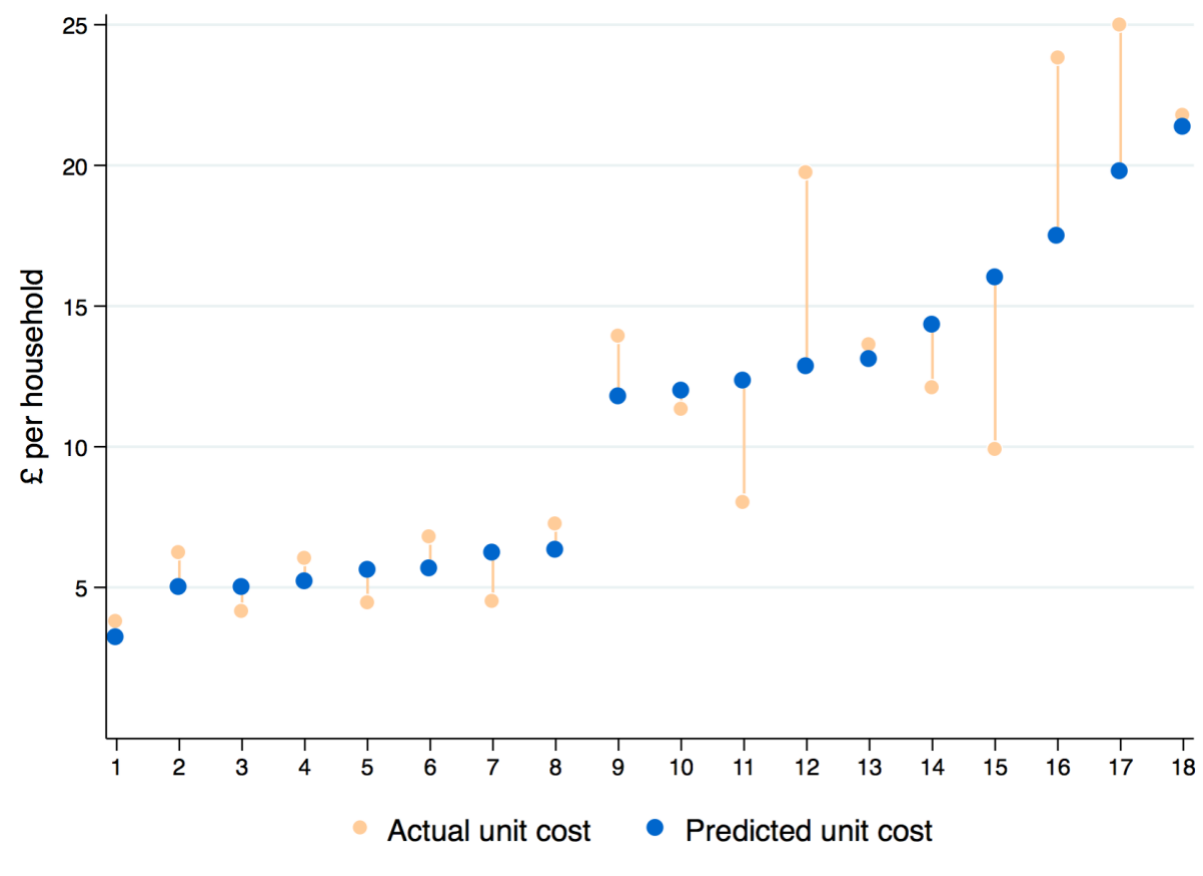
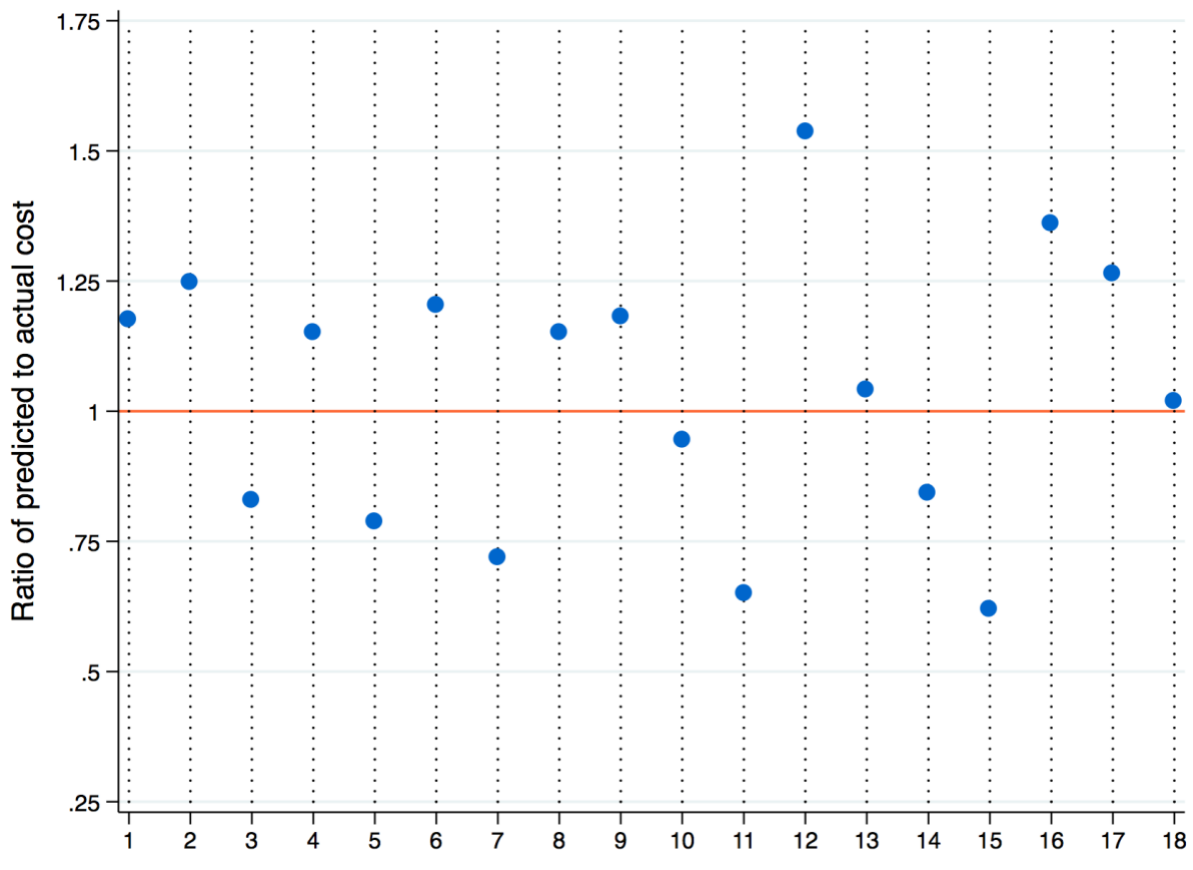


Figure 8 Ratio of predicted to actual bad debt costs, model BD1_d1 (2013/14 to 2016/17)



Discussion

96. The set of results presented above points to two over-arching findings.

97. First, controlling for differences in deprivation levels across the areas served by water companies helps explain some of the observed variation in companies' bad debt costs. We found this to be the case across the set of alternative measures of deprivation that we explored, and across the estimation results from the exercise of varying the dataset on which the model is estimated (which we carried out by dropping, in turn, the observations for a given company or the observations for a given year, and estimating the model on the smaller dataset).

98. We have calculated the effect on costs that the models predict for variations in deprivation levels across companies. Across the set of the models of type BD1, where the dependent variable is expressed as the logarithm of bad debt costs per household, the models predict that, other things equal, a change in the deprivation level from the industry-average to the highest value would increase unit bad debt costs by 19 to 26 per cent, depending on the model. Across the set of models of type BD2, such a change in deprivation level is predicted to increase the ratio of bad debt costs to revenue by 0.006 to 0.01, depending on the model. For the average company, this represents an increase in the ratio of bad debts costs to revenue of 18 to 25 per cent.

99. The predicted effect is large. Bad debt costs account for a significant share of total retail operating costs (37 per cent on average across companies). The models predict that the company serving the most deprived area will have bad debt costs that are (around) 20 to 25 per cent higher than if it served an area with average levels of deprivation.

100. It merits highlighting that the predicted size of the effect of deprivation on bad debt costs is broadly similar across the set of models. Across the two types of models we considered – type BD1, where the dependent variable is the logarithm of bad debt cost per household, and type BD2, where the dependent variable is the ratio of debt costs to revenue – and across the six different deprivation measures we drew on, we found that the magnitude of the effect of varying levels on deprivation was similar.

101. A second over-arching finding from our bad debt cost models concerns the importance of controlling for revenue per household when comparing bad debt costs. The interpretation of the results from across the set of models of type BD1 suggest that a 10 per cent increase in the billed revenue per household has an effect of increasing bad debt costs by around 11 per cent. This size of effect makes intuitive sense.

102. The results from the models indicate that there is greater estimation uncertainty in terms of the estimated scale of effects of deprivation on bad debt costs compared to the scale of the effects of average bill size on bad debt costs. This is understandable given the greater complexities associated with concepts of deprivation and how they feed through to bad debt, compared to the relationships between bill size and bad debt costs.

103. The finding from the BD1 models, that the estimated relationship between average revenue per household and bad debt costs is not far from a 1:1 relationship, is consistent with the logic for the structure of the BD2 models, which implicitly imposed the assumption that bad debt costs vary 1:1 with average revenue per household. That said, our work does not show a strong case for supplementing the BD1 models with the BD2 models. The BD2 models are more restrictive than the BD1 models, in not allowing for bad debt costs to vary with average revenue per customer to a different degree than implied by a 1:1 relationship. The BD1 models are more compatible with the view that higher levels of water and wastewater bills increase bad debt costs not just by increasing the amount of money at stake in instances of late payment and non-payment, but also by increasing the likelihood of late payment and non-payment (as more customer may struggle to pay of bills are higher). Furthermore, from the estimation results we obtained, we did not identify any countervailing advantages of the BD2 models compared to the BD1 models.

Section 5: Models of remaining residential retail operating costs

104. This section presents our analysis of models of residential remaining retail operating costs. These are defined as the total residential retail operating costs excluding the costs of debt management and the costs of bad debt.

105. The section is structured as follows:

(a) We give an overview of how residential remaining retail operating costs vary across water companies.

(b) We discuss the specification of the models explored.

(c) We present the main results from our estimation of the models.

(d) We present analysis of the sensitivity of those results to variations in the dataset.

(e) We present the outcome of statistical diagnostic tests on the models.

(f) We compare companies' actual costs with those predicted by the models.

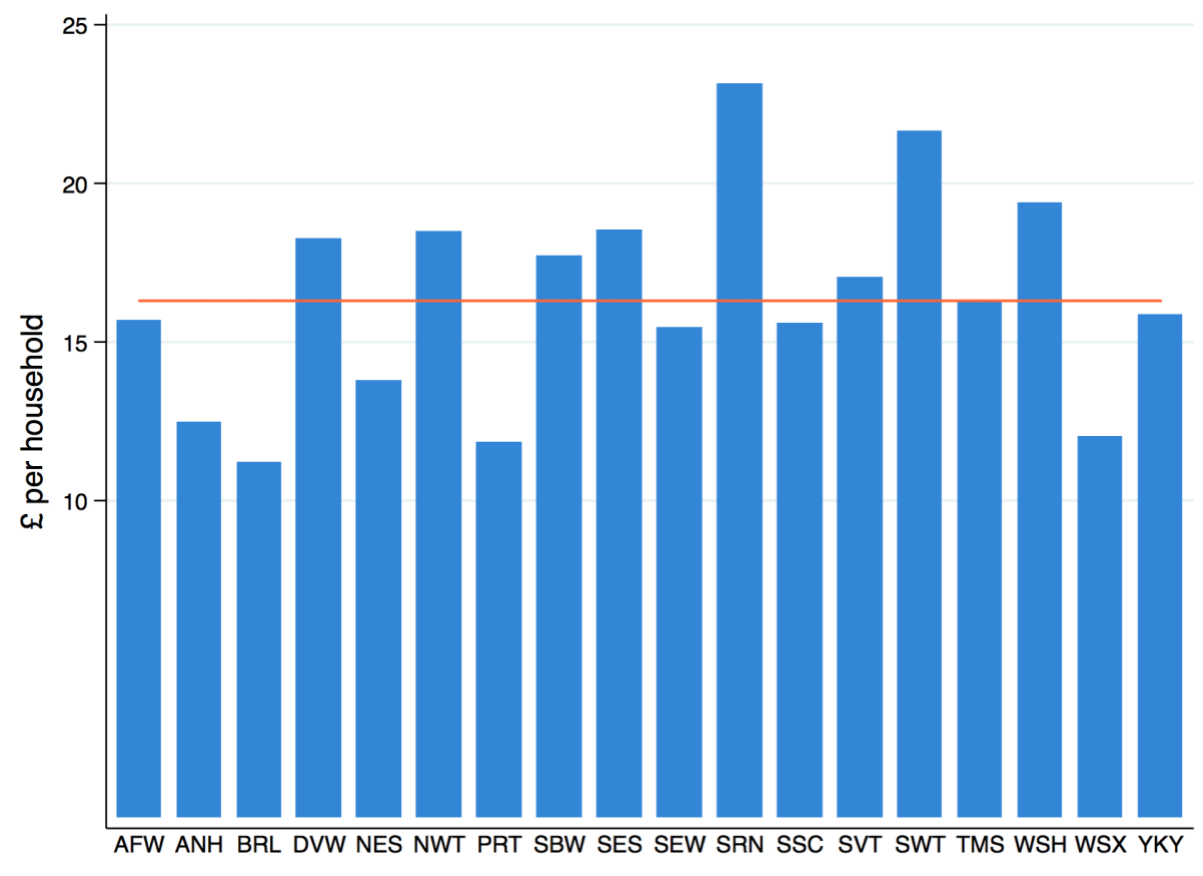
(g) We discuss the findings.

Overview of residential remaining retail operating costs

106. Across the period 2013/14 to 2016/17 and across all companies, the average annual remaining retail operating cost per household was £16.[18]

[18] Values expressed in 2016/17 prices. For Bournemouth Water and South West Water, we used data from the separate companies for years 2013/14 to 2015/16, and for the combined company for 2016/17.

Figure 9 Average annual remaining operating cost per household (2013/14 to 2016/17)



Model specification

107. The dependent variable in each of the models of residential remaining retail operating costs we explored is defined as the remaining retail operating cost per household. The dependent variable is expressed in terms of £ per household, and is not transformed by taking the natural logarithm of it.

108. We explored models that included measures of meter penetration and of the proportion of households that are dual service customers. We discussed the rationale for these in Section 3.

109. We constructed the cost driver relating to meter penetration as the ratio of the number of metered services provided to the number of households served. This implies that a metered household which takes both water and wastewater services from the same company counts as being provided with two metered services by that company. Constructing a meter penetration measure in this way recognises that, where there is a different water and wastewater provider, meter reading costs and potentially other meter related costs are shared (in some way) between the water only company and the water and sewerage company, whereas that is not the case when the same company provides both services. Calculating a meter penetration rate as we have done seeks to deal with that. That would not be achieved if, instead, we had calculated a measure of meter penetration based simply on the proportion of households that are metered. We

consider that our approach makes more sense from an intuitive perspective, given how costs are reported in the regulatory accounting data.

110. In initial phases of the work, we explored too the inclusion within the set of explanatory variables of measures relating to quality of service in our models, namely in the models of remaining retail costs and of total operating costs. For this purpose, we used data on companies' overall SIM score. Our initial results had indicated a positive figure for the estimated coefficient on this measure of quality of service. This suggested that the coefficient was capturing factors other than the underlying relationship between the quality of service provided and the costs of provision (perhaps reflecting positive correlations between cost efficiency and service quality across companies' residential retail operations).

111. We considered that further work on the development of econometric benchmarking models that capture differences in the quality of service between companies was not as promising, in terms of developing econometric models. We did not pursue this it at this stage.

112. Table 3 lists the explanatory variables (further to the time dummies) included within each model.

Table 14 Explanatory variables in models of residential remaining retail operating costs

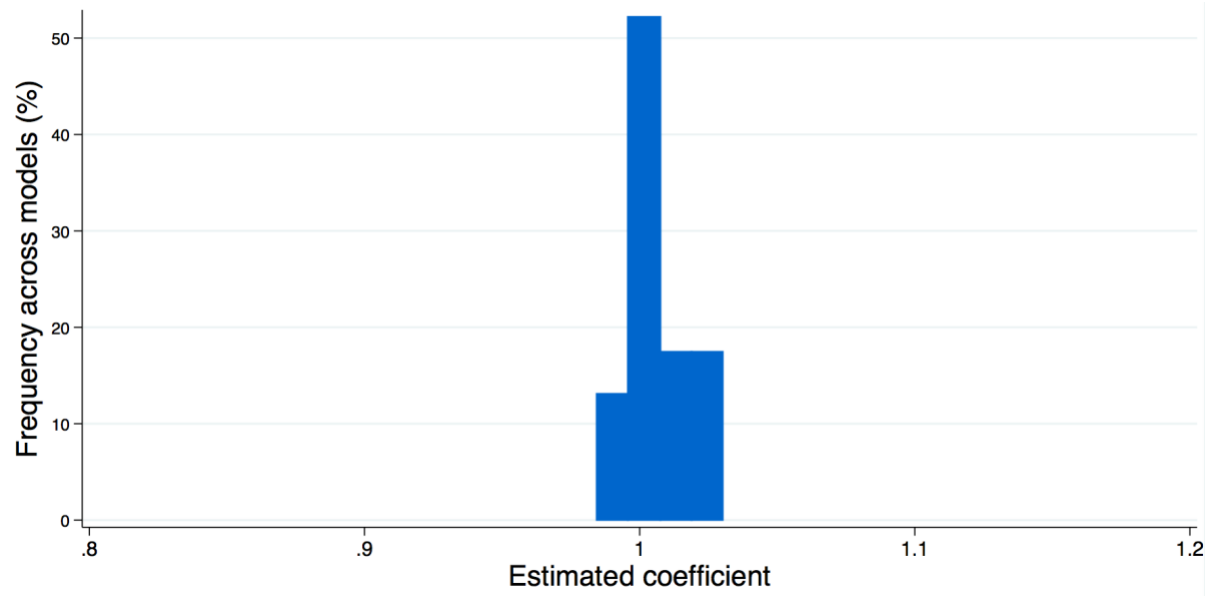
Ref	Explanatory variables
RR1	• None (i.e. this is a simple unit cost model, useful as a benchmark)
RR2	• Number of metered services per household • Proportion of households that are dual service customers
RR3	• Proportion of household customers that are dual service customers
RR4	• Number of metered services per household

113. As outlined in the table, the set of models we considered are linear models of unit cost where we have not included scale variables within the set of factors being controlled for. Implicit in these specifications is the assumption that there are no economies of scale. We carried out some analysis of whether such an assumption was reasonable, given the data. We found that it was. Our analysis pointed to there not being identifiable economies of scale, that is to say, that the cost per household does not vary with the number of households served. We observed this by running a set of models where we regressed companies' residential remaining operating costs, expressed in aggregate and not on a per household basis, against the number of households served. In such models, when both the dependent variable and the number of customers is expressed in logarithmic terms, the economies of scale can be read from the estimated coefficient for (the logarithm) of households. An estimated coefficient of 1 indicate that there are no economies of scale, i.e. that an *x* per cent change in households served is associated with an *x* per cent change in aggregate costs. Conversely, an estimated coefficient above 1 would be a signal of diseconomies of scale, and a value below 1 of economies of scale. We found the

estimated coefficient for the variable on number of households to be very close to 1, for a series of model estimations from an exercise of varying the dataset by dropping company- or year-specific data points.

114. Figure 10 illustrates this. The figure shows the histogram for the estimated coefficients on the logarithm of households that are obtained when a simple model that regresses the logarithm of remaining operating costs on the logarithm of the number of household customers is run across variations to the dataset.[19]

Figure 10 Histogram of ln(Households) in an aggregate remaining operating cost model



115. As shown in the histogram, the estimated coefficient on the variable relating to the (logarithm) of the number of households is consistently very close to 1. This lends support to the choice of exploring models of unit cost, with no requirement to include variables to allow for economies of scale.

Estimation results

116. Table 15 shows the results for models RR2, RR3 and RR4 of residential remaining operating costs.

[19] The dataset was varied by (a) for each company, dropping observations for that company; and separately, (b) for each year, dropping the observations for that year. This produces 22 variants of the dataset. The model was estimated using each of those variants of the dataset.

Table 15 Company-wide models of residential remaining operating costs

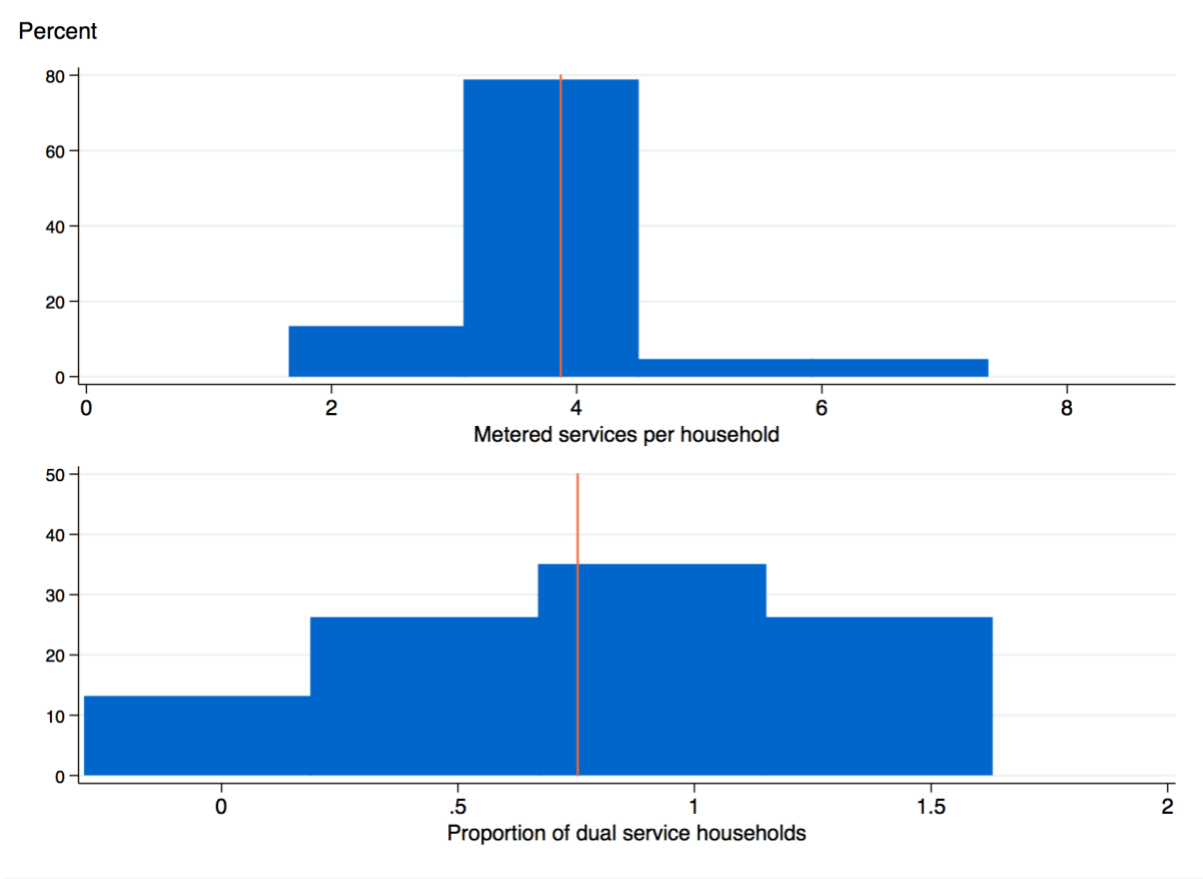
Model Ref.	RR2	RR3	RR4
Dependent variable	Remaining operating cost per household		
Explanatory variable			
Number of metered services per household	3.87		4.469
	(1.043)		(1.559)
Proportion of dual service households	0.753	2.714	
	(0.389)	(1.727)	
R-squared	0.179	0.111	0.175
Observations	71	71	71

Sensitivity to changing dataset

117. As with our analysis of the models of bad debt costs, we examined the sensitivity of the results to variations in the sample period over which the models are estimated. We re-estimated each of the models after dropping, in turn, the observations for each of the four years between 2013/14 to 2016/17. We also examined the sensitivity of the results to dropping, in turn, observations for each of the 18 companies and re-estimating the model. This meant that we re-estimated each of the model for 22 different variations of the dataset.

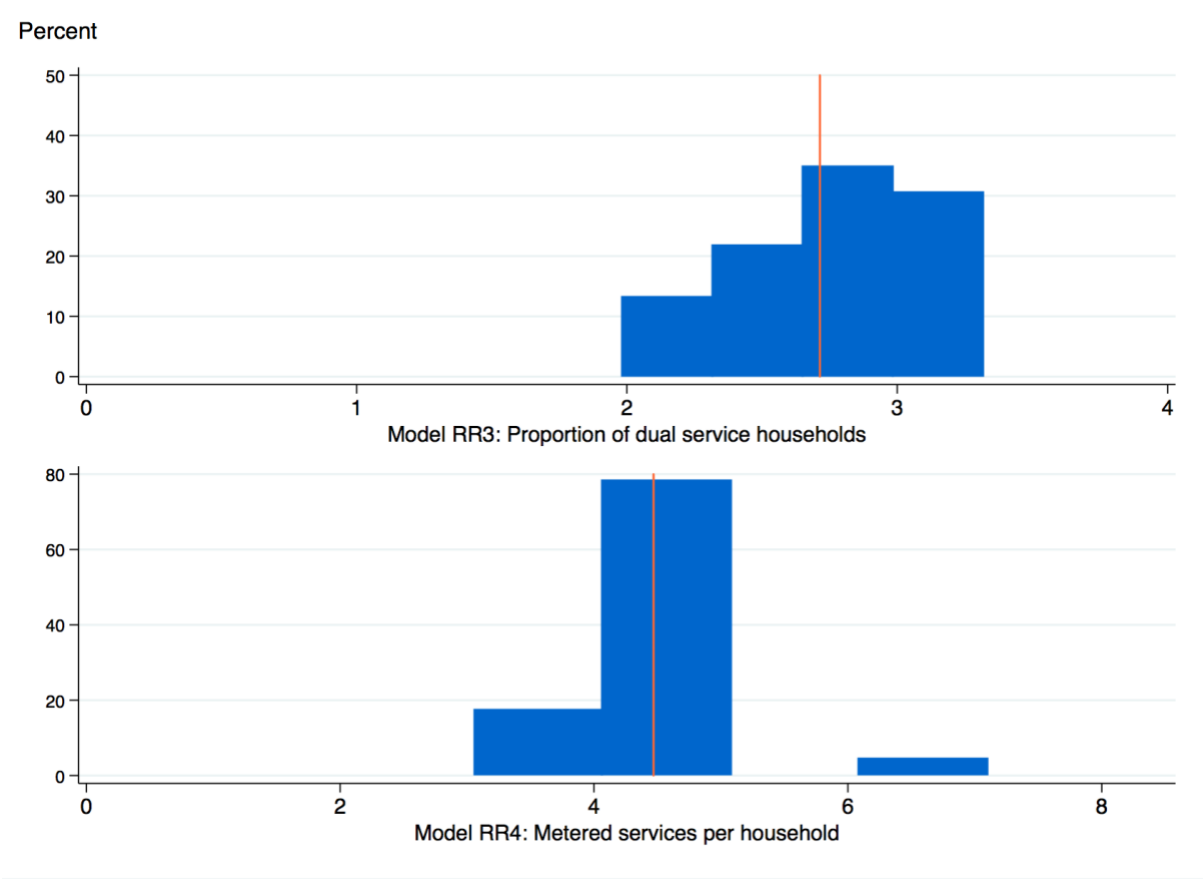
118. We find that there is considerable variation in the estimated coefficients for the two cost drivers for estimations of model RR2, where both are included within the set of explanatory variables, as shown in Figure 11. In the case of the variable relating to the proportion of dual service households, the range of the estimated coefficients across variations to the dataset include negative values. The vertical orange line in each of the histograms marks the value of the estimated coefficient when the model is run on the complete dataset.

Figure 11 Histogram of estimated coefficients for model RR2



119. Where only one of the two variables is included in the set of explanatory variables – model RR3 in the case of the number of metered service per household, and model RR4 in the case of the variable on proportion of dual service households – the range of estimated coefficients is narrower, and is consistently in the positive quadrant. This is shown in Figure 12.

Figure 12 Histogram of estimated coefficient for cost driver in models RR3 and RR4



Diagnostic tests

120. As discussed earlier in Section 3, we ran a series of statistical diagnostic tests on the models. These tests can help in the model development stage, potentially identifying modelling issues that would need closer attention. Table 16 reports the outcome of these tests on each of the models shown above. The table reports the p-values of the tests. We have colour coded the cells in the table using a 5 per cent significance level as the threshold (taking each test independently of one another).

121. The light-green colouring of all cells in the table indicate that none of these diagnostic tests raised issues in relation the three models reported.

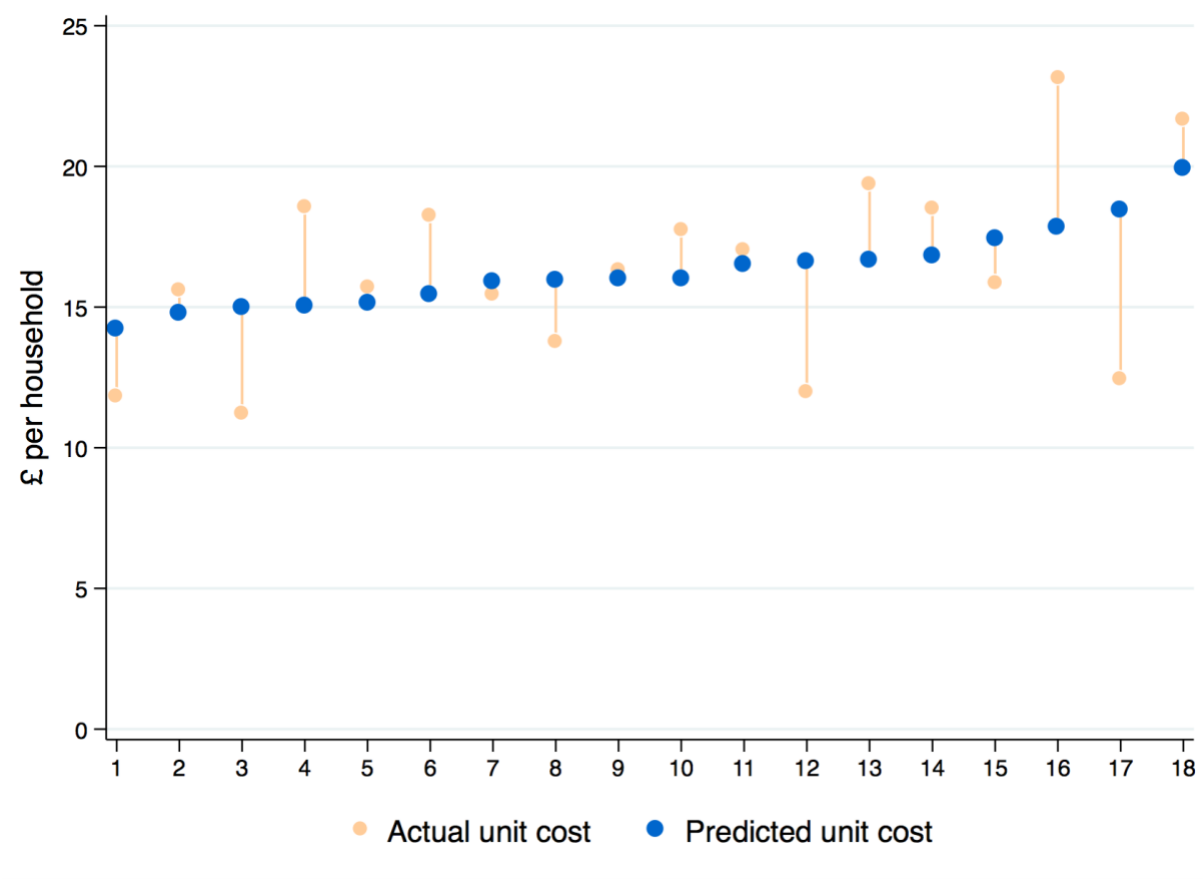
Table 16 Summary of diagnostic tests: models on remaining operating costs

Model Ref.	RR2	RR3	RR4
Ramsey RESET test for model specification	0.857	0.942	0.122
Linktest model specification test	0.801	0.146	0.553
Shapiro-Wilk test for normality residuals	0.358	0.364	0.249

Comparison of actual and predicted costs

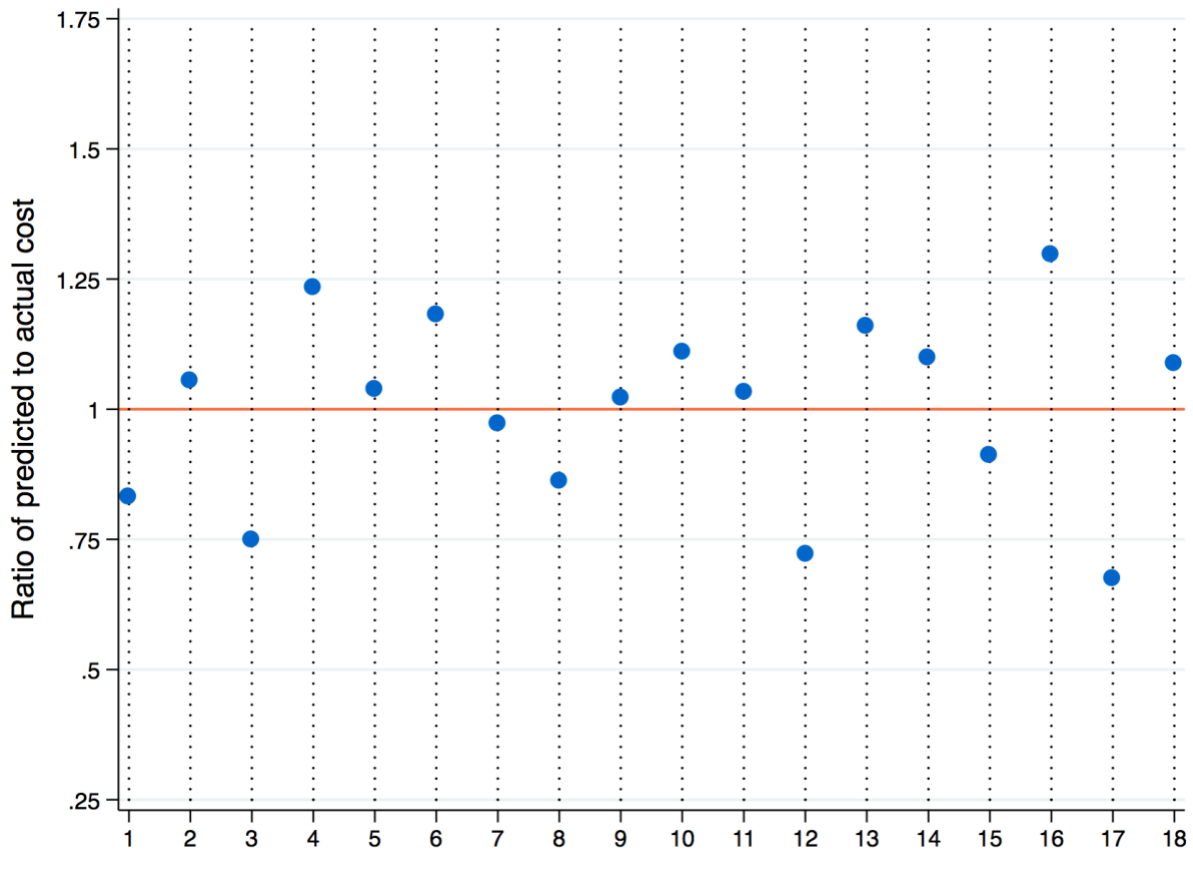
122. We have calculated the costs that the models predict for each company and contrasted these with companies' actual costs. Figures 13 and 14 show this comparison for one of the models, namely model RR2 which controls for meter penetration and for the proportion of dual service customers. Figure 13 compares, for each company, the actual and the predicted remaining operating costs per household, averaged over the four-year period from 2013/14 to 2016/17. We have anonymised the companies and have numbered companies 1 to 18 in ascending order of predicted remaining operating costs per household.

Figure 13 Actual and predicted remaining operating costs per household, model RR2 (2013/14 to 2016/17)



123. Figure 14 compares the actual and predict costs of each company by showing the ratio between the two.

Figure 14 Ratio of predicted to actual remaining costs, model RR2 (2013/14 to 2016/17)



Discussion

124. The set of results reported above points to the finding that there is a positive association between remaining operating costs per household and the measure of the average number of metered services per household. The estimated coefficient on that variables in models RR2 and RR4 suggest that, keeping other things constant, providing one additional metered service increases companies' aggregate remaining operating costs by £3.87 according to model RR2 and by £4.47 according to model RR4. This estimate is greater than the average, across companies, of their meter reading cost per metered service: for the period from 2013/14 to 2016/17 that average was around £1.90 per metered service. However, that average figure relates to the cost of meter reading alone, whilst the estimated coefficient in models RR2 and RR4 may be capturing other costs that are also driven by meter penetration. These may include, for example, costs of more frequent billing of metered customers or more query and complaint handling and processes associated with metered households.

125. With regard to the variable that relates to the proportion of dual service households, there is a wider difference between the effects predicted by the two models that

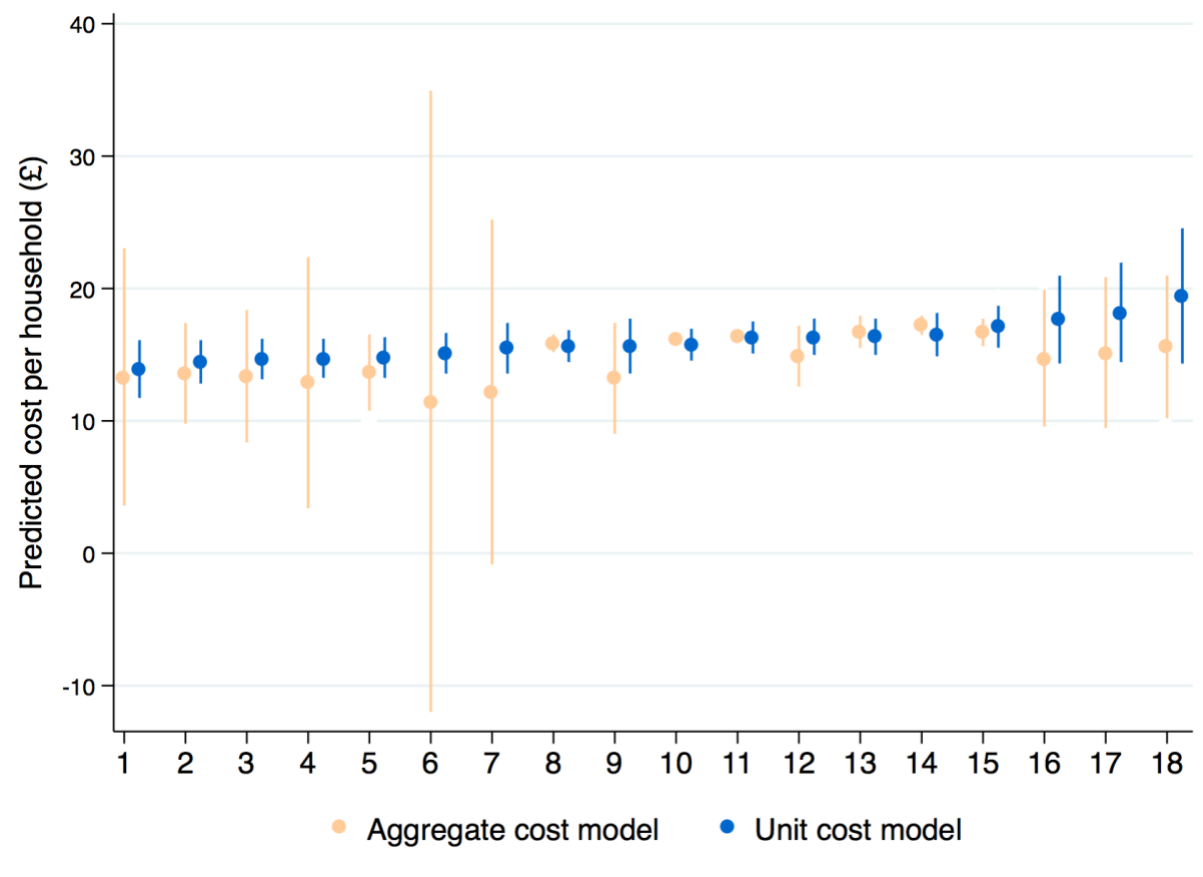
include that variable, models RR2 and RR3. The former, predicts that an additional dual service household will increase aggregate remaining operating costs by £0.75, whilst the latter predicts that the incremental cost would be £2.71. We found, however, that there was more statistical imprecision around the estimated coefficient for this variable, particularly for model RR2.

126. The results in Table 15 show that the value of the R-square statistic across the three models to be between 0.1 and 0.2. The statistic is a measure of the goodness of fit of the model. These values indicate that the models explain less than a fifth of the observed variation in companies' remaining operating costs per household. In itself, this does not detract from the validity of the findings made in relation to the estimated relation between remaining operating costs per household and the measures of meter penetration and of the proportion of dual service customers.

127. As a further check, we re-specified the three models reported above, changing their specification so that they were models of aggregate cost, rather than cost per household. In each case, we revised the explanatory variables so that these too were aggregate measures rather than "per household" measures. For example, in the revised aggregate cost version of model RR2, the dependent variable was the aggregate remaining operating cost and the set of explanatory variables included the number of metered services provided and the number of dual service households, as well as year-specific dummy variables. Across those three aggregate models we find that the R-squared statistic is close to 0.9, suggesting that those models fit the data better. This is not surprising: the scale of a company is the key factor in explaining its aggregate costs and, in aggregate models the set of explanatory variables includes a measure that reflect companies' scale. In contrast, in the unit cost version of the models, we are, by construction, controlling for the differences in scale and the model is instead, focusing on explaining the variations in the cost per household across companies.

128. We found that the aggregate cost version of the models produced estimates of predicted expenditure that tended to be less precise than those derived from the unit cost models. In particular, we found that, on average across the 18 companies, the standard deviation of the predicted expenditure was higher in the aggregate cost models, though for some companies the converse was the case. Figure 15 illustrates this. It compares the 95 per cent confidence interval for expenditure per household predicted for 2016 by model RR2 when this is specified in its "unit cost" version, and when it is specified in its "aggregate cost" version. A similar picture emerges for the predicted values for other years, and for a comparison of unit cost and aggregate cost versions of the other models.

Figure 15 **95 per cent confidence interval of predicted cost of unit cost and of aggregate cost version of model RR2 (2016)**



129. On average, across the 18 companies, the standard deviation of the predicted values shown in the figure was 1.6 for the unit cost model and 1.8 for the aggregate cost model. For the purpose of cost assessment, the precision on the predicted expenditure should be a criterion of interest: other things equal, we prefer a model that yields more statistically precise estimates of predicted expenditure. In the light of this, and of our earlier discussion, we do not consider that the R-squared statistic for models R2 to R4 being between 0.1 and 0.2 provides a reason to reject these models in favour of aggregate cost models, or in favour of a non-econometric approach to benchmarking. Our overall modelling results indicate that the vast majority of differences in residential retail costs between companies (leaving aside bad debt costs) can be explained by differences in the number of households supplied and that part of the remainder can reasonably be attributed to differences in the levels of metering and in dual service provision. The results of the models with regard to the cost of providing one additional metered service seem reasonable and are consistent with our prior expectations, grounded on companies' metering costs. At the same time, the low R-squared indicates that there may be opportunities in the future to improve on the design of the models for the remaining residential retail costs.

Section 6: Models of total residential retail operating costs

130. This section presents our analysis to develop econometric models to benchmark measures of water companies' total operating costs relating to the provision of retail services to households.

131. The section is structured as follows:

(a) We discuss the specification of the models explored.

(b) We present the main results from our estimation of the models.

(c) We present analysis of the sensitivity of those results to variations in the dataset.

(d) We present the outcome of statistical diagnostic tests on the models.

(e) We compare companies' actual costs with those predicted by the models.

(f) We discuss the findings.

Model specification

132. The dependent variable in each of the models of total residential retail operating costs is the natural logarithm of total residential retail operating cost per unique customer.

133. Tables 17 and 18 list the explanatory variables (further to the time dummies) included in each model. These draw on explanatory variables introduced and discussed in Sections 3, 4 and 5.

Table 17 Explanatory variables included in models of total residential retail operating costs

Ref	Explanatory variables
RT1	• None (i.e. this is a simple unit cost model, useful as a benchmark)
RT2	• A deprivation measure (from set in Table 18 below) • Natural logarithm of revenue per household
RT3	• A deprivation measure (from set in Table 18 below) • Proportion of households that are dual service customers • Number of metered services per household • Variable constructed to capture difference between a company's average revenue per household customer and the average revenue per household it would have if it had industry-average bills, taking account of its mix of water-only, wastewater-only and dual service customers
RT4	• A deprivation measure (from set in Table 18 below) • Proportion of households that are dual service customers • Variable constructed to capture difference between a company's average revenue per household customer and the average revenue per household it would have if it had industry-average bills, taking account of its mix of water-only, wastewater-only and dual service customers

Table 18 Deprivation measures included in models of total residential retail operating costs

Suffix	Deprivation measure
_d1	Proportion of revenue from households served in LSOAs within top-20 per cent, as measured by RGC102
_d2	Proportion of revenue from households served in LSOAs within top-20 per cent, as measured by IMD predicted
_d3	Average RGC102 score
_d4	Average IMD (predicted) score

134. Two of the candidate cost drivers, household revenue per household and the proportion of dual service customers are highly correlated; the correlation coefficient is 0.91, based on data averaged over 2013/14 to 2016/17. All the same, we considered it appropriate to explore models that included both as the potential routes through which they may have an impact on operating costs differ:

(a) Household revenue per household affects total retail costs through its influence on debt costs, as discussed in our analysis of the models of debt costs.

(b) The proportion of dual service customer may have an effect on total retail costs through its potential effect on the costs of dealing with customers' queries.

135. To deal with the multicollinearity, we have constructed a variable that is designed to capture the differences between companies in their average bill, having first controlled for differences in the mix of customers they serve. This measure, which we have called *billratio*, is constructed as follows:[20]

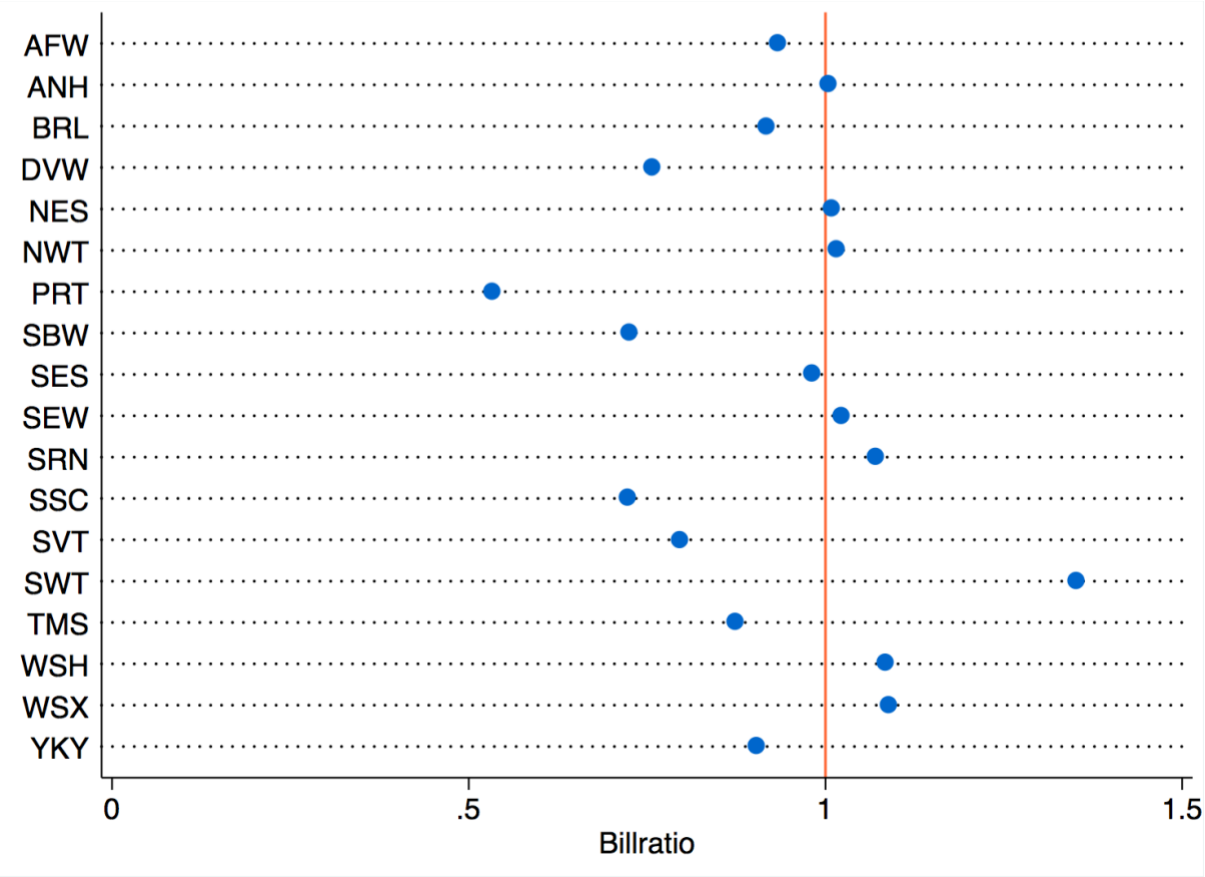
Bill ratio = Company's revenue per household / (p_water * Industry average water only bill + p_waste * Industry average waste only bill + p_dual * Industry average dual service bill)

where p_service is the proportion of company's household customers that are provided with the relevant service.

136. The definition of *billratio* is such that it will be 1 for a company for which, for each customer group, its average revenue per household, is the same as the industry average. The value of *billratio* will be above 1 for a company that tends to have higher bills than industry average, and be below 1 if it tends to have lower bills than the average. Figure 16 shows the value of *billratio* for each of the companies, averaged over the 2013/14 to 2016/17 period.

[20] In calculating this variable, we took account of the £50 government contribution to the bills of customers of South West Water.

Figure 16 Comparison of variable *Billratio*, average 2013/14 to 2016/17



Estimation results

137. The following three tables report the results for the range of models estimated.

Table 19 Retail household total operating cost model: Models RT2

Model Ref.	RT2_d1	RT2_d2	RT2_d3	RT2_d4
Dependent variable	Logarithm of total operating cost per household			
Explanatory variables				
Logarithm of revenue per household	0.521	0.508	0.512	0.498
	(6.125)	(6.328)	(6.067)	(6.235)
Deprivation measure	0.289	0.456	-0.009	1.13
	(0.569)	(1.04)	(-0.857)	(1.29)
R-squared	0.638	0.652	0.645	0.660
Observations	71	71	71	71

Table 20 Retail household total operating cost model: Models RT3

Model Ref.	RT3_d1	RT3_d2	RT3_d3	RT3_d4
Dependent variable	Logarithm of total operating cost per household			
Explanatory variables				
Metered services per household	0.105	0.135	0.137	0.136
	(0.378)	(0.533)	(0.526)	(0.558)
Proportion of dual service customers	0.29	0.254	0.212	0.21
	(1.091)	(1.221)	(0.905)	(1.049)
Billratio	0.738	0.73	0.804	0.772
	(2.082)	(2.339)	(2.274)	(2.505)
Deprivation measure	0.624	0.751	-0.02	1.885
	(0.707)	(1.274)	(-1.153)	(1.595)
R-squared	0.665	0.681	0.678	0.692
Observations	71	71	71	71

Table 21 Retail household total operating cost models: Models RT4

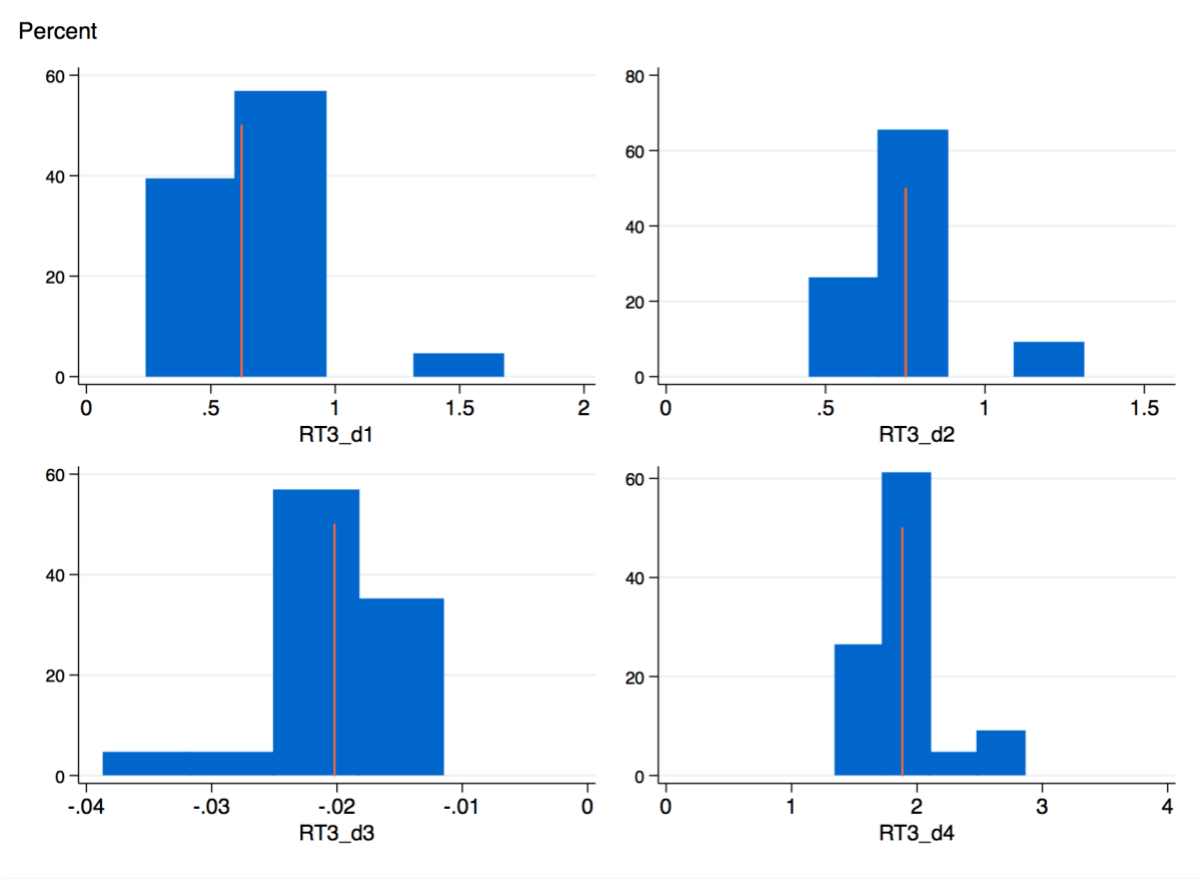
Model Ref.	RT4_d1	RT4_d2	RT4_d3	RT4_d4
Dependent variable	Logarithm of total operating cost per household			
Explanatory variables				
Proportion of dual service customers	0.36	0.323	0.296	0.278
	(1.812)	(2.011)	(1.494)	(1.698)
Billratio	0.772	0.81	0.861	0.859
	(2.317)	(3.011)	(2.643)	(3.267)
Deprivation measure	0.439	0.606	-0.016	1.62
	(0.546)	(1.09)	(-0.905)	(1.405)
R-squared	0.662	0.676	0.672	0.686
Observations	71	71	71	71

Sensitivity to changes in dataset

138. We examined the sensitivity of the results to variations in the sample period over which the models are estimated. We re-estimated each of the models after dropping, in turn, the observations for each of the four years between 2013/14 to 2016/17. We also examined the sensitivity of the results to dropping, in turn, observations for each of the 18 companies and re-estimating the model.

139. Figure 17 shows a set of histograms of the estimated coefficient across the different runs of the models of the deprivation measures for each of the RT3 models, which besides controlling for deprivation, also control for the number of metered services per household, for the proportion of dual service customers and for *billratio*. The vertical line in each of the histograms marks the value of the estimated coefficient when the model is estimated on the full dataset.

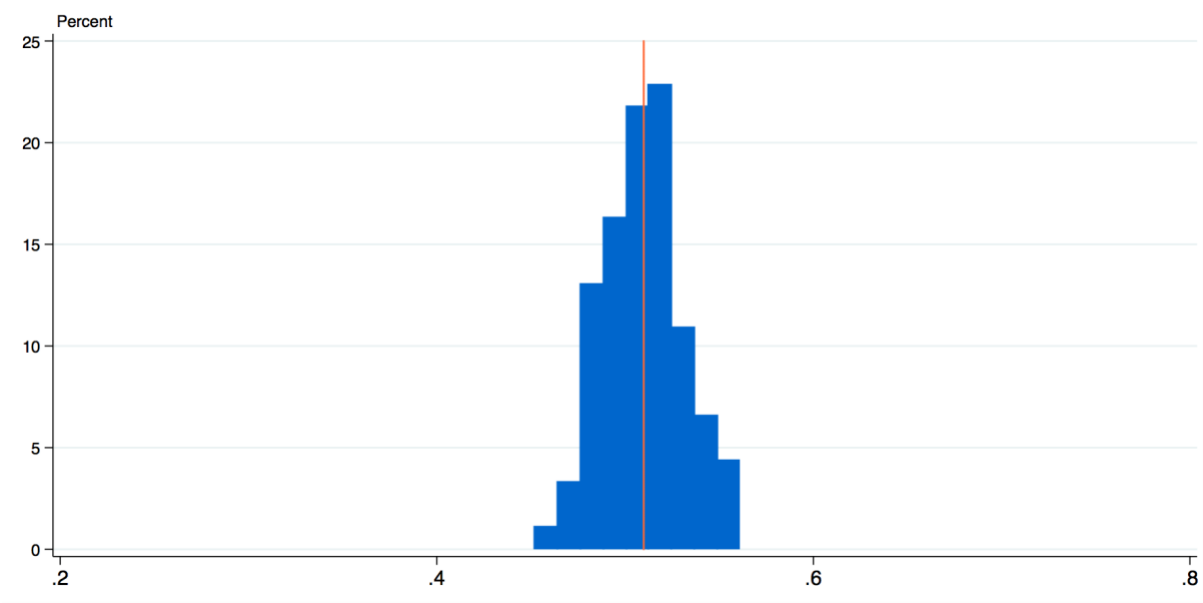
Figure 17 Histogram of estimated coefficients on deprivation measures for RT3-type models



140. The set of histograms in Figure 17 shows that, as would be expected, there is some variation in the estimated coefficient of the deprivation measures across the different runs of the same model. All the same, the set of histograms shows that consistently, across the variations to the datasets, there is a positive association between deprivation and total operating cost per household. (The estimated coefficient on the deprivation measure for model RT3_d3 is negative, but that is because the deprivation measure in that model is the Equifax variable RGC102 and this is constructed such that higher values are associated with lower arrears risk/socio-economic derivation, and so the negative sign is as expected.) We find similar patterns across the analogous set of histograms for the models RT2 and RT4. (The single exception concerns RT2_d1, where the estimated coefficient for the relevant measure of deprivation included is a (very small) negative number in one of the runs of that model).

141. We find that in the models of type RT2, the estimated coefficient on the (logarithm) of household retail revenue per household is consistently within a fairly narrow range, between around 0.45 and 0.55. This is shown in Figure 18.

Figure 18 Histogram of estimated coefficient for logarithm of revenue per household in models of type RT2



142. We also find that the estimated coefficient on the variable *billratio* is also consistently positive, both in the models of type RT3 and in the models of RT4, although there is some variation across the runs of the models.

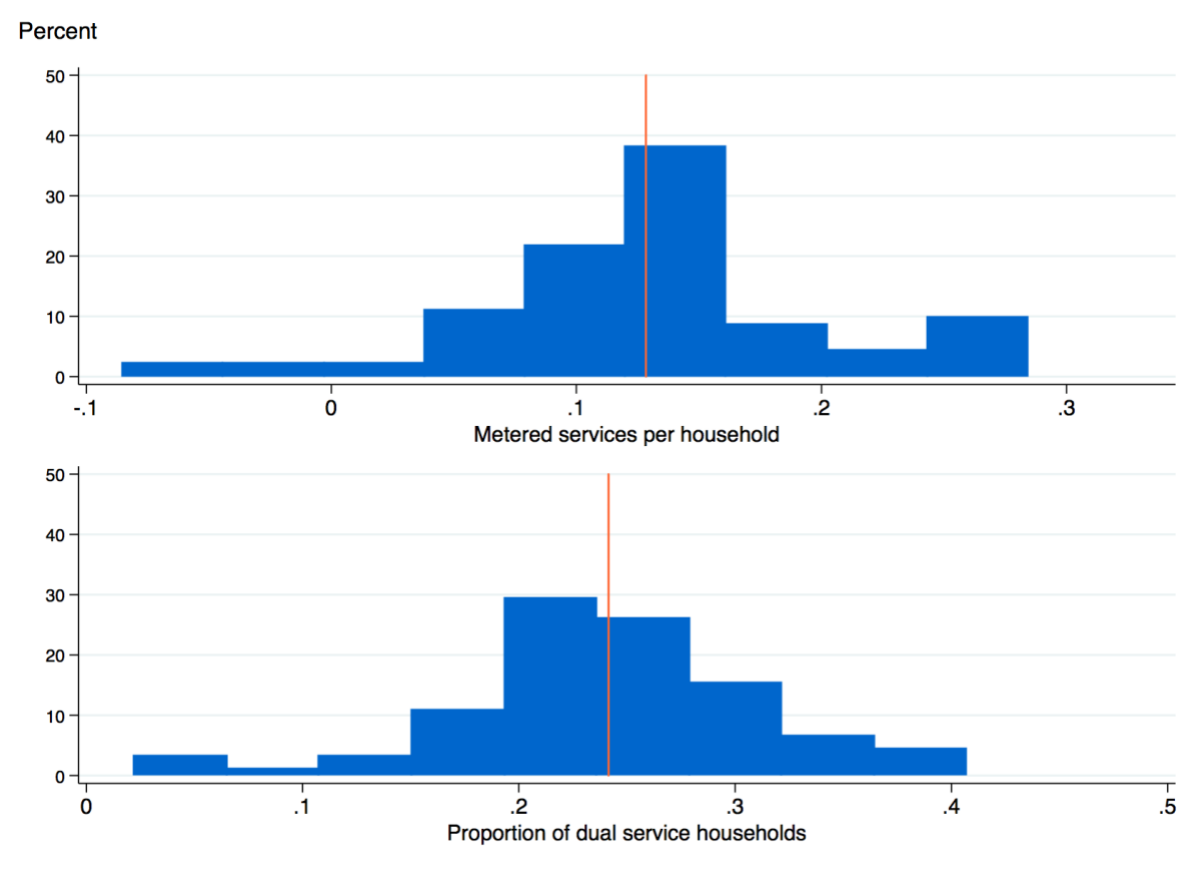
143. With regard to the variation in the estimated coefficients of the other variables explored:

    (a) We find considerable variation in the estimated coefficient on the variable relating to number of metered services per household, relating to model type RT3. For some (albeit a small number) of the runs of the data, the estimated coefficient was negative.

    (b) With regard to the variable relating to the proportion of dual service households, we find that the estimated coefficient for that variable is consistently positive across models of type RT3 as well as across models of type RT4.

144. Figure 19 illustrates the variation of the estimated coefficients for those two variables across the models of type RT3. The histograms in the figure reflect the distribution of the estimated coefficients for those variable across the variations to the dataset, as well as across the different choice of measure used to capture deprivation.

Figure 19 Histogram of estimated coefficient for cost drivers in models of type RT3



Diagnostic tests

145. We ran a series of diagnostic tests on each of the models presented here, as discussed earlier in Section 3. These tests can help as part of the model development process, identifying potential statistical issues that would need closer attention. The three tables below report the p-values of the diagnostic tests on each of the models of total operating costs discussed above. As with the reporting of the outcome of these tests for the other set of models, we have colour coded the cells in the table, applying a 5 per cent significance level as the threshold.

146. We can see that some issues are indicated by these tests, particularly for the models with the d1 deprivation measure (proportion of revenue from households served in LSOAs within top-20 per cent, as measured by RGC102) but this is for a minority of cases.

Table 22 Summary of diagnostic tests: models RT2

Model Ref.	RT2_d1	RT2_d2	RT2_d3	RT2_d4
Ramsey RESET test for model specification	0.037	0.080	0.037	0.271
Linktest model specification test	0.043	0.064	0.050	0.078
Shapiro-Wilk test for normality of residuals	0.046	0.067	0.063	0.110

Table 23 Summary of diagnostic tests: models RT3

Model Ref.	RT3_d1	RT3_d2	RT3_d3	RT3_d4
Ramsey RESET test for model specification	0.107	0.169	0.161	0.232
Linktest model specification test	0.336	0.371	0.403	0.477
Shapiro-Wilk test for normality of residuals	0.028	0.063	0.053	0.060

Table 24 Summary of diagnostic tests: models RT4

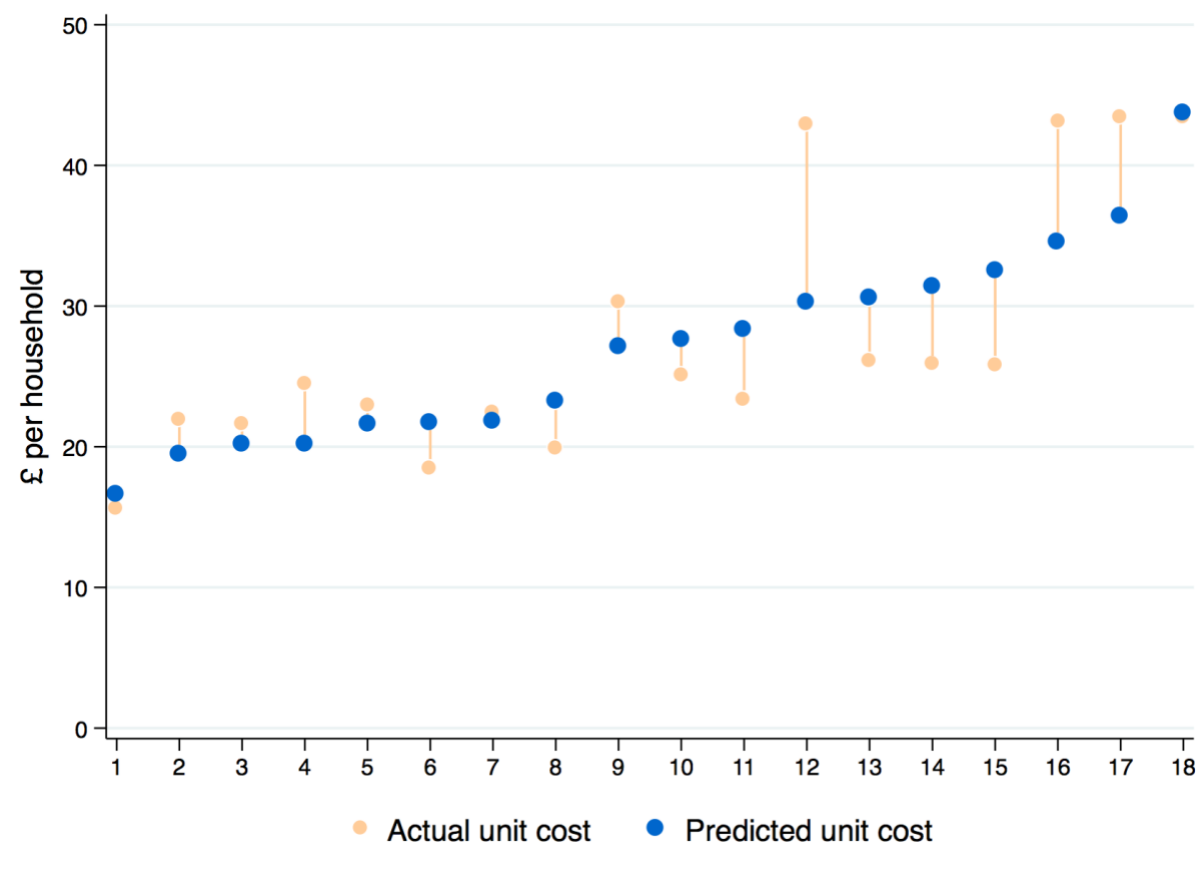
Model Ref.	RT4_d1	RT4_d2	RT4_d3	RT4_d4
Ramsey RESET test for model specification	0.060	0.208	0.148	0.615
Linktest model specification test	0.161	0.205	0.192	0.303
Shapiro-Wilk test for normality of residuals	0.031	0.067	0.071	0.095

Comparison of actual and predicted costs

147. We have calculated the costs that the models predict for each company and contrasted these with their actual costs.

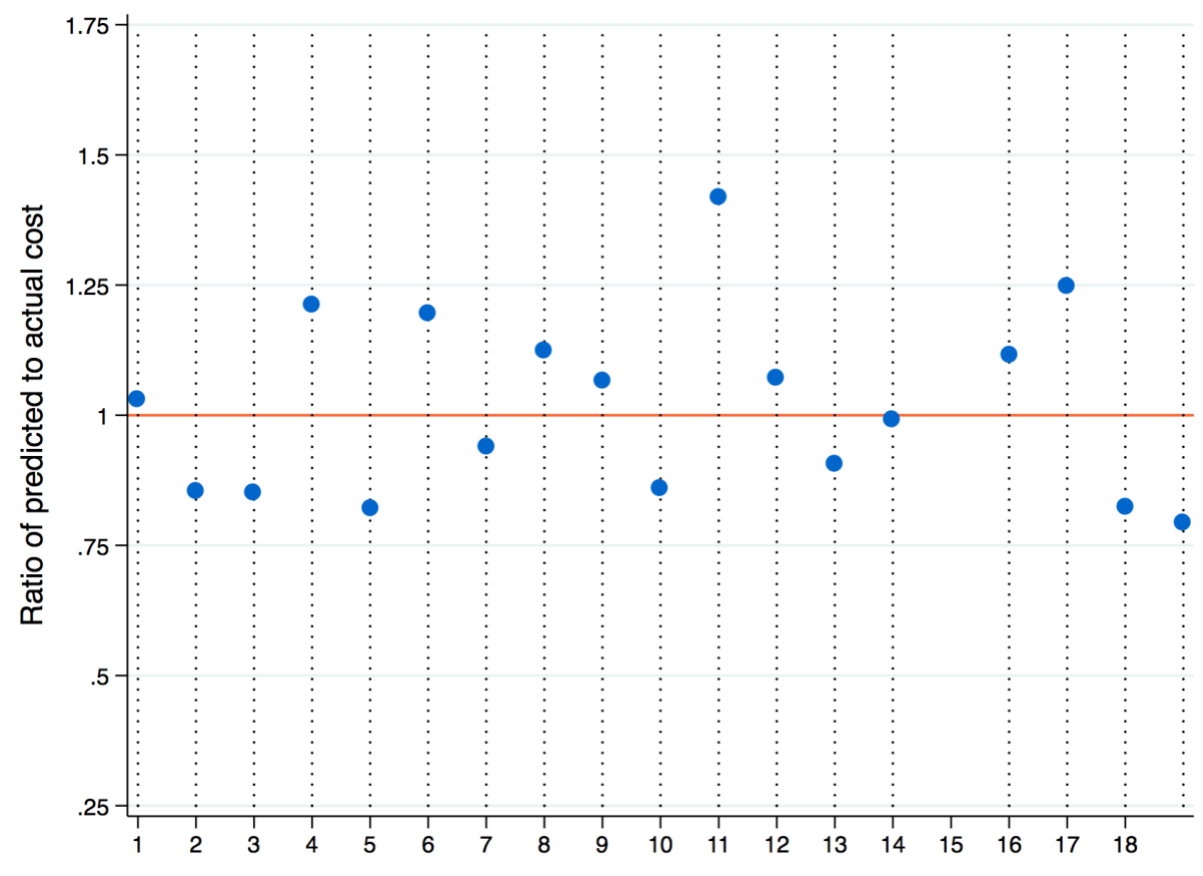
148. Figures 20 and 21 show this comparison for one of the models, namely model RT3_d1. Figure 20 compares, for each company, the actual and the predicted total operating costs per household, averaged over the four-year period from 2013/14 to 2016/17. As with the comparison of the analogous figures for debt cost and for remaining operating cost models, we have anonymised the companies. We have numbered companies 1 to 18 in ascending order of predicted total operating costs per household.

Figure 20 Actual vs predicted total operating costs per household, model RT3_d1 (2013/14 to 2016/17)



149. Figure 21 compares the actual and predict costs of each company by showing the ratio between the two.

Figure 21 Ratio of predicted to actual total operating costs, model RT3_d1 (2013/14 to 2016/17)



Discussion

150. The set of results for the models of total operating costs echoes some of the findings from the models of debt costs and of remaining operating costs.

151. The differences in deprivation levels help explain some of the observed variation in companies' total retail operating costs. The estimated coefficients on the alternative deprivation measures used point to a consistently positive relation between the levels of deprivation and the measure of total retail operating costs, although there is a degree of imprecision around those estimates. We found this to be the case across the set of alternative measures of deprivation that we explored, and across the results obtained from the exercise of varying the dataset over which the models are estimated.

152. We calculated the size of the effects predicted by the models concerning the impact of differences in deprivation levels on company's total retail operating costs. The magnitude of that effect varies across the three different types of model specifications, and to a less extent, across the choice of measure used to control for deprivation. Across the models of type RT2, which only control for deprivation levels and revenue per household, a change in deprivation level from the average to the highest level across the industry is predicted to increase unit total operating costs by 5 to 9 per cent,

depending on the deprivation measure. For models of type RT3, the predicted effect ranges from 10 to 16 per cent, and for models of type RT4 from 7 to 14 per cent. The size of this predicted effect is large, although, in percentage terms, it is smaller than that which we had derived for the models of bad debt costs (which we had estimated to be in the range of 19 to 26 per cent).

153. The set of models of type RT2 include the (logarithm) of revenue per household as an explanatory variable. The estimated coefficients for that variable across the set of RT2 models are all close to 0.5. The interpretation of these estimates is that a 10 per cent, say, increase in revenue per household is predicted to increase unit total operating costs by 5 per cent, all other things equal.

154. In the remaining set of models, we included the variable *Billratio* to capture companies' average bill, whilst controlling for differences in the mix of customers served. In those models, we estimated the coefficient for this variable to be around 0.8. This result can be interpreted to say that if a company whose average bills were the same as those of the industry average were to increase its average bill by 10 per cent, the models predict that its unit total operating costs would increase by 8 per cent (calculated as the exponential of 0.8 * 0.1). The estimated size of that effect would be smaller if, instead, the company's starting position was that it had lower average bills than the industry average, and would be larger if the opposite were the case.

155. Overall, we found greater sensitivity and imprecision in the estimated coefficients for the models of total retail operating costs presented in this section, than for the models that take debt costs and remaining operating costs separately (as set out in Sections 4 and 5). This was particularly so for the variables relating to meter penetration and dual service. At least from the work presented in this report, the more granular models, which differentiate between debt costs and remaining operating costs, seem a better basis for econometric benchmarking of residential retail costs than the models that are intended to cover the whole of residential retail operating costs. However, it is possible that there is some countervailing benefit of the latter type of models.

Section 7: Findings on capturing retail cost drivers in econometric models

156. Ofwat has said that for its PR19 residential retail cost assessment it plans to use econometric models for benchmarking costs across companies, rather than use benchmarks based on simple unit cost comparisons.

157. The work we have carried out, to develop and test econometric models of companies' residential retail costs, shows that it should be possible for Ofwat to use econometric models for PR19, rather than a unit cost approach. An econometric approach would allow the benchmarking analysis to take account of the multiple drivers of residential retail costs in a systematic and evidence-based way. The key question is not whether such models are perfect, but how they compare against alternative feasible approaches to cost assessment.

158. We consider that our model development work has achieved significant progress, and provides valuable insight, in terms of the way that a number of residential retail cost drivers could be treated as part of the specification of econometric models to be used for Ofwat's benchmarking purposes.

159. Below, we take a number of different candidate cost drivers in turn, and pick out the main findings we have drawn from the more detailed model development and estimation work we have carried out (the results of which are presented in Sections 4, 5 and 6 above).

160. Ahead of that, a more general comment. Our modelling results reinforced our view on the benefits of using models defined at a more granular level than total residential retail costs. We produced separate models for (a) costs relating to bad debt and debt management; and for (b) other retail costs. Our view is that these models are better able to capture (or approximate) relationships between costs and cost drivers than an approach based on models for total residential retail operating costs. We would certainly not rule out the consideration of models of total residential retail costs, and these models may, in some circumstances, have some countervailing benefits. For example, they may have a role if there are concerns about potential inconsistencies or limitations in companies' approaches to allocating costs.[21] All the same, models of total residential retail costs do not seem the best starting point for Ofwat's development and estimation of econometric models for residential retail costs, and risk becoming a distraction.

161. The points below are based on our work to-date, and are not intended to provide a final word on any aspect of retail cost model development.

Number of households supplied with retail services

162. The number of households supplied with retail water and wastewater services is clearly a cost driver for residential retail costs.

[21] Furthermore, the more aggregate models could provide a check on the use of any frontier or upper quartile benchmarks derived from taking the results from more granular models in isolation. However, in line with Ofwat's approach for PR14, we consider that the proper approach to combining results across different models is first to combine them to produce estimated costs in £, and then to make any assessments of the upper quartile, rather than assessing the upper quartile individually for each model.

163. The models we developed capture this cost driver by specifying the dependent variable as a measure of costs per household (other than for the special category of models that use a measure of bad debt costs divided by a measure of billed revenue).

164. This approach imposes an assumption that, within the sample of water companies covered by the analysis, there are no significant economies of scale with respect to customer numbers (or at least, none that should be allowed for as part of Ofwat's cost assessment). We found that this was not an unreasonable assumption to make. In particular, we carried out some exploratory analysis of models in which the dependent variable was a measure of aggregate costs and which included household numbers as one of the explanatory variables. The estimated coefficient for the variable on household numbers in such aggregate cost models suggested that our modelling assumption of no significant economies of scale across the sample dataset was reasonable. This finding echoes those reported in Anglian Water's recent analysis of benchmarking models for residential retail services.[22]

Measures of economic deprivation and arrears risk within geographic areas served

165. We consider that one of the important drivers of differences between water companies, in terms of their residential retail costs, is the substantial variation across different parts of England and Wales in the levels of economic and social deprivation.

166. For PR14, Ofwat made some allowances for differences in deprivation via special cost factor adjustments. For PR19, Ofwat plans to use econometric modelling and this provides an opportunity to take account of these differences between companies, in the levels of deprivation in the areas they supply, through the explanatory variables in the econometric models. This would also help Ofwat to achieve its aim of adopting a more symmetric approach to special cost factors for PR19.

167. We carried out a wide range of modelling using variables derived from Equifax data. Our results are consistent with the view that economic deprivation is a significant driver of companies' bad debt costs (and hence total retail operating costs of which bad debt is a major component). We had also found support for this in the analysis we carried out in our May 2017 working paper where we used data from United Utilities at the LSOA level to develop models of bad debt costs.[23]

168. The deprivation variables we developed based on Equifax data address the key shortcomings, for the purposes of Ofwat's retail cost assessment, in the available data on deprivation from the DCLG and Statistics for Wales.

169. Our modelling indicates that our explanatory variables on deprivation and arrears risk, derived from Equifax data, can be incorporated into models of either total retail operating costs or bad debt costs and give reasonable results.

---

[22] Anglian Water (2017) "Water industry cost modelling. Anglian Water's approach and initial results". See Figure 3 in Annex 12 which presents the results of those models where the dependent variable is the logarithm of aggregate retail operating costs and where the explanatory variables include the logarithm of the number of households. For the models of total retail operating costs where that specification is used, Anglian Water reports the estimated coefficient on the logarithm of the number of households to be very close to 1, indicating no economies of scale.

[23] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities.

170. Our results indicate some degree of imprecision and sensitivity in the model estimation results. This is not a surprise given the small sample size and the likely complexity of the underlying relationships between forms of economic deprivation and risks of late- or non-payment of water bills. This is more of an issue for the models covering total retail operating costs than models focused on bad debt costs. This does not detract from the strength of evidence that, consistent with intuition, differences across England and Wales in the extent of economic deprivation are a driver of residential retail costs. It seems important for Ofwat to make allowances for this variation in deprivation as part of its residential retail cost assessment. Doing so through the econometric models provides a way for Ofwat to make evidence-based and systematic allowances for the effects of this variation on companies' efficient levels of retail costs (this is preferable to an approach involving ad hoc and asymmetric adjustments through a special cost factor process).

171. We have explored a shortlist of variables for deprivation and arrears risk. At this stage, we would not suggest a strong need to cut down this shortlist. We provide some comparisons of modelling results for the different variables in Sections 4 and 6. It may be possible to carry out further analysis in the future to assess the relative merits of the different variables we used.

Size of household retail bills

172. There is a strong intuitive and theoretical case for seeking to include an explanatory variable for the size of bills in models that compare measures of bad debt costs (or total residential retail costs) between companies. First and foremost, the higher the bills, the more money is at risk from any instances of late- or non-payment. There may also be effects of higher bills increasing the risks of late- or non-payment.

173. We consider that this cost driver should be incorporated into the explanatory variables in models in which the dependent variable is based on either a measure of the debt costs divided by the number of households supplied or the total retail operating costs divided by the number of households supplied. The models we have developed use a measure of average retail bills as an explanatory variable. We have defined this measure of average retail bill as the retail and wholesale revenue divided by the number of household customers. Our modelling analysis indicates that this is an important driver of retail costs.

Proportion of dual service customers

174. The proportion of customers that take both water and wastewater services affect the average revenue per customer, which we identified above as an important cost driver. However, we consider it more accurate and more logical to capture that effect through variables relating to average bills or average revenue per household supplied, rather than the more indirect measure of the proportion of dual service customers.

175. All the same, the proportion of dual service customers may affect retail costs other than through the impacts on average bills, for example if dual service customers generate more customer enquiries.

176. Our modelling results suggests some basis for including an explanatory variable for the proportion of dual service customers, particularly in models of residential retail operating costs excluding bad debt costs. However, our analysis raised questions about the accuracy of the estimated coefficients and the sensitivity of results. As things stand, we do not have a strong view on whether or not this cost driver should be included and this may depend on other aspects of model specification.

177. If the proportion of dual service customers is included in a model of total retail operating costs, there are risks of the model estimation being adversely affected by multi-collinearity arising from the correlation between this variable and the cost driver above for the size of average bills. It is possible to tackle this issue through the way that these two cost drivers are incorporated as explanatory variables in the models. See Section 6 for further details on a potential approach.

The transiency of households within the area served (i.e. rate of occupancy changes)

178. We reviewed available data on population transiency, namely data reported within the Equifax dataset and data published as part of ONS statistics on migration flows. We identified some concerns with these data. All the same, we carried out some initial analysis on models incorporating a measure of transiency based on that data and found that it did not perform well.

179. We considered that further work on the identification or development of data sources for transiency was likely to reveal fewer insights on the development of econometric models than other aspects of our analysis and we did not pursue it for this report.

Quality of service provided to customers

180. While the provision of a higher quality of service should, all else equal, entail higher expenditure requirements, our initial modelling results (and qualitative analysis) indicated that it may be difficult to use econometric benchmarking models to estimate the relationship between measures of retail quality of service and expenditure requirements.

181. Our exploratory modelling results had indicated a positive figure for the estimated coefficient on a measure of quality of service. This suggested that the coefficient was capturing factors other than the underlying relationship between the quality of service provided and the costs of provision (perhaps reflecting positive correlations between cost efficiency and service quality across companies' residential retail operations).

182. We considered that further development of econometric benchmarking models that capture differences in quality of service between companies was likely to reveal fewer insights than other aspects of our analysis and we did not pursue it further for this report.

183. It may be possible to use the calibration of ODIs, rather than the cost assessment process, as a means to ensure that companies are appropriately remunerated for the quality of service that they provide. Alternatively, it may be possible to make allowances for variations in service quality using more bottom-up cost estimates rather than econometric benchmarking models.

Meter penetration rate for customers

184. There seems to be a strong case for Ofwat's cost assessment to take account of differences between companies in meter penetration.

185. The costs reported by water companies for meter reading reflect about 5 per cent of their total retail operating expenditure. While a small proportion of retail costs, they are not immaterial. Metered customers may give rise to higher costs in areas besides meter reading (e.g. more customer enquiries). Furthermore, not remunerating water companies for the incremental costs of metered customers could artificially discourage increases in meter penetration.

186. We identified two different ways to capture metering penetration through explanatory variables in our models:

    (a) Measures of the proportion of households supplied who are metered.

    (b) Measure of the average number of metered services provided per household supplied (where a customer supplied with metered water and wastewater services is treated as being supplied with two metered services).

187. While the first approach may be more familiar, we believe that the second approach makes more sense from an economic perspective. This is because companies carrying out meter reading for water-only customers will tend to charge some costs/commission to the wastewater supplier which means that meter reading costs are in effect split between water services and wastewater services. A water retail company's meter reading costs would be higher for customer supplied both water and wastewater rather than only one service.

188. Our modelling results suggests some basis for including an explanatory variable for meter penetration in econometric models of residential retail costs, at least for models that exclude bad debt costs. However, our analysis identified a degree of sensitivity and imprecision in the estimation results (more so than for the deprivation variables). This was particularly so for models of total retail operating costs.

189. At this stage, we would also suggest the exploration of other methods to capture retail costs associated with metered customers, such as non-econometric unit cost metrics for meter reading costs. Such models could ultimately be used alongside or instead of econometric methods for capturing the additional costs of metered customers. The point above in relation to the two different ways to capture metering penetration would also apply to the denominator for unit cost metrics for meter reading costs.

Summary of findings

190. Table 25 provides a summary of the above discussion on the inclusion of the different candidate cost drivers in econometric models to benchmark residential retail operating costs.

Table 25 Summary of findings on incorporation of cost drivers in econometric models

Candidate factor	Does our work provide evidence on materiality of cost driver?	Should cost driver be included in econometric models of residential retail costs?	Are there other ways to capture *without* using econometric models of retail costs?
Number of households supplied	Yes	Yes We can specify as denominator in the dependent variable in model	Yes (e.g. assumption of proportional relationship between number of households and total residential retail costs)
Size of household retail bills	Yes:	Our work supports use of this cost driver in explanatory variables for econometric models	Could make assumption on relationship between bills and bad debt but risk that assumption lacks evidence
Proportion of dual service customers	Some indication of material positive relationship but not strong evidence	Our work provides some support for this cost driver, but results quite sensitive to model specification	Without Ofwat collecting more granular cost data, it seems difficult to capture this driver without econometric approach
Measures of economic deprivation and arrears risk	Yes	Our work supports use of this cost driver in explanatory variables for econometric models	Difficult to estimate financial impact of varying levels of deprivation on bad debt costs without econometric modelling
Transiency of households within the area served (i.e. rate of occupancy changes)	Cost driver not fully considered in the models due to data limitations. Drawing on the data that are available data, no strong evidence of positive relationship.	Not based on the work to date. May be possible to overcome with better measure of transiency.	Difficult to estimate financial impact of varying levels of transiency on retail costs without econometric modelling
Quality of customer service	No (due to modelling issues: does not imply this is not a cost driver in practice)	Not based on work to-date (though may be possible to overcome modelling issues in the future)	There seem to be other ways that companies could be remunerated for the costs of providing higher-quality services
Meter penetration rate for customers	Some indication of material positive relationship but not strong evidence	Our work provides some support for this cost driver, but results quite sensitive to model specification	Meter reading unit cost metric approach seems feasible (either as alternative or complement) though may not capture all meter-related costs

Appendix 1: Data sources

191. This appendix sets out information on the compilation of the dataset used in the analysis. There were a number of strands to this:

(a) Compiling data on company performance (e.g. on households served, costs and revenue).

(b) Compiling data on deprivation measures.

(c) Compiling other data (e.g. on measures of population transiency, on CPI).

(d) Reflecting the merger of South West Water and Bournemouth Water in the dataset.

192. We consider each of these in turn.

Data on company performance

193. We extracted data from:

(a) Regulatory accounts. This refers to companies' published regulatory accounts, from 2013/14 to 2016/17.

(b) Ofwat PR14. This refers to the Excel file "Household retail average cost to serve final determination populated model.xlsx"" published by Ofwat. [24] We drew on this for information on a small number of cost drivers (namely number of households broken down by service and metered/unmetered) for 2013/14 and 2014/15.

(c) Ofwat SIM. We extracted data on companies' overall SIM score for 2013/14 to 2015/16 from Ofwat publications, available from its website.

194. The table below lists the source(s) of the data we drew on to compile a panel dataset, covering the years from 2013/14 to 2016/17. Where relevant, the table describes how we filled in gaps in the data (e.g. where data items for more recent years are not available due to changes to regulatory reporting requirements). In the table, we identify the source as "Derived from input data" for those variables that were derived from data on other variables reported in the table.

Table 26 Sources of data: company-specific

Ref	Variable	Source	Comment
dv101 – dv106	Households connected (000), split by service provided and by metered/unmetered	Ofwat PR14: 2013/14 and 2014/15 Reg. Accounts: 2015/16 and 2016/17	For WSX, data for 2014/15 were extracted from its 2015/16 Reg. Accounts

[24] Accessed from https://www.ofwat.gov.uk/publication/household-retail-average-cost-to-serve-final-determination-populated-model/.

Ref	Variable	Source	Comment
dv107	Households connected (000)	Derived from input data	Sum of dv101 to dv106
dv108	Proportion of metered customers	Derived from input data	Derived from dv101 to dv106
dv109	Proportion of dual service customers	Derived from input data	Derived from dv101 to dv106
dv142	Metered services per connected households	Derived from input data	Derived from dv101 to dv106
dv111	Household revenue (£m)	Reg. Accounts: 2013/14 to 2016/17	For SWT, we deducted £50 per connected household, to reflect the £50 Government payments benefitting SWT customers.
dv143 – dv145	Revenue for water only, wastewater only and dual service households	Reg. Accounts: 2015/16 to 2016/17	Data used to construct revenue per household by type of service provided which, in turn, is used to construct variable Billratio, as explained in main body of report For 2013/14 and 2014/15, for which data on revenue by type of service is not reported, and for purpose of calculating Billratio, we set the revenue per household for each service equal to value in 2015/16.
billratio	Billratio, a measure of companies' average bill compared to industry average, controlling for mix of services provided	Derived from input data	Variable derived as explained in main body of the report
dv120	Debt management cost (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv121	Doubtful debts (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv122	Debt costs (£m)	Derived from input data	Calculated as sum of dv120 and dv121
dv123	Meter reading costs (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv124	Customer services (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv125	Other operating expenditure (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv127	Depreciation (£m)	Reg. Accounts: 2013/14 to 2016/17	For 2013/14 and 2014/5, reported in accounts as "Total capital maintenance – Total" For 2015/16, reported in

Ref	Variable	Source	Comment
			accounts as "Depreciation – Total" For 2016/17, calculated as sum of "Depreciation – tangible fixed assets, total" and "Amortisation – intangible fixed assets, total"
dv128	Total operating costs (£m)	Reg. Accounts: 2013/14 to 2016/17	
dv129	Retail remaining operating costs (£m)	Derived from input data	Calculated as dv128 – dv122
SIM_overall	SIM score	Ofwat: 2013/14 to 2015/16 Datashare: 2016/17	

Data on deprivation measures

195. Appendix 2 further below outlines how we drew on Equifax data to derive the measures of deprivation and arrears risk at the company level. The working paper we produced in May 2017 contains more detailed information on the approach followed and data sources used. [25] The only change to the approach set out in that paper concerns an improvement made in the mapping of LSOAs to sewerage service areas.

Other data

196. The table below shows two further data sources we used.

Table 27 Sources of data: other data

Ref	Variable	Source	Comment
onstransience	Ratio of sum of population flow into and out of LAD to mid-year population	ONS, Migration Indicators Tool: 2013 to 2016	Variable calculated for each LAD in England and Wales.
cpiavg	Average CPI in financial year	ONS, series "CPI All Items Index: Estimated pre-97 2015=100", released 12 Sept 2017	Average CPI in financial year calculated as average of CPI in the relevant quarters

Dealing with the merger of South West Water and Bournemouth Water

197. In compiling the data for analysis, we dealt with the merger of South West Water and Bournemouth Water from 1 April 2016 as follows:

(a) We retained separate data points for the two companies for the years 2013/14, 2014/15 and 2015/16.

[25] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities.

(b) We used data for the merged entity for 2016/17, as data (e.g. on expenditure) were not reported for each of the two areas separately.

198. It follows that our panel dataset contained 71 observations: annual observations on 18 companies in the three years from 2013/14 to 2015/16, and observations on 17 companies for 2016/17.

199. We cross-checked the results of some of the models estimated on the basis of the panel dataset with those estimated on the basis of a cross-sectional dataset, where the data for a company was averaged over the four-year period 2013/14 to 2016/17. For that purpose, we created two versions of the cross-sectional dataset: one, where we grouped together the data on South West Water and Bournemouth Water even in the years for which separate data were reported; and a second version where we kept the data for the two companies separate and, for each, took the average values over the three-year period 2013/14 to 2015/16 (and not the full four-year period).

Appendix 2: Development of measures of economic deprivation and of arrears risk

200. This appendix gives an overview of our earlier work to develop measures of economic deprivation and arrears risk. A fuller account of that work was set out in our May 2017 working paper.[26]

201. This appendix covers the following:

(a) We set out the motivation for developing measures of economic deprivation and arrears risk in the context of comparing water companies' residential retail operating costs and, in particular, those costs relating to bad debt.

(b) We summarise the main elements of the work we carried out to develop measures of economic deprivation and of arrears risk.

Capturing differences in economic deprivation across parts England and Wales

202. Each water company's supply of water and/or wastewater services to household customers is made within a defined geographic area. There is no opportunity for households to switch supplier. Likewise, companies cannot compete to take on household customers from outside the area in which they have historically operated. As a consequence, any factors which mean that customers in one geographic area are, on average, more costly to serve than customers in another geographic area, will tend to feed through to create differences in retail costs between water companies (even if they were equally efficient).

203. Costs relating to bad debt account for close to 40 per cent of water companies' costs of providing retail services to households. These include the costs of managing cases of late payment and non-payment and the costs of debt that is written off as not economically recoverable.

204. In this context, we consider that one of the important drivers of differences between water companies, in terms of their residential retail costs, is the substantial variation across different parts of England and Wales in the levels of economic and social deprivation.

205. There is no unique concept of economic and social deprivation; it may be manifest in different ways. The Index of Multiple Deprivation (IMD) published by the Department for Communities and Local Government (DCLG) attempts to cover a range of different elements of deprivation, including deprivation in relation to incomes, employment, education and skills, health and disability, crime, housing and services, and living environments.[27]

206. Intuitively, the degree of deprivation in a particular part of England or Wales, particularly the more economic aspects of deprivation, seems likely to affect the

[26] Reckon (2017) "Capturing deprivation and arrears risk in household retail cost assessment", working paper for United Utilities.

[27] See https://www.gov.uk/government/collections/english-indices-of-deprivation. In our May 2017 working paper we wrongly attributed the publication of these indices to the Office of National Statistics.

ability of households in that area to pay their water bills. In more deprived areas, a greater number of customers may struggle to pay their water bills, with greater risks to the company concerned of late payment and non-payment.

207. Ofwat has recognised the relevance of differences in deprivation levels across England and Wales for the purposes of assessing companies' retail costs. In its PR14 final determinations, Ofwat reflected, with some modifications, proposals from several companies for upward adjustments to the residential retail cost allowances to take account of greater levels of deprivation in their areas of appointment relative to other companies. For example, in its assessment of United Utilities' proposal, Ofwat concluded that "*United Utilities provided sufficient and convincing evidence that deprivation (especially extreme deprivation as measured by the 10 per cent most deprived households) affects United Utilities in a materially different way to other companies*".[28] The evidence base for United Utilities' proposal drew on the Index of Multiple Deprivation (IMD) measure of deprivation referred to above.

208. As indicated above, for PR19 Ofwat said that it will consider deprivation levels when assessing bad debt costs, and indicated that this would be done as part of an econometric approach rather than through company-specific adjustments. This raises the question of data availability for capturing deprivation or those aspects of deprivation that matter most to risks of bad debt.

209. The data published by the DCLG in the context of calculating the IMD provides a rich characterisation of deprivation in different areas, at a geographically granular level. The use of this data for econometric modelling of retail costs is, however, affected by three issues. First, the data only covers England and does not cover Wales. Statistics for Wales publishes similar deprivation measures, but these are not consistent with those produced by the DCLG. We are not aware of other similar published data that captures deprivation at the local level, on a consistent basis across England and Wales. Second, the IMD for England is published only every few years; the last publication was 2015 and prior to that it was 2010. The next publication is due in summer 2019. Third, these statistics are specific measures of the underlying concept of deprivation, and are not necessarily the best available means of understanding differences in the types of economic deprivation that affect the risks that households struggle to pay their water bills.

210. For PR19, there is an opportunity to address this gap in the available data. United Utilities has been working with Equifax to identify additional sources of relevant data, which could help tackle some of the limitations of the deprivation data available from the DCLG and Statistics for Wales, and which could bring a wider perspective on the interactions between economic deprivation and risks of non-payment of water bills.

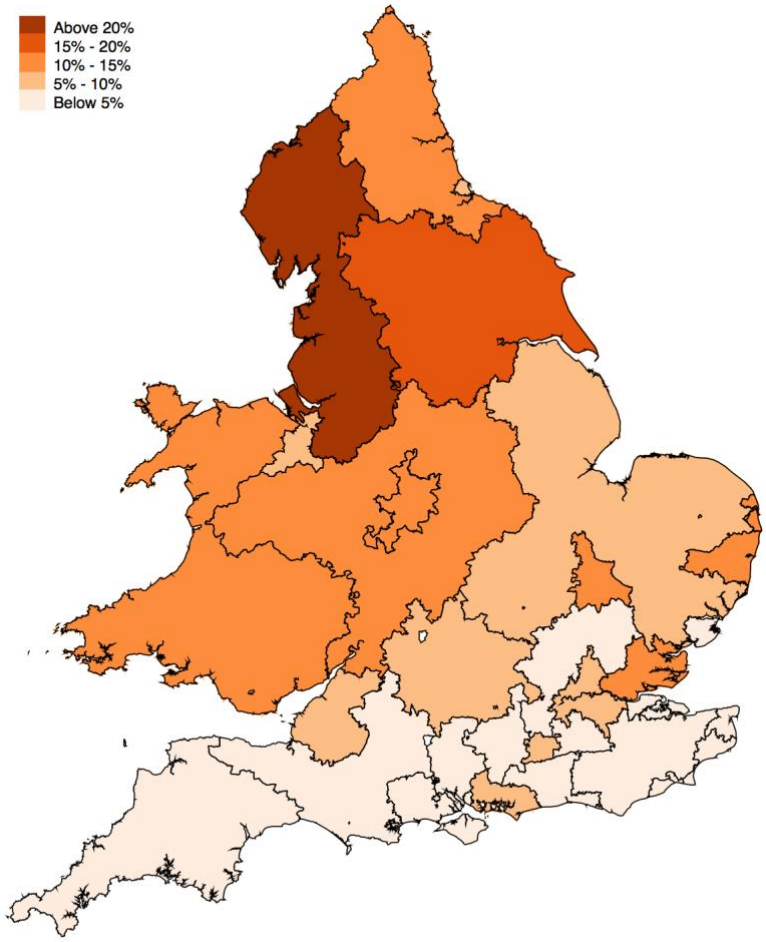
211. Reckon supported United Utilities with analysis of the data provided by Equifax. We have sought to identify good quality candidate variables to reflect differences between companies in deprivation.

[28] Ofwat (2014) *Draft price control determination notice: company-specific appendix – United Utilities*, page 33

212. One of the major contributions of our work, in conjunction with United Utilities and Equifax, has been the development of a dataset on measures of deprivation that could be used for econometric modelling of water companies' residential retail costs.

213. These measures are derived from data provided by Equifax. The raw data from Equifax data is at the level of postcodes and covers England and Wales. We used this data to produce variables that correspond to the geographic areas served by each of the main English and Welsh water companies. Figure 1 below provides an illustration of company-level variables that are based on an Equifax measure of arrears risk (RGC102) which we found to be highly correlated with the English IMD measures of deprivation. The map shows, for each water company's area of supply, the share of households that are within the 10 per cent of English and Welsh LSOAs (lower layer super output areas) with the highest arrears risk.

Figure 22 Share of households in top decile of arrears risk, as measured by RGC102 (2016)



214. The Equifax data that we used is not publicly available and there are restrictions on access to it. We have worked with United Utilities and Equifax to better understand what variables derived from Equifax data could be shared with other parties. United Utilities and Equifax agreed that the company-level variables relating to deprivation

and arrears risk that we had derived from the raw Equifax data could be shared with Ofwat and other water companies. United Utilities provided an initial dataset to Ofwat in September 2017 and an updated dataset in November 2017. The datasets provided so far are development datasets. We envisage a further layer of review and audit before the dataset could be used for Ofwat's PR19 determinations.

Main elements of work to develop measures of deprivation and arrears risk

215. In earlier phases of our work for United Utilities, we developed alternative measures of deprivation and arrears risk of the area served by each water company. These drew on the Equifax dataset that was made available to us. Our working paper published in May 2017 set out the details of that work. Here, we confine ourselves to outlining its main elements.

216. The Equifax dataset contains data at the postcode level. In an early phase of our work we aggregated these data to the level of LSOAs, calculating weighted-average values of each Equifax variables across the relevant set of postcodes. We used either population or household numbers as weights, depending on how the relevant Equifax variable was defined.

217. We combined the Equifax dataset, aggregated to the LSOA level, with a dataset containing data on the 2015 Index of Multiple Deprivation (IMD) for England, as well as the scores for the income and employment deprivation domains. We used that dataset to develop econometric models that enabled us to construct proxy or predicted values of the English IMD for LSOAs across England and Wales and across different years. This allowed us to remedy the coverage issues due to the incompatibility of the measures of deprivation in England and the analogous measures produced for Wales.

218. We drew on the Equifax dataset and on the "predicted" IMD we developed, to construct measures of deprivation and arrears risk for the areas served by each of the water companies. This involved mapping LSOAs to companies' area of service, and then aggregating each of the deprivation measures to the company-level. In doing so, we considered alternative approaches of aggregating across LSOAs to deal with the fact that water and sewerage companies offer different services over different parts of the area they serve. We chose to do so by constructing, for each deprivation measure, the weighted average of that measure across all LSOAs where a company provides water or wastewater services, or both, with the weights given by the product of household numbers (or population numbers, depending on the deprivation measure) and average bill of the services provided in each LSOA. The weights used to aggregate the LSOA-level measure to the company-level measure are a proxy of the relative contribution that each LSOA makes to a company's revenue. This ensures that for a water and wastewater company the deprivation level of an LSOA that is within its water supply area and within its sewerage service area has a greater weight than that of an LSOA that is only within its water supply area.

Appendix 3: Interpreting estimated coefficients

219. This appendix is concerned with the interpretation of the estimated coefficients from across the set of models presented in the report.

220. All of the models reported are ones where we specified a linear relation between the dependent variable and each of the explanatory variables. Within this type of models, it is possible to model relations between the underlying cost drivers and the dependent variable that are not linear. This can be done by transforming the cost driver variable and then including that transformed variable within the set of explanatory variables. For example, if we thought there was a relation between the square of a given variable and expenditure, we could construct a transformed variable equal to the square of that variable, and then include that transformed variable within the set of explanatory variables in a linear model.

221. For some of the models we explored, a transformation that we carried out on the dependent variable or on some of the explanatory variables, or on both, was to take their natural logarithm. The appropriateness of doing so depends on the view about the relationship between the dependent variable and the relevant explanatory variable.

222. Having regard to that aspect of the model specification alone, the set of models we report on in this paper fall under one of the following two specifications:

Type 1 specification: $Ln(Y) = a + b1 * Ln(X) + b2*Z + error$

Type 2 specification: $Y = a + b1 * X + b2 * Z + error$

223. In type 1 specification, the dependent variable is expressed in natural logarithms, and the set of explanatory variables may include some that are also expressed in natural logarithms and others that are not. The range of debt cost models we labelled as BD1, or those that we considered for total operating costs are of this type. In type 2 specification, the dependent variable is not subject to the logarithmic transformation, nor are any of the explanatory variables. The range of models we labelled as BD2, and those that we considered to model remaining operating costs are of this type.

224. Table 28 summarises the interpretation that can be given to each of the estimated coefficients for the two types of specification listed above.

Table 28 Interpretation of estimated coefficients

Specification: Ln(Y) = a + b1 * Ln(X) + b2*Z	
Interpretation of b1	A *d* per cent change in X leads to a (*d**b1) per cent change in Y
Interpretation of b2	A change in Z by *d* units leads to a change in Y of (exp(*d**b2) – 1) per cent If (*d**b2) is small in absolute terms, say below 0.1, this is approximately equal to (*d**b2) per cent

Specification: Y = a + b1 * X + b2*Z	
Interpretation of b1	A change in X by *d* units leads to a change of b1**d* units in Y
Interpretation of b2	A change in Z by *d* units leads to a change of b2**d* units in Y

225. Leaving aside whether a variable was included in the model after being first "logarithmically" transformed, the variables we considered varied with respect to how they were constructed. In particular, for some of the models we explored the dependent variable was a unit cost (e.g. in the models of retail remaining operating costs), and in other cases the dependent variable was expressed as the ratio of a measure of cost to revenue. Similarly, amongst the set of explanatory variables we considered, some were expressed as proportions (e.g. the proportion of dual service customers, or the proportion of households served which are in the top decile of most deprived LSOAs) some were ratios more generally (e.g. the average number of metered services provided per household), and some were "raw" measures, (e.g. the Equifax deprivation measure RGC102.

226. The different "formulations" of the variables do not alter how their estimated coefficients should be interpreted from that which we described in Table 28. The differences matter to the extent that they guide us when considering what is a reasonable "thought experiment" to consider in the interpretation of the estimated coefficients. For example, in interpreting the estimated coefficient on the variable relating to the proportion of dual service customers, it would be reasonable to consider what the model predicts a change in the proportion of dual service customers by say, 0.1, would be than to consider a change in the proportion of, say, 1 (which would not be sensible).