Reckon LLP
Data analysis | Economic regulation | Competition law

Review of Severn Trent's sewerage cost models

13 February 2018

Introduction

1. This report presents our review of a set of econometric models developed by Severn Trent to benchmark wastewater base expenditure and wastewater network plus base expenditure.

2. On 20 October 2017, Severn Trent shared with us a note describing a set of models it had developed, as well as supporting data and Stata script files. We submitted a first review of those models on 13 November 2017, and discussed this with Severn Trent. Subsequently, on the 16 November, Severn Trent shared with us a revised set of those initial models, and on the 24 November a further set of additional models.

3. The complete group of models proposed by Severn Trent, reflecting the iteration following our initial feedback, comprised 10 models. This included models that varied across a number of dimensions. It included models with three different forms: (i) models with translog variables; (ii) models with a form of Cobb-Douglas structure; and (iii) models in which the dependent variable is a measure of unit costs. It included models that covered all of wholesale wastewater base expenditure and models focused on base expenditure for wastewater "network plus" activities. There was some variation in the explanatory variables used. And some models were estimated using pooled ordinary least squares (OLS) and others were random effects models estimated using generalised least squares (GLS).

4. The scope of our review is concerned with addressing a number of questions concerning Severn Trent's compilation of the data it used and the econometric models it has put forward.

Have the data been processed accurately and in accordance with the data definitions?

5. We recreated for ourselves the dataset used by Severn Trent. We compiled this dataset drawing on input data provided to us by Severn Trent, and supplemented it with data published by Ofwat (historical June returns) and ONS data on the Retail Price Index. Auditing the dataset in this way provided a, more effective than an alternative approach of checking data items on a sample basis.

6. We compared our dataset with that which Severn Trent had compiled and used in its initial modelling. We found some discrepancies between the two. We discussed and resolved these, and Severn Trent revised its dataset accordingly.

Have the key findings of the wastewater cost driver report, completed by Jacob, been appropriately included in the models, or been adequately rejected?

7. Jacobs prepared a report on cost drivers relevant to wastewater treatment. Severn Trent drew on this to formulate its models. We reviewed the Jacobs report (draft dated 1 August 2017) with a view to identifying the key findings from it that are pertinent to the development of models to benchmark expenditure, and to assess whether these had been adequately considered by Severn Trent.

8. We found that the way in which Severn Trent's models seek to capture economies of scale at the sewage treatment works level and the effect of treatment complexity is broadly consistent with the findings put forward by Jacobs.

9. Two aspects where Severn Trent's modelling choices diverge from the views expressed in Jacobs' report concern the choice of density measure (Jacobs suggest that length of pipe per property is not appropriate) and the choice of load received as scale variable (Jacobs proposes that load removed is preferable). On both of these points, however, we consider Severn Trent's choices to be reasonable; on the grounds that its modelling is concerned with sewage collection too, and not just treatment, and on the grounds that data on load removed are not readily available. Severn Trent told us that it had considered both of those points, and that it is currently pursuing work to explore the use of data on load removed.

What are the main points from the review of the models?

10. We have reviewed Severn Trent's econometric models in terms of their potential contribution to expenditure benchmarking analysis for wastewater activities, within the context of Ofwat's price control regulation.

11. As a preliminary point, we do not consider that there is any perfect or true model: the models can only be approximations. We identified a number of detailed observations on the models and on the approach taken to their development, which we have discussed with Severn Trent.

12. The models developed by Severn Trent have a number of desirable features for the purpose of Ofwat's expenditure benchmarking analysis. For instance, they are models of base expenditure and, appropriately, do not cover enhancement expenditure.

13. We agree with the emphasis that Severn Trent placed on the engineering and economic basis for developing model specifications, rather than treating this as a purely statistical exercise. The range of specifications in Severn Trent's models stem from the view that these should control for several key factors (scale of wastewater operations at the company level, economies of scale at the treatment works level, complexity of wastewater treatment and population density). This is a reasonable basis from which to develop models for wastewater.

14. We also identified a number of limitations and areas for potential improvement in the initial set of models shared with us. We had particular concerns with the inclusion of some translog variables although we recognise that this was a prominent feature of Ofwat's PR14 models and may be retained for PR19. Severn Trent told us that in its

initial model development it had given greater consideration to the choice of cost drivers and estimation method (OLS or GLS random effects) than to the question of capturing or not translog variables. Following our discussions, Severn Trent expanded its set of models to consider more alternatives to the translog approach.

15. As with other models we have seen for wholesale cost assessment in the water industry, the models developed by Severn Trent show the difficulty of developing models that capture all key cost drivers. Particularly with the small sample size available, results can be quite sensitive to model specification. This is not a criticism of Severn Trent's specific models but is a feature of these and other models.

16. Given that all such models are approximations and are likely to involve limitations, it is a good idea to consider results across a wide range of models. The final set of models that Severn Trent provided to us usefully captures a number of different modelling and estimation approaches, each of which may have strengths and weaknesses. At the same time, there seem further opportunities to explore models with different cost drivers (or which capture the same underlying cost drivers through different explanatory variables).

Provide considered opinion of the extent to which the Severn Trent models are likely to improve the robustness of AMP7 totex estimates relative to the approach used at PR14.

17. As highlighted above, we agree with the emphasis that Severn Trent placed on engineering and economic considerations in the model development and specification, with less weight placed on purely statistical metrics compared to the model development approach used by Ofwat for PR14.

18. There are several specific aspects of the Severn Trent models that offer significant improvements to the models developed at PR14. In particular:

    (a) The Severn Trent modelling seeks to better capture economies of scale at the treatment work level and to capture treatment complexity.

    (b) All of Ofwat's models of wastewater base expenditure at PR14 included translog variables. The use of translog models for benchmarking water companies was criticised by the CMA in its *Bristol Water* inquiry. For benchmarking across the ten water and sewerage companies in England and Wales, the translog approach poses high risks of misleading results, from an attempt to fit a relatively complex model structure onto a limited data sample. Whilst the set of Severn Trent models also included some models with translog variables, it also covered two alternative types of models: models with a Cobb Douglas form and unit cost models. These alternatives avoid the issues associated with translog models and contribute significantly to the overall analysis.

    (c) Severn Trent refined its models to remove the explanatory variable for wages, following on from its identification of concerns with the estimation results for this variable. While we recognise that regional wages are a relevant cost driver, it seems better that the econometric modelling omits some potential cost drivers rather than including variables for which estimated coefficients are highly sensitive and counterintuitive (as was the case for wages in Ofwat's PR14 models).

(d) We also consider that Severn Trent's approach to modelling "time effects" through the inclusion of year-specific dummy variables to be preferable to Ofwat's approach at PR14 of imposing a time trend.

19. We consider that Ofwat's model development of models to replace (or complement) its PR14 could benefit from the work that Severn Trent has done in these areas.

Appendix: supporting tables

This appendix sets out five tables, supporting our review and providing further detail:

- Table 1 lists the set of files Severn Trent sent us relevant to our review of the data used and models developed.
- Table 2 describes the set of 10 models developed by Severn Trent which we reviewed.
- Table 3 sets out our comments on key benefits of Severn Trent's approach.
- Table 4 reports on questions raised and resolved with Severn Trent during the review process, resulting from a first review of an initial set of models it shared with us (file SVT1 listed in Table 1). The points we raised in that initial review were taken on-board in Severn Trent's subsequent further developments of its models.
- Table 5 set out our comments regard to potential limitations of the models.
- Table 6 identifies potential extensions to develop wider set of models.
- Table 7 sets out our comments on the write up of the initial set of models Severn Trent shared with us, particularly in regards to expectation about the magnitude of the expected coefficients. We provide these for completeness, as they are not comments on the statistical performance of the models themselves.

Table 1 Files sent by Severn Trent

Ref.	Date received	File name	Content
SVT1	20 Oct. 2017	“Write-up of waste models-final.docx”	Write-up of rationale for and presentation of initial set of models
SVT2	20 Oct. 2017	“For Reckon.do”	Stata script file with commands to run regressions for initial set of models
SVT3	20 Oct. 2017	“For Reckon.dta”	Stata data file containing dataset on which initial set of models were estimated
SVT4	20 Oct. 2017	“Latest Masterfile PR14 and June Returns data.xlsx”	Excel file with data on costs and on cost drivers; basis for Severn Trent compiling its dataset
SVT5	20 Oct. 2017	“B1958940_DOC-001_Rev2_Final Issue.pdf”	Jacobs’ report “PR19 Assurance Support”, version dated 1 August 2017
SVT6	24 Oct. 2017	“High density and sparcity – for Reckon.xlsx”	
SVT7	24 Oct. 2017	“Master waste hc 20161209.xlsx”	Set of Excel files containing data on cost drivers used by Severn Trent to construct its dataset (and not reported within SVT4)
SVT8	30 Oct. 2017	“RPI deflator.xlsx”	
SVT9	2 Nov. 2017	“2016 Master waste hc.xlsx”	
SVT10	16 Nov. 2017	“Update of SVT models (16Nov2017)(004).xlsx”	Excel file reporting updated set of Severn Trent models, reflecting revisions to dataset following our initial review of the data compiled by Severn Trent
SVT11	24 Nov. 2017	“Cobb-Douglas and Unit Cost models.docx”	Description of a set of Cobb-Douglas and unit cost models, developed in the light of Reckon’s initial comments concerning the inclusion of translog variables

.

Table 2 Severn Trent models reviewed

Model ref.	1	2	3	4	5	6	7	8	9	10
Estimation technique										
OLS	■			■	■	■		■	■	■
Random effects (GLS)		■	■				■			
Dependent variable										
Ln(Wastewater botex)	■	■	■		■					
Ln(Network plus botex)				■		■	■			
Ln(Wastewater botex / Total load])								■	■	
Ln(Network plus botex / Total load)										■
Explanatory variables										
Ln(Load)	■	■	■	■	■	■	■			
Ln(Number of STWs)	■	■	■	■	■	■	■			
(Ln(Load))^2	■	■		■						
(Ln(Number of STWs))^2	■	■		■						
Ln(Load)*Ln(Number of STWs)	■	■		■						
Ln(Number of tight BOD and N3 works)		■	■		■	■	■			
Ln(Number of large tertiary works)	■			■						
Ln(Properties per length of mains network)	■	■	■	■	■	■	■	■	■	■
Ln(No. of STWs/Total load)								■	■	■
Tight N3 works/Total load								■		■
Tight BOD load/Total load									■	
Ln(Length of mains network/Total load)								■	■	■
Prop. of customer base in high density LADs										■
Set of year-specific dummy variables	■	■	■	■	■	■	■	■	■	■

Table 3 Comments on key benefits of approach adopted

Topic	Comments
Exclusion of enhancements	We support the approach of excluding enhancement expenditure and estimating econometric models of base expenditure only. It may be possible to extend the models to cover some elements of enhancement expenditure (certainly not all) but it seems reasonable to estimate models that focus on base expenditure
Economic and engineering emphasis in model development	We agree with the emphasis placed on engineering and economic considerations in the model development and specification, rather than treating the model development process as primarily a statistical exercise. It is good that Severn Trent has not sought to apply (or justify its model by reference to) the kind of general-to-specific approach that CEPA used as part of its work for Ofwat for PR14
Review of sign and magnitude of coefficients	We support the emphasis placed on considering whether the sign and magnitude of estimated coefficients are consistent with what one might expect from an engineering and economic perspective. We did have some comments concerning the stated expectations about the magnitude of certain coefficients provided in the initial write=up of the modelling (see table 5) but these do not detract from this point.
Intuitive sense	Leaving aside the questions about translog and Cobb-Douglas aspects of the model specification, the core elements of the Severn Trent model specifications make intuitive sense: • Load as the key scale variable driving wastewater costs • A variable (the number of STWs) to take account of differences between companies in economies of scale at the treatment works level • A variable to capture issues relating to treatment complexity • A variable to capture the density of properties connected to the sewerage systems (total properties divided by total sewer length) • A variable to take account of the differences in wage rates across England and Wales We have some comments on the details, as set in the tables further below
Capturing time	We agree with Severn Trent's preference for time dummies rather than time trends.

Topic	Comments
	It may be useful to also consider time trend models, as a sensitivity, because Ofwat favoured this for PR14.
OLS/GLS(RE) vs SFA	Some commentators argue for the application of stochastic frontier analysis for water company wholesale cost assessment, rather than the OLS or GLS random effects approaches that Severn Trent has used We support Severn Trent's focus on OLS and GLS random effects, and not SFA
Random effects models vs OLS models	As reported in SVT10 and SVT11, Severn Trent has considered both random effects versions of models, estimated using GLS, and cross-sectional time-series versions estimated using OLS. OLS is simpler and less assumption-heavy approach. Ofwat used random effects models for PR14, and it seems reasonable to include consideration of such models.
Smoothed capital expenditure	Severn Trent has followed Ofwat's approach of smoothing capex over a five-year period This approach seems reasonable, especially given variations in capex over the five-year price control cycle We note that it does raise a bit of a question about the totex approach (opex is not smoothed which means that opex and capex are treated differently for cost assessment purposes)
Robust standard errors (and corresponding p values)	We support the use of cluster robust standard errors, which we can see from Severn Trent's Stata code in SVT2 has been used The use of cluster robust standard errors is consistent with the approach used by the Competition and Markets Authority in its Bristol Water determination in 2015 (see Appendix 4.2 of the CMA report).[1]

[1] Competition and Markets Authority (2015) "Bristol Water plc: A reference under section 12(3)(a) of the Water Industry Act 1991", available from https://assets.publishing.service.gov.uk/media/5627995aed915d101e000001/Appendices_1.1_-_4.3.pdf.

Table 4 Points raised and addressed during review process

Topic	Comment
Further consideration of Cobb-Douglas and unit cost models	The first set of models in SVT1 covered four translog specifications, and one Cobb-Douglas specification. We raised concerns about some properties of the translog versions and suggested that simpler versions of the models be explored. Severn Trent took this suggestion on board and subsequently reported on a set of Cobb-Douglas and unit cost models (SVT11).
Wages	We agree with the concerns raised by Severn Trent in its explanation of the models (SVT1) about the inclusion of a variable reflecting wages in its initial set of models. We discussed these concerns with Severn Trent. In its updated models reported in SVT10, Severn Trent did not control for wage differences. We agree with the approach taken. It seemed reasonable to try to control for regional wage differences, to see how it works, given the use of it at PR14 and the intuitive argument for doing so. However, in the light of the high sensitivity of the estimated coefficient on that variable to model specification there was a high risk that the inclusion of the wage variable could be making the model worse rather than better. Indeed, the model estimation results reported in SVT1 cast doubt on whether the estimated coefficient on the wage variable is approximately underlying cost drivers: • The wage coefficient in the two random effects models of over 1, which seems suspiciously high. We would generally expect regional wage differences (and related factors such as regional input prices) to affect significantly less than 100% of costs, such that a 10% increase in wages should lead to significantly less than 10% increase in costs. • The estimated coefficient for the wage variable is highly unstable across the four models reported, taking the values 0.15, 0.43, 1.09 and 1.16. In this light, we agree with Severn Trent's decision to not control for regional wage differences in its revised set of models (SVT10). This says nothing about regional wage differences as an underlying cost driver, but reflects: • The deficiencies in the Ofwat PR14 regional wage variable. • The difficulty in capturing all material cost drivers in an econometric model of only 10 water and sewerage companies.
Capturing time	The first set of models in SVT1 included one model where time was captured through the inclusion of a time trend rather than through the inclusion of year-specific dummy variables, as was the case for the other three models presented in that initial set. Severn Trent's revised set of models, based on an amended dataset following our review of the data compilation, addressed this issue and adopted a consistent approach of time dummy variables.

Topic	Comment
Stata code: density variable for OLS model	In the Stata code provided to us by Severn Trent (SVT2), the OLS version of the network plus base expenditure translog model includes "density" rather than "lndensity)" as an explanatory variable. This seemed inconsistent with the intended treatment of the density variable (all other explanatory variables seem to be in logs). We confirmed with Severn Trent that this was just a problem in the State file SVT2, and not in the model estimation used to produce the results reported in SVT1 or SVT10.

Table 5 Limitations of models presented

Topic	Comment
Translog model	Severn Trent's consideration of translog models makes good sense in the light of the weight that Ofwat placed on the translog modelling approach for PR14. However, our starting position is that Ofwat's PR14 translog approach was ill-conceived and had bad consequences for the quality of Ofwat's overall cost assessment. We can see some potential arguments for elements of the translog approach (e.g. square terms). But with a small sample, there is a high risk of a complex model structure such as the translog producing misleading results. We have a concern that the translog models are "overfitted", in the sense that they contain too many explanatory variables given the size and nature of the data sample, and estimated coefficients may be reflecting differences in efficiency and noise in the data rather than underlying relationships between costs and cost drivers. While the sample size in the revised models (SVT10) cover 60 observations, there is a lack of variation over time in many of the explanatory which means that the power of the model is less than it would be from, say, a hypothetical cross section of 60 companies. The translog approach also makes it difficult to interpret the implied relationship between costs and cost drivers. For instance, in the translog OLS botex model presented by Severn Trent, the load treated by each company enters the model through three different explanatory variables: Ln(load); (Ln(load))^2 and Ln(load)*ln(Number of STWs). Furthermore, since the number of customers served will affect both load and number of works, there are interrelationships between these variables. We observe that the results of the translog models produced some counter-intuitive results with regard to the prediction of how costs vary with scale of activity (we shared Excel spreadsheet with Severn Trent that illustrated the concerns).
Use of load received as a scale variable	In their report, Jacobs put forward suggestion that "load removed" is a preferable scale variable than "load received" (page 87 and elsewhere).

Topic	Comment
	However, as highlighted by the Jacobs report, data on load removed are not readily available. In this light, it seems reasonable for Severn Trent to have used the data on load received.
Construction of variable on treatment complexity	In its translog and Cobb-Douglas models, Severn Trent included a variable that measures the number of tight consents (relating to ammonia or to BOD) across a company's works. Two implications of using this measure as a candidate cost driver are that: (i) the incremental costs of having a tight BOD consent are similar to those for a tight ammonia consent; and (ii) the incremental costs of having a tight BOD consent and a tight ammonia consent are seen as twice as the incremental costs of a tight consent on only of those parameters. We were not able to distil from the Jacobs report the reasonableness, or otherwise, of these implications. We discussed the point with Severn Trent. In its unit cost models, Severn Trent controlled for treatment complexity differently; it considered separate measures to reflect the proportion of load treated at works with tight ammonia consents, and the proportion treated at works with tight BOD consents.

Table 6 Potential extensions for wider set of models

Topic	Comment
Scope of models considered	There may be a case for considering the results implied by a wider set of models that contains more variety in modelling approaches rather than placing reliance on any one approach. By way of example: • **Granularity**: There is merit in considering models that take sewage collection separately from sewage treatment (and separately from sludge). The CMA criticised Ofwat's approach to water service models for PR14 for overlooking the benefits of more granular models, and Ofwat's PR19 draft methodology seems to take this point on board. • **Range of explanatory factors**. Where there are different ways to capture the same underlying cost driver through the specification of explanatory variables, and each has imperfections, it can be better to specify a set of alternative models that make use of different explanatory variables than to try to find a single best approach. We pick up on some of these issues in comments further below.
Addressing multicollinearity	Severn Trent's explanatory note (file SVT1) identified problems experienced with multicollinearity. We understand that there were alternative model specifications or variables that Severn Trent wanted to consider but which did not work well due to multicollinearity.

Topic	Comment
	The extent of correlations between variables is, in large part, a symptom of the overall approach of using a Cobb-Douglas or translog structure in a setting the dependent variable is an aggregate cost measure. One option is to include only one scale variable in the model, ensuring that other explanatory variables are expressed relative to this (e.g. number of STWs per unit of load). Another option we identified is the unit cost approach. This can help reduce correlations between explanatory variables and make it less difficult for the model estimation to isolate the effects of different underlying cost drivers. This is particularly relevant in a small sample size. For example: • Dependent variable could be the natural logarithm of expenditure per customer or expenditure per unit of load • If the former, the load variable could become load/customer (this might be dropped altogether if the load/customer ratio is very similar across companies). • The treatment complexity variables could be recast along the lines of proportion of load going through large tertiary works / works with tight ammonia consents, etc. Severn Trent's unit cost models presented in SVT11 are in this mould. A unit cost approach can also make it easier to interpret the implications of the coefficient on explanatory variables for the implied relationships between costs and cost drivers.
Use of number of STWs as an explanatory variable	We understand the rationale for consideration of models that make use of data on the number of STWs: there seems good logic for seeking to take account of differences between companies in ability to benefit from economies of scale at treatment works level. Whilst we understand that there could be concerns about the use of this measure as a cost driver– because it is not exogenous to management decisions – we do not consider that this is sufficient to warrants not considering it as an explanatory variable. It seems wrong to exclude variables just because there is some degree of management control. Many other candidate variables could be affected by this argument (e.g. length of sewers, proportion of water from reservoirs, etc.). Ultimately there is a trade-off between the benefits of using a particular variable and the drawbacks/risks. There remains a concern about placing too much reliance on models that are based on a variable capturing the number of works. One approach would be to also consider models that include variables that reflect "more exogenous factors", i.e. that capture the underlying factors that may lead to differences between companies in the extent to which they benefit from economies of scale in treatment works. For instance, variables relating to customer/population density within the area of appointment. Such models may not perform as well in terms of perceived accuracy, but are worthy of consideration as they tackle the issues above.

Topic	Comment
Number of connected properties as a potential variable	Whilst the costs modelled include sewage collection costs, the number of customers supplied (e.g. number of connected properties) is a relevant cost driver that does not feature in the models. It may be worth considering whether there is a basis for covering models that take account of the number of properties served, either as part of the set of explanatory variables or by capturing this within the dependent variable (e.g.by modelling cost per property) and checking whether this makes any material difference.
Reflecting volumes at peak period	We think it would be useful to explore the inclusion in the models of a cost driver to reflect volumes at peak periods, such as peak to average load. We would expect "peak capacity" to be a factor in sizing the network and treatment capacity of the system as a whole, and companies may differ in the extent of their peak to average ratio.

Table 7 Comment on initial write-up of models

Topic	Comment
Expected coefficient on load variable	Severn Trent's explanatory note (SVT1) said on page 1 *"The main costs that vary with a change in load are power and chemical costs which make up 40%-50% of Severn Trent's operating costs at STW's. Given that the coefficient on the treatment complexity variable should account for the increased power and chemical costs that the more stringent treatment consents entail, we would expect an upper limit of around 0.5 on the load variable"* We are unsure about the basis for this statement in the context of models of base expenditure that cover the whole wastewater service or wastewater network plus: • We expect load to affect capital maintenance expenditure as well as operating expenditure (e.g. power and chemicals), particularly for sewage treatment and for bioresources, even allowing for some economies of scale at the sewage treatment centre level. • The dependent variable in all of Severn Trent's models covers not just sewage treatment expenditure but also sewage collection expenditure. As such, information on the cost drivers at STWs would only relate to a subset of modelling costs. Load may not be a direct cost driver for the sewage collection part of the business; the total number of connected properties may be a more relevant cost driver. But the dataset shows that total load and total connected properties are highly correlated. We might expect a relationship between load and sewage collection base expenditure of approximately 1:1.

Appendix: supporting tables

Topic	Comment
Expectation on coefficient on variable reflecting the number of STWs	Severn Trent states that it expects the coefficient on the variable capturing the number of STWs to be 0.2, adding a caveat that this assumes a company only uses large Band 6 work. The basis for the figure is a comparison of the costs of two of Severn Trent's Band 6 treatment works that are similar but for their capacity; the capacity of one is (roughly) twice the other and has a unit cost that is roughly 80 per cent of the other. We think that the expectation concerning the minimal size of the coefficient is unwarranted. A comparison of two of Severn Trent's works provides little guidance on the relation that might be identified from data across the industry. Further, the comparison done of the two works was concerned with operating expenditure, and did not include capital expenditure relating to the maintenance of the assets.
Write-up of rationale for models	The write-up of the initial set of models (presented in SVT1) was based on a dataset that (following our review of the data) was subsequently revised.