









NWG
living water

IMPROVING THE OUTCOME AND DELIVERY INCENTIVE REGIME

A Water 2020 paper by Northumbrian Water



OUTCOMES AND DELIVERY INCENTIVES
A WATER 2020 PAPER BY NORTHUMBRIAN WATER

Introduction

The concept of Outcomes, Performance Commitments (PCs) and Outcome Delivery Incentives (ODIs) was introduced at PR14 for the first time in a water industry price review. This marked a significant change in approach from previous regimes and has raised a number of important questions.

Is the use of comparative industry performance, on a consistent basis, a prerequisite for setting PCs and ODIs? Should PCs be set with reference to upper quartile industry performance or based on what companies' customers say they want? How might the method of setting PCs interact with setting the totex baseline?

What role should willingness to pay (WTP) research have in cost benefit assessment for business planning and determining the quantum of ODI penalties and rewards? Is its use for each of these purposes conceptually compatible? Has WTP research sufficient credibility with all stakeholders?

Is it acceptable to have a regime where some companies might attract a reward for a particular level of service while another company receives a penalty for the same service? How can each company's customers be safeguarded in this respect?

How can legitimising PCs and ODIs with customers be achieved? How might ODIs be addressed by acceptability research at PR19?

The potential disaggregation of wholesale price caps also brings with it the question of how PCs and ODIs will work in this environment.

This paper looks at the considerations that need to be taken into account when addressing the questions set out above. It has been produced by Northumbrian Water in support of the Water 2020 consultation process. It is one of a suite of papers to be produced by companies in the water industry during 2015 and is intended to stimulate debate rather than advocate a specific viewpoint.

We would like to express our thanks to Southern Water, who assisted in the preparation of this paper by providing comments and challenge as well as insights from their own PR14 experience. The contents and the views expressed in the paper are those of Northumbrian Water alone.

Comments should be sent to:
Water 2020, Economic Regulation, Northumbrian Water Ltd, Boldon House, Wheatlands Way, Pity Me, Durham, DH1 5FA or by email to crawford.winton@nwl.co.uk

Background

The concept of Outcomes, PCs and ODIs was introduced at PR14. The approach actually adopted at the final determination reflected a significant evolution of the initial business plan guidance. This is because questions and issues arose over the course of the periodic review process as companies and Ofwat got to grips with the practicalities and implications of applying the new approach.

Key considerations that resulted in the significant change in approach to setting PCs and reward thresholds were how to ensure that:

- customers only pay rewards for excellent service; and
- there is some degree of consistency and fairness in target setting across companies.

To address these questions, Ofwat introduced, for a number of key Measures of Success (MoS), the use of industry comparative data to set PCs and ODIs relative to the upper quartile level.

This approach included comparative assessments for a number of measures where it was possible to make a reasonable comparison of PCs and ODIs across the sector. Using this information, Ofwat challenged companies to deliver an upper quartile level of performance. The purpose of this was

described as being required to protect customers and ensure that any outperformance is rewarded only when reward thresholds are genuinely stretching.

This marked a significant departure from the initial approach where companies proposed service levels and expenditure based on customer views and willingness to pay (WTP) information.

Ofwat said about this change in methodology:

'We recognised that companies and their customers and CCGs had not previously had access to comparative data on companies' revised outcomes proposals, as each company had developed its own set of outcomes together with its customer representatives with limited or no visibility of other companies' proposals.

'However, we explained that in future we would expect companies to engage with their customers on their comparative performance and ODIs, and that we had designed our approach on the comparative checks to encourage this.'

Given the above, there are important considerations to be made in terms of the approach to be adopted for PR19 in terms of:

- should PCs and ODIs reflect each company's customers' expressed preferences for service levels and prices, albeit with customers' making choices taking into account comparative performance (a 'bespoke' approach); or
- should PCs and ODIs be set purely on an upper quartile comparative basis (a 'comparative' approach); or
- some hybrid of the above?

These questions are discussed in section 1.

Any approach that involves comparisons of performance across the industry requires the data for a range of core measures to be produced on a robust and consistent basis. This is discussed in section 2.

Further questions that arose before and during the PR14 process concern:

- the best timing for incentive payments;
- the correct use of WTP information for setting penalties and rewards and whether this is consistent with the use of WTP information in cost benefit assessment; and
- how to deal with Measures of Success (MoS) that span more than one review period?

These questions are discussed in section 3.

After the final determination, Ofwat[1] suggested that comparative PC's and ODIs could be dynamic in PR19 (2020-25). This possibility is discussed in section 4.

Finally, the potential for separate price controls within the wholesale business brings questions about how PCs and ODIs will work for the individual price control areas. This is discussed in section 5.

How PCs are set, and incentives determined, are key questions ahead of PR19. It is hoped that this paper facilitates a full discussion of the options.

[1] Jonson Cox, Policy Exchange 11th March 2015

1 How should PCs and ODIs be set?

1.1 Key questions

Do customers want PCs that are set relative to industry upper quartile (purely comparative) or would they prefer them bespoke based on their preferences (purely bespoke)?

Should we have a core set of comparative outcomes covering the main services, with others bespoke?

Are the current six comparative measures the right ones? Should we aim for more?

Do we need consistent MoS definitions for consistency and comparability?

If we have consistent comparative PCs and ODIs, does it follow that consistent reward and penalty rates across the industry? What does that then mean for different regional willingness to pay data?

Should any company that improves upon upper quartile performance receive a reward and any company that falls short receive a penalty? Should there be a deadband like in SIM? Would that mirror a competitive market?

1.2 Advantages and disadvantages of comparative measures

The use of comparative measures to set PCs and ODIs encourages excellence in service provision. PCs and ODIs can be designed using comparative performance data to ensure rewards are only paid in appropriate circumstances (i.e. there are no anomalies where companies are being rewarded for performance that falls short).

The use of publicised comparative measures also encourages outcome performance competition between companies, a very powerful incentive from a reputational perspective.

ODIs based on comparative measures can be perceived by customers as fair – customers only pay a premium for excellent service, similar to how an unregulated competitive market would work.

The current approach to totex baseline assessment does not take into account differing starting service levels. This potentially argues for consistency in setting PCs across companies (it could be seen as unfair if a company received totex on the same basis as other companies but had poorer PCs and reward thresholds).

However, using comparative measures to define ODIs, without reference to what each company's customers say they want in terms of the price/service package, could be seen as reducing the responsiveness of both companies and regulators to customers' wishes and priorities. Is it appropriate to assume that we know that customers want ever-increasing service levels based on upper quartile (and to pay more for these through rewards)? Should we ask customers if this is what they want?

1.3 Advantages and disadvantages of bespoke measures

Bespoke measures would take account of local customers' particular service and price preferences, rather than override them with industry targets that may not be supported at a local level. Customers may not value upper quartile performance across the board, but instead prefer more improvements in a particular area (or areas) or, simply, lower bills.

However, without any comparative checks, bespoke measures could result in some companies receiving rewards and others penalties for similar performance levels – this is difficult to justify.

Bespoke measures might also require some form of adjustment to totex to reflect differing PC base service levels in companies (matching expenditure with the service provision).

1.4 Other potential approaches?

An approach that combines elements of comparative and bespoke approaches might ensure that customers continue to drive service priorities while also ensuring they receive excellent comparative service.

The 'default' PC representing 'base' performance could be based on upper quartile industry performance. This could be overridden for specific MoS if there is robust customer support for a lower or higher service level than indicated by upper quartile. This would require a consequent adjustment to the totex baseline to reflect the cost of activities associated with the variation in base service.

The extent to which companies should then endeavour to improve on PCs (and earn ODI rewards) would be determined with reference to specific research confirming that customers actually want this, using straight-forward questions. This would help legitimise the ODI regime. But to protect customers, it would be important to ensure that reward thresholds were set so that rewards were only paid for upper quartile (or better) performance.

In this approach, WTP valuations of changes in service would only be used to determine the quantum of penalties and rewards. It would not be used as customer evidence that ODI rewards and penalties were actually supported by customers.

2 Trust and confidence in comparative performance data

In the interests of transparency, it could be argued that water customers should have access to robust comparative performance data on a consistent basis, whatever approach is adopted for setting PCs and ODIs. This enables customers and other stakeholders to understand whether they are being served well by their company or less well, and hold the company to account.

For any approach to setting PCs and ODIs that uses comparative data, it is critical that information is robust and based on consistent, unambiguous definitions. This would require a common list of measures to be identified with clear definitions. Periodic horizontal audits, in a similar way as for the SIM measure, in addition to companies' own assurance, could help ensure that the data is produced on a robust and consistent basis.

By and large, water companies provide similar services to their customers. A suitable approach might be to agree a set of core measures covering the main outcomes for which companies would be required to produce data on the basis of the same, clear definitions. This would need to start as soon as practicable so that a data history could be built up for use at PR19.

There is no suggestion here that the FD14 would be amended. These proposals relate to comparative measures that would be used to set future AMP7 PCs and ODIs at PR19.

3 Other issues that arose before and during the PR14 process

3.1 Timing of incentive payments

Three companies were allowed in-period outcome delivery incentives in PR14 that could be paid during AMP6. Others have incentive payments that are applied at PR19.

Making a clear link between the performance improvement and bill impacts is important in terms of customer legitimacy. All financial rewards for performance are paid for through customer bills, so it is helpful if customers can make the connection between the improved service received and the bill effect. That said, service improvements can often affect a small minority of customers and many may not notice the change in service. However, they will notice the increase in their bill.

The minimisation of financing risk is also important. The timing of rewards needs to consider the profile of expenditure for improved performance where significant totex is required. Any mismatch in the timing of financial payments may reduce the incentive to spend to improve sevice.

A difficulty arises where there is a step change in performance, requiring expenditure, where the service improvement will last for the long term. For this kind of investment to be properly incentivised, appropriate rewards might need to be paid for an extended period. This would ensure the company is adequately rewarded but also reduces the bill impact by spreading the incentive payment over time. This is further considered in 3.2 below.

The potential for in-period incentives paid yearly to vary from year to year (possibly in different directions) could bring the ODI regime into disrepute and needs to be taken into account.

The choice of payment timing would have to be carefully thought through should dynamic comparative outcomes be introduced (see section 4).

3.2 Use of Willingness to Pay (WTP)

Companies have used valuations of service increments from stated preference WTP research to inform if they would propose service level changes in business plans. Typically customers would indicate how much they would be willing to pay each year on their bill for a service improvement and cost benefit analysis was used to identify the economic level of service. Adjustments to opex and capex in FDs were made for cost beneficial service improvements supported by customers. The use of upper quartile industry performance to inform PCs and ODIs has somewhat sidelined the use of this form of analysis for PR14.

Instead WTP was used to inform the rates for ODI penalties and rewards, a purpose many companies' WTP research (which had been undertaken earlier in the process) was not specifically designed to support. There was a consequential disconnect between customers' consideration, in customer research, of the bill impact of the base package compared to the bill impact taking into account potential ODI rewards and penalties.

There was also concern, from many CCGs, that WTP research is complex and produced significantly different values from company to company. However, academics support the approach as the best available.

The following considerations need to be taken into account when developing the approach for PR19. Some aspects may affect the regulatory methodology and some the way companies go about their planning.

The ODI regime is quite complex. If it is to drive significant performance improvement companies will need to feel the available rewards are sufficient to make a meaningful impact on returns. The WTP framework could limit the available reward and marginalise the significance of ODI incentives for investors.

Is there an inconsistency between how WTP is used to inform service improvements in cost benefit assessment versus how it is used in ODI valuation and payment of ODI incentives?

If 'base' service levels are set based on upper quartile industry performance, perhaps ODI rewards, based on WTP, could be used to provide funding for any service increases above 'base' levels that are a customer priority. In developing this approach consideration needs to be given as to whether the ODI is simply a cost recovery mechanism for additional improvements beyond base case or a stronger incentive with a larger reward element.

It will be necessary to consider if WTP research can be designed to apply directly and appropriately to ODI valuation rather than to inform a cost benefit assessment. These are two very different applications and it is likely that the research would be configured differently in each case.

It would be necessary to carefully consider the timing of incentive payments, and how long they would last, so that this can be properly reflected in the WTP research. Can the ODI reward regime be amended to properly incentivise a step change in service that requires significant one-off investment and continuing operating costs. That is to say, ODI that are large enough, through time, to compensate for a large up front cost, which results in on-going customer benefit. This would need to identify what WTP information would be required and how would it be used. Such an approach would require the company to receive ODI rewards beyond the end of a review period to reflect the ongoing benefit received by customers.

For example, is the customer paying a one-off payment for one year of service improvement in that year? This has shortcomings as it is unlikely that an incentive on this basis would finance a permanent service improvement requiring significant totex – even when the totex incentive is taken into account. Or will the customers' bill be increased on a longer term basis for a permanent increase in service (more akin to a traditional cost benefit assessment using WTP to determine the benefits)?

On a separate point, is using regional WTP data compatible with ODIs that have been set using national performance data (based on upper quartile performance)? The range of service level increments explored in the regional research may be quite different to the ODI performance range for rewards and penalties based on national comparative data. This area is complicated further by the potential impact of diminishing marginal utility for customers already receiving upper quartile performance i.e. regional research by a company which has leading performance may potentially produce a lower WTP value for a further unit of improvement than research conducted by a company with more ground to make up.

How can we ensure business plans and FDs incorporate plans that use WTP appropriately, in conjunction with acceptability research, to ensure customers fully support proposals including ODIs? Is there a potential to explore WTP for an incremental improvement in the league table ranking (not resulting from a deteroration in others performance) rather than incremental improvements in the unit of service? The different purposes of WTP and acceptability research need to be clearly identified and methodologies configured to suit.

Can the way WTP research is undertaken be improved and stakeholder confidence restored? A method that is looked upon with suspicion by stakeholders may undermine future price reviews.

3.3 MoS that take more than one regulatory period to deliver

It is important that, when companies set out on a significant programme of work that spans two regulatory periods, it has certainty over how this will be treated in terms of incentives.

Option 1 – regular short-term incentives – (PR14 approach)
Assess delivery when companies reach each MoS milestone and apply incentives that reflect their performance over the period defined by the milestone. Companies would receive either a reward or penalty when they reach each milestone.

Option 2 – one-off assessment at end of agreed timescale for outcome delivery
Only assess delivery and apply an incentive at the end of the agreed timescale for MoS delivery. Ofwat would not apply any interim incentives; any milestones the company sets would only be for the monitoring of delivery.

Option 3 – short-term incentives with end of outcome true-up
Assess delivery as option 1. Ofwat would, however, also assess delivery at the end of the agreed period for achieving the MoS as option 2 and apply an incentive that reflects performance over the entire period. This final, long-term incentive would be adjusted (trued-up) to account for any over or under-delivery earlier in the outcome period. Clarity up front about the incentives that would apply and the process to be employed would be necessary.

4 How might dynamic comparative outcomes work?

4.1 Introduction

There has been a suggestion that the ODI regime could be 'dynamic' in the future. Depending on how this was implemented, this could either be relatively simple or very complex. Material unintended consequences would need to be avoided.

It would be necessary to identify what the dynamic data would apply to (e.g. only PCs, only ODIs or both). Would it apply to both penalties and rewards? What process would be applied over what timeframe? Robust comparative performance data on a consistent basis would be of the utmost importance and how would this be achieved (see also section 2)? How could the approach be configured so that performance excellence is continually incentivised without unrealistic expectations of companies?

The interaction between WTP and dynamic upper quartile benchmarking merits further research. It may be expected that as performance levels improve the WTP for further service improvement will decline (that is that there are declining marginal benefits). This needs to be better understood in order to inform the design of incentives.

4.2 Dynamic comparative outcomes - possible considerations

PCs, ODIs or both?

PCs or ODIs (or both) could be varied in-period based on dynamic industry performance data. A key consideration is, perhaps, that companies should not be penalised for failing to meet a performance target it was unaware of, or had insufficient time to react to achieve.

Revising base performance requirements embodied in PCs relative to industry data in-period would leave companies without any understanding about its absolute delivery commitments over the coming five years. This would increase uncertainty and risk and make planning more difficult.

A possible approach could be to apply dynamic upper quartile performance to revise ODI reward thresholds only. Companies would need to keep up, in-period, with the better performing companies before rewards could be earned.

Under this approach, PCs and penalty thresholds could remain as set at the FD. This would ensure that companies would not be penalised relative to dynamic performance data changes in service areas requiring significant lead-in times to achieve step changes in performance, possibly requiring planning time and investment. It would introduce a little certainty into a regime where companies would not be able to assess their reward potential in advance (as other companies' future actions and performance is unknown).

Consideration would need to be given to a scenario where upper quartile industry performance worsened (e.g. because of adverse weather). Would ODI reward thresholds be revised downwards in this case?

Timing?

In a market, companies do not know in advance how their competitors will perform in terms of levels of service. Indeed, they will only find out in retrospect how well they have done. However, they are unlikely to have to wait for a whole year to do so. So to mimic this market, the publishing of assured quarterly or half yearly industry performance data would be

necessary. On the down side, this would represent a significant resource burden and part-year data may be misleading due to seasonal trends.

Applying retrospective dynamic targets would be challenging for water companies that may need a significant lead-in time to make investments to deliver higher levels of service.

Incentives to improve service (especially where expenditure is required) would also be blunted if there was no assurance regarding the outcome. A company may make an investment decision and find that by the time the scheme has been delivered it has not attained the upper quartile level of performance applying at that later time (and, therefore, will not earn the expected incentive for the investment made). An uncertain incentive may be perceived by company management as not an incentive at all.

Making PCs or ODIs dynamic implies in-period recalculation of upper quartile industry performance levels. Assuming this is done annually, the first in-period industry data would be available in July 2021 (2020-21 data).

Using the above scenario as an example, there are a number of options on how to use the 2020-21 industry performance levels in a dynamic way:

- Use it to reset the 2020-21 upper quartile ODI reward level.
  This would likely be seen as retrospective and inappropriate.

- Use it to set the 2021-22 upper quartile ODI reward level.
  This could be seen as the approach that was closest to a dynamic competitive market, as improved performance fed into higher future customer expectations, although it could still be seen as retrospective. Companies would have insufficient time to adjust (less than one year).

- Use it to set the 2022-23 upper quartile ODI reward level.
  This would give companies a year to assess and respond to dynamic industry performance. This might be considered as more reasonable.

Amending PCs or ODIs based on a single year's data may result in swings in requirements in either direction, on an annual basis, should upper quartile performance vary from year to year. This may be undesirable. Another approach might be to make adjustments based on two or three years' rolling averages to reduce volatility.

4.3 Comparisons with the Service Incentive Mechanism (SIM)

The SIM is an example of a comparative outcome that was applied in PR14. It was comparative, asymmetric and included a deadband where no reward or penalty applies. It had a strong reputational incentive, as published annual data was quickly converted into league tables.

The SIM is dynamic in the sense that the performance levels required to achieve rewards change as companies deliver in competition with each other. It is an ex-post mechanism - companies only find out their final rewards or penalties after the event, although there are annual assessments to allow their progress to be assessed.

The key difference between SIM and other outcomes is that SIM improvements are typically achieved through relatively low levels of investment – often through management, cultural or process change. Improvements in wholesale outcomes will often require a higher level of upfront investment, hence the proposals for incentives for their delivery.

5 Internal outcomes between price control areas

For PR19, it is likely there will be price controls within the current wholesale services control, so when Outcome responsibility spans sub-services (e.g. a Water quality outcome) it is likely there will need to be Service Level Agreements (SLAs) between sub-services. This is demonstrated in the table below.

Water quality outcome						
Resources	SLA	**Treatment**	SLA	**Distribution**	ODI	**End Customer**
Catchment Management		Treatment works compliance		Mains flushing & renewal		Contacts regarding Water Quality, Compliance with DWI standards

Each sub-service in the table has a potential part to play in delivery of he outcome. As all financial incentives for outcome delivery will ultimately come from end customers, who will have little interest in or knowledge of the sub-services of their supplier, it will be up to the companies to translate end customer service valuations into the various parts of the upstream supply chain, probably via SLA rewards / penalties.

This may be relatively easy within an integrated company, but applying an ODI would need careful thought if there were several different organisations responsible for its delivery.