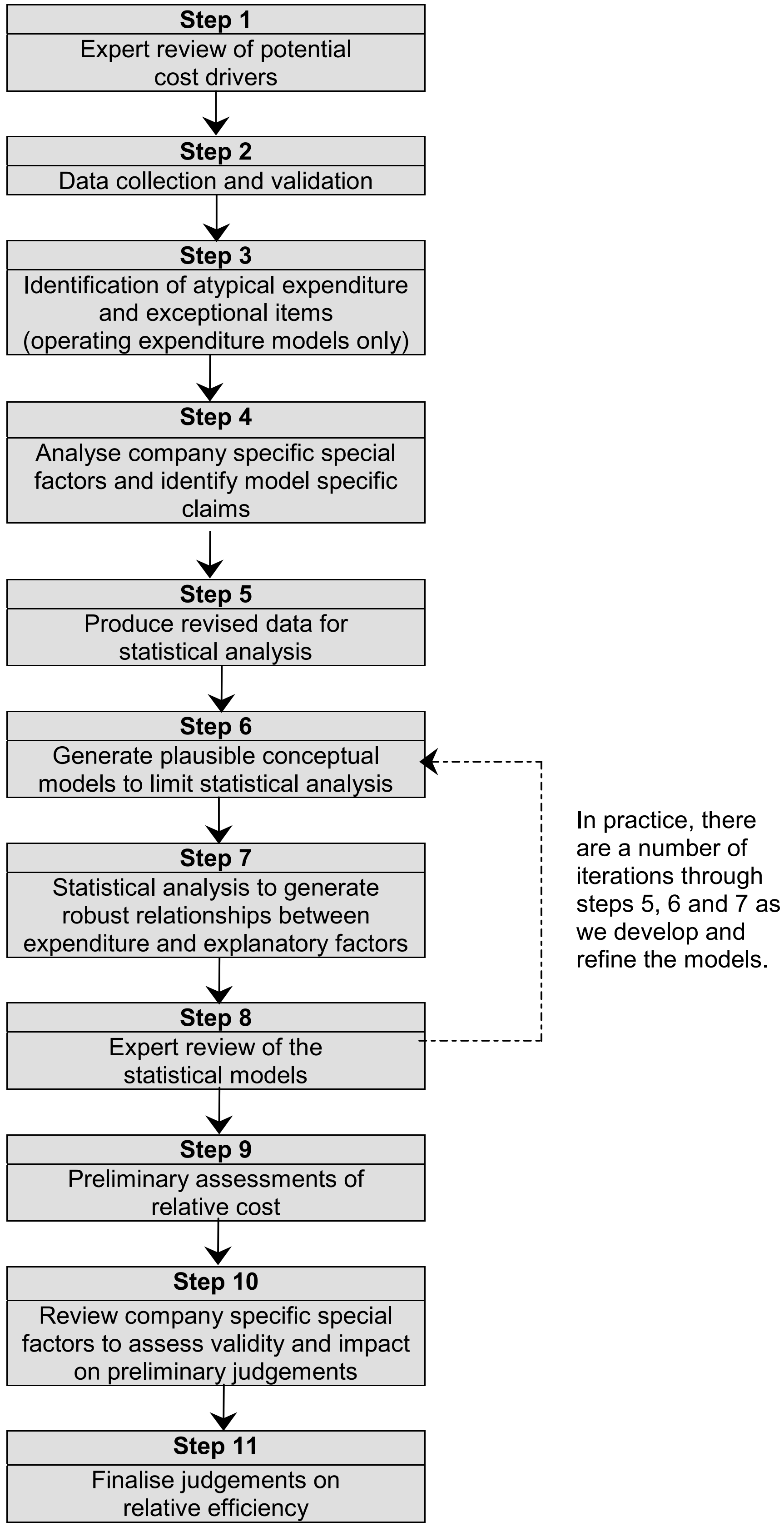
Appendix 1: Efficiency models

Step-by-step approach used to derive the statistical models



Operating expenditure – water models

We assess operating expenditure relative efficiency for the water service using four econometric models. The form and coefficients of the models is set out later in this appendix.

Water distribution

This is a log model expressed in unit cost form with resident population as the scale variable. The proportion of large mains to small mains is the cost driver in this model. We use this as a proxy for urbanisation. Repairs, maintenance and inspection of large mains incur much greater costs than the same work on small mains.

Water resources and treatment

This is a linear model expressed in unit cost form with resident population as the scale variable. This year we made a change to one of the variables in the model. We now include the proportion of supplies from boreholes, instead of the proportion of supplies from rivers. We also take into account the explanatory variables of number of sources and distribution input. These variables ensure that we take into account economies of scale at source level (costs will be lower if fewer sources are used) and the difficulty of treatment (borehole supplies will generally be cheaper to treat).

Water power

This is a log linear model. For most companies, power expenditure is almost entirely for pumping, although there are some water treatment processes that use a lot of energy. The model considers the effects of terrain (companies in hilly areas will require more power to move water around) and the significant economies of scale associated with high power consumption. The cost drivers in this model are average pumping head and distribution input.

Water business activities

This is a log linear model. Business activities include customer services and scientific services, and the charge for doubtful debts. The cost driver is number of billed properties. The model takes into account the economies of scale associated with high volume billing and customer service activities.

Operating expenditure – sewerage models

Three of the operating expenditure models for the sewerage service are unit cost models for:

- small sewage treatment works (and sea outfalls);
- sludge treatment and disposal; and
- business activities.

The remaining two operating expenditure econometric models for the sewerage service are set out below.

Network including power

This is a log linear model expressed in unit cost form with the total length of sewer as the scale variable. The explanatory variables used in the network model are:

- sewer length;
- area of sewer district;
- resident population; and
- holiday population.

In simple terms, the model takes account of the density of the sewerage network and the population it serves, and of the higher costs associated with the sewer capacity required to serve additional summer populations.

Large sewage treatment works

This is a log linear model. It uses a number of explanatory variables that take account of the total load, the type of treatment used, and the nature of effluent consents. These explanatory variables affect costs (it is more expensive to meet tight effluent consents, for example).

Capital maintenance expenditure – water models

The capital maintenance assessment consists of one unit cost comparison (water resources and treatment) and three econometric models, which are set out below.

Water distribution infrastructure

This is a log unit cost model expressed in unit cost form with length of main as the scale variable. The key cost driver in this model is the log of connected properties per length of main.

Water distribution non-infrastructure

This is a log unit cost model expressed in unit cost form with total booster pumping station capacity as the scale variable. We take into account water tower and service reservoir capacities in this model. The ratio of storage capacity to pumping station capacity helps to explain the higher costs of companies with relatively greater storage capacity.

Water management and general

This is a log model expressed in unit cost form with the total number of billed properties as the scale variable. The key cost driver in this model is the proportion of billed properties that are non-household. The model explains the higher unit costs incurred by companies with more business customers due to metering and billing requirements.

Capital maintenance expenditure – sewerage models

For capital maintenance expenditure, we assess the sewerage service using unit cost models for:

- sewerage non-infrastructure;
- sludge treatment and disposal; and
- management and general.

The remaining two econometric models are set out below.

Sewage treatment

This is a log unit cost model with the total load treated by works as the scale variable. The explanatory factor in this model is the load treated per sewage treatment works. Economies of scale at the works level are taken into account, because fewer larger works will be less costly to maintain, as are the higher costs associated with post-primary levels of treatment.

Sewerage infrastructure

This is a log unit cost model with the total length of sewers as the scale variable. The number of combined sewer overflows is the key cost driver in this model. It is an indicator of the extent of combined sewers in the network. Combined sewers are generally larger and more costly to replace than foul sewers. This cost driver therefore helps to explain the higher maintenance costs incurred by companies with a greater number of combined sewer overflows.

Definition of the terms and abbreviations in the econometric model descriptions

We have used the following terms and abbreviations in tables a-r.

Model significance (F test): We use this to determine whether there is an association between the modelled cost and the other variables used in a model. The number in our tables is the probability that there is no association between the variables. Therefore, a small number (we usually say less than 0.05) means that there is a statistically significant association between the variables.

R^2: This shows how closely the variables in the model are related to each other. It lies between 0 and 1; an R^2 of 1 means there is a perfect linear relationship and an R^2 of 0 means there is no linear relationship. However the R^2 value can be misleading if used on its own. Other statistics should be considered as well to assess how well a model performs.

Model standard error: This is a measure of the overall variation in a model. It measures how spread out the data is around the regression line. The further away the data points are from the regression line the greater the value of the standard error. It is based on the standard deviation of the data points – the smaller the number relative to the mean of the data the better the fit of the regression line.

In: The logarithm to base e, also called the natural logarithm.

BOD_5: Five-day biochemical oxygen demand (BOD_5) is a standard way of measuring the polluting effect of wastewater. It is a measure of how much oxygen is consumed in five days at a temperature of 20ºC by any micro-organisms or organic matter present in the water.

Table a Operating expenditure – water distribution model

Water service:	Distribution expenditure	
Data:	June returns 2006	
Modelled cost:	In (distribution functional expenditure excluding power expenditure [£m], divided by resident winter population [000s])	
Explanatory variables:	Coefficient	Standard error
Constant	-5.118	0.169
Length of main greater than 300 mm diameter [km] divided by total length of main [km]	4.534	2.118
Form of model:	Modelled cost = -5.118 + 4.534 x $\frac{\text{length of main > 300 mm diameter}}{\text{total length of main}}$	
Statistical indicators:	Number of observations: 22	R^2: 0.186
	Model standard error: 0.232	Model significant (F test): 0.045

Table b Operating expenditure – water resources and treatment model

Water service:	Resources and treatment expenditure	
Data:	June returns 2006	
Modelled cost:	Resources and treatment functional expenditure [£m], less power expenditure [£m], less service charges [£m], divided by resident winter population [millions]	
Explanatory variables:	Coefficient	Standard error
Constant	5.199	0.927
Number of sources divided by distribution input [Ml/day]	24.082	6.957
Proportion of supplies derived from boreholes	-5.124	1.844
Form of model:	Modelled cost = 5.199 + 24.082 x $\frac{\text{number of sources}}{\text{distribution input}}$ – 5.124 x (proportion of supply from boreholes)	
Statistical indicators:	Number of observations: 22	R^2: 0.389
	Model standard error: 2.010	Model significant (F test): 0.009

Table c Operating expenditure – water power model

Water service:	Power expenditure	
Data:	June returns 2006	
Modelled cost:	ln power expenditure [£m]	
Explanatory variables:	Coefficient	Standard error
Constant	-8.326	0.254
ln (distribution input (DI) [Ml/day] multiplied by average pumping head)	0.908	0.023
Form of model:	Modelled cost = -8.326 + 0.908 x ln (DI x average pumping head)	
Statistical indicators:	Number of observations: 22	R^2: 0.987
	Model standard error: 0.145	Model significant (F test): 0.000

Table d Operating expenditure – water business activities model

Water service:	Business activities expenditure	
Data:	June returns 2006	
Modelled cost:	ln (business activities expenditure [£m] including doubtful debts [£m])	
Explanatory variables:	Coefficient	Standard error
Constant	-3.820	0.213
ln of number of billed properties [000s]	0.945	0.033
Form of model:	Modelled cost = -3.820 + 0.945 x ln (number of billed properties)	
Statistical indicators:	Number of observations: 22	R^2: 0.976
	Model standard error: 0.186	Model significant (F test): 0.000

Table e Operating expenditure – sewerage network (including power) model

Sewerage service:	Sewerage network expenditure	
Data:	June returns 2006	
Modelled cost:	ln (sewerage network functional expenditure [£m], less service charges [£m], less BWB charges [£m], per kilometre of sewer, for each area)	
Explanatory variables:	Coefficient	Standard error
Constant	-5.802	0.531
ln (area sewer district [km^2] per kilometre of sewer)	0.184	0.042
ln (resident population [000s] per kilometre of sewer)	0.737	0.283
Holiday population divided by resident population [000s]	1.761	1.250
Form of model:	Modelled cost = - 5.802 + 0.184 x ln (area of sewer district per km of sewer) + 0.737 x ln (resident population per km of sewer) + 1.761 x $\frac{\text{holiday population}}{\text{resident population}}$	
Statistical indicators:	Number of observations: 63	R^2: 0.314
	Model standard error: 0.307	Model significant (F test): 0.000

Table f Operating expenditure – large sewage treatment works model

Sewerage service:	Cost of sewage treatment at large works	
Data:	June returns 2006	
Modelled cost:	ln (functional expenditure on sewage treatment at large works [£000s] less service charges [£000s] less terminal pumping costs [£000s])	
Explanatory variables:	Coefficient	Standard error
Constant	-1.670	0.244
ln of total load[1] [kg COD/day]	0.806	0.027
Activated Sludge[2]	0.390	0.051
Tight effluent consent for both suspended solids and BOD_5[3]	0.080	0.046
Form of model:	Modelled cost = -1.670 + 0.806 x ln (total load) + 0.390 x activated sludge + 0.080 x tight effluent consent for both suspended solids and BOD_5	
Statistical indicators:	Number of observations: 392	R^2: 0.745
	Model standard error: 0.456	Model significant (F test): 0.000

Table g Operating expenditure – small sewage treatment works unit cost model

Sewerage service:	Cost of sewage treatment at small works
Data:	June returns 2006
Unit cost model:	We used a unit cost approach for modelling sewage treatment at small works, using ten treatment categories and five size bands. For each treatment type we compared each company's average annual expenditure (direct costs [£000s], less service charges [£000s], less sludge costs [£000s], plus general and support costs [£000s]) with each company's estimated expenditure (weighted average industry unit cost multiplied by each companies load [kg BOD_5/day]).

	Weighted average industry unit cost £000s/(kg BOD_5/day)									
Treatment type	Primary	Secondary activated sludge	Secondary biological	Tertiary A1	Tertiary A2	Tertiary B1	Tertiary B2	Sea outfall preliminary	Sea outfall screened	Sea outfall unscreened
Size band 1	0.69	0.98	1.04	1.08	1.94	1.06	1.38	3.17	0.04	0.49
Size band 2	0.21	0.76	0.51	0.87	0.40	0.64	0.55	0.00	0.00	0.00
Size band 3	0.19	0.48	0.31	0.48	0.40	0.37	0.44	0.07	0.14	0.00
Size band 4	0.26	0.27	0.19	0.25	0.26	0.19	0.22	0.00	0.27	0.00
Size band 5	0.00	0.16	0.13	0.15	0.17	0.11	0.13	0.02	0.00	0.02
Number of observations:	500									

Notes:
1. Total load in this model is estimated as population equivalent x 120.
2. Activated sludge includes secondary and tertiary treatment (variable value is 0 if not used, 1 if used).
3. The tight effluent consent is based on the levels of suspended solids (mg/l) and BOD_5 (mg/l) and on the ammonia levels (mg/l) (variable value is 0 if tight consent does not apply, and 1 if the tight consent does apply).

Table h Operating expenditure – sludge treatment and disposal unit cost model

Sewerage service:	Cost of sludge treatment and disposal							
Data:	June returns 2006							
Unit cost model:	We used a unit cost approach for modelling the treatment and disposal of sludge. For each disposal route, we compared each company's average annual expenditure (sludge functional expenditure [£000s], less service charges [£000s]) with each company's estimated costs (weighted average industry unit cost multiplied by each company's load [ttds]).							
Weighted average industry unit cost £000s/(thousand tonnes of dry solids)								
Disposal route	**Farmland** (untreated)	**Farmland** (conventional)	**Farmland** (advanced)	**Incineration**	**Landfill**	**Composted**	**Land reclamation**	**Other**
£000s/ttds	185.3	161.4	228.6	213.9	127.3	180.7	180.6	244.1
Number of observations:	80							

Table i Operating expenditure – sewerage business activities unit cost model

Sewerage service:	Sewerage business activities expenditure
Data:	June returns 2006
Unit cost model:	We used a unit cost approach for modelling business activities, based on the number of billed properties. We compared each company's average annual business activities expenditure (total business activities [£m], plus doubtful debts [£m], divided by the number of billed properties) with the weighted average industry cost.
£/billed property	Weighted average industry unit cost 12.90
Number of observations:	10

Table j Capital maintenance expenditure – water distribution infrastructure model

Water service:	Water distribution infrastructure expenditure	
Data:	Capital maintenance econometric return, June returns 2000-2005	
Modelled cost:	ln (annual average water distribution infrastructure functional expenditure [£m], divided by length of main [km])	
Explanatory variables:	Coefficient	Standard error
Constant	-5.103	0.661
ln (total number of connected properties [000s], divided by total length of main [km])	0.739	0.244
Form of model:	Modelled cost = $-5.103 + 0.739 \times \ln\left\{\frac{\text{no. of connected properties}}{\text{total length of main}}\right\}$	
Statistical indicators:	Number of observations: 22	R^2: 0.314
	Model standard error: 0.260	Model significance (F test): 0.007

Table k Capital maintenance expenditure – water distribution non-infrastructure model

Water service:	Water distribution non-infrastructure expenditure	
Data:	Capital maintenance econometric return, June returns 2000-2005	
Modelled cost:	ln (annual average water distribution non-infrastructure functional expenditure [£m], divided by pumping station capacity [kW])	
Explanatory variables:	Coefficient	Standard error
Constant	-5.739	0.530
ln (water service reservoir and water tower storage capacity [Ml/d] divided by pumping station capacity [kW])	0.941	0.206
Form of model:	ln modelled cost = $-5.739 + 0.941 \times \ln\left\{\frac{\text{storage capacity}}{\text{pumping station capacity}}\right\}$	
Statistical indicators:	Number of observations: 22	R^2: 0.510
	Model standard error: 0.565	Model significance (F test): 0.000

Table l Capital maintenance expenditure – water management and general model

Water service:	Water management and general expenditure	
Data:	Capital maintenance econometric return, June returns 2000-2005	
Modelled cost:	In (annual average management and general expenditure [£m], divided by billed properties [000s])	
Explanatory variables:	Coefficient	Standard error
Constant	-5.543	0.255
Proportion of billed non-household properties	9.165	3.324
Form of model:	Modelled cost = -5.543 + 9.165 x proportion of non-household properties	
Statistical indicators:	Number of observations: 21	R^2: 0.286
	Model standard error: 0.209	Model significance (F test): 0.013

Table m Capital maintenance expenditure – water resources and treatment model

Water service:	Water resources and treatment expenditure
Data:	Capital maintenance econometric return, June returns 2000-2005
Unit cost model:	We used a unit cost approach for the water resource and treatment model. We divided each company's average annual expenditure by the total connected properties and compared this with the weighted average industry cost.
£/connected properties	Weighted average industry cost 8.973
Number of observations:	22

Table n Capital maintenance expenditure – sewerage infrastructure model

Sewerage service:	Sewerage infrastructure expenditure	
Data:	Capital maintenance econometric return, June returns 2000-2005	
Modelled cost:	ln (average annual sewerage infrastructure expenditure [£m], divided by the total length of sewer [km])	
Explanatory variables:	Coefficient	Standard error
Constant	-6.141	0.211
ln (the number of combined sewer overflows (CSOs) divided by the total length of sewer [km])	0.385	0.060
Form of model:	Modelled cost = -6.141 + 0.385 x ln $\left\{ \frac{\text{number of CSOs}}{\text{total length of sewer}} \right\}$	
Statistical indicators:	Number of observations: 63	R^2: 0.399
	Model standard error: 0.471	Model significance (F test): 0.000

Table o Capital maintenance expenditure – sewage treatment model

Sewerage service:	Sewage treatment expenditure	
Data:	Capital maintenance econometric return, June returns 2000-2005	
Modelled cost:	ln (average annual sewage treatment functional expenditure [£m], divided by the total load received at sewage treatment works [kg BOD_5/day])	
Explanatory variables:	Coefficient	Standard error
Constant	-7.849	0.300
ln (the total number of works divided by total load received at sewage treatment works [kg BOD_5/day])	0.204	0.044
Form of model:	Modelled cost = -7.849 + 0.204 x ln $\left\{ \frac{\text{total number of works}}{\text{total load received at works}} \right\}$	
Statistical indicators:	Number of observations: 60	R^2: 0.270
	Model standard error: 0.535	Model significance (F test): 0.000

Table p Capital maintenance expenditure – sewerage non-infrastructure model

Sewerage service:	Sewerage non-infrastructure expenditure
Data:	Capital maintenance econometric return, June returns 2000-2005
Unit cost model:	A unit cost approach resulted from modelling sewerage non-infrastructure data. We compared each company's average annual expenditure [£m] divided by the total number of pumping stations with the weighted average industry cost.
£m/number of pumping stations [000s]	Weighted average industry cost 3.304
Number of observations:	10

Table q Capital maintenance expenditure – sludge treatment and disposal model

Sewerage service:	Sludge treatment and disposal expenditure
Data:	Capital maintenance econometric return, June returns 2000-2005
Unit cost model:	We used a unit cost approach for the sludge treatment and disposal model. We compared each company's average annual sludge expenditure [£000s], divided by the total weight of dry solids disposed of [ttds] with the weighted average industry cost.
£000s/thousand tonnes of dry solids	Weighted average industry cost 70.566
Number of observations:	10

Table r Capital maintenance expenditure – sewerage management and general model

Sewerage service:	Sewerage management and general expenditure
Data:	Capital maintenance econometric return, June returns 2000-2005
Unit cost model:	We adopted a unit cost approach for sewerage management and general expenditure, comparing each company's average annual expenditure [£m] per billed property [millions] with the weighted average industry cost
£/billed property	Weighted average industry cost 6.768
Number of observations:	10