Appendix 3.5.B – Third party report on analysis of RoRE for outcome delivery incentives

Wessex Water

September 2018

Wessex Water
YTL GROUP

Business plan section		Supporting document	
	Board vision and executive summary		
1	Engaging customers		
2	Addressing affordability and vulnerability		
3	Delivering outcomes for customers	3.1	Performance commitment overview
		3.2	Triangulation methodology
		3.3	Cost benefit analysis – methodology and results
		3.4	Calculation of incentive rates including enhanced incentives
		3.5	Inputs to RORE for outcome delivery incentives
		3.6	Sharing success – Proposals for a Wessex Water community foundation
4	Securing long term resilience		
5	Markets & innovation: wholesale		
6	Markets & innovation: open systems & DPC		
7	Markets & innovation: retail		
8	Securing cost efficiency		
9	Aligning risk and return		
12	Securing trust, confidence and assurance		
13	Data tables and supporting commentaries		

August 2018

REVIEW OF WESSEX WATER'S RORE RANGE SPREADSHEET TOOL

An assurance note prepared for Wessex Water

Wessex Water is in the process of finalising its package of performance commitments (PCs) and outcome delivery incentives (ODIs) ahead of PR19. Ofwat has laid out an expectation that companies' proposals for ODIs should have a financial impact in an indicative range of ±1% to ±3% of Return on Regulatory Equity (RoRE). Wessex has developed a RoRE range spreadsheet tool to summarise what its package of PCs & ODIs implies in RoRE terms. Frontier has been commissioned to carry out a review of the spreadsheet tool.

1. INTRODUCTION

Wessex has produced a spreadsheet tool to estimate the Return on Regulatory Equity (RoRE) impact of its proposals for PCs & ODIs. We have been commissioned to review this tool and to:

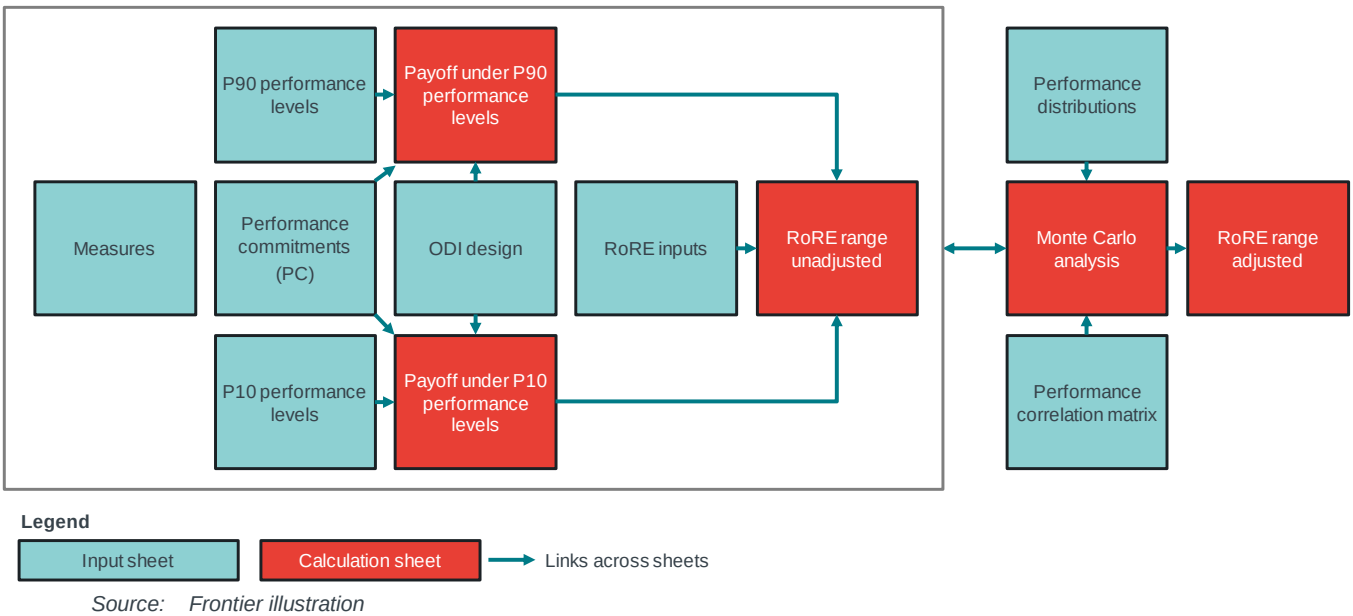
Comment on the extent to which we believe it satisfies Ofwat's guidance;

Assure the accuracy of the calculations; and

Comment on the overall results in terms of the RoRE range.

The diagram below provides a high level illustration of the logical flow of the tool:

Figure 1 Overview of Wessex's RoRE range tool



Source: Frontier illustration

We describe the tool in more detail on a sheet-by-sheet basis in the rest of this note. It is split into three parts:

The calculation of the 'unadjusted' RoRE range;

The calculation of the 'adjusted' RoRE range which is based on Monte Carlo simulation analysis; and

Our views on the reasonableness of the unadjusted and adjusted ranges.

The distinction between the unadjusted and adjusted RoRE ranges is to reflect Ofwat's guidance as set out in its "Final guidance on business plan data tables"[1], and in particular the distinction between:

Table App1 Performance commitments (PCs) and outcome delivery incentives (ODIs); and

Table App26 RoRE scenarios.

App1 calls for companies to report P10 and P90 performance levels[2] and the corresponding financial impacts at the level of individual measures. However, App26 calls for companies to model aggregate level P10 and P90 scenarios across all measures. This would need to reflect that in reality, and probabilistically, it would be unlikely for companies to perform at P10 or P90 performance levels for all measures at the same time. This is either because in reality, some measures may be positively correlated (such as supply interruptions and mains bursts) or negatively correlated (such as hosepipe bans and flooding), and more generally that individual risks may be independent of each other. Therefore, we would expect the financial impact of the aggregate P10 or P90 scenario to be lower than the sum of financial impacts for each individual P10 or P90 scenarios. In this regard, Ofwat has provided the following guidance:

> *"Performance against each individual ODI will be driven by different factors, therefore, the P10 and P90 values for each ODI cannot all be assumed to occur together and some statistical analysis of all ODIs should therefore be undertaken. This could for example involve Monte Carlo analysis using both variances for individual ODIs and covariance's between ODIs."*[3]

Therefore, Wessex has followed Ofwat's guidance by carrying out Monte Carlo analysis to produce more robust estimates of the financial impact under the aggregate P10 and P90 scenarios.

The scope of this project was to focus specifically on Wessex's approach to calculating the RoRE range, and the extent to which we believe that it satisfies Ofwat's guidance. As described in the rest of this note, this draws upon various inputs such as PCs, P10s, P90s and ODIs. However, for the purposes of this review, we have taken these inputs largely as a given, and instead have focussed on the calculations and method that have been used to estimate the RoRE range. Separately, Wessex has commissioned various reports to assure the appropriateness and robustness of the inputs themselves.

---

1 https://www.ofwat.gov.uk/wp-content/uploads/2017/12/PR19-Final-guidance-on-business-plan-tables.pdf

2 P10 relates to a level of poor performance such that there is only a 10% chance of observing a worse performance level. P90 relates to an excellent level of performance where there is only a 10% chance of observing better performance level.

3 Final guidance on business plan data tables. Page 28.

2. UNADJUSTED RORE RANGE

In this section we describe the part of the spreadsheet tool which is designed to calculate the 'unadjusted' RoRE range. This is the sum of financial impacts under the P10 and P90 scenarios for individual measures.

3. Measures – input sheet

In this sheet, the model user inputs the list of measures to be included in the overall analysis, along with units of measure, and other categorisations such as the outcome, and the direction of improving performance.

4. Performance commitments – input sheet

In this sheet, the model user inputs the performance commitment for each measure for each year of AMP7. We understand that Wessex has followed a clear and objective approach to setting PCs, which is in line with Ofwat's guidance – although we have not specifically assured this work. In particular, for each measure it has challenged itself against Ofwat's six proposed approaches to setting PCs. This covers:

Cost benefit analysis;

Comparative information;

Historical information;

Minimum improvement;

Maximum level attainable; and

Expert knowledge.

5. P10 and P90 performance levels – input sheet

In these sheets, the model user defines how Wessex would be expected to perform under the P10 and P90 scenarios for each individual measure separately and for each year of AMP7. We understand that this is based on various techniques which include:

An assessment of Wessex's own historical performance. For example, for some measures, the P10 scenario could be informed by Wessex's worst ever performance on record.

An assessment of comparative information, where appropriate and applicable. For example, for some measures, the P10 and P90 scenarios could be informed by the worst and best ever performances by any company on record.

Expert judgement. For some measures, especially those relating to low probability, high impact incidents, historical information may not be useful if it is the case that those incidents have not happened recently – e.g. hosepipe bans. In these instances, it may be appropriate to consider expert judgement. Similarly, some measures are more bespoke to Wessex and require expert judgement due to the lack of availability of benchmarks.

The assessment of P10 and P90 levels needs to take account of efficient management mitigation and response. This will, to some extent, be reflected in the historical performance data but it is also another reason to apply expert judgement to the assessment of reasonable best case and worst case outcomes.

6. ODI design – input sheet

In this sheet, the model user defines the ODI design for each individual measure. This includes populating the following parameters:

ODI type: ODIs can be reputational only, or financial, with a further distinction between underperformance only incentives, outperformance only incentives, and both under- and outperformance incentives.

Caps, collars and deadbands: Companies can propose caps, collars and deadbands for individual measures:

A cap imposes a level where performance better than this level does not result in any additional outperformance payments.

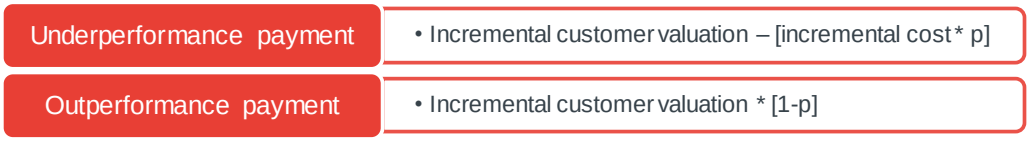
A collar imposes a level where performance worse than this level does not result in any additional underperformance payments.

A deadband introduces a range around the performance commitment level where within that range no outperformance or underperformance payments are earned.

We note that Wessex does not have any caps or collars, and has five deadbands.

Incentive rates: Where measures have financial incentives, the model user must input the underperformance and outperformance incentive rates. As a default, Ofwat has been clear that companies should use the formulas below to set incentive rates.[4] In most instances, Wessex applies the standard formulas. However, in some instances Wessex has deviated from the standard approach – which we have reviewed as part of a separate assurance report for Wessex on setting ODIs.

Figure 2 Ofwat formula for ODI payments



Enhanced incentives: For PR19, Ofwat is allowing companies to introduce 'enhanced' incentives for a selection of common measures. The purpose of enhanced incentives is to encourage industry leading companies to continue to shift the frontier to set new benchmarks for excellent performance. This means that companies intending to propose enhanced incentives will need to specify:

Enhanced 'multipliers' – i.e. precisely how enhanced the enhanced rates should be relative to the standard rates (e.g. whether they should be twice as large as the standard rates or ten times as large); and

Enhanced 'thresholds' – i.e. the performance levels beyond which enhanced outperformance payments would be earned, and the performance levels below which enhanced underperformance payments would be incurred.

[4] In the formulas, 'p' is the cost sharing rate in the totex sharing mechanism. Ofwat's guidance is that companies should assume this to be 50%, unless there is a good reason to use an alternative figure.

Finally, we note that all of Wessex's incentives are in-period, although this does not have a direct impact on the RoRE calculations[5].

7. Payoff under P90 performance levels – calculation sheet

This is a calculation sheet which draws in inputs from the following sheets:

Performance commitments;

P90 performance levels; and

ODI design.

In this sheet, the tool calculates the financial payoffs under the P90 performance levels for each measure. This is based on:

Calculating the number of units of outperformance – i.e. the difference between the P90 level of performance and PC in each year;

Summarising for how many of those units of outperformance Wessex would earn outperformance payments, if relevant, i.e. excluding any reputational or 'underperformance only' measures. This also controls for:

Any units of outperformance which lie within a deadband; and

Whether enhanced outperformance payments are applicable.

Multiplying the applicable units of outperformance by the corresponding outperformance payment rates, and taking into account whether for some of those units, Wessex would earn enhanced outperformance payments.

The output of this sheet is therefore the financial payoff for each individual measure for each individual year of AMP7.

8. Payoff under P10 performance levels – calculation sheet

This sheet is similar to the sheet described above, but estimates the financial payoff under the P10 scenarios for each individual measure.

9. RoRE inputs – input sheet

In this sheet, the model user inputs the financial inputs required to estimate the 'baseline' RoRE. This covers:

Average RCV;

Gearing;

Debt;

WACC; and

Cost of debt.

In line with Ofwat guidance the RoRE calculations are based on a notional capital structure.

---

5 There would though be a discounting impact for delayed incentive payments.

10. RoRE range unadjusted – calculation sheet

This is a calculation sheet which draws in inputs from the following sheets:

Payoff under P90 performance levels;

Payoff under P10 performance levels; and

RoRE inputs.

In this sheet, the tool:

Estimates the baseline RoRE;

Estimates the sum of the net present value (NPV) of payoffs under the P10 and P90 performance levels for all measures, discounted by the WACC; and

Expresses the values above as percentages of RoRE.

The output of this sheet is therefore the unadjusted RoRE range – i.e. the RoRE range before any Monte Carlo analysis is carried out.

Based on our review of the spreadsheet tool, we believe that Wessex is estimating the unadjusted RoRE range in line with Ofwat's guidance.

11. ADJUSTED RORE RANGE

In this section we describe the part of the spreadsheet tool which is designed to calculate the 'adjusted' RoRE range. This reflects Ofwat's guidance as set out in its "Final guidance on business plan data tables" where App26 calls for companies to model aggregate level P10 and P90 scenarios across all measures.

Ofwat has suggested that companies could carry out "Monte Carlo analysis using both variances for individual ODIs and covariances between ODIs." Wessex has followed this guidance and has used the @risk add-in in Excel to carry out Monte Carlo simulation analysis.

12. Performance distributions – input sheet

In this sheet, the model user defines the probability distribution for each individual measure. This is anchored by the fact that Wessex has already defined the P10 and P90 performance levels for each individual measure for each individual year of AMP7. The model user can then select different probability distributions around the P10 and P90 performance levels for each measure – such as the uniform distribution or the normal distribution.

We understand that for some measures, Wessex already has relatively well defined probability distributions. However, for others, especially where information is relatively scarce, it has needed to apply judgement.

13. Performance correlation matrix – input sheet

In this sheet, the model user defines the correlation between each pair of measures. A correlation of +1 implies that the two measures are perfectly and positively correlated. A correlation of -1 implies that the two measures are perfectly and negatively correlated. A correlation of 0 implies that the two measure are completely independent such that the performance of one measure has no impact on the performance on the other, and vice versa.

We understand that Wessex has largely relied on applying judgement to populate this matrix.

We consider Wessex's approach to be reasonable. Specifically we do not consider that there is sufficient historical data to identify robust correlations and that therefore the exercise of expert judgement is a valid and important element of this process. In addition we note that there are reasons why measures that do not appear to have any correlation in exogenous drivers may exhibit some correlation in practice. First, performance across all measures may be driven by the overall efficiency and innovation of the company management and operations. Second, a large negative exogenous shock to one measure may require significant management resources to mitigate and this may divert resources from other measures, with a knock-on impact on performance.

14. Monte Carlo analysis – calculation sheet

This is a calculation sheet which draws in inputs from the following sheets:

ODI design;

Performance distributions; and

Performance correlation matrix.

In this sheet, the tool carries out Monte Carlo analysis based on the inputs provided in the previous two sheets. The tool performs thousands of model runs and carries out the following calculations:

Random performance generator – based on the performance distributions and correlation matrix, the tool randomly generates Wessex's performance for each individual measure for each individual year of AMP7;

Payoff calculation – similar to the calculations in the sheet 'Payoff under P90 performance levels' the tool calculates the number of units of outperformance or underperformance relative to the PC in each year for each measure. It then calculates the number of units of standard outperformance, standard underperformance, enhanced outperformance and enhanced underperformance for each measure, taking into account deadbands. And it then multiplies the outperformances and underperformances by the relevant incentive rates, where applicable.

NPV of payoffs – the tool then calculates the NPV of the payoffs for the different model runs.

Aggregate P10 & P90 – the tool then ranks all model runs by the NPV of their payoffs. It then selects the aggregate P90 outcome and the aggregate P10 outcome.

The output of this sheet is therefore the aggregate P10 and P90 financial payoffs.

15. RoRE range adjusted – calculation sheet

This is a calculation sheet which draws in inputs from the following sheets:

Monte Carlo analysis; and

RoRE inputs.

In this sheet, the tool expresses the aggregate P10 and P90 financial payoffs as a percentage of RoRE. The output of this sheet is therefore the adjusted RoRE range – i.e. post-Monte Carlo analysis.

Based on our review of the spreadsheet tool, we believe that Wessex is estimating the adjusted RoRE range in line with Ofwat's guidance and in a way which we believe is fit for purpose for including in App26.

16. REVIEW OF RESULTS

In this section we comment on the results in terms of Wessex's unadjusted and adjusted RoRE ranges. The table below summarises the results:

Figure 3 Summary of results

		P10	P90
Figure 1: range	Unadjusted RoRE	-2.7%	3.1%
Figure 2: range	Adjusted RoRE	-1.2%	1.5%

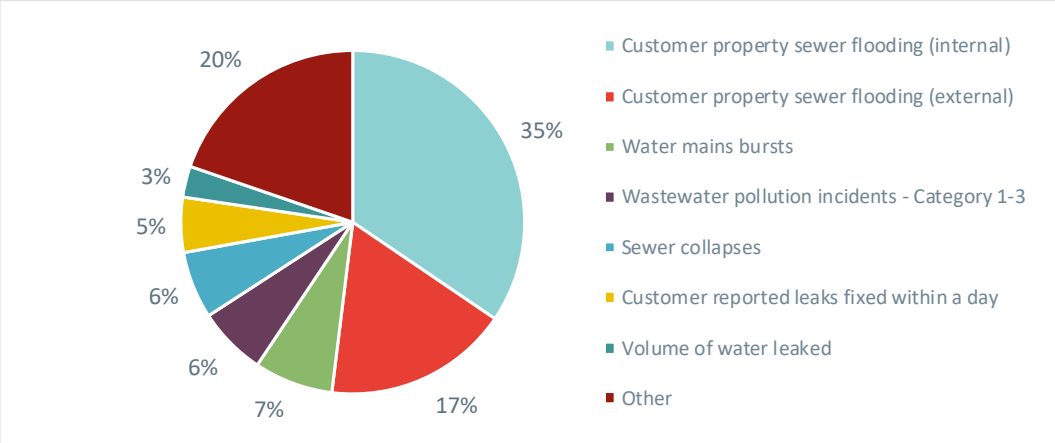
Source: Wessex

Note: This includes WaterWorCX, but does not include Wessex's managing uncertainty measure relating to the WINEP.

First, Ofwat has laid out an expectation that companies' proposals for ODIs should have a financial impact in an indicative range of ±1% to ±3% of RoRE. Therefore, the adjusted RoRE range (which is a better reflection of likely performance than the unadjusted RoRE) lies within this indicative range. In this respect, Wessex's package of ODIs can be considered sufficiently powerful and in line with Ofwat's guidance.

We have carried out a sense check of the composition of Wessex's RoRE range. The chart below shows the breakdown of the downside to Wessex's unadjusted RoRE range:

Figure 4 Breakdown of Wessex's unadjusted RoRE range downside

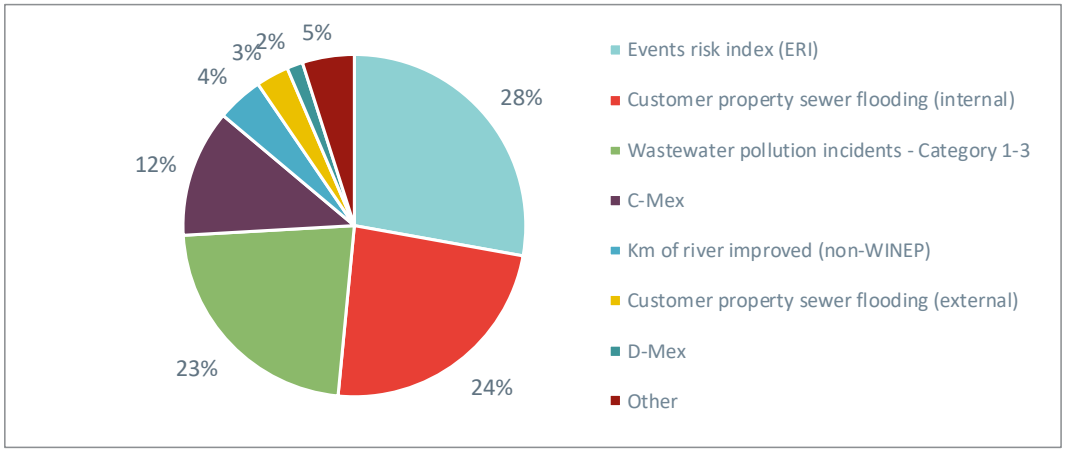


Source: Wessex

It shows that internal sewer flooding is the measure with the largest financial impact. This appears to be reasonable because it is a common measure with typically a very large customer valuation. There is also scope for Wessex to incur enhanced underperformance payments. Ofwat's common measures also represent a significant proportion of the overall downside. The breakdown of the results does not appear unreasonable.

The chart below shows the breakdown of the upside to Wessex's unadjusted RoRE range:

Figure 5 Breakdown of Wessex's unadjusted RoRE range upside



Source: Wessex

It shows that ERI is the measure with the largest financial upside. Ultimately, this is driven by how Wessex has decided to design the ODI for ERI. In particular:

Wessex identified that the most important outcome for its customers is drinking water quality. CRI is a common measure for drinking water quality but there is no scope for companies to earn outperformance payments. Therefore, Wessex decided to include ERI as another measure for drinking water quality and to include the possibility to earn outperformance payments. Wessex recognises that it is challenging to engage with customers on willingness to pay for technical index-type measures like ERI. Therefore, it decided to adopt a more top down approach. In particular, it compared the potential RoRE range upsides under the P90 scenarios for each of its outcomes, and sought to make sure that the upside for the drinking water quality outcome was in line with the largest upside of any other outcome, where the difference is allocated to ERI. This implies that if Wessex were to perform excellently on all measures – with P90 levels of performance across the board – the total payoff for the drinking water quality outcome (which is the number one outcome for customers) would be the joint largest. In this light, due to ERI's ODI design, this result does not appear unreasonable.

Internal sewer flooding and pollution incidents collectively represent nearly half of the RoRE range upside. As two of Ofwat's common measures, we consider this to be broadly sensible. This is because customer valuations tend to be high because they are emotive measures, leading to relatively high outperformance payment rates. Also, Wessex is industry leading for both measures, and the PC for both is the forecast UQ – i.e. there is scope for Wessex to earn outperformance payments, including possibly enhanced outperformance payments, for both measures.

Following the Monte Carlo analysis, Wessex's adjusted RoRE range narrows. The downside falls by nearly 60% and the upside falls by 50%. We consider this 'narrowing' of the RoRE range to be reasonable. First, we would expect the range to narrow because probabilistically speaking it is unlikely to achieve P10 and P90 levels of performance across all measures simultaneously. However, a large proportion of the unadjusted RoRE range relates to a handful of measures, and some, including internal and external sewer flooding are strongly correlated. This means that probabilistically speaking, it is not unreasonable to expect that an aggregate P10 scenario could imply very poor performances on a small number of measures, some of which are strongly correlated, which would ultimately lead to large financial impacts. Similarly, it is plausible that an aggregate P90 performance would involve performing well on

a small number of important measures, some of which are positively correlated. Therefore, we consider the results to appear reasonable.