Water & Sewerage Services

Price Control 2015-21

Final Determination - Annex Q

Alternative Efficiency Modelling

December 2014

Contents page

Water and Sewerage Services Price Control 2015-21

Final Determination Annex Q Alternative Efficiency Modelling

1.0 Introduction

1.1. Problems

1.1.1 The efficiency gap has been assessed using the corrected ordinary least squares (COLS) models. The findings suggest a gap estimate of 22%. This means that NI Water would have to reduce costs by this amount to be a frontier company.

1.1.2 Within the business plan, the company has cited a number of concerns with COLS. These include:

a) Models have become outdated and less robust over time;

b) The method used to allow for real opex changes may introduce bias as each companies costs do not move uniformly;

c) The frontier companies may now be different due to changing opex; and

d) Inclusion of PPP costs increases uncertainty of the models.

1.1.3 Since June Return data stopped being published, benchmarking has become harder to do. The COLS models are a cross-sectional comparison at a particular point in time. Unfortunately, the present approach means comparing updated NI Water data with older England and Wales information.

1.1.4 The Utility Regulator (UR) recognises the issue and has tried to address it to some extent within the current methodology. Analyses of regulatory accounts help to provide changes in company and industry opex. These changes are reflected by adjusting average and frontier positions.

1.1.5 However, it is accepted that an issue still exists. The models do not include updated explanatory data. Some bias may also be introduced by simply amending predicted cost by the industry change.

1.1.6 To correct for these issues, the UR has completed further efficiency modelling.

1.2. Solution

1.2.1 The PR14 August data submission has provided updated (2012-13) cost and physical data for the England and Wales companies. The submission is not as detailed as June Returns. They do however allow for high-level comparison between companies.

1.2.2 In order to test the robustness of the COLS findings, the UR has developed some total opex (topex) models. These can be used to triangulate around an efficiency figure given that no one method is perfect.

1.2.3 To help improve model precision, the latest data has been pooled with historic June Returns. The UR has used figures from 2008-09 to 2012-13 for the ten WaSCs. This gives 50 observations against which NI Water can be compared.

1.2.4 The purpose of this annex is to explain the 'pooled' models, their rationale and findings. By considering updated models, the UR can test if the COLS problems are adversely affecting the efficiency gap analysis.

1.2.5 The benefit of developing new models is that they provide an independent check on the COLS gap. This can help verify the results or identify areas where concern may arise.

1.3. Model types

1.3.1 A number of potential models are investigated. The options reflect the type of analyses that could be undertaken to support the COLS findings and aid target setting.

1.3.2 In this paper, eight alternative options are presented, though any number of variations exists. The options include:

a) **Properties Unit Cost –** This compares NI Water's cost per property against the England and Wales industry average;

b) **Volume of Water/Wastewater Unit Cost** – The second method is again a simple unit cost comparison. This method differs only in the use of volumes rather than properties;

c) **Population Topex Model** – A log regression using total opex as the dependent variable. Population served is the single explanatory factor;

d) **CSV Topex Model A –** This option regresses log opex against a composite scale variable. The variable differs between the water and sewage models but includes factors considered key cost drivers. Each is given an equal weighting;

e) **CSV Topex Model B –** A similar regression to Model A, but explores the use of different weights;

f) **CSV Upper Quartile** – A repeat of Model B. The difference being the frontier used is the 2nd and 3rd ranked WaSCs, rather than a chosen company;

g) **CSV Standardised Variable Weights –** The model uses individual variable coefficients to determine the weights used for the composite variable. The variables are first converted onto a standard scale; and

h) **NI Water Inclusive** – Repeats the standardised variable model. The difference is that NI Water observations form part of the regression.

1.3.3 The models were chosen for a variety of reasons. The unit cost method has appeal in that it is a relatively simple approach, which is easy to grasp.

1.3.4 The denominators (properties and volumes) are similar to those previously used by Ofwat to compare companies' opex costs in their annual *Water and Sewerage Service Unit Costs and Relative Efficiency* reports.

1.3.5 The population regression estimates the relationship between costs and the population served. It is anticipated the two variables will be closely correlated.

1.3.6 CSV model A uses a composite of variables. Combining them has the benefit of including their impact without introducing a correlation problem. Multicollinearity issues often exist with a number of related independent variables.

1.3.7 CSV model B uses the same approach but varies the weights allocated to each input. The upper quartile (UQ) takes these same results but compares to a different frontier. Such a model helps to determine if the current chosen frontier companies bias results.

1.3.8 The CSV Standardised Variable Weights model involves two steps. First, individual variables are changed to a standard scale. This ensures that coefficient weights are not arbitrarily distorted by the different units of measurement of each variable.

1.3.9 Standardisation is done by subtracting the mean from each observation, then dividing by the standard deviation. The result is that each variable has the same standard unit.

1.3.10 Secondly, a regression is run using the *individual* standard variables. Their coefficients then determine the weights used for the CSV. This approach follows on from work submitted by NI Water in their consultation response.

1.3.11 The NI Water Inclusive model repeats the standardised variable approach. On this occasion, NI Water is included in the regression data.

1.3.12 The company raised this as an issue within the consultation response. As they are no longer an outlier, NI Water believes it beneficial to be included in the sample. The UR has considered it prudent to look at the impact this might have.

1.4. Assumptions

1.4.1 The models have been developed using five years of pooled data, from 2008-09 onward. Nominal cost is inflated to 2012-13 prices. The adopted approach provides more observations. This should help improve model specifications and accuracy.

1.4.2 In the interest of simplicity and comparability with the COLS models, a number of assumptions have been made. These include:

- Total opex for all models excludes business activity and other removed costs e.g. rates, service charges, third party costs etc;
- NI Water observations are excluded from the calculation of the unit cost industry average and the topex regressions. The final model is an exception;

- Special factors and atypical costs are the same as those used in the COLS analysis;
- Residual adjustments remain the same as COLS (10% for water models and 20% for sewerage); and
- Frontier companies remain as Yorkshire Water and Wessex Water for the water and sewage areas. The exception is the use of the upper quartile for comparison.

1.4.3 At this stage, the assumptions are for the benefit for simplicity. They may not however reflect what a more robust approach might look like.

2.0 Unit Costs – Properties

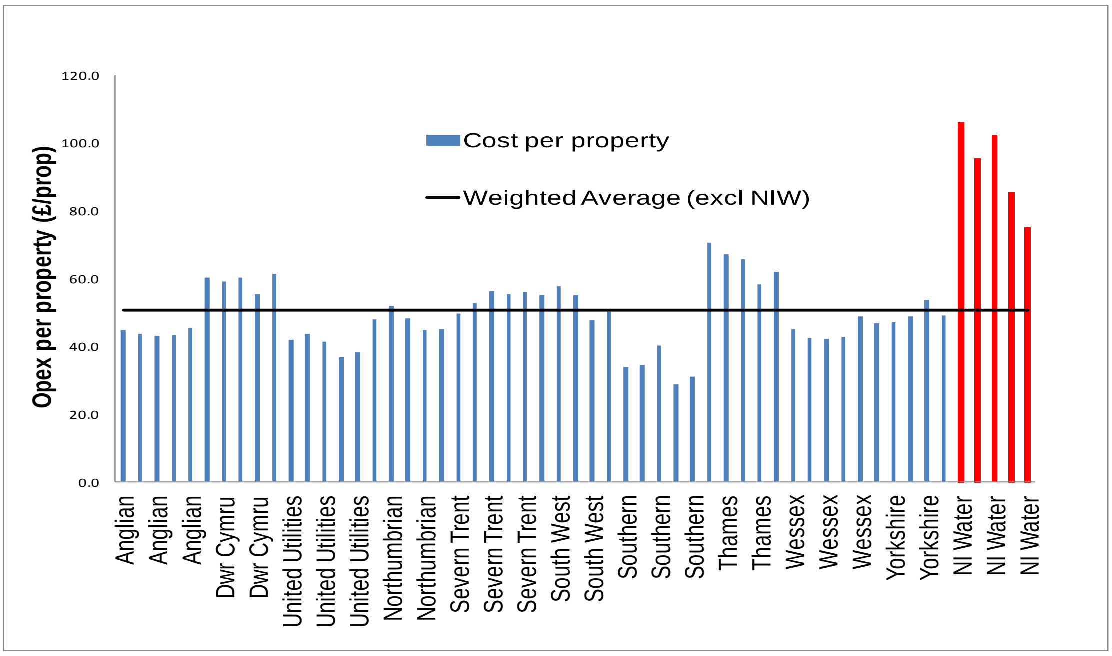
2.1. Water properties

2.1.1 The simplest method of benchmarking is unit costs. This involves dividing costs by a variable that drives expenditure. In this case, connected properties is the chosen variable as billed property data is restricted.

Table 2.1 – Water service cost per connected property

Water Service:	**Total opex (excluding business activities)**
Data:	June Returns and PR14 August submission
Unit cost model:	The unit cost reflects the weighted industry opex cost per connected property. Comparison is made of functional expenditure less service charges and business activities against predicted costs (connected properties multiplied by the weighted average industry unit cost).
£/property	Weighted average industry unit cost: £50.64[1]
Number of observations	50

Figure 2.1 – Water service cost per connected property



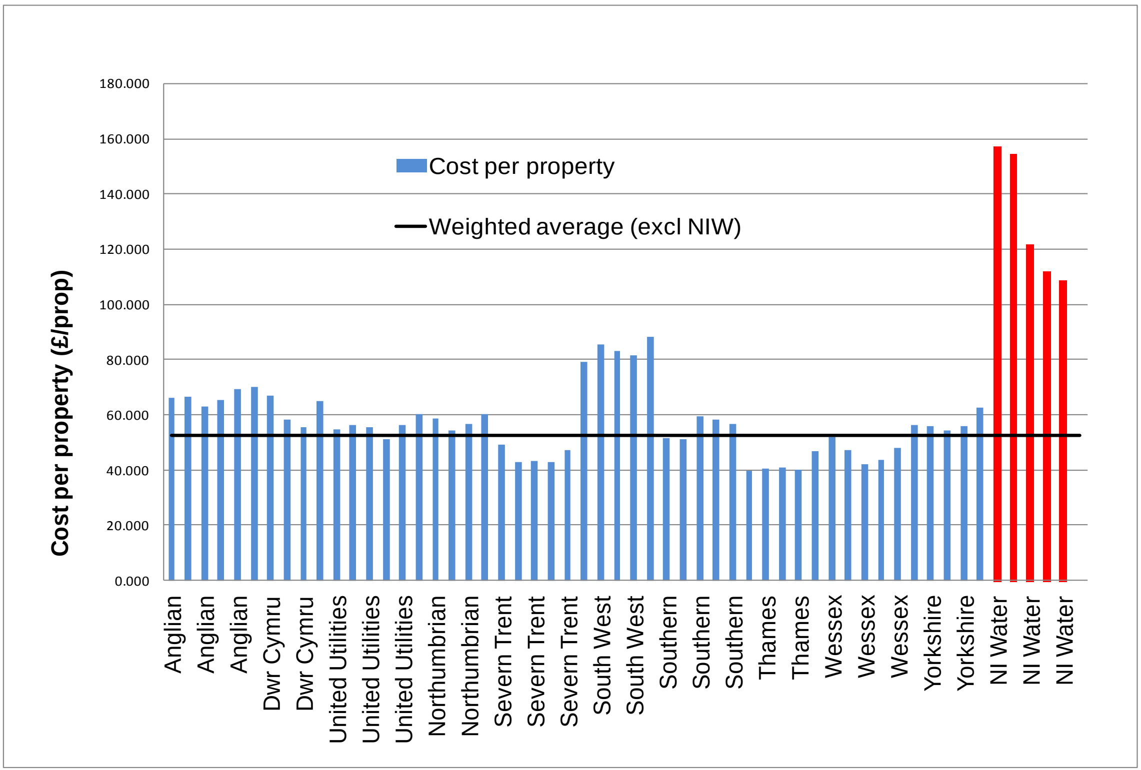
[1] All financial figures in this report are given in 2012-13 prices unless otherwise stated.

2.1.2 Comparisons show NI Water improvement over time. They also indicate much higher spending levels than comparators.

2.2. Sewage properties

2.2.1 Billed properties are used as the denominator for the sewage model. The story is similar to water, though the gap is greater.

Figure 2.2 – Sewage service cost per billed property



2.3. Unit cost efficiency gap

2.3.1 Calculation of the efficiency gap follows the same process as the COLS models. NI Water predicted cost is found by multiplying the industry average cost by its own property data.

Table 2.2 – Predicted cost calculations for water and sewage

Special Factor Claim	Water 2012-13	Sewage 2012-13
NI Water properties (000's)	818.00	618.38
Industry average unit costs (£/prop)	50.64	52.64
NI Water predicted cost (£m)	41.43	32.55

2.3.2 Residual adjustments are then applied alongside a frontier correction. This gives an efficiency gap as follows:

Table 2.3 – Properties unit cost efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)					
	Category	**Process Rule**	**Water (£m)**	**Sewerage (£m)**	**Total (£m)**
A	NI Water actual cost		61.52	67.34	128.86
B	Less Adjustments		-4.12	-4.46	-8.58
C	**Modelled Cost**	**A – B**	**57.40**	**62.88**	**120.28**
D	Predicted Cost (average)		41.43	32.55	73.98
E	Difference	C – D	15.97	30.33	46.30
F	Adjustment Factor (%)		10%	20%	
G	Residual Adjustment	E * F	1.60	6.07	7.66
H	**New Predicted Costs**	**D + G**	**43.02**	**38.62**	**81.64**
I	Frontier Adjustment (%)		-2.51%	-7.22%	
J	**Frontier Predicted Costs**	**H * (1 + I)**	**41.94**	**35.83**	**77.78**
K	Efficiency Gap (to average)	C – H	14.37	24.27	38.64
L	**Efficiency Gap % (to average)**	**K / C**	**25.04%**	**38.59%**	**32.12%**
M	Efficiency Gap (to frontier)	C – J	15.45	27.05	42.51
N	**Efficiency Gap % (to frontier)**	**M / C**	**26.92%**	**43.02%**	**35.34%**

Figures may not sum due to rounding.

2.3.3 The findings suggest inefficiency levels quite a bit larger than the COLS analysis. Not much faith is placed in these results. While special factors are included to the same extent as the COLS models, simple unit costs do not properly account for (dis)economies of scale.

2.3.4 The selection of the explanatory variable is also crucial. The next models highlight the difference by replacing properties with volumes as the denominator.

3.0 Unit Costs – Volumes

3.1. Water volumes

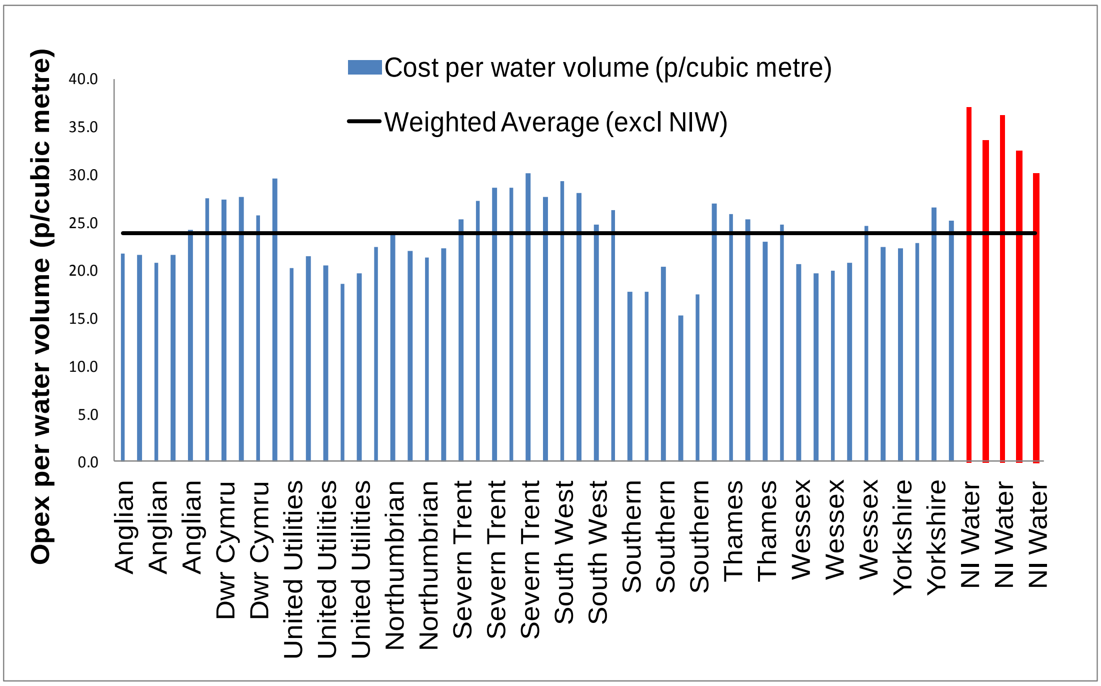
3.1.1 This approach mirrors the property method, the only difference being the denominator. In this instance, costs are divided by the volume of water entering the system. The form of the model is as follows:

Table 3.1 – Cost per water volume

Water Service:	Total opex (excluding business activities)
Data:	June Returns and PR14 August submission
Unit cost model:	The unit cost reflects the weighted industry opex cost per distribution input. Comparison is made of functional expenditure less service charges and business activities against predicted costs (distribution input multiplied by the weighted average industry unit cost).
pence /m³	Weighted average industry unit cost: 23.87p
Number of observations	50

Figures may not sum due to rounding

Figure 3.1 – Cost per water volume



3.1.2 The graph shows annual improvement for NI Water. These comparisons show the company much closer to the industry average than before.

3.2. Sewage volumes

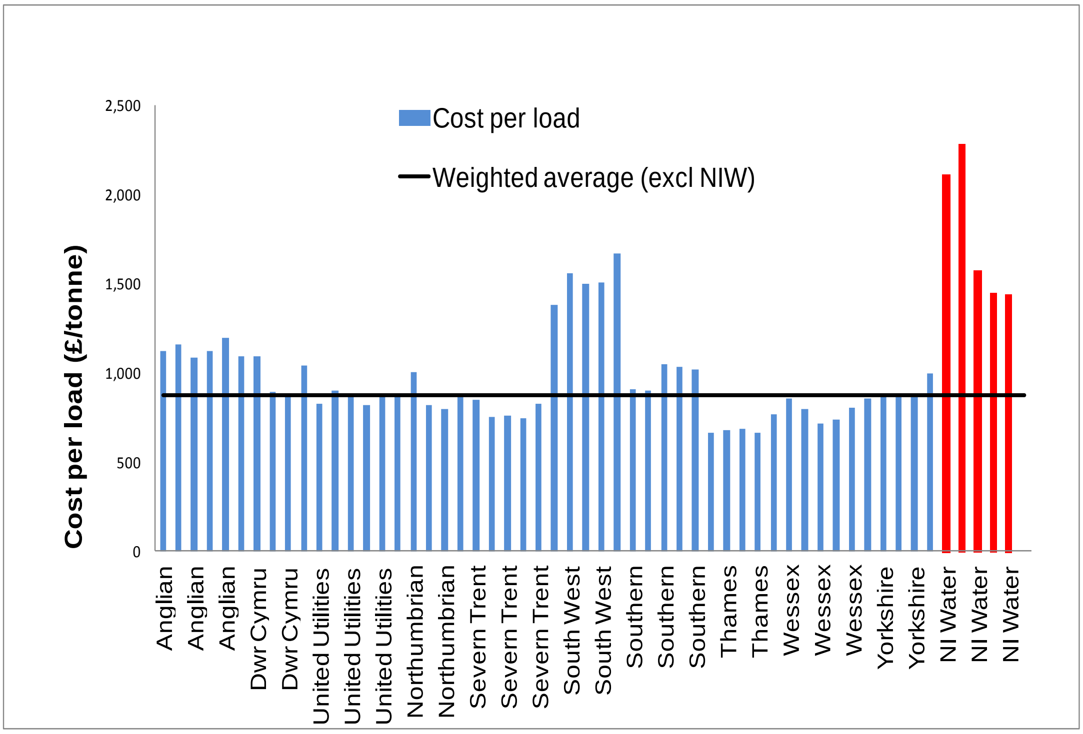
3.2.1 The total load entering the system explains sewage costs.

Table 3.2 – Cost per sewage load

Sewage Service:	Total opex (excluding business activities)
Data:	June Returns and PR14 August submission
Unit cost model:	The unit cost reflects the weighted industry opex cost per tonne of sewage load. Comparison is made of functional expenditure less service charges and business activities against predicted costs (the total load multiplied by the weighted average industry unit cost).
£/tonne BOD5	Weighted average industry unit cost: £872.13
Number of observations	50

Figures may not sum due to rounding

Figure 3.2 – Cost per sewage load



3.2.2 Whilst the sewage load has not varied greatly for NI Water over the years, costs have constantly been reducing.

3.3. Unit cost efficiency gap

3.3.1 In 2012-13 the efficiency gap using this unit cost is:

Table 3.3 – Volume unit cost efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)					
	Category	**Process Rule**	**Water (£m)**	**Sewerage (£m)**	**Total (£m)**
A	NI Water actual cost		61.52	67.34	128.86
B	Less Adjustments		-4.12	-4.46	-8.58
C	**Modelled Cost**	**A – B**	**57.40**	**62.88**	**120.28**
D	Predicted Cost (average)		48.70	40.91	89.61
E	Difference	C – D	8.70	21.98	30.68
F	Adjustment Factor (%)		10%	20%	
G	Residual Adjustment	E * F	0.87	4.40	5.27
H	**New Predicted Costs**	**D + G**	**49.57**	**45.30**	**94.87**
I	Frontier Adjustment (%)		4.90%	-6.11%	
J	**Frontier Predicted Costs**	**H * (1 + I)**	**51.99**	**42.54**	**94.53**
K	Efficiency Gap (to average)	C – H	7.83	17.58	25.41
L	**Efficiency Gap % (to average)**	**K / C**	**13.64%**	**27.96%**	**21.13%**
M	Efficiency Gap (to frontier)	C – J	5.40	20.35	25.75
N	**Efficiency Gap % (to frontier)**	**M / C**	**9.41%**	**32.26%**	**21.41%**

Figures may not sum due to rounding.

3.3.2 The efficiency level is similar to the COLS findings. However, there does appear to be a slight anomaly for the water models. The gap to the average is larger than the frontier. This illustrates the problem of maintaining the current benchmark company.

3.3.3 Using volume is arguably a better predictor than properties. Volumes measure actual activity so should be closely linked with cost. However, similar criticisms apply. No allowance is made for (dis)economies of scale and many of the explanatory variables are excluded.

4.0 Population Topex Regression

4.1. Water and sewage population

4.1.1 An alternative to the unit cost approach is a total opex (topex) regression. The benefit of regression is that scale is taken into consideration. Unlike unit costs, it can also account for the impact of more that one variable.

4.1.2 In this case, costs are modelled against the population served (water) and population connected (sewage). Any number of other variables could have been used e.g. properties, network size, volume etc. Results are below.

Table 4.1 – Water topex population model

Water Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (water functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.265	0.346
Ln (population served)	1.054	0.042
Form of Model:	ln (modelled cost) = -4.265 + 1.054 * ln {population served}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.930
	Standard error = 0.184	F test = 0.000

Table 4.2 – Sewage topex population model

Sewage Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (sewage functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-1.664	0.309
Ln (connected population)	0.753	0.037
Form of Model:	ln (modelled cost) = -1.664 + 0.753 * ln {connected population}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.898
	Standard error = 0.158	F test = 0.000

4.1.3 Both models give good statistical results. Explanatory variables are strongly significant. The regressions appear to be a good fit for the data as they suggest that population explains around 90% of opex spend.

4.2. Population topex model efficiency gap

4.2.1 Model results are given below:

Table 4.3 – Population model efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)					
	Category	**Process Rule**	**Water (£m)**	**Sewerage (£m)**	**Total (£m)**
A	NI Water actual cost		61.52	67.34	128.86
B	Less Adjustments		-4.12	-4.46	-8.58
C	**Modelled Cost**	**A – B**	**57.40**	**62.88**	**120.28**
D	Predicted Cost (average)		38.42	46.80	85.22
E	Difference	C – D	18.98	16.09	35.06
F	Adjustment Factor (%)		10%	20%	
G	Residual Adjustment	E * F	1.90	3.22	5.12
H	**New Predicted Costs**	**D + G**	**40.32**	**50.02**	**90.33**
I	Frontier Adjustment (%)		0.61%	-19.62%	
J	**Frontier Predicted Costs**	**H * (1 + I)**	**40.56**	**40.20**	**80.76**
K	Efficiency Gap (to average)	C – H	17.08	12.87	29.95
L	**Efficiency Gap % (to average)**	**K / C**	**29.76%**	**20.46%**	**24.90%**
M	Efficiency Gap (to frontier)	C – J	16.83	22.68	39.52
N	**Efficiency Gap % (to frontier)**	**M / C**	**29.33%**	**36.07%**	**32.85%**

Figures may not sum due to rounding.

4.2.2 The gap is materially larger than the COLS. It would be expected that population will influence cost. Population is however closely correlated with other key variables i.e. company size, usage, network connections etc. It is therefore not clear if population is the crucial factor.

4.2.3 Use of only one variable also fails to take proper account of particular company circumstances e.g. different network needs depending on how the population is distributed. This means that results must be treated with caution.

5.0 CSV Topex Model A

5.1. Water CSV model A – equal weights

5.1.1 The composite scale variable (CSV) is an approach used by modellers. It is a useful technique when there are a low number of data points. It is further helpful when there are a number of variables that might be expected to impact on costs, but are themselves highly correlated.

5.1.2 For the water service, the explanatory variable is a combination of mains length, population and distribution input. Each variable is given an equal weight (33/33/33) for ease of calculation. The model is in natural logs format.

5.1.3 Any number of variables or weights might have been employed. The chosen variables are already used in the COLS models and seemed the most obvious option.

5.1.4 Results are as follows:

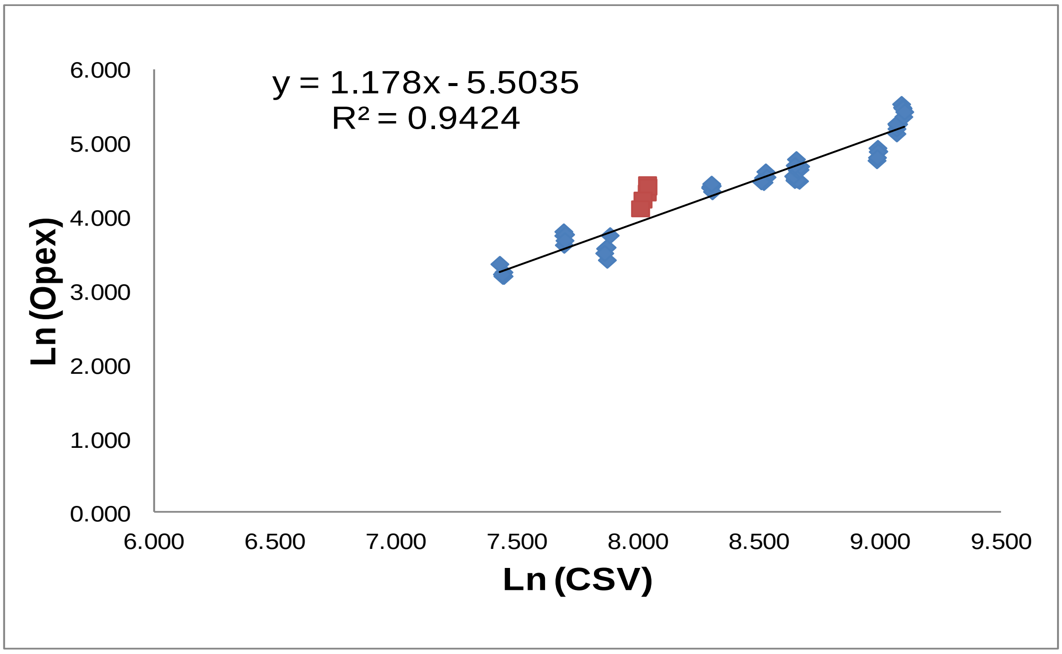
Table 5.1 – Water topex CSV model – equal weights

Water Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (water functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-5.503	0.355
Ln (CSV)	1.178	0.042
Form of Model:	ln (modelled cost) = -5.503 + 1.178 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.942
	Standard error = 0.166	F test = 0.000

5.1.5 This model is a very good fit for the data. The regression suggests that the composite variable explains almost 95% of water opex. As the model includes a number of variable impacts, findings are more robust than the single variable model.

5.1.6 In graphical form, the model looks as follows (NI Water observations in red):

Figure 5.1 – Water topex CSV model using equal weights



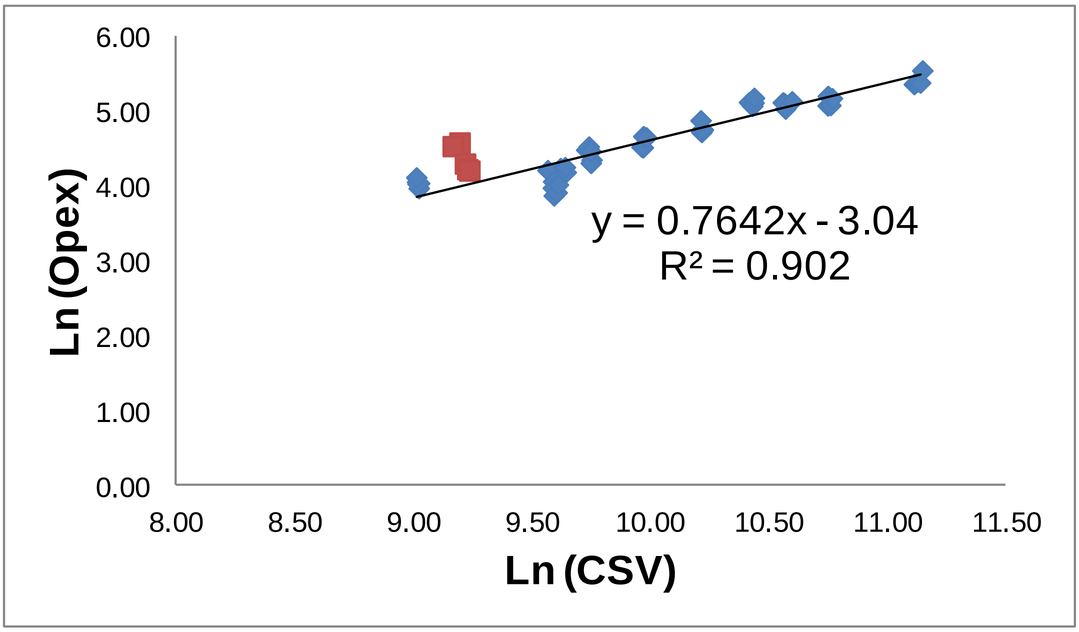
5.2. Sewage CSV model A – equal weights

5.2.1 The sewage composite variable is derived from the length of sewers, total load and the connected population. Again, all elements have an equal weighting.

Table 5.2 – Sewage topex CSV model – equal weights

Sewage Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (sewage functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.040	0.368
Ln (CSV)	0.764	0.036
Form of Model:	ln (modelled cost) = -3.040 + 0.764 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.902
	Standard error = 0.155	F test = 0.000

Figure 5.2 – Sewage topex CSV model – equal weights



5.2.2 The graphic illustrates NI Water in red. Whilst individual years cannot be observed, findings show the company moving toward average performance.

5.3. CSV topex model efficiency gap – equal weights

5.3.1 Model results are below:

Table 5.3 – CSV model A efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	**Water (%)**	**Sewerage (%)**	**Total (%)**
A	Efficiency Gap % (to average)	10.22%	9.30%	9.74%
B	Efficiency Gap % (to frontier)	8.45%	27.91%	18.62%

Figures may not sum due to rounding.

5.3.2 The topex model gap is similar to the COLS findings, though slightly lower. Again, some adjustment may be required for water models as the imposed frontier is below average performance.

6.0 CSV Topex Model B

6.1. Water CSV model B – changed weights

6.1.1 This model approach uses the same variables to generate a CSV. The only difference is that the weights are amended.

6.1.2 For water costs, there is a much higher correlation[2] with volumes and population than with network length. This may be due to the different sizes and location of mains, meaning they cannot be treated uniformly.

6.1.3 Since the correlation with cost is lower, the assumption is that this element of the CSV should be given a lesser weight. A 20/40/40 split has been adopted, with mains length given a reduced impact.

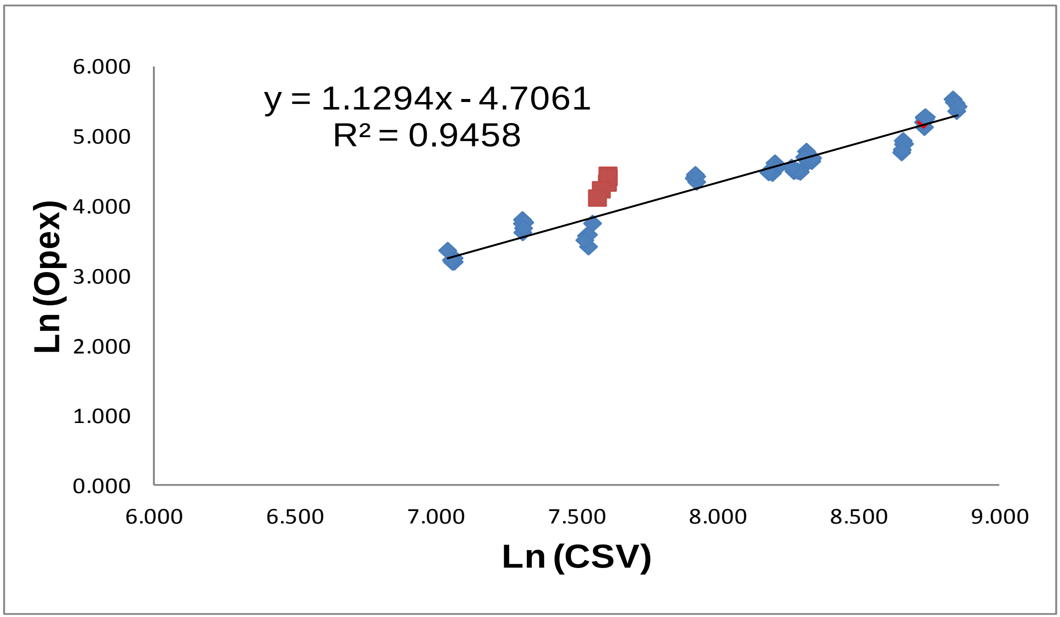
Table 6.1 – Water topex CSV model – different weights

Water Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (water functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.706	0.316
Ln (CSV)	1.129	0.039
Form of Model:	ln (modelled cost) = -4.706 + 1.129 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.946
	Standard error = 0.161	F test = 0.000

6.1.4 Whilst similar to the previous CSV model, statistical results are slightly improved. The goodness-of-fit of the model suggests that omitted variables are not an issue. The improved results indicate that mains length does have a lesser impact on costs.

[2] Reference here is to the Pearson correlation coefficient. This measures the strength of the linear relationship between two variables. Values range from -1 to 1 for negative and positive relationships respectively. A value close to zero indicates that a linear relationship does not exist.

Figure 6.1 – Water topex CSV model – different weights



6.2. Sewage CSV model B – changed weights

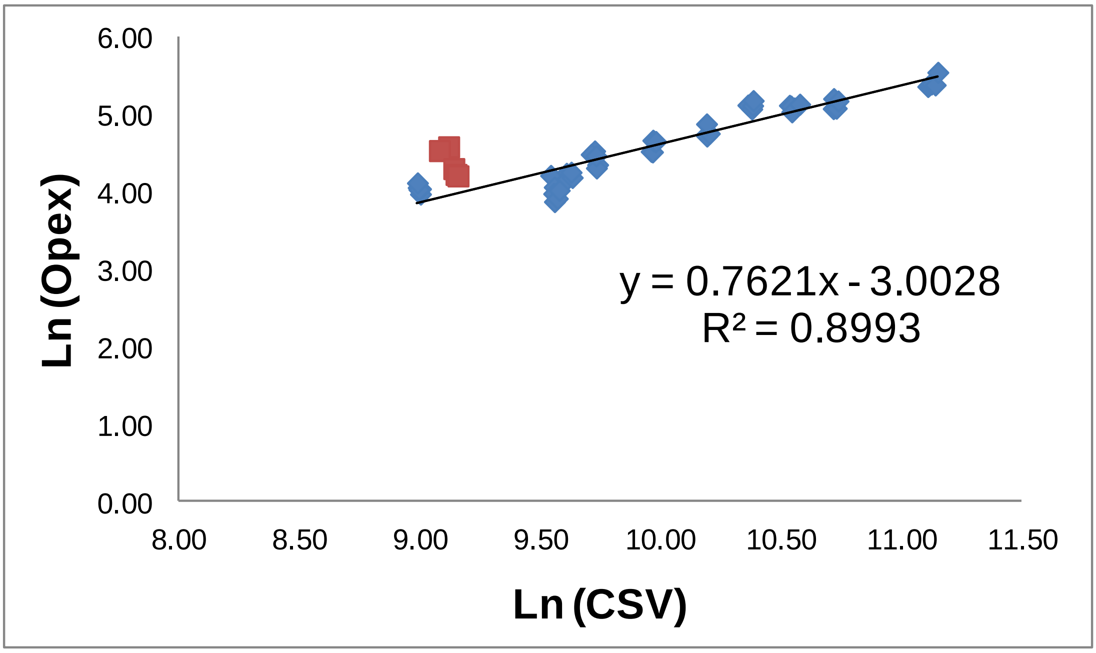
6.2.1 The difference in correlations does not exist on the sewerage models. The different split regression has been run for consistency purposes.

Table 6.2 – Sewage topex CSV model – different weights

Sewage Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (sewage functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.003	0.372
Ln (CSV)	0.762	0.037
Form of Model:	ln (modelled cost) = -3.003 + 0.762 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.899
	Standard error = 0.157	F test = 0.000

6.2.2 Results are very similar to the equal weight model. In this case, changing the weights has not improved estimations.

Figure 6.2 – Sewage topex CSV model – different weights



6.3. CSV topex model efficiency gap – different weights

6.3.1 Model results are below:

Table 6.3 – CSV model B efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	**Water (%)**	**Sewerage (%)**	**Total (%)**
A	Efficiency Gap % (to average)	16.66%	12.17%	14.31%
B	Efficiency Gap % (to frontier)	14.53%	29.80%	22.51%

Figures may not sum due to rounding.

6.3.2 The model gap is very close to the COLS findings. Again, it can be seen that the frontier may be incorrect because of our imposed assumptions.

7.0 Upper Quartile Models

7.1. Rationale

7.1.1 In their consultation response, NI Water raised a concern about the choice of frontier companies. The Regulator shares this concern. Changing opex profiles mean a shift in efficiency over time. Naturally, this results in movements in company rankings.

7.1.2 An alternative to choosing a company is to look at the upper quartile. The Regulator has done so in order to see if it has a material impact on the efficiency gap.

7.2. UQ topex model efficiency gap

7.2.1 There are a number of ways to decide upon what the upper quartile is. For the purpose of this report, the UR has used data from the 2nd and 3rd ranked companies (out of ten).

7.2.2 The models do not change from those used for CSV model B. The gap to the average will remain unchanged. The difference is movement in the frontier position. Results are as follows:

Table 7.1 – Upper quartile efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	Water (%)	Sewerage (%)	Total (%)
A	Efficiency Gap % (to average)	16.66%	12.17%	14.31%
B	Efficiency Gap % (to frontier)	29.07%	13.75%	21.06%

Figures may not sum due to rounding.

7.2.3 In comparison to Table 6.3, it can be seen that the efficiency gap is now quite different by service area. The total gap is slightly less, though not materially different from COLS models.

7.2.4 Findings suggest that no specific overall bias is found in 2012-13. This provides assurance as to the robustness of the COLS gap.

7.2.5 When looking at the time trend there can be a fair divergence between frontier and UQ results (in both a positive and negative way). This suggests that consideration should be given to this issue in the future.

8.0 Standardised Variable Weights

8.1. Rationale

8.1.1 This model developed from work submitted in the consultation response. NI Water (with Frontier Economics) proposed that individual variable coefficients could be used to determine the weights in a CSV.

8.1.2 The Utility Regulator's advisors (CEPA) agreed this to be a viable option. They did however suggest that the variables be standardised first. As a result, the UR ran a model using this approach.

8.1.3 The standardised weights involve two steps. First, individual variables are changed to a standard scale. This ensures that coefficient weights are not arbitrarily distorted by the different units of measurement of each variable.

8.1.4 Standardisation is done by subtracting the mean from each observation, then dividing by the standard deviation. The result is each variable has the same standard unit. The normalised variables all measure the distance of the observation from the mean.

8.1.5 Secondly, a regression is run using the *individual* standard variables. Their coefficients then determine the weights used for the CSV. This approach follows on from work submitted by NI Water in the consultation response.

8.1.6 Within this model the same variables are used as in the previous CSV regressions. This differs from NI Water who developed a specific way of selecting variables.

8.2. Water standardised weight model

8.2.1 Running a regression on opex against the standardised variables provides coefficient values. These are then summed and a weighting found for each variable. For the water regression, the weights are as follow:

Table 8.1 – Regression weights

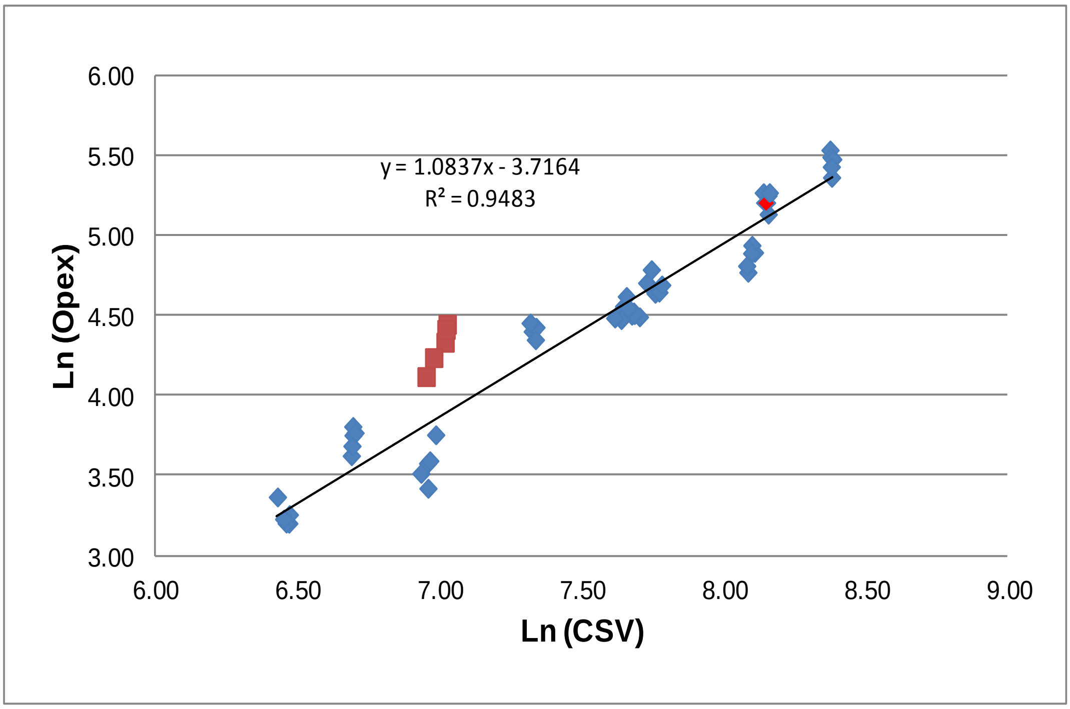
Variable	Weighting
Mains Length	8.2%
Population Served	26.1%
Distribution Input	65.7%

8.2.2 Using these weights for the CSV gives statistical results which improve on the previous approaches.

Table 8.2 – Water topex CSV model – standardised variable weights

Water Service:	Total opex (excluding business activities)	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (water functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.716	0.275
Ln (CSV)	1.084	0.037
Form of Model:	ln (modelled cost) = -3.716 + 1.084 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.948
	Standard error = 0.158	F test = 0.000

Figure 8.1 – Water topex CSV model – standard variable weights



8.2.3 The choice of variables remains in question. However, the improved statistical result does indicate merit in the method.

8.3. Sewage standardised weight model

8.3.1 On the sewage side the weights changed somewhat. Greater emphasis was placed on the size of the network. Population also became the most important variable.

Table 8.3 – Regression weights

Variable	Weighting
Sewer Length	29.6%
Sewage Load	26.0%
Connected Population	44.4%

Table 8.4 – Sewage topex CSV model – standardised variable weights

Sewage Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (sewage functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.799	0.355
Ln (CSV)	0.763	0.036
Form of Model:	ln (modelled cost) = -2.799 + 0.763 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 50	R^2 = 0.903
	Standard error = 0.154	F test = 0.000

8.3.2 Again, the statistical results appear to show a slight improvement in the model.

8.4. Standardised variable weight efficiency gap

8.4.1 Using this model gives the following results:

Table 8.5 – Standardised variable weights CSV efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	**Water (%)**	**Sewerage (%)**	**Total (%)**
A	Efficiency Gap % (to average)	18.78%	11.08%	14.76%
B	Efficiency Gap % (to frontier)	15.23%	29.22%	22.54%

Figures may not sum due to rounding.

8.4.2 The UR has some issues with the approach.[3] However, it does have certain favourable points. These include:

a) Better statistical results;

b) Removes the more arbitrary way of assigning CSV weights; and

c) Removes potential for weights to be skewed simply by their unit of measurement.

8.4.3 The final gap of 22.5% is closely aligned with the COLS figure for 2012-13.

[3] These are discussed in more detail in the NI Water Consultation chapter.

9.0 NI Water Inclusive Model

9.1. Rationale

9.1.1 The final model repeats the standardised weighting method. However, on this occasion NI Water is included in the regression.

9.1.2 NI Water raised the point that they are no longer an outlier. As such, the company state:

"it is beneficial to include NI Water in the sample in order to maximise the sample size and improve the applicability of the models for NI Water."

9.1.3 The UR considers it prudent to investigate the impact of including company data.

9.2. Water inclusive model

9.2.1 For the water model, variable weights did not change greatly after incorporating NI Water data.

Table 9.1 – Water model – standardised weights including NI Water

Water Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (water functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.323	0.342
Ln (CSV)	1.028	0.045
Form of Model:	ln (modelled cost) = -3.323 + 1.028 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 55	R^2 = 0.907
	Standard error = 0.202	F test = 0.000

9.2.2 Whilst still good, the statistical results are worse than before. R-squared has fallen as has the significance of the composite variable.

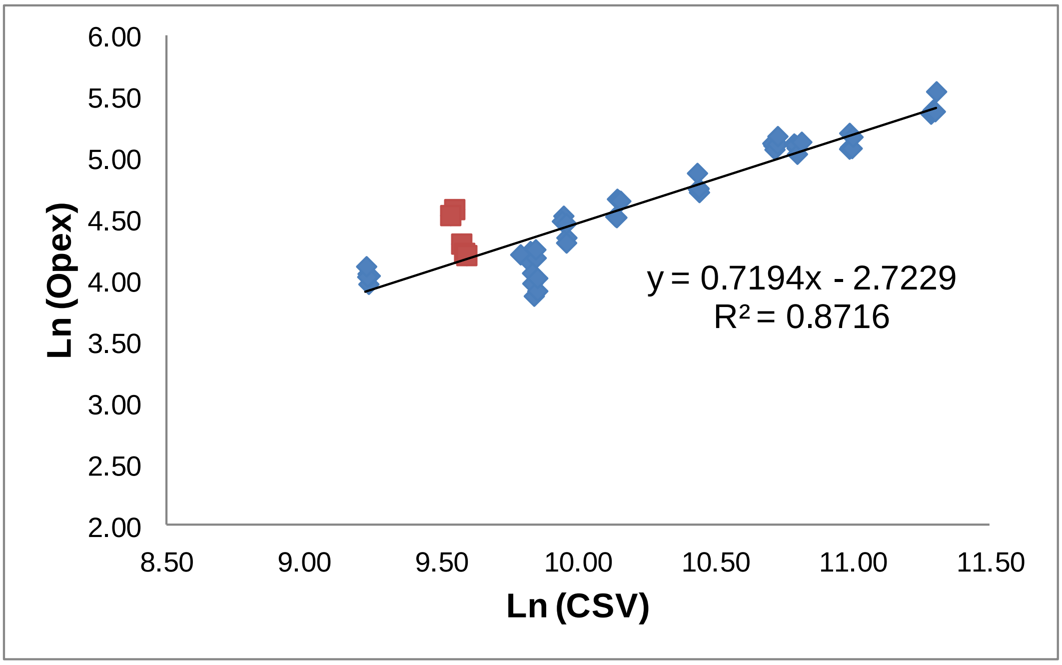
9.3. Sewage inclusive model

9.3.1 In this regression, including NI Water has a large impact on variable weights. Sewer length has gone from a 30% weighting to over 60%. Results are given below:

Table 9.2 – Sewage model – standardised weights including NI Water

Sewage Service:	**Total opex (excluding business activities)**	
Data:	June Returns and PR14 August submission	
Modelled cost:	ln (sewage functional expenditure less business activities, rates and service charges [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.723	0.390
Ln (CSV)	0.719	0.038
Form of Model:	ln (modelled cost) = -2.723 + 0.719 * ln {CSV}	
Statistical Indicators:	Nr. of observations = 55	R^2 = 0.872
	Standard error = 0.173	F test = 0.000

Figure 9.1 – Sewage model – standardised weights including NI Water



9.3.2 Whilst a good model, the outcome is the same as the water regression. Statistical results are not as good as the previous standardised weights which excluded NI Water.

9.4. NI Water inclusive model efficiency gap

9.4.1 The loss of significance in these models suggests they may be less reliable. The efficiency results show a fall in the assessed gap.

Table 9.3 – NI Water inclusive model efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	**Water (%)**	**Sewerage (%)**	**Total (%)**
A	Efficiency Gap % (to average)	14.50%	-2.89%	5.41%
B	Efficiency Gap % (to frontier)	13.19%	22.23%	17.92%

Figures may not sum due to rounding.

9.4.2 Simply including NI Water data gives rise to a fall from 22.5% to 17.9%. This is quite a material shift.

9.4.3 Whilst it is true to say that NI Water is no longer an outlier, the 'pooled' regressions contain historic observations. These may well be subject to outlier status.

9.4.4 Their impact on a relative small sample can be seen to be quite important. This may influence both CSV weights and predicted costs.

9.4.5 Since NI Water is no longer an outlier, including them in single year cross-sectional comparisons does not seem unreasonable. However, the UR would be concerned about including NI Water where 'pooled' historic data is used.

9.4.6 The fact that the statistical results of both service area models reduced is important. This suggests that exclusion may be a better option in trying to ascertain efficient costs.

10.0 NI Water Consultation Response

10.1. Issues and responses

10.1.1 As part of their consultation response, NI Water raised some issues with the alternative modelling presented at draft stage. The table below highlights these concerns and makes response.

Table 10.1 – Modelling issues raised by NI Water

Issues and Responses
Issue 1 – Given differences in cost reporting, it is unlikely that NI Water's costs are fully consistent with comparators.
Response 1 – The UR does not see this as a material issue in the current analysis. The company did their own assessment of wholesale costs. Once rates and service charges are removed (as they are in UR models), differences appear marginal. However, this issue, as well as cost reporting will be examined going forward.
Issue 2 – It is misleading to include simple unit cost models at all. There are legitimate cost drivers which invalidate these models.
Response 2 – The Regulator would agree with this point. No reliance has been placed on the unit costs for the purpose of triangulation. The models are simply there to illustrate commonly used benchmarking techniques.
Issue 3 – The population topex regression is not materially better than unit cost models.
Response 3 – Agreed. Again no reliance is based on these findings.
Issue 4 – The CSV models indicate a difference in the impact of scale by service area. For water the models suggest diseconomies as a company grows. For sewage there is strong evidence of economies of scale. There is no justification for this inconsistency.
Response 4 – The regressions do indicate these findings. However they may not be entirely unexplainable. Diseconomies of scale may be associated with the higher cost of running and maintaining a water network in urban areas. Sewage services may be less affected due to offsetting power generation capacity at large works etc. Unexpected results do not automatically invalidate the models. The Regulator would also point out that NI Water's alternative composite scale variables have similar outcomes.
Issue 5 – The CSV models do not include operational cost drivers but simply focus on different combinations of scale.
Response 5 – Size of the network will affect fixed costs. In theory, no further variable costs will be incurred if more people use it (assuming the network stays the same size). In this sense it might be said that the variable is not an operational cost driver. The same cannot be said about population and volumes. Although they represent a scale

factor, they also drive opex spend. In this respect the issue seems unwarranted. Choice of variable is however a key consideration in these models. Further debate is required to settle this problem.
Issue 6 – The way in which weights are varied for CSV Model B is arbitrary.
Response 6 – Weightings are not totally arbitrary in this model. The decision to lower the mains length weight is based on its lower correlation with opex. That said, it is by no means an exact methodology. In response to this concern, the UR adopted NI Water's proposed alternative. This manifests itself in the standardised variable weight model. The result is weightings based on individual variable coefficients.
Issue 7 – NI Water has not been included in the sample despite no longer being an outlier in terms of efficiency.
Response 7 – The UR has run a further model to test the impact of this. Analysis shows the models are less robust and the efficiency gap changes quite a lot. Findings suggest it may be dubious to include NI Water data where 'pooled' historic information is used. This may not be the case for single year cross section regressions.

10.1.2 Extra modelling in the final determination is designed to further address the company concerns. The Regulator also believes that it helps validate the robustness or otherwise of the COLS efficiency gap.

10.2. NI Water alternative modelling

10.2.1 As part of the consultation response, NI Water submitted a paper from Frontier Economics. This set out a new approach to both selecting and weighting variables for CSV modelling.

10.2.2 Models focused on wholesale costs including rates and service charges. Regressions were run using a large sample of explanatory variables. Those found to be insignificant or having unforeseen impacts were excluded.

10.2.3 Remaining variable coefficients were then used to determine the weights in the CSV model. The Regulator adopted a similar method in the standardised variable weight analysis.

10.2.4 Numerous models were investigated. Final results were based on an average of two water models and three sewage regressions.

10.2.5 The conclusion of the work was that the scale of the efficiency gap could potentially be much lower than the UR assumes.

10.2.6 When comparing to the very best companies, the gap is actually larger than the COLS findings. However, the situation is somewhat different when using the Regulator's chosen frontier performers or the upper quartile.

10.2.7 Findings are shown below.

Table 10.2 – NI Water alternative model efficiency gap for 2012-13

Efficiency Gap Calculation (2012-13 prices)				
	Category	**Water (%)**	**Sewerage (%)**	**Total (%)**
A	Efficiency Gap % (to 1st ranked company)	18.5%	34.7%	26.9%
B	Efficiency Gap % (to UR frontier companies)	-11.4%	34.7%	12.4%
C	Efficiency Gap % (to upper quartile)	-3.1%	21.1%	9.4%

10.2.8 Taken on a comparable basis, the alternative models show a gap of 12.4%. This is much lower than the Regulator's position. NI Water is therefore concerned about the COLS assessment.

10.3. UR views on alternative models

10.3.1 It is clear that alternative efficiency models need to be developed in the future. In light of this, the UR welcomes the paper submitted by NI Water. The approach adopted is comprehensive and adds valuable insight to the issue.

10.3.2 As regards the analysis, certain concerns remain. These include:

a) Multicollinearity in the initial models may be excluding valid variables;

b) Standardising variables may help to avoid distorted weightings;

c) Whether and how adjustment factors might be applied;

d) How negative coefficients should be treated; and

e) Whether rates should be included in any efficiency analysis.

10.3.3 All these points need further discussion as a new efficiency approach is developed. For the 2012-13 analysis, the main issue for the UR is rates.

10.3.4 Even though they are subject to challenge, rates have long been excluded from the efficiency analysis. The reason is not because they are uncontrollable. Rather, they are removed as they are not influenced by explanatory variables.

10.3.5 Including rates presents a particular problem at this time. Non-domestic valuations have not been re-valued in Northern Ireland since 2003. The result is a very low charge compared to England and Wales. This skews efficiency results.

10.3.6 If the Regulator was to accept the inclusion of rates, this would present serious issues going forward. If as expected, NI Water incurs a substantial rates increase in 2015-16, their efficiency gap would deteriorate considerably. This does not seem correct if operational practices are unchanged.

10.3.7 The UR is also aware that rates in England and Wales may also change substantially. This could be either as a result of revaluation of company rates bills and/or any wider review of business rates post the 2015 General Election.

10.3.8 By way of a sensitivity test, the UR ran NI Water models. The only difference was the exclusion of rates.

Table 10.3 – NI Water alternative models (excluding rates)

Efficiency Gap Calculation (2012-13 prices)				
	Category	Water (%)	Sewerage (%)	Total (%)
A	Efficiency Gap % (to 1st ranked company)	28.9%	37.2%	33.3%
B	Efficiency Gap % (to UR frontier companies)	-1.3%	37.2%	18.9%
C	Efficiency Gap % (to upper quartile)	11.3%	23.9%	17.9%

10.3.9 Comparison with Table 10.2 illustrates the impact rates can have. The Regulator does not consider it appropriate to include rates. The lack of re-valuation will distort results for 2012-13. Timing differences between N Ireland and England and Wales valuations are also likely to cause problems in future years.

10.3.10 Excluding rates, the scale of the NI Water model gap is not that different from the UR's COLS model results.

10.3.11 The Regulator welcomes further engagement on all these issues as an alternative approach is created.

11.0 Summary of Findings

11.1. Comparing approaches

11.1.1 A summary of the alternative model findings is given below.

Table 11.1 – Efficiency gap estimates using different methods

Methods	2012-13 Efficiency Gap (To Frontier)
COLS Models	21.6%
Unit Costs – (Properties)	35.3%
Unit Costs – (Volumes)	21.4%
Topex Regression (Population)	32.9%
Topex Regression A (CSV – 33/33/33)	18.6%
Topex Regression B (CSV - 20/40/40)	22.5%
Topex Regression (CSV - 20/40/40 - Upper Quartile)	21.1%
Topex Regression (CSV - Standardised Variables)	22.5%
Topex Regression (CSV - Standardised - Including NI Water)	17.9%
Topex Regression (NI Water alternative excluding rates)	18.9%

11.1.2 More work is required to develop efficiency models going forward. In spite of a reasonably large degree of variability, it can be stated with certainty that a gap exists and that this gap has been falling over the past five years.

11.1.3 Calculation of the efficiency gap is not an exact science. This is why different models have been investigated. Whilst useful in showing annual changes, the simple unit cost models are not considered robust. Neither they, nor the population regression has been used to verify findings.

11.1.4 The UR places most value on the CSV models. What these show is that the efficiency gap estimate is close to the COLS results. They are perhaps even a little underestimated given the problems with the water frontier company.

11.1.5 Since findings are similar, the UR feel justified in using COLS to set efficiency targets. The CSV approach also uses updated cost and asset data. This helps provide assurance that the COLS approach of using old model comparisons has not introduced a bias that adversely affects NI Water.

11.1.6 It is also worthwhile noting that the catch-up challenge is 80% of the COLS gap. This is roughly a 17% challenge over the price control. Such a challenge is lower than all the alternative estimates. This provides the UR with assurance that the efficiency target for PC15 does not go beyond the frontier on any of the current measures.

11.1.7 Further work will be required when the COLS models are retired. However, the alternative models help address company concerns. They further improve regulatory certainty by triangulating around a number of robust approaches. Such new alternative models will use ‘up-to-date’ data to produce robust efficiency gap estimates.

11.1.8 The Regulator is convinced continued dialogue and engagement with the company will offer the opportunity to develop a new set of models. These will inform annual reporting of NI Water’s progress during PC15.

11.1.9 It is hoped to develop such a new approach to at least inform the next price control of NI Water at PC21. Ideally this will also be in place by next year’s Cost and Performance Report.