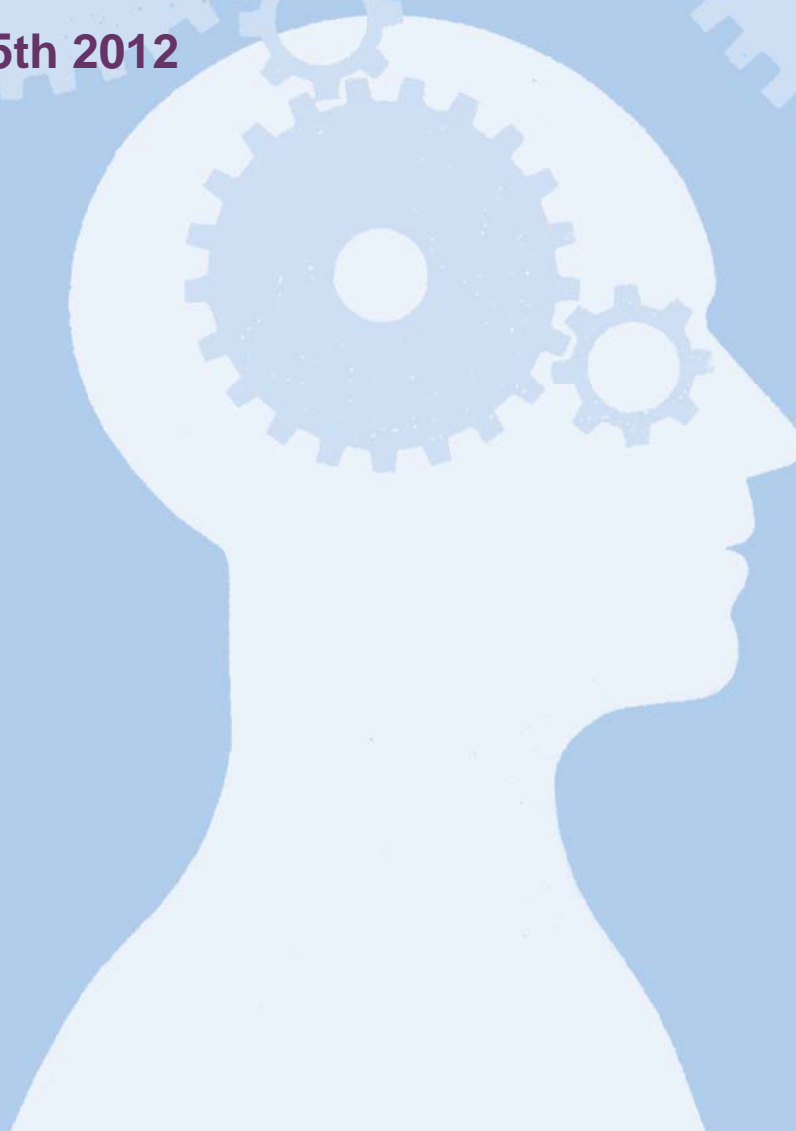


Outcome delivery incentive

**Options in setting future price limits
in the England and Wales water industry**

**An independent report
prepared for Ofwat by Oxera**

August 15th 2012



Oxera Consulting Ltd is registered in England No. 2589629 and in Belgium No. 0883.432.547. Registered offices at Park Central, 40/41 Park End Street, Oxford, OX1 1JD, UK, and Stephanie Square Centre, Avenue Louise 65, Box 11, 1050 Brussels, Belgium. Although every effort has been made to ensure the accuracy of the material and the integrity of the analysis presented herein, the Company accepts no liability for any actions taken on the basis of its contents.

Oxera Consulting Ltd is not licensed in the conduct of investment business as defined in the Financial Services and Markets Act 2000. Anyone considering a specific investment should consult their own broker or other investment adviser. The Company accepts no liability for any specific investment decision, which must be at the investor's own risk.

© Oxera, 2012. All rights reserved. Except for the quotation of short passages for the purposes of criticism or review, no part may be used or reproduced without permission.

Executive summary

Ofwat's proposals for future price limits (FPL)¹ in the England and Wales water sector indicate that companies and customers should seek to propose mutually acceptable outcomes and incentives to the regulator in their business plans. As Ofwat is about to consult on the detail of the wholesale price control, it has asked Oxera to review the incentive options available for outcome delivery and the key considerations when determining the appropriate package to be applied. As part of this review, Oxera has:

- engaged industry stakeholders;
- considered the merits of various incentives and how they relate to outcomes;
- reviewed a set of principles that could be applied;
- put forward recommendations for next steps.

Incentive options and key design considerations

There are a number of incentive options—enforced standard, procedural, reputational, and financial—which can be used on their own or as part of a package.

Of these four, financial incentives ultimately have the greatest impact on behaviour. However, each of them may have a role to play, depending on the time and the circumstances in which the company finds itself and the characteristics of the outcomes, the delivery of which is being incentivised.

It appears to be important to consider characteristics such as the relevance of the outcome, its measurability, the extent to which it is under the company's control, and the scope for innovation to deliver the outcome when deciding on the type of incentives. As part of this study, a high-level assessment of four hypothetical outcomes, each with differing characteristics, indicated that a reputational incentive may also be important. In particular, the reputational incentive may be important when combined with financial incentives, thereby strengthening the overall outcome-delivery incentive.

The suitability or otherwise of an incentive option may be assessed, provided that full consideration is given to key design issues such as the use of one- or two-sided incentives and allowing for trade-offs between measures of success or outcomes when assessing successful delivery. Against this backdrop, as part of its wholesale incentives consultation, Ofwat is consulting on the merits of adopting financial incentives focusing on four key options that differ in their use of penalty only or penalty and reward as well as trade-offs between measures or outcomes.

Penalty-only incentives may result in companies being more cautious when deciding on the outcomes and to be less innovative in their choice of activities undertaken to achieve those outcomes. The incentive package may therefore need to be two-sided and, in some cases, symmetric. Penalty-only incentives, however, may still be appropriate in certain circumstances (eg, when the baseline level of service is to be maintained or legislative requirements are to be achieved).

The objectives of outcome-based regulation (eg, companies focusing on what customers want, taking a long-term perspective and introducing more innovation) may be better incentivised by taking a holistic view, evaluating success or failure as a package, ideally over the medium to long term. By allowing scope for trade-offs between the outcomes as part of

¹ See for example, Ofwat (2012), Future of price limits-statement of principles

an overall outcome package, the associated ‘portfolio effect’ could reduce risk and the need for intervention.

To an extent, the portfolio effect may also arise when relating incentives to individual outcomes. However, the more that the incentives are focused on achieving specific short-run measures of success, the less able companies will be to trade off within a package.

There needs to be a recognition that external ‘events’ (such as a drought or flooding) will always play a role. Often these events are largely offsetting, with the net effect being frequently immaterial to either investors or customers. In such circumstances it may not be necessary to apply a reward or penalty of any kind (for example through the use of the ‘deadbands’).

Process

For an incentive package to be appropriate, it not only needs to have the right economic properties, but must also be capable of influencing management and investor decision-making. To this end, the incentive package must be measurable, transparent, simple, consistent and material.

The incentive package may seek to influence decisions taken by management and Boards, and should not prejudice the financeability of the business. Ultimately, therefore, financial incentives may need to be determined in relation to the resulting returns.

The outcomes and the outcomes package need to focus on what matters most to customers. Success or failure could be judged, where possible, by their local representatives.

As the outcomes and the package need to be measurable, scores or points should be attributed to each element. While this would be akin, to some extent, to the former overall performance assessment (OPA) approach, the exact scores and weights may need to reflect customers’ preferences.

Developing a practical proposal

Having considered the principles from which an incentive package could be developed and the key options available, the way in which that package might be implemented needs to be explored. Oxera will work with Ofwat and industry stakeholders to develop further the details of a practical proposal in the coming months based on Ofwat’s own consultation proposals.

While, in principle, each type of incentive can be applied, further consideration needs to be given to the relevance and practicality of each. For example, if the service package is to reflect local priorities, it is not immediately clear that reputational incentives—for example, in the form of comparative performance tables—are either workable or appropriate. Moreover, thought needs to be given to what procedural incentives—for example, in the form of regulatory investigation of certain aspects of performance—can be warranted in what circumstances.

Additionally, thought needs to be given to the exact form of the financial incentive. For example, how the companies’ proposed incentive (ie, the level of the financial penalty/reward established before price-setting) would be translated into adjustments to price and the time over which under-/outperformance can be retained by the company (ie, how that level of incentive is administered to the company considering other incentives after price-setting).

Activities

As part of the next steps that Ofwat could consider to ensure that its proposed incentive mechanism is feasible and practical, it could explore the implications of the options being considered in the wholesale incentives consultation. In this context, the following steps could be undertaken:

- engagement with companies, UKWIR and their advisers and testing of ideas;
- continuous assessment to ascertain whether mechanisms are likely to work as expected;
- assessment of whether individual incentive mechanisms are really necessary, or whether one simple mechanism could achieve the same goal.

Contents

1	Introduction	1
2	Outcomes delivery incentive: the journey thus far	2
2.1	The Future Price Limits journey	2
2.2	Company responses	4
3	Engagement with the industry	6
3.1	The engagement process	6
3.2	The emerging themes	6
3.3	Issues to be considered: a summary	13
4	High-level assessment of incentive options	15
4.1	Introduction	15
4.2	High-level overview of incentive types	16
4.3	Assessment of incentive types	17
4.4	Assessing outcomes: some examples	20
4.5	High-level assessment of options: summary	33
5	Incentive design	35
5.1	Potential incentive strategies	35
5.2	Specifics of design	37
5.3	Incentive process issues	48
5.4	Conclusions on incentive design	49
6	Next steps	52
6.1	Overview	52
6.2	The incentive design in practice	52
A1	Ofgem experience	56

List of tables

Table 3.1	Outcomes/measures of success characteristics	13
Table 4.1	Measure of success for reliable water supply	22
Table 4.2	Outcomes/measures of success characteristics	23
Table 4.3	Measure of success for good quality of water that is safe to drink	25
Table 4.4	Outcome/measures of success characteristics	26
Table 4.5	Measures of success for sewer flooding	28
Table 4.6	Outcome/measures of success characteristics	29
Table 4.7	Sewerage service clean aquatic environments	31
Table 4.8	Outcome/measures of success characteristics	31
Table 4.9	Key considerations to decide on the use of a financial penalty	34
Table 5.1	Summary of discussion	49
Table A.1	Incentive types adopted by Ofgem	57

List of figures

Figure 2.1	Ofwat's proposed statement of commitments	4
Figure 4.1	Incentive options	17

Figure 4.2	Assessment criteria	18
Figure 4.3	Evaluation matrix	20
Figure 4.4	A reliable water supply	22
Figure 4.5	Evaluation matrix	24
Figure 4.6	Good quality of water that is safe to drink	25
Figure 4.7	Evaluation matrix	27
Figure 4.8	Sewerage service supply: sewer flooding	28
Figure 4.9	Evaluation matrix	30
Figure 4.10	Clean aquatic environments	30
Figure 4.11	Evaluation matrix	32
Figure 5.1	Targeted incentives (an example)	39
Figure 5.2	Possible functional forms for incentive mechanisms	44
Figure 6.1	An illustration of possible implementation of two-sided financial only incentive mechanism	53
Figure 6.2	An illustration of possible implementation of a two-sided financial and procedural/reputational incentive mechanism	55

1 Introduction

Ofwat is now at a crucial stage in the evolution of its overall approach to setting price controls, embodied in its work on Future Price Limits (FPL), which is expected to move away from the type of regulatory approach of incremental complexity over time. The key FPL changes will increasingly involve customers in outcomes-setting, introduce a greater focus on incentives, change the approach to cost assessment, and simplify and reduce the regulatory burden of the regulatory process.

A key element of the FPL changes will be to introduce an outcome-based approach to regulation.

Ofwat is now preparing to consult on the key issues associated with setting a separate wholesale price control, and, in particular, on the package of incentives. In this context, it has asked Oxera to review the options available for providing companies with incentives to deliver the outcomes they commit to achieve as part of their regulatory package; and to set out the main considerations that would need to be taken into account when developing specific incentive design.

To assist Ofwat to develop insights into feasible incentive types and an overview of the merits of different design options, Oxera has therefore set out to:

- engage industry stakeholders to inform the key considerations when developing incentives, as perceived by the industry;
- compile a set of assessment criteria and appraise the key incentive options available to Ofwat;
- provide insights into a number of design aspects; and
- put forward recommendations on the next steps to be undertaken to develop more detailed incentive proposals.

This report is structured as follows.

- Section 2 gives a high-level overview of Ofwat's journey thus far on outcomes and incentives.
- Section 3 presents the main considerations emerging from the engagement with the industry.
- Section 4 illustrates the high-level assessment of incentive options.
- Section 5 reports the insights on incentive design.
- Section 6 puts forward some recommendations on the potential next steps for Ofwat.

It is worth noting that this report sets out the key issues and considerations, and initial proposals for the incentive mechanism. More detailed recommendations will be provided over the autumn period.

2 Outcomes delivery incentive: the journey thus far

2.1 The Future Price Limits journey

This section describes how the rationale for, and design of, outcomes delivery incentives has been discussed and has evolved in various Ofwat discussion and consultation papers over the course of the FPL programme of work.

In ‘Beyond limits—how should prices for monopoly water and sewerage services be controlled?’, Ofwat outlined for the first time that, in future price controls, it was ‘keen to focus more on incentivising outcomes, rather than outputs or inputs’.² This outcome-focused approach was confirmed and further developed in its March 2011 discussion paper ‘Inputs, outputs and outcomes – what should price limits deliver?’:

we [Ofwat] think that economic regulation must focus on incentivising the companies to deliver efficiently the outcomes that customers and society value. The regulator should only concern itself with inputs or outputs to the extent that it is necessary to incentivise outcomes.³

In the same paper, Ofwat defined what it meant by outcomes and provided a first set of potential outcomes, as suggested by industry stakeholders at an FPL workshop:

Outcomes are **the higher-level objectives that company actions, activities and achievements are intended to help deliver**. They represent what customers and society really value...stakeholders suggested that outcomes could include a reliable service to provide clean drinking water and take away wastewater; customer satisfaction; value for money; and environmental sustainability.⁴ [emphasis added]

Several explanations were provided for why an outcome-focused approach might be in the interests of customers. In particular, it was noted that such an approach could grant the companies greater freedom to innovate and to improve processes, could lead to a longer-term focus within the industry (with associated dynamic efficiency benefits), would represent more proportionate and better targeted regulation, and might facilitate a reduction in the regulatory burden for Ofwat and the water companies.⁵

The paper also indirectly noted some of the potential difficulties in applying an outcome-based regime, including the potential non-controllability of outcomes and the difficulties in measuring them owing to their multi-dimensionality and long-term focus, as well as the potential need for trade-offs between outcomes. In particular, Ofwat noted the impact that the non-controllability of outcomes might have in an outcome-focused regime:

There may be an element of ‘rougher justice’ for the companies. Because outcomes are less under company control than the delivery of outputs, the companies may experience pain or gain (that is be penalised or rewarded) as a result of factors not entirely within their control.⁶

² Ofwat (2010), ‘Beyond limits—how should prices for monopoly water and sewerage services be controlled’, July, p. 23.

³ Ofwat (2011), ‘Inputs, outputs and outcomes – what should price limits deliver? A discussion paper’, March, p. 7.

⁴ *ibid.*, p. 8.

⁵ *Ibid.*, pp. 14–17.

⁶ *Ibid.*, p. 16.

Finally, Ofwat acknowledged that delivery incentives might be necessary to incentivise outcomes. It outlined a set of principles that could underpin the outcome-focused framework, as set out in Box 2.1.

Box 2.1 Ofwat principles of an outcome-focused approach

Ofwat proposed the following potential principles underlying an outcome-focused approach.

- As a default position, incentives would relate to the delivery of (or contribution to) outcomes.
- Incentives would only relate to delivering outputs or inputs where:
- Ofwat had reason to consider delivery of an outcome to be high risk;
- incentivising that output or input would significantly reduce the risk of delivery of the outcome; and
- this was consistent with the risk-based framework.
- Information on inputs and outputs should only be collected and used where this was linked to the delivery of an outcome.’

Source: Ofwat (2011), ‘Inputs, outputs and outcomes – what should price limits deliver? A discussion paper’, March, p. 20.

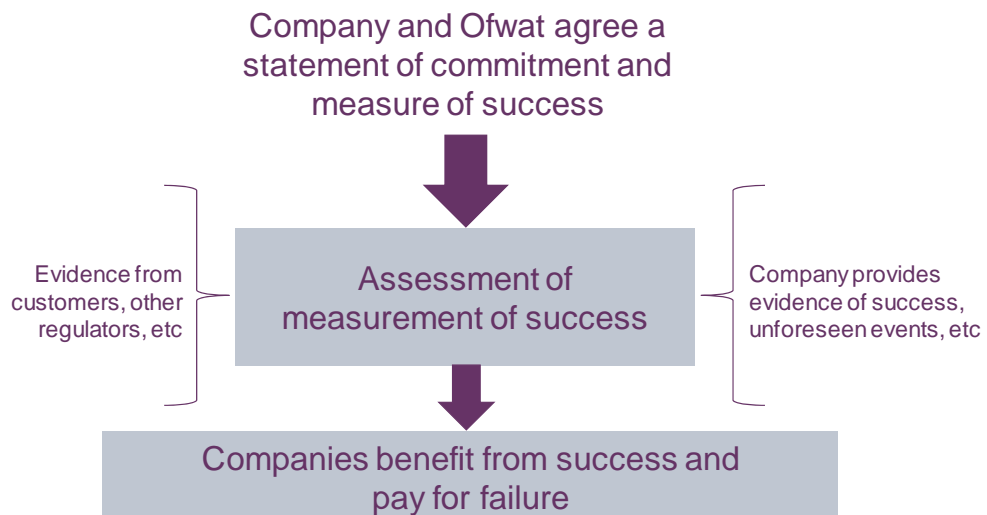
In its FPL framework consultation, Ofwat noted that, following the discussion paper, the outcome-focused approach had received support from companies and industry groups, and had been endorsed in the Gray review.⁷ The regulator proposed a process by which:

- companies would first propose a set of outcomes, along with their associated costs;
- Ofwat and the companies would then agree on outcome-delivery incentives for further consideration by customers;
- the companies would publish a statement of commitments comprising a definition of outcomes with agreed indicative costs; measures of success, timescales and milestones against which delivery would be assessed; and the agreed outcome-delivery incentives;
- performance would then be assessed against the statement of commitments and, if the company were deemed not to have delivered an outcome, cost adjustments could be made with reference to the original indicative cost;
- additional procedural incentives might also be implemented relating to the level of scrutiny that would be applied to a company’s future plans.

This proposed process is illustrated in Figure 2.1.

⁷ Ofwat (2011), ‘Future price limits – a consultation on the framework’, November, p. 19.

Figure 2.1 Ofwat’s proposed statement of commitments



Source: Ofwat (2011), 'Future price limits – a consultation on the framework', November, p. 19.

Most recently, Ofwat acknowledged in its FPL statement of principles that stakeholders were concerned with the proportionality and lack of symmetry in the proposed outcome-delivery incentives mechanism.⁸ In particular, companies expressed the need for a symmetrical incentive to avoid incentivising conservative, risk-averse behaviour with no incentives to outperform. However, Ofwat continued to consider that the delivery incentives should incorporate only penalties, not rewards:

many companies expressed reservations about the proposed delivery incentive mechanisms. Their main concern was that these should be symmetrical—that is, the companies should earn greater rewards if they over-deliver on outcomes. We do not think additional rewards are appropriate...If a company delivers more of an output than customers have expressed a willingness to pay for, then it would appear inappropriate that customers pay more for the thing they stated they did not want.⁹

Ofwat noted that it would carry out further work on the delivery incentive mechanism before deciding on the final make-up of the outcome-delivery incentives.

2.2 Company responses

Following the publication of Ofwat’s FPL consultation and discussion papers, companies were given the opportunity to provide their feedback on the proposals relating to outcomes delivery. These responses, which have played a role in shaping the evolution of Ofwat’s thinking on outcomes regulation, are summarised below. In particular, this section details the comments made by companies (in February 2012) on Ofwat’s November 2011 FPL framework consultation.¹⁰

While stakeholders were generally in favour of an outcome-based approach, particularly with regard to the potential for innovation, they were less favourable towards the proposed delivery incentive mechanism. As noted above, a key issue for stakeholders concerned the asymmetry of the regime, and the potential for it to create incentives for companies to be risk-averse. For example:

⁸ Ofwat (2012), 'Future price limits – statement of principles', May.

⁹ Ibid, p. 21.

¹⁰ Stakeholders’ responses to the consultation are available at: http://www.ofwat.gov.uk/consultations/pap_con201111fpl.pdf?view=responses

[Severn Trent] do not consider that delivery incentives should be expressed solely in terms of penalties for failure. The asymmetric nature of the regulatory regime has been a major factor in encouraging companies to be risk-averse. There should be rewards for delivery of improvements beyond those incorporated in the business plan, which could be agreed as part of the customer engagement process.¹¹

[South West Water] would encourage a balanced approach to rewards and penalties, as a focus on penalties...may make companies less willing to be innovative.¹²

A particular concern in this regard has been that the delivery incentive framework should not preclude companies from delivering more of an outcome, or delivering the outcome earlier than planned, if this is valued by customers. The temporal issue—of whether a company faces incentives to deliver outcomes as early as possible—was seen as one area in which rewards could be introduced to balance the penalties:

If rewards were introduced for delivering the outcomes earlier then this would make the incentive mechanism more balanced between reward and penalty so encouraging more ambition and innovation.¹³

Companies also stated that greater detail was needed on how outcomes would be measured, noting that the assessment was likely to involve a great deal of subjectivity, and thus regulatory risk. Moreover, it was noted that the non-measurability of some outcomes might lead to a return to measuring outputs (ie, a return to the current approach):

A concern...is it will require a significant amount of regulatory judgement in whether an outcome has been achieved and what the appropriate level of penalty may be. This may increase regulatory uncertainty and may not help to stimulate innovative approaches.¹⁴

We foresee the biggest difficulty being defining success criteria against which delivery can be measured...the acceptance that some outcomes will need to be considered as complete by the delivery of the composite outputs, while being a return to the current approach, seems pragmatic.¹⁵

Some companies noted that unless Ofwat could show that its proposals were risk-neutral, there should be a higher cost of capital for PR14.

Other potential issues with the proposed delivery incentive mechanism arose from Oxera's engagement with companies and other industry stakeholders during the course of this project. The main emerging themes from this engagement with companies are included in section 3.

¹¹ Severn Trent Water (2012), 'Future price limits – a consultation on the framework, Severn Trent Water response', February 15th, p. 10.

¹² South West Water (2012), 'Future price limits – a consultation on the framework, Response from South West Water', February 22nd, p. 8.

¹³ Environment Agency (2012), 'Response to Ofwat consultation, Future Price Limits framework', February, para 3.6.

¹⁴ South West Water (2012), op.cit., p. 8.

¹⁵ Bristol Water (2012), 'Future price limits – a consultation on the framework, Consultation question responses', February 15th, p. 7.

3 Engagement with the industry

3.1 The engagement process

As part of the project, Oxera has engaged with water companies to help identify the issues to consider when selecting types of incentive and designing the incentive mechanism. In this context, representatives have been interviewed from a sample of water companies in England and Wales that had expressed an interest either directly to Oxera or via Ofwat. This sample comprises seven water and sewerage companies (WASCs) and three water-only companies (WOCs), or around half of the industry. Furthermore, a WOC provided written comments and information to Oxera on the topics proposed for discussion.

The questions/areas for discussions were as follows.

- The work carried out to date, on defining the outcomes, milestones and measures of success. Discussions in this area were principally aimed at gauging what companies understand by outcome. This information has been used as the basis for developing stylised outcomes (ie, outcomes of a hypothetical company used for the purpose of assessment but not actual outcomes).
- The issues, if any, emerging from the definition of outcomes, milestones and measures of success, such as:
 - definitional aspects—eg, difficulties in defining an outcome;
 - process/governance aspects—eg, the need to balance the objectives and aspirations of stakeholders involved (including the water company);
 - measurement aspects—eg, difficulties in identifying the measures of success/achievement of outcomes or intermediate milestones;
 - any other issues that companies felt needed to be raised.
- Companies' preliminary thinking around how outcome-delivery incentives may be designed and the issues that need to be taken into account in this process. This was aimed at understanding the various aspects involved in developing an incentive, as perceived by the companies themselves.

In addition, Oxera has engaged with the Consumer Council for Water (CCWater), to understand its views in relation to the process to decide on outcomes, incentives and monitoring.

3.2 The emerging themes

The points made by companies and related issues emerging—in particular with regard to the companies' thinking around the incentives—are summarised below according to the following themes: outcomes characteristics, incentive design, and process.

3.2.1 Outcomes characteristics

Water companies' control of the outcome

The degree of control over outcomes may vary between companies. Two factors that determine this are the influence of:

- **other agents**—for example, farmers and other stakeholders in a catchment area could have an impact on the achievement of an outcome, such as healthy aquatic

environments. Another example could be the achievement of good-quality bathing water—eg, using Blue Flag status as a measure—where other parties (ie, not the water company) would need to act in order for the status to be achieved;

- **events outside the company’s control**—for example, the impact of extreme weather events (eg, drought or flooding) on the effectiveness of activities undertaken by companies, and therefore on their ability to deliver.

The extent to which companies have control over an outcome may determine their response to an incentive mechanism. For example, if a company has little control over the outcome, linking a financial penalty to delivery of that outcome increases the risk faced by the company. A company might respond to this increased risk by being less innovative on the inputs it chooses, instead opting for tried-and-tested solutions.

By contrast, however, it was pointed out that the lack of controllability should not be used as the basis to avoid introducing financial penalties. Indeed, if lack of controllability were acknowledged through the introduction of non-financial penalties only, overall this could give the company an opportunity to reduce effort and enjoy any financial benefits arising. It might be tempted to spend less than what it was funded for the delivery of the specific outcome, and hence gain from the difference between allowed and actual costs. At the same time, because of the lack of controllability, it could blame its failure to deliver on external factors. If the company were penalised financially, unless it could demonstrate that it had spent in good faith to achieve the outcome, this might provide a stronger incentive to the company and its management to deliver.

A related point made was the need for the incentive mechanism to be ‘smart’, and to penalise a company for lack of action and not for ‘bad luck’. Conversely, it could be argued that, if a two-sided incentive tool were used, the incentive would need to be designed so as not to reward companies for ‘good luck’.

Measurability of the outcome

Some companies considered that it would not be appropriate to adopt powerful incentives (eg, financial incentives) if the outcome is not easy to measure. For example, an outcome could consist of reducing the company’s contribution to climate change. Some measures may be suitable for this (eg, reductions in annual CO₂ emissions). However, the information provided by these measures about the outcome may be only partial, and the lack of comprehensive measures of success fully covering the outcome make it difficult to measure.

Elaborating on the point made by the companies, it should be noted that the lack of measurability would make it more difficult to monitor companies’ progress and objectively assess their delivery or otherwise. Therefore, the decision on whether a company has failed to deliver the outcome could become somewhat arbitrary. If a powerful incentive is attached to such an outcome, it could expose the company to the possibility of being penalised on a basis that is neither transparent nor predictable. To the extent that these penalties affect a company’s financial performance, a financial incentive in this case could result in increased variability in that performance (ie, it would increase the company’s risk).

Importance of the outcome

When designing an incentive—in particular, when deciding on its power (eg, financial penalties as opposed to reputational incentives; high financial penalties as opposed to lower financial penalties, etc)—one of the many considerations is the importance of the outcome.

Some companies noted that if the outcome is important for customers then it would be understandable (and legitimate) to introduce powerful incentives, such as high financial penalties for non-delivery. Similarly, if the costs allowed for delivering the outcome are high, it may be appropriate to use a powerful incentive to penalise under-delivery.

Oxera's understanding of this point made by companies is that a high financial penalty may need to be adopted for under-delivery because the more money that is funded for the delivery of the outcome, the stronger the temptation is for the company to postpone delivery or to under-deliver, thereby incurring lower costs and, in turn, gaining from retaining, at least temporarily, the difference between allowed and actual costs.

3.2.2 Incentive design

What to link the incentive to

An important decision for the design of the incentive is whether this is to be linked to the overall outcome (ie, assessing the company on its performance in delivering the outcome), or to specific measures of success. In general, it was felt that linking the incentive to a measure of success might induce companies to focus on the measure (ie, delivering on the specific metric) rather than on the outcome. This could have the unintended consequence that companies' response to an incentive linked to a measure of success may be similar to that adopted in the context of output-based regulation. There may be cases when measures are higher-level in nature and closely related to the outcome, such that incentivising the measure or the outcome would result in companies behaving similarly. Elaborating on the above discussion, it may be appropriate to assess the suitability of linking the incentive to the measure of success on a case-by-case basis.

In particular, where more than one measure of success is used to establish delivery of an outcome or milestone, it may be appropriate for companies to be allowed flexibility and scope to trade off delivery among those measures (eg, the company could under-deliver on one measure and over-deliver on another, and overall be judged to be delivering the outcome).

Oxera understands that the work carried out to date by Frontier Economics Ltd (hereafter 'Frontier') for the UK Water Industry Research (UKWIR) has focused on setting out a framework for defining outcomes and measures of success. It is further understood that Frontier has not firmed up specific recommendations as to what to link the incentive to. Its initial thinking seems to favour linking the incentives (financial or otherwise) to measures of success. This may be appropriate in some cases, although to the extent that this approach is adopted for all outcome/measures, in some cases this may result in a regulatory framework with similar incentive properties to those of output-based regulation.

The potential for trade-offs among outcomes was also considered in a few discussions. It was recognised that this could be a means to allow companies more flexibility to adjust to changed circumstances. If a degree of substitutability between outcomes delivery were allowed, it was felt that companies might consider the possibility of being penalised for failing to deliver an outcome, milestone or even a measure of success less likely. In turn, this might lead them to choose innovative solutions for which effectiveness is less clear, on the understanding that failing to meet the commitments in one area could be balanced by accelerated or over-delivery in another area.

Targeted nature of incentive

A point related to the above section is whether to consider incentives incrementally—ie, to target them solely to those outcomes or components of the outcomes for which there are not existing incentives. For example, the delivery of an outcome such as 'providing healthy aquatic environments' may, at least in part, already be incentivised by existing tools, such as enforced standards and/or penalties for pollution incidents, consent compliance, etc.

In this context, it may be appropriate to incentivise those areas/components of an outcome that are not already subjected to an incentive (ie, incentives are added to those already existing). It was noted that this might result in incentivising individual measures of success rather than the overall outcome. Indeed, when companies adopt outcomes with measures of success that have incentives attached to them (eg, number of pollution incidents) and those that do not (eg, progress on delivering local environmental objectives), the incremental approach would result in incentives being introduced only in relation to the latter. The

potential drawbacks of linking incentives to measures of success were illustrated in the previous section.

Risk/reward balance

In general, companies felt that the introduction of outcomes might create more risk, in that the higher-level nature of outcomes might result in companies being assessed on their performance in delivering something less easy to measure and monitor. The somewhat less specified goals may also make it difficult for companies to adjust their planning processes and procedures in response to the shift to outcome-based regulation and other important concurrent changes (eg, split price-setting and planning between wholesale/retail). The novelty of the framework and difficulties affecting the planning/decision process may increase the risk of not delivering. It may be appropriate to recognise this context when designing the delivery incentives. In particular, in recognition of the difficulties in the transition to outcome-based regulation, it may be appropriate to use less powerful incentives (eg, keep the financial penalty relatively small), at least for the first time that this new regulatory framework is applied. In other words, while both regulator and regulated companies, as well as other stakeholders, will improve their understanding of the implementation of outcome-based regulation, it may be appropriate to avoid incentives with large financial implications.

However, if financial incentives are small in a transition period, it could be argued that it might be difficult to assess whether the incentive mechanism works or would work in the longer term if the financial implications were larger.

Companies noted that they might also find it difficult to propose a package of outcomes and relative incentives until Ofwat clarifies other aspects of the regulatory framework, such as cost recovery and menus. This is required as the company would need to understand the potential risk/reward of the overall package before finalising and committing to a set of outcomes, milestones and incentives.

An additional point raised was that the balance of rewards and penalties should be considered in light of the level of innovation that is desired. In other words, if, in a particular area, greater innovation is to be incentivised, this would need to be reflected in the balance of rewards and penalties.

One- versus two-sided incentives

Companies' view is that the use of a one-sided incentive (assumed during the discussions to be penalties only) may be potentially detrimental for innovation.

It was suggested that companies would tend to 'play it safe'. Indeed, a riskier strategy to delivery would not pay off (ie, companies would not receive a reward), even if it delivered a better outcome (eg, higher quantity or quality, or delivered earlier than scheduled). At the same time, it could cost companies' money if the strategy proved to be ineffective and resulted in under- or late delivery. Oxera's understanding of this issue arising from the use of a one-sided approach, as explained by companies, is reported in Box 3.1.

Box 3.1 Potential unintended consequences of one-sided incentives

The line of argument put forward by companies is that a one-sided, financial tool that penalises companies for not delivering or delivering late may result in higher expected costs of delivery. This is because the companies would add to the planned cost of delivery the expected costs of a penalty (ie, the probability of being penalised times the value of the penalty). The expected costs of a penalty would not be offset by any reward for outperforming the agreed levels of the outcome, as would happen if the incentives were two-sided. Overall, this could result in higher expected costs. Companies may seek to reduce these costs by lessening the probability of being penalised, for example. Discussions with companies reveal that this could be done in one of two ways:

- **taking a cautious approach to outcomes-setting**—the company aims to set outcomes and measures of success in a conservative way so as to increase the level of controllability and, more generally, reduce the level of effort required to deliver the

outcome;

- **having a bias towards solutions that are more certain**—companies may decide to adopt less innovative solutions (eg, demand management in water supply and demand balance), where the likelihood of delivery is greater than for other, more innovative, solutions.

Companies may choose to do a mixture of the two. Overall, the result may be a less dynamic and innovative industry in terms of both setting the outcomes and deciding how to achieve them.

Indeed, widespread use of a one-sided (penalty-only) incentive may lead to unintended consequences, as highlighted by the companies.

Companies expressed the view that it might be possible to use a one-sided incentive when they have to meet legislative requirements, for example, and the outcomes are broadly under their control. However, where there is lack of controllability, a pre-defined maximum penalty may be needed in order to limit the financial risk to the company.

It was also noted that commitments to maintain base service may be better incentivised using one-sided (penalty-only) tools because companies should not be rewarded for providing their base services, but may be rewarded for any improvements above that basic level of service.

Building on these points, it may be appropriate to use a two-sided incentive for incentivising improvements to services (ie, over and above the base level).

The value of a financial incentive

Companies did not provide a consensus view on what the value of financial incentives should be based; some indicated that the financial incentive may need to be related to the money allocated to the outcome—ie, the capital and operating expenditure (CAPEX and OPEX) allocated by the companies to the delivery of the outcome.

There could be issues surrounding the possibility of costs being incurred across more than one outcome (ie, joint costs). For example, an upstream measure (eg, catchment-sensitive farming) may result in improvements to aquatic environments (ie, an environmentally driven outcome) and water quality (ie, a water supply-related outcome). In this instance, when it comes to designing the incentive, where would the cost be allocated?

A few companies mentioned that it may be appropriate in some cases to link the financial incentive to the value loss by customers from non- or late delivery (assuming that this is well-defined).

A final point relates to the possible need to scale the financial penalty to the degree to which the company is failing to meet the outcome or milestone.

A package of incentive options

While discussions initially focused on individual types of incentive (eg, financial or reputational), they went on to explore different types of incentive. The following points emerged from these discussions.

- **Reputational incentives are not just league tables**—discussions seemed to indicate that, even if a financial incentive only were chosen, there would be a reputational effect incentivising companies and their management. For example, if a company is fined for not delivering and this becomes publicly known, there would be a further negative reputational effect. The senior management teams of water companies will want to avoid this; hence, the threat of a financial penalty in itself is likely to move management to act.

- **Scope to use more than one incentive type.** The possibility of using more than one incentive type for different outcomes was discussed. In general, companies were supportive of the principle that different incentive types might be used in different circumstances. In particular, discussions with some companies indicated their support for an approach based on the adoption of procedural and reputational incentives in the first instance and then financial incentives.
- **Simplicity and transparency.** During discussions some of the companies stressed the need that, independently of the choice of the type of incentives, key to a successful implementation (ie, for companies' managements to understand and respond to the incentive and for customers to be effectively involved) is the need for the incentive to be as simple and transparent as possible.

In practice, the framework that companies seemed to envisage for a company that is not delivering an outcome could be broadly as follows.

- **Procedural implications.** Once it has been determined that the milestone has not been met and the outcome is not on track, the company could be subjected to more in-depth scrutiny (eg, the regulator would require more information from the company). This would give it the opportunity to explore the reasons for its non-delivery and demonstrate, if relevant, that this is due to factors outside its control.
- **Further procedural implications.** The non-delivery of one milestone could also result in the company being subjected to more in-depth scrutiny at the following price control review. In other words, missing the milestone in one period would result in the company being 'slow-tracked' at the next review.
- **Reputational implications.** The regulator or other nominated body could publicly signal that the company is failing to meet the outcome (this would have a reputational impact).
- **Financial implications.** Ultimately, based on the evidence, financial incentives could be applied, as appropriate.

While there may be merits in such a gradual approach (eg, reducing the risk of a company being financially penalised for events outside its control), certain issues might arise as a result of adopting a framework of this kind:

- where the company does not deliver and shows that this is mainly due to external circumstances, the customers will be the bearers of the full risk of non-delivery. This in turn would result in customers suffering a loss, as a result of the non-delivery, for a number of years (potentially in excess of five years depending on how often milestones are set and assessed). It is unclear that customers are better placed than companies to bear the risk and face the consequences of non-delivery;
- companies might focus their effort on trying to show that non-delivery is due to factors outside their control, rather than trying to manage these external factors and still make progress towards delivery. This would be an undesirable unintended consequence of the framework;
- related to the above point, the implementation of the framework could dampen the overall power of the incentives being adopted, as companies would see the scope for 'talking themselves out of' paying a financial penalty (ie, convincing the regulator or the designated body that external circumstances underpin the under-delivery);
- the adoption of the framework may become cumbersome, complex to implement and intrusive for companies and, therefore, not in line with Ofwat's intention, as set out in FPL, to shift towards lighter-touch regulation.

There are merits in having an incentive framework that mitigates the risk for companies arising from events outside their control and that, in doing so, relies on different incentive types. However, such a framework will need:

- **to be easily understood and implemented**—this is important for companies to respond appropriately to the incentives, and for stakeholders, who may have a role in monitoring the companies to ensure that they carry out their activities properly;
- **to be based on transparent rules**—this is to limit, as far as possible, the discretion and potential arbitrariness that could result in an increased risk faced by the companies;
- **to provide companies with incentives to manage/mitigate against external events**—companies should be left with a credible prospect of facing powerful (ie, financial) incentives if their failure to deliver outcomes or meet milestones is noticeable. This may provide them with incentives to seek to manage any unforeseen circumstance and keep delivery of outcomes and its pace at a level that is acceptable to customers.

3.2.3 Process/governance (ie, who does what)

Most companies expressed the view that greater clarity is needed on the process and governance, including the role of the CCG, for example. Indeed, the governance needs to be clarified with regard to the period preceding final determinations—for example, the interaction between CCG and the company, and which body decides on what the incentives should be.

Even more important is to clarify which body will be monitoring outcome delivery and assessing whether a penalty needs to be applied. Moreover, whichever party is responsible for this monitoring needs to have the resources and capabilities to assess the company's performance.

It was suggested during the interviews that this might be beyond the skills of the local CCG, and could be dominated by the Consumer Council for Water (CCWater) and the regulatory bodies, which might not be consistent with targeting outcomes that local customers want. A related point is that there could be an issue around institutional set-up and continuity. The CCG is set up to assist in the business planning stage, with the process resulting in the company's commitments; it is not set up for the monitoring phase. If the remit of the CCG were expanded to give it a role in the monitoring phase, the members might need to be recruited on a different basis to the current practice, giving them a full-time job. This might also allow those who help to establish, alongside the companies, certain objectives to monitor the delivery of those objectives and judge whether an incentive mechanism (eg, penalty) needs to be triggered.

Some companies suggested that it may be appropriate for customers to oversee delivery and contribute to the decision on whether incentives should be triggered. In this context, it may be appropriate for companies to move towards customer engagement on a continuous basis, not just in the run-up to a price control review.

The balance between centrally determined (by Ofwat) and bespoke/locally determined outcomes, measures of success and incentives was also discussed. Some companies acknowledged that a process guided by Ofwat might help to ensure consistency of approach within the industry. However, there were suggestions that any centrally developed guidance would need to be applied on a bespoke basis.

It is noted here that a consistent approach may be of help in developing reputational incentives based on league tables/comparisons.

There were doubts/concerns around the possibility of Ofwat reining back and having more control over the overall outcomes-setting and incentives at this stage—ie, once companies

have invested substantially in developing the engagement process, defining what is expected of members of the CCGs, etc.

3.3 Issues to be considered: a summary

Building on discussions with the companies, the following issues/themes have been identified to be considered when designing the incentives.

Outcomes and measures of success characteristics

The characteristics of the outcomes and/or measures of success to consider when assessing the type of incentive to use are reported in Table 3.1.

Table 3.1 Outcomes/measures of success characteristics

Characteristic	Definition	Key implications
Relevance of the measure	The extent to which the measure of success exhaustively captures the outcome	Incentives linked to measures of success that are not high level and close to the outcome may provide companies with incentives akin to those in output-based regulation
Controllability	The extent to which the company can influence the achievement of the measure of success and outcome; either directly, by carrying out a particular activity, or indirectly, by influencing the behaviour of other individuals or organisations	Incentives based on measures of success and outcomes that are not within a company's control may increase regulatory risk, particularly when financial incentives are applied and depending on the financial impact of the incentive. In relation to financial incentives, for example, a company may benefit from windfall gains owing to 'good luck' or could be adversely affected by penalties as a result of 'bad luck'
Measurability	The extent to which the measure of success and achievement of the outcome can be observed at the time of the assessment	Measures of success for which the metric is not easily defined may lead to an increase in regulatory uncertainty at the time when performance is assessed
Importance	The extent to which customers or society attaches a high value to the measure of success and the achievement of the outcome, as elicited using stated-preference studies, for example; and/or the delivery of the outcome has a material impact on customers' bills	Implication for incentives: financial incentives may be appropriate for outcomes that are important to customers. Two-sided incentives, in particular, may be appropriate for the delivery of desirable improvements—ie, improvements that are not driven by legislation, and for outperformance in areas for which customers have expressed a willingness to pay
Scope for innovation	The extent to which the company has the flexibility to adopt innovative approaches to deliver the measure of success and to achieve the outcome, while ensuring that this is not at the cost of undue risks to the customer or society	Strong financial penalties may lead to an increase in risk aversion and a bias against innovative and alternative solutions. Two-sided incentives may help to offset this risk aversion

Note: Comparability of outcomes and measures of success will also be relevant to the extent to which reputational and comparison-based incentives can be used. However, comparability will depend on the extent to which outcomes or measures are set centrally (eg, by Ofwat). This issue is therefore touched upon in the section focusing on process issues (see the next page).

Source: Oxera on the basis of discussions with a sample of England and Wales water companies.

Incentive design issues

The characteristics of the incentive design to consider further are as follows.

- **Individual or multiple benchmark**—the objective is to assess whether the incentive should be linked to the delivery of multiple measures of success, feeding into an overall judgement as to the delivery of the overall outcome (in principle, this could allow trade-offs between performance on different measures);

- **trade-offs and risk**—the objective is to establish how the balance of risk is changed by assessing the delivery by companies considering multiple outcomes;
- **targeted incentives**—the objective is to assess whether the incentive should be linked to areas (eg, measures of success or parts of outcomes) currently not subjected to any form of incentive, or to the overall outcome;
- **the level of financial incentives**—the objective is to illustrate the issues to consider when setting a financial penalty or reward;
- **the symmetry of incentives**—the objective is to consider the incentive properties and implications of a one- or two-sided incentive; and
- **the balance of risk**—the objective is to assess how risk allocation and risk mitigation are addressed by the incentives.

Process issues

Building on the companies' comments, practical application of the incentives raises issues relating to the balance between general or bespoke incentives for different outcomes and different companies, and the role and responsibilities of the regulator, companies, and stakeholders.

Taking the issues forward

The remainder of this report provides some insights into the issues emerging in discussions with the companies. First, the type of incentives to be used for a generic outcome is assessed, and then specific insights are provided into outcomes with differing characteristics (see section 4). Second, an overview of the considerations in relation to incentive strategies and design is given, along with insights regarding the implications of the process and associated choices (see section 5).

4 High-level assessment of incentive options

4.1 Introduction

Ofwat has noted that, as an economic regulator, incentives are the most important tool available to it to discharge its responsibilities.¹⁶ The challenges of incentive design have their provenance in asymmetries of information between regulators and the companies they regulate. As Church and Ware (2000)¹⁷ noted:

A perfect regulatory scheme would provide the firm with incentives to attain the same outcomes that a regulator would *direct* if the regulator had the same information as the firm

Much of the literature on incentive design is concerned with identifying means to overcome this asymmetry, and is reflected in preoccupations with theoretical mechanisms intended to permit regulators to discriminate between firms of different type (such as high- and low-cost firms).

On the one hand, private information exacerbates the friction between regulators and firms by increasing the scope for hidden actions that firms can take to maximise their profits at the expense of social welfare (and hence the need for incentives and incentive schemes capable of aligning diverging interests). On the other hand, private information can be harnessed in incentive schemes by regulators willing to offer financial rewards to firms prepared to supply private information. Moreover, incentive regulation takes advantage of the fact that the regulated firm is likely to have a better understanding of how to reduce its costs than the regulator. The realisation of cost savings is driven by—and to a large extent relies on—the pursuit of profit (and profit outperformance) by the regulated business:

the regulator delegates certain pricing decisions to the firm and...the firm can reap profit increases from cost reductions. Incentive regulation makes use of the firm's information advantage and profit motive. The regulator thus controls less behaviour but rather rewards outcomes.¹⁸

Consequently, the greater the extent of private information, the greater the scope (and need) for incentive schemes.

At a high level, at least two relevant lessons emerge from this literature. The first is that regulation in the presence of asymmetric information is inherently about trade-offs. For example, it may not be possible to distinguish easily between high- and low-cost firms, and the difficulty of making this distinction may mean that some firms earn higher returns than they would under full information. The second, related, point is that practical issues in an unavoidably imperfect world cannot be overlooked.

Both of these themes are explored in this and the next sections, which are concerned with the design of incentives that are in principle capable of supporting Ofwat's objectives with respect to outcome regulation.

¹⁶ Ofwat (2010), 'The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper', p. 4.

¹⁷ Church, J. and Ware, R (2000), *Industrial organization*, Boston, McGraw Hill, p. 839.

¹⁸ Vogelsang, I. (2002), 'Incentive Regulation and Competition in Public Utility Markets: A 20-year Perspective', *Journal of Regulatory Economics*, 22:1, p.6.

This section provides an overview of the incentive types that can be expected to incentivise the delivery of a generic outcome, and how the outcome might change when considering specific outcome examples, in order to illustrate how particular aspects of the outcome may affect the choice of incentive. The section reflects and builds on work that Ofwat has done on incentives as part of the FPL programme. The next section focuses on the incentive strategies and specific design.

4.2 High-level overview of incentive types

To identify the most suitable incentive types, this section begins by reporting out the options available to Ofwat. An appraisal framework is then illustrated to assess the merits of using one of these options. Finally, an assessment is presented of the types of incentive for a generically defined outcome using the framework.

The results from the appraisal of the unspecified outcome have been tested using hypothetical outcomes with differing characteristics. These examples are described and the assessment results reported. Finally, the combined or sequential use of incentive types is considered further.

Five broad types of incentive that could be applied to incentivise outcome delivery were identified as part of this study. These build on the three categories of incentives identified by Ofwat in its paper on good incentive design¹⁹ and on Ofgem RIIO regulatory precedent.²⁰ The Ofgem regulatory precedent is particularly important because, in that context, incentives are aimed at companies to achieve high-level outputs, which are broadly equivalent to outcomes as defined by Ofwat.

The types of incentive identified are:

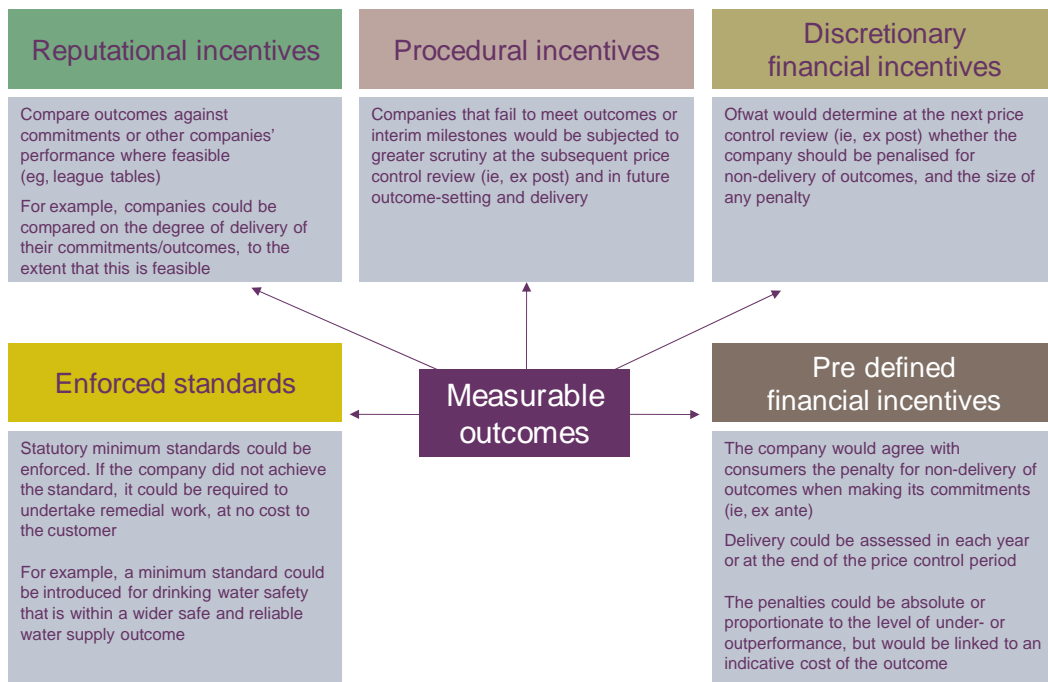
- **financial**—where the companies stand to gain or lose financially depending on their delivery of the incentivised outcome. Two types of financial incentive are outlined, according to whether the level of the financial penalty is pre-defined or subject to regulatory discretion;
- **reputational**—where companies, and their management, stand to gain or lose reputationally from their delivery of the incentivised outcome (ie, with customers and, from a careers perspective, with future employers);
- **procedural**—where companies are subjected to greater or lesser regulatory scrutiny depending on their delivery of the incentivised outcome.

A fourth option—enforcing additional minimum standards—is also considered (see Figure 4.1 below).

¹⁹ Ofwat (2010), 'The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper', October, p. 5.

²⁰ Ofgem (2011), 'Decision on strategy for the next gas distribution price control-RIIO-GD1 Outputs and incentives'; and Ofgem (2011), 'Strategy for the next transmission price control-RIIO-T1 Outputs and incentives'. See, in particular, table 1.1 in both documents.

Figure 4.1 Incentive options



Source: Oxera.

These options are not mutually exclusive—indeed, they may be inherently linked (such that a procedural intervention may incur financial or reputational costs for the company), or they may be explicitly used in combination to incentivise the delivery of a single outcome. As an example of how this can work in practice, the overall performance assessment (OPA) carried the threat of financial penalties (in the form of a revenue adjustment), but also created strong reputational incentives for companies (by ranking them against one another).

If Ofwat were to use bundles of options, these could, in principle, be applied simultaneously (eg, it could enforce a minimum standard and apply a financial penalty at the same time), or sequentially (eg, a league table could be published annually and a company that consistently underperformed could then be penalised financially). The use of bundles of options could also depend on whether it would be proportionate to have more than one option without creating undue regulatory burden.

4.3 Assessment of incentive types

4.3.1 Assessment criteria

Oxera has developed a set of criteria against which incentive options can be assessed. This is a bespoke set of criteria incorporating Ofwat's guidance on the characteristics of a good incentive,²¹ its broader FPL principles,²² and general regulatory best practice in the implementation of incentives. The criteria are described in Figure 4.2, together with what they are intended to cover.

²¹ An incentive should a) have a clear aim; b) be proportionate to its aim; c) be clear and transparent; d) minimise unintended consequences; and e) be robust to change. See Ofwat (2010), 'The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper', October, pp. 11–12.

²² Price control regulation should a) be targeted; b) be proportionate; c) incorporate effective incentives; d) give companies ownership of and accountability for the delivery of outcomes; e) be flexible and responsive to change; and f) be transparent and predictable. See Ofwat (2012), 'Future Price Limits – Statement of Principles', May, pp. 10–12.

Figure 4.2 Assessment criteria

Effectiveness	How likely is it that the approach will encourage companies to achieve the outcomes?
Efficiency	Are the outcomes likely to be achieved without any unnecessary costs?
Proportionality	Are the penalties that might result from implementation of the framework commensurate with the benefits/losses to the consumers?
Balance of risk	Is the resulting allocation of risk likely to be optimal (ie, are the risks borne by those best placed to manage them)?
Practicality	How practical is the framework to deliver the aims, given the skills and resources required, and the associated cost and/or time implications?
Unintended consequences	Does the framework incentivise behaviour that is neither intended nor desirable?
Overall incentive package	How does the proposed framework interact with the wider package of incentives (eg, the total expenditure (TOTEX) approach)?
Adaptability	Is the incentive framework flexible such that it can be altered to reflect changes in circumstance?

Source: Oxera.

4.3.2 High-level assessment of incentive types

In the first instance, the incentive options introduced above are assessed against their ability to incentivise the delivery of a generally defined outcome. Thus, it is assumed that the outcome considered can be identified and measured. Section 4.3 extends the assessment to individual outcomes to account for the fact that the characteristics (eg, measurability, controllability, etc) of the outcomes and the related measures of success may result in different types, or different combinations, of incentives being more suitable for incentivising certain outcomes.²³

From assessment against the above criteria, it would appear that a financial incentive is the most suitable option for outcome delivery. In particular, financial incentives have the following advantages.

- **Effectiveness**—financial penalties would be likely to create stronger incentives for companies to deliver outcomes than reputational or procedural options because they have a clearer, more direct implication for the company’s profit.²⁴ Senior company management would therefore be likely to be under pressure from shareholders to deliver on the agreed milestones, measures of success and outcomes, and thereby avoid the penalty.
- **Proportionality**—in principle, financial penalties can be set in line with the expected forgone benefits/welfare loss arising from non-delivery, and could thus be considered to be proportionate.

²³ In other words, it may be preferable to incentivise some outcomes financially, while, in other cases, reputational and procedural incentives may be needed to substitute for, or complement, financial incentives.

²⁴ Schemes that leave the firm a large share of profits as a result of behaving in accordance with societal preferences are known as high-powered incentive schemes. See, for example, Laffont, J.J. and Tirole, J. (1993), *A Theory of Incentives in Procurement and Regulation*, The MIT Press, p. 11.

- **Balance of risk**—financial penalties shift the risk around delivery onto the companies, which in most instances, to the extent that the outcome measure focuses on the controllable aspects, are likely to have a better control over outcomes delivery.
- **Interaction with the wider incentive package**—financial incentives are likely to interact with the menu incentives such that outcomes are delivered efficiently, and only true efficiency is rewarded.
- **Adaptability**—discretionary financial penalties would be more adaptable than pre-defined financial incentives, as Ofwat would be able to take account of changing circumstances before applying penalties. This would allow the regulator to make judgements about whether non-delivery was the result of non-controllable factors before imposing penalties.

On the other hand, as noted in section 3, one concern with financial incentives may be that companies could be reluctant to innovate if they stand to face a large financial penalty for not delivering an outcome. This would be a particular concern for outcomes where there is a significant scope for innovation, such that the unrealised benefits from innovation are high. Similarly, loss-averse companies might inefficiently overspend on outcome delivery so as to ensure that they were not penalised, should financial penalties be applied. Finally, financial penalties may not be suitable where there is a large degree of non-controllability in the targeted outcome. This is captured, for example, in Ofgem’s criteria for when it is appropriate to adopt a financial incentive as an option, as shown in Box 4.1.²⁵

Box 4.1 Ofgem guidelines on the use of financial incentives

As set out in the RIIO handbook, Ofgem deems it to be appropriate to apply financial incentives only when:

- there is clarity on the primary outputs to be delivered;
- there is confidence in the data used to measure performance;
- it considers delivery of the primary output to be important;
- there are not already incentives in place on the network company through other schemes or obligations.

Furthermore, Ofgem’s guidelines require that, when applying financial incentives, ‘the strength of any incentive should take account of the network company’s degree of control over the output.’

Source: Ofgem (2010), ‘Handbook for implementing the RIIO model’, October 4th., p.76; Ofgem (2010), ‘Consultation on strategy for the next transmission and gas distribution price controls – RIIO-T1 Outputs and incentives’, December 17th, p. 53.

Reputational incentives have been introduced by regulators on the basis that senior management is unlikely to want to be held to account in the media if the company’s performance is below expectation, or poor relative to other industry participants. The reputation of the company may therefore be valued by senior management, who would aim for high scores/relative performance. Moreover, some companies may have pay incentive schemes such that employees’ rewards are linked to the company’s performance. This may result in strong incentives for the employees to work towards the delivery of the desired outcomes.

On the other hand, reputational incentives may be less effective than financial incentives to the extent that the link to the company’s profit is less direct. Indeed, in a monopoly market, reputational incentives are likely to hold less sway than in a competitive market, or where there is the threat of entry, where customers can choose between suppliers. One potential

²⁵ Greater detail on Ofgem’s use of delivery incentives can be found in Appendix 1.

way to overcome this might be to couple reputational incentives with financial incentives. Experience from the OPA, for example, showed that combining a reputational component with a financial one can strengthen the overall effectiveness of the incentive.²⁶

The use of **procedural incentives** (eg, fast-tracking) in regulation is more recent. As such, it is not yet entirely clear how great an impact such incentives have on the behaviour of regulated companies. Procedural incentives could generate reputational impacts (ie, customers and investors will know whether a company is being subjected to greater scrutiny) and they have the potential to affect costs (eg, less regulatory scrutiny might reduce Business Plan costs). One factor that may limit the strength of procedural incentives is that they involve a temporal trade-off. This is because the benefits (penalties) from delivery (non-delivery) in the current period are not realised (incurred) until some point in the future (eg, failing to deliver against a milestone might not be penalised until the next round of outcomes-setting).

The matrix shown in Figure 4.3 summarises a high-level evaluation of the options against the assessment criteria. Equals signs (=) indicate an effect equivalent to the status quo; plus signs (+) an improvement relative to the status quo (ie, with no outcome-delivery incentives); and minus signs (–) a worsening. The evaluation is carried out also considering how options perform relative to each other (ie, the matrix should be read across rows, as the options are also assessed relative to one another). The overall result is an un-weighted summary of the two sets of considerations. It is important to note that the specific characteristics of each option can affect the assessment against each of the criteria, and hence it should be interpreted as indicative only.

Figure 4.3 Evaluation matrix

	Enforced standards	Reputational incentives	Procedural	Discretionary financial	Pre-defined financial
Effectiveness	=	=	+	+	++
Efficiency	=	=	=	+	+
Proportionality	=	=	=	+	++
Balance of risk	=	–	+	+	++
Practicality	+	+	+	–	–
Unintended consequences	=	–	+	–	+
Interactions with overall wholesale incentives	=	=	+	+	++
Adaptability	–	=	+	++	--

Source: Oxera.

4.4 Assessing outcomes: some examples

Hypothetical examples of four outcomes have been developed to illustrate the characteristics that are likely to be important when designing the outcome-delivery incentive mechanism. The examples set out the outcome, the broad dimensions of the outcome, and the measures of success. They also indicate activities that could be undertaken to achieve those measures

²⁶ See for example Ofwat (2010), Ofwat response to the call for evidence of the Ofwat review, para 75

of success and deliver the outcomes. While these examples are not intended to provide a comprehensive assessment of each outcome—nor, indeed, an exhaustive set of measures of success—they are constructed on the basis of material provided by, or discussions with, companies. As such, the examples could be deemed broadly representative of the thinking of a sample of companies to date. It is acknowledged that the examples are closely linked to what have traditionally been identified as the key services provided by the water industry, although this does not necessarily imply that the outcomes and measure of success are optimally defined. There are exceptions, however—indeed, in a limited number of instances, companies have sought to define outcomes at a higher level such that they are somewhat de-linked from the traditional definition of a water company’s services.

The first step was to define the outcome and identify the broad areas that it covers—ie, its dimensions. Such an approach aims to identify measures of success that relate, as closely as possible, to the outcome. In particular, it was assumed that a company would choose and propose measures of success according to pre-defined criteria, such as measurability (the ease with which milestones can be set) and completeness (the extent to which the measures of success cover the whole dimension of the outcome). In most instances, however, more than one measure of success for each outcome has been identified. It is understood that this is consistent with the work carried out to date for UKWIR by Frontier. Since one of the objectives of this study is to assist in the process of developing recommendations on the next steps in relation to developing an outcome-delivery incentive, and more generally the wholesale incentive package, this work is aimed at complementing that being carried out for the water industry to move from hypothetical to practical, specific methodology options for Ofwat to develop.

Once the measures of success have been identified, the types of activity that a company may undertake to achieve those outcomes are assessed. This approach is useful in understanding the suitability of different incentive mechanisms and, in particular, ensuring that the incentive mechanism targets the desired type of behaviour. Through this process, for example, it would be possible to understand the scope for innovation in delivery.

The next section presents examples of:

- the outcomes, types of activity undertaken to achieve them, and measures of success for water supply and sewerage services;
- the key characteristics of the outcomes;
- the implications for the assessment of incentive types for a generic outcome, as reported in the preceding sections.

4.4.1 Assessment of incentive types for water services supply outcomes: reliable water supply

The outcome relating to the reliable supply of drinking water covers three broad areas: acceptable water pressure; secure supply; and minimising interruptions (see Figure 4.4).²⁷

²⁷ Oxera started with an overall water supply outcome—ie, reliable supply of high quality and safe water. However, the thought process carried out to assess this overall outcome identified that reliable water supply, high quality and safe water had different characteristics and needed to be split and assessed separately. Moreover, It is understood that the split between reliable and safe water supply has been made in the work carried out by Frontier for UKWIR.

Figure 4.4 A reliable water supply

Outcomes:		A reliable supply			
	Acceptable pressure		Secure	Interruptions minimised	
Measure(s) of success	% of properties with persistent inadequate pressure	% of properties with temporary inadequate pressure measure weighted towards length of time	Security of supply headroom (expressed as a %)	Resilience or asset health index (eg, based on serviceability)	% of properties affected by supply interruptions
Activities/Inputs	Enhancement activities to remove properties with persistent problems Maintenance activities Operating activities (eg, fixing leaks, etc)		'Supply-side' intervention eg. development of new resources, bulk supply; and leakage control 'Demand-side' intervention (eg, water metering, water efficiency advice)	Maintenance activities (eg, mains relining, renewal)	Operating activities (eg, fixing leaks, etc)

Source: Oxera.

In relation to the achievement of this outcome, Oxera has identified five measures of success, as shown in Table 4.1.

Table 4.1 Measure of success for reliable water supply

Outcome dimension	Measure of success	Rationale
Acceptable pressure	% of properties with persistent inadequate pressure	Covers properties that have persistent problems with inadequate pressure
	% properties with temporary inadequate pressure	Covers properties that have had a temporary problem with inadequate pressure
Secure supply	% security of supply headroom	Provides an overall view of security of supply that should reflect performance in other areas (eg, hosepipe bans)
Interruptions minimised	Resilience or asset health index	Provides a long-term view of the condition of the water network at a point in time
	% properties affected by supply interruptions	Provides a view of properties that have been affected by interruptions to supply

Source: Oxera.

Figure 4.4 also summarises the **activities** that the company can undertake to achieve the outcome. Some of these, for example, could be done to improve the security of supply; these can be divided into supply- and demand-side activities. The former may include the development a new water resource, the agreement of a bulk supply transfer, and active leakage control; the latter may include any activity undertaken to encourage customers to use less water, such as educating them on the importance of reducing water use, etc.

The activities vary in the extent to which they are likely to contribute to achieving the required measure of success, and, in turn, deliver the outcome. In light of the potential for innovation in demand-side measures, it would be important the package of incentives does not lead to a bias against a particular type of solution. It is also important to note that, for this particular

example, there may be wider objectives that should be taken into account, such as reducing consumption to 130 litres per person per day by 2030.

In the context of a separate wholesale price control, the demand-side type of activities may be of less relevance, as the wholesaler would have limited or no control over them and it would be for the retailer to promote them. In this context, the considerations around the scope for innovation may need to be tempered. This example also highlights the importance of designing schemes that take into account the interactions with other elements of the regulatory regime (eg, enforced compliance standards for drinking water).

Key characteristics

Table 4.2 presents a description of the measures of success and outcomes in this example using the characteristics identified earlier in the report: relevance, controllability, measurability, importance, and scope for innovation.

Table 4.2 Outcomes/measures of success characteristics

Characteristic	Description
Relevance of the measure	The measures of success identified are all relevant to the outcome
Controllability	<p>External factors: extreme weather events could lead to interruptions to supply or losses in water pressure. The impact of extreme weather could be measured against a baseline (eg, average performance over a regulatory control period)</p> <p>External factors: other individuals or organisations, such as customers changing their consumption pattern, could influence security of supply. Although companies are already responsible for water resource management, they may need time to adapt to more frequent prolonged dry spells. As such, non-financial incentives may be more appropriate, at least initially</p> <p>The measures of success covering water pressure and interruptions to supply seem controllable to a large extent</p>
Measurability	In general, it appears feasible to establish robust measures of success that are relatively easy to monitor. Asset resilience, however, may not be as easily measured
Importance	<p>Although a reliable supply of drinking water has a direct impact on customers, they may not attach the same importance to individual measures of success. Secure supply, for example, may be considered more important to customers than adequate water pressure—assuming that customers prefer some supply than none at all</p> <p>The measures of success are likely to be important to customers</p>
Scope for innovation	<p>There is scope for innovation in relation to ensuring a reliable supply of drinking water</p> <p>There are also many opportunities to identify alternative approaches to improve security of supply, such as bulk supply transfers and water efficiency measures. Such approaches may have a different risk profile to conventional solutions—eg, developing a new water resource, etc</p>

Source: Oxera.

Assessment

The results of the assessment of the outcome using the criteria illustrated in Figure 4.2. The key findings emerging are as follows.

- **Effectiveness.** It is assumed that security of supply is highly important to customers. As such, non-delivery in relation to security of supply would be likely to impair a company’s reputation. This is likely to increase the effectiveness of the reputational incentives.
- **Efficiency.** Companies have historically reported on these parameters. As a result, the use of reputational tools could be as efficient as, for example, financial incentives.

- **Proportionality.** In this instance, the performance of the company against measures of success is likely to be highly visible to the stakeholders. The impact on the reputation of the company is therefore likely to be proportionate to stakeholders' benefits/disbenefits arising from the company's performance.
- **Balance of risk.** In the assessment of a generic outcome, it was considered that customers would bear the risk associated with non-delivery of the outcome if a relatively less powerful incentive (eg, a reputational one) were adopted. As this outcome is easy to monitor and non-delivery by companies would be detected, customers' risk exposure seems less of an issue than for some of the other measures.
- **Practicality.** Given that the measures to report on for this outcome are similar to those required in the past, it is likely to be more practical to introduce a reputational incentive than, for example, a procedural one, which is likely to require more effort to set up and make it work.
- **Unintended consequences.** Reliability of supply is assumed to be an area where water companies are likely to have relatively more scope for innovating. Incentives need to be designed to consider this potential for innovation. In this context, reputational incentives may be suitable for reducing the risk of companies adopting a risk-adverse strategy and being less innovative than otherwise.
- **Interactions.** As reputational incentives are likely to be effective in this case, there may also be positive interactions with other components of the incentive package aimed at promoting companies' innovation (eg, TOTEX).

The assessment for this outcome, incorporating the above points, is summarised in Figure 4.5. It is worth noting that ultimately the assessment would depend on the form that the incentive would take (ie, specific design) and therefore the evaluation below is illustrative. The same comment applies to the remaining matrices in this section.

Figure 4.5 Evaluation matrix

	Enforced standards	Reputational incentives	Procedural	Discretionary financial	Pre-defined financial
Effectiveness	=	+	+	+	++
Efficiency	=	+	=	+	+
Proportionality	=	+	=	+	++
Balance of risk	=	+	+	+	++
Practicality	+	++	+	-	-
Unintended consequences	=	+	+	-	+
Interactions with overall wholesale incentives	=	+	+	+	++
Adaptability	=	=	+	++	--

Source: Oxera.

The relevance to the outcome suggests that financial and non-financial incentives may be appropriate.

Overall, in delivering this outcome, companies are likely to have substantial scope for innovation or alternative approaches, particularly in their management of water resources.

While financial incentives appear to be appropriate, reputational incentives may also be appropriate; while a combination of both may also be appropriate in this instance.

4.4.2 Assessment of incentive types for water services supply outcomes: good-quality water that is safe to drink

This outcome relates to the quality of drinking water: its look and taste and compliance with public health standards.

Figure 4.6 Good quality of water that is safe to drink

Outcome:	good-quality drinking water that is safe to drink	
	Looks and tastes good	Complies with public health standards
Measure(s) of success	Aesthetic quality and taste based on an assessment of customer satisfaction or customer complaints	% compliance with DWI standards (legislative)
Activities/Inputs	'End-of-pipe' solution: potentially more certain to achieve outcome (eg, enhanced treatment facilities) 'At-source' solution: potentially less certain to achieve outcome (eg. catchment management)	

Source: Oxera.

There are two measures of success related to the achievement of the outcome (see Table 4.3):

- aesthetic quality and taste based on an assessment of customer satisfaction or customer complaints; and
- compliance with Drinking Water Inspectorate (DWI) standards (legislative).

Table 4.3 Measure of success for good quality of water that is safe to drink

Outcome dimension	Measure of success	Rationale
Looks and tastes good	Aesthetic quality and taste based on an assessment of customer satisfaction or customer complaints	Covers the taste and the appearance of drinking water
Complies with public health standards	Percentage compliance against DWI standards	Covers the quality of drinking water as monitored by the DWI on a basis that is already established

Source: Oxera.

Figure 4.6 also summarises the activities that the company could undertake to achieve the outcome, some of which improve the quality of drinking water, including CAPEX to enhance water treatment facilities at a particular works, or solutions that involve working with landowners in a local water catchment to reduce pollution at source. These activities may

differ in their likelihood of achieving the required outcome. As such, this is an important consideration in the design of the outcome-delivery incentive mechanism.

Key characteristics

Table 4.4 describes the measures of success and outcomes in this example in the context of characteristics identified earlier in the report: relevance, controllability, measurability, importance, and scope for innovation.

Table 4.4 Outcome/measures of success characteristics

Characteristic	Description
Relevance of the measure	The measures of success are relevant to the achievement of the outcome
Controllability	Water companies are not the sole users of water courses; however, the companies are responsible for meeting the standards for drinking water quality as set out by the DWI
Measurability	It is likely that robust measures of safe drinking water can be established that are easy to monitor A measure for the look and taste of drinking water may have to be developed (eg, based on customer satisfaction). An approach may need to be agreed to ensure that this measure is assessed on a consistent basis across periods
Importance	The measures of success are important to public health and are likely to be considered important to customers. The safety of drinking water, however, is likely to be more important to be customers than its look and taste. Improvements to drinking water quality comprise a large portion of the capital investment programme and, as such, have a material impact on customers' bills
Scope for innovation	The scope for innovation may be limited owing to risks to public health. However, there may be scope for adopting alternative approaches on an incremental basis when the risk of non-compliance can be managed (eg, catchment management)

Source: Oxera.

Assessment

The assessment of water quality is similar to that reported for water supply. The main differences are in the following areas:

- **effectiveness**—the impact on reputation may be greater if a company fails to provide safe drinking water to customers, as opposed to failing to provide a secure supply. Reputational incentives may therefore be more effective for this outcome;
- **unintended consequences**—one of the objectives of the outcome approach is to promote innovation. In this context, it is of note that because companies are likely to be risk-adverse and given the potentially large reputational impact of a water safety issue, a reputational incentive may lead to a more conservative approach in the solutions adopted, which may result in less innovation.

The assessment for this outcome, incorporating the above points, is summarised in Figure 4.7.

Figure 4.7 Evaluation matrix

	Enforced standards	Reputational incentives	Procedural	Discretionary financial	Pre-defined financial
Effectiveness	=	++	+	+	++
Efficiency	=	+	=	+	+
Proportionality	=	+	=	+	++
Balance of risk	=	+	+	+	++
Practicality	+	++	+	-	-
Unintended consequences	=	=	=	-	+
Interactions with overall wholesale incentives	=	+	+	+	++
Adaptability	=	=	+	++	--

Source: Oxera.

The relevance of the measures of success to the outcome suggests that both financial and non-financial incentives may be appropriate.

Overall, with regard to its characteristics, the measures of success for this particular outcome are likely to be relevant, controllable, important and visible. The scope for innovation may be more limited given the potential risk to public health. As such, a package of financial and non-financial incentives may be more appropriate. As will be specified more extensively in section 5, a one-sided incentive (in this case, assumed to be a penalty-only) may also be more appropriate given that standards, as set out in legislation, will either be met or not met.

Incentives may need to be introduced just to address the quality aspect of the outcome, as DWI may already be enforcing standards and penalties; this is assuming that the incentives are introduced only incrementally to the existing ones (see section 5 for a more detailed explanation of this point).

4.4.3 Assessment of incentive types for sewerage services supply outcomes (sewer flooding)

In this example the outcome is defined as being minimisation of the risk of sewer flooding (see Figure 4.8 below).

Figure 4.8 Sewerage service supply: sewer flooding

Outcome:	customers do not experience sewer flooding in their homes,		gardens or local communities
	Minimise internal sewer flooding incidents		Minimise external sewer flooding incidents
Measures of success	Number of properties at risk of internal sewer flooding	Number of properties affected by internal sewer flooding	Number of external sewer flooding incidents
Activities	Enhancement solutions (eg, increasing sewer capacity)	Maintenance activities (eg, sewer replacement, rehabilitation)	Operating activities, cleaning sewers, etc
	Sustainable drainage systems (SUDS)	Customer side-solutions (eg, customer education about oils, fats etc)	

Source: Oxera.

In this example, the outcome is that customers do not experience sewer flooding in their homes, gardens and local communities. It covers two broad areas: minimising internal sewer flooding incidents and minimising external sewer flooding incidents. There are three measures of success related to achieving this outcome (see Table 4.5):

- minimising internal sewer flooding incidents: the number of properties at risk of sewer flooding; and the number of internal sewer flooding incidents;
- minimising external sewer flooding incidents: the number of external sewer flooding incidents.

Table 4.5 Measures of success for sewer flooding

Outcome dimension	Measure of success	Rationale
Minimise internal sewer flooding incidents	Number of properties at risk of sewer flooding	Covers properties that are at risk of sewer flooding
	Number of properties affected by internal sewer flooding	Covers properties that have been affected by sewer flooding during the period
Minimise external sewer flooding incidents	Number of external sewer flooding incidents	Covers all external sewer flooding incidents from the sewerage network

Source: Oxera.

Figure 4.8 summarises the activities that can be undertaken by the company to achieve the outcome. Many types of activity could be undertaken, including capital enhancements to remove properties from the register of properties at risk, operating solutions (cleansing, etc) and alternative solutions, such as educating customers on what they should and should not put down the drain (eg, oil, fat, etc). Also included are systems being developed to reduce the volume of rainwater draining into the sewer in order to increase capacity—ie, rainwater-harvesting. Such activities may have an impact on the design of incentives.

Key characteristics

Table 4.6 below describes the measures of success and outcomes in this example in the context of characteristics identified earlier in the report: relevance, controllability, measurability, importance, and scope for innovation.

Table 4.6 Outcome/measures of success characteristics

Characteristic	Description
Relevance of the measure	The outcome is narrow in scope. As such, all measures of success are relevant to the outcome
Controllability	Heavy rainfall could increase the risk of sewer flooding incidents but, to some extent, could be taken into consideration through the basis of measurement (eg, average performance over a regulatory control period, etc)
Measurability	In general, it appears feasible to establish robust measures of success that are relatively easy to monitor
Importance	As sewer flooding directly affects the service customers receive, it is likely to be considered important. Arguably, external flooding may be less important than internal flooding, which affects customers directly The measures of success and, in turn, the outcome is likely to be important to customers
Scope for innovation	There could be scope for innovation to reduce sewer flooding incidents without putting disproportionate risk on customers and the environment There are also opportunities to identify alternative approaches to increase the capacity of the sewerage network—eg, capturing rainfall for non-consumption usage and reducing blockages by educating customers on the substances that should and should not be put into the sewer, etc. Such approaches may have a different risk profile to conventional solutions

Source: Oxera.

Assessment

The relevance of the measures of success to the outcome suggests financial and non-financial incentives may be appropriate.

The key characteristics of this outcome entail that its assessment is akin to the generic outcome. The only differentiator may be the effectiveness of reputational incentives in this case. As the effects of sewer flooding have great resonance with customers, a reputational tool may be effective. However, since the impact on customers and their awareness about sewer flooding are likely to be localised, any reputational implications may be more limited than, for example, those arising from water safety issue. It may therefore be appropriate to use a financial incentive and potentially to strengthen its power by using a reputational incentive as well.

The assessment for this outcome, incorporating the above points, is summarised in Figure 4.9.

Figure 4.9 Evaluation matrix

	Enforced standards	Reputational incentives	Procedural	Discretionary financial	Pre-defined financial
Effectiveness	=	+	+	+	++
Efficiency	=	=	=	+	+
Proportionality	=	=	=	+	++
Balance of risk	=	-	+	+	++
Practicality	+	+	++	-	-
Unintended consequences	=	-	+	-	+
Interactions with overall wholesale incentives	=	=	+	+	++
Adaptability	-	=	+	++	--

Overall, with regard to its characteristics, the measures of success for this particular outcome are likely to be controllable, important, and measurable. There also appears to be some scope for innovation. As such, it may be appropriate to introduce a package of two-sided financial and non-financial incentives. However, financial rewards based on outperformance could only be made if customers are willing to pay for it.

4.4.4 Assessment of incentive types for sewerage services: clean aquatic environments

In this case an outcome could be defined as being clean aquatic environments (see Figure 4.10).

Figure 4.10 Clean aquatic environments

Outcome:	contributing to a clean water environment that meets the standards that customers and society value			
	Delivering base services well		Improving services	
Measure(s) of success	Discharge compliance (percentage)	Number of pollution incidents weighted towards high-impact incidents	Progress on delivering legislative improvements	Progress on delivering other improvements
Input/ activities	Maintenance/ operating activities		Enhancement to achieve the company's contribution to the standard	Enhancement
	Working with other stakeholders; land management agreements, etc		Working with other stakeholders to achieve an improvement	Maintenance/ operating activities

Source: Oxera.

The outcome relates to contributing to a clean water environment that meets the standards that customers and society value, and covers two areas: delivering base services well, and improving services. The outcome has been separated these two areas as they may warrant applying different types of incentive. There are four measures of success related to the achievement of the outcome (see Table 4.7):

- delivering base services well: discharge compliance (percentage); number of pollution incidents weighted towards high-impact incidents;
- improving services: progress on delivering legislative improvements; and progress on delivering desirable improvements.

Table 4.7 Sewerage service clean aquatic environments

Outcome dimension	Measure of success	Rationale
Delivering base services well	Discharge compliance	Covers wastewater discharges as monitored by the Environment Agency
	Number of pollution incidents	Covers pollution incidents into water bodies as monitored by the Environment Agency
Improving services	Progress on delivering legislative improvements	Covers improvements that need to be made to comply with standards set out in legislation
	Progress on delivering other improvements	Covers all other improvements. Desirable improvements that are valued by customers or society

Source: Oxera.

Key characteristics

Table 4.8 below describes the measures of success and outcomes in this example in the context of characteristics identified earlier in the report: relevance, controllability, measurability, importance, and scope for innovation.

Table 4.8 Outcome/measures of success characteristics

Characteristic	Description
Relevance of the measure	As this outcome is broader than other possible outcomes, it may be more challenging to establish comprehensive measures of success. Therefore, the measures of success may be more diverse than in other areas, making the judgement about the overall delivery more complex and potentially more subjective (eg, how to bring together many different indicators)
Controllability	The activities of other stakeholders (eg, passenger ferries, fisheries, oil refineries and heavy industry) have an impact on the water environment. However, all measures of success for delivering base services relate to the water company's contribution to a clean water environment In relation to improving services, in a small number of cases, other organisations may have part responsibility for the achievement of a standard set out in legislation (eg, the Water Framework Directive). In those cases a company should have an incentive to achieve its contribution
Measurability	In most instances, it appears feasible to establish robust measures of success that are relatively easy to monitor. However, as it may be difficult to establish high-level measures of success, the assessment may need to rely on different and not necessarily consistent metrics
Importance	Compared with the other outcomes, a cleaner water environment may have less of a direct impact on customers. Nevertheless, it may be one of the outcomes valued most by society as a whole Investment in environmental improvements comprises a significant portion of the capital investment programme The measures of success are likely to be important to society.
Scope for innovation	There may be some opportunities for innovation without putting disproportionate risk on customers or the environment

Source: Oxera.

Assessment

The key features illustrated in the above table are likely to have a bearing on the following criteria.

- **Effectiveness.** The use of financial incentives may have limited impact on the companies' behaviour, to the extent that companies and management feel that the probability of being penalised (or rewarded) is somewhat out of their control. As such companies may feel they would be better off just not acting, retaining the gains and 'hoping for the best'. However, a financial penalty may be still effective if, for example, it is applied to those companies that cannot prove their 'good faith' (ie, that they tried to achieve the outcome/milestone but that adverse events or concurrent behaviour by others has resulted in failing to achieve the outcome/milestone).
- **Efficiency.** Because of the lack of controllability, risk-averse companies might choose more expensive solutions (eg, CAPEX solutions) even when other (cheaper) solutions may be as effective, which could lead to an overall inefficient delivery.
- **Proportionality.** It may be difficult to structure the incentive financial implications to be proportionate to each aspect of the outcome (eg, the value of a transitory impact versus a cumulative and permanent degradation of aquatic environments, impacts on coastal versus inland waters, etc). In this context, financial penalties/rewards may be set disproportionately high or low.
- **Balance of risk.** A financial penalty in this case may shift too much risk on to the companies, which are not necessarily the most appropriate party to bear it.
- **Practicality.** For similar reasons to those illustrated for proportionality (ie, potentially a highly complex financial penalties/rewards scheme design), the financial incentives in this particular instance may be less practical than assessed for a generic outcome.

The assessment for this outcome, incorporating the above points, is summarised in Figure 4.11.

Figure 4.11 Evaluation matrix

	Enforced standards	Reputational incentives	Procedural	Discretionary financial	Pre-defined financial
Effectiveness	=	=	+	+	+
Efficiency	=	=	=	-	-
Proportionality	=	=	=	-	-
Balance of risk	=	-	+	-	-
Practicality	+	+	++	--	--
Unintended consequences	=	-	+	-	+
Interactions with overall wholesale incentives	=	=	+	+	++
Adaptability	-	=	+	++	--

Source: Oxera.

Overall, with regard to its characteristics, the measures of success for this particular outcome are likely to be controllable (in part), important to wider society, and measurable (again in part). There may also be some scope for innovation. As such, financial incentives seem less appropriate in this instance.

It is worth noting a few points here specific to this hypothetical outcome. First, although non-financial incentives may be more appropriate in this case, financial incentives may remain appropriate for non-compliance in relation to base services or the non-delivery of a standard set out in legislation (ie, a one-sided financial incentive may be relevant). Second, as noted earlier, wider use of one-sided financial incentives may lead to risk aversion by the companies, and a bias against innovation and alternative-based solutions, especially if highly powered (ie, a high penalty) incentives are adopted. Two-sided financial incentives could be used, provided that customers are willing to pay for outperformance.

4.5 High-level assessment of options: summary

The assessment of different incentive tools against a set of criteria for a generic outcome would indicate that an incentive of a financial nature may be the most appropriate tool that Ofwat may want to consider introducing when deciding on or providing guidance on the outcome delivery incentive.

The analysis of the four hypothetical outcomes shows that the above conclusion may vary depending on the type of outcome being considered. For example, the clean aquatic environments outcome case shows that factors such as lack of controllability by the company may make delivery riskier for the company, while the potential use of diverse measures (ie, measurement issues) could make the assessment of the company performance somewhat arbitrary. Therefore, in these instances, reputational incentives or a combination of reputational and financial incentives may be more suitable than financial-only incentives if Ofwat's aim is to avoid loading the company with regulatory risk while incentivising it to deliver the outcomes.

As specified already, the assessment of the outcome examples is not aimed at providing recommendations as to the types of incentive that should be used by companies for similar outcomes. The thought process considered above, however, may help to identify which characteristics of the outcome are likely to have an impact on the types of incentive and their power.

In particular, it is important to take into account how the characteristics considered throughout the section (relevance for the outcome, controllability, measurability, importance, and scope for innovation) affect the different incentive options and, accordingly, the key considerations for the decision as to the use or otherwise of a financial incentive can be set out (see Table 4.9).

Table 4.9 Key considerations to decide on the use of a financial penalty

Characteristic	Considerations
Relevance of the measure	<p>A measure of success closely related to the outcome (ie, that covers the outcome in its entirety) could be used directly for the assessment of the outcome/milestone delivery or otherwise (ie, the company is judged only on that measure).</p> <p>Where individual measures are not fully representative of the outcome, more than one measure is to be used. In this case, the assessment of delivery or otherwise may become more discretionary and less transparent</p> <p>A financial incentive applied to a company through a non-transparent process may increase regulatory risk. Therefore, the use of a financial penalty may be preferred in outcomes where a higher-level measure of success is available</p> <p>Financial incentives may be used provided that the process to judge progress on the delivery is set out clearly—ie, through rules to determine whether the company has passed or failed are transparent</p>
Controllability	<p>The level of controllability of the outcome and related measure of success by the company is an important consideration in deciding whether a financial incentive is suitable for that outcome. The less control the company has over that outcome, the less advisable it is to introduce a financial incentive</p>
Measurability	<p>The extent to which the measure of success can be assessed on the basis of a robust and transparent metric that is easy to monitor is important because it reduces any scope for ambiguity and discretion, and thus any risk relating to measurement. The more measurable the outcome/measure of success, the more suitable this is to the use of financial incentives</p>
Importance	<p>If an outcome is highly valued by customers and society, a highly powered incentive, such as a financial incentive, may be desirable</p>

Source: Oxera.

When considering whether a financial incentive may be used for a particular outcome, it might be the case that considerations in relation to some characteristics support the use of financial incentives while in relation to others they suggest otherwise. For example, an outcome that may be robustly measured and important from a customer perspective could point to the use of a financial incentive; but if that outcome is only partially under the control of the company, then the use of a financial incentive would not be advisable.

In these instances where a ‘clear verdict’ cannot be reached, it may be appropriate to consider if and how the use of a specific design may increase the suitability of a financial incentive (eg, reducing companies’ risk exposure by using deadbands). The specific incentive design is explored in the next section.²⁸

²⁸ It is worth noting that some of the characteristics (eg, controllability) may also be used in deciding on the strength of a, for example, financial incentive (eg, high or low penalty)

5 Incentive design

This section first focuses on the issues around incentive design as examined in the academic literature, and then outlines the advantages and disadvantages of different options as applicable to the context of outcome-delivery incentive design.

5.1 Potential incentive strategies

5.1.1 The scope of regulatory incentive strategies

A firm's decision processes involve finding a balance between the net benefits associated with the different actions that are available to it at a particular point in time, and in the context of the particular set of circumstances that it faces. This notion is captured by incentive compatibility, which is a characteristic of incentive mechanisms that essentially means that firms (or other entities) will act rationally by choosing courses of action that produce the greatest net benefit, given the actions available to it.

The incentive-compatibility *constraint* reflects how the set of choices that a firm faces may be restricted or expanded to meet certain objectives. In particular, incentive regulation attempts to act on the incentive-compatibility constraint of firms by inducing them to behave in a way that they would otherwise not have done. This requires a connection between the efforts of the firm being rewarded and the desired outcome, which, in turn, depends on the 'transmission mechanism'—the mechanism by which incentives produce outcomes.²⁹

Ofwat's high-level characterisation of a transmission mechanism in the 2010 discussion paper describes a sequential process that reflects how companies take decisions in responding to an incentive, and captures the essence of how an incentive affects the incentive-compatibility constraint in different ways.

A number of barriers may hamper the realisation of the optimal outcome, and these give rise to the different types of incentive design discussed in section 6. Examples of these possible barriers include the following.

- **Controllability: weak correlation between effort and outputs**—where the relationship between the company's effort and its outputs is weak, the company has little control over the final outcome and the incentive scheme may have little or no impact.
- **Perverse targets: imperfect incentive scheme design**—an incentive scheme may encourage a company to focus its efforts on actions that contribute only marginally to the overall welfare of the society. For example, if one outcome is desired (eg, clean drinking water) but a less desirable outcome is produced as a consequence (eg, the water company simply complies with targets for clean water in a mechanistic way but does not respond to the 'big picture' of what is being incentivised). At the same time the company's resources may be diverted from areas where they could have been used more efficiently.
- **Participation constraint**—a company may choose not to participate in the market if it perceives the risks imposed by the incentive schemes to be inadequately remunerated.

In addition to these high-level barriers, other factors could be both enablers and barriers. Ofwat's 2010 report notes the importance of company culture, which reflects factors such as

²⁹ Ofwat (2010), 'The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper', p. 13.

investor priorities, corporate experience and memory, accounting policies, and wider ownership group influence. This could be one reason to introduce some incentives on a company-specific basis, a topic considered below in the context of process issues.

Additionally, the interaction between incentive mechanisms could be both an enabler and a barrier to the operation of a transmission mechanism. This possibility is relevant for a number of reasons, and is considered in various places in the text below.

5.1.2 Elements of incentive design

Given regulatory objectives (incentivising outcomes) and general and specific barriers or constraints that are relevant to the decision processes of regulated water companies, this section enumerates different types of incentive design.

High- or low-powered incentives mechanisms

The 'power' of an incentive for a commercial, profit-oriented organisation is defined as the link, provided by the incentive, between the firm's prices and its cost or profit performance. More specifically, the power of incentive schemes is proportionate to the amount of profit that the company is allowed to retain. Schemes that leave the firm a large share of profits as a result of behaving in accordance with societal preferences are known as high-powered incentive schemes, whereas those that leave the agent a smaller share are known as low-powered incentive schemes.

Absent other considerations, such as the reputational effect of poor performance under a particular incentive mechanism, high-powered incentives will be more likely to deliver desired outcomes. However, low-powered incentives are not irrelevant, although their impact needs to be considered in the context of wider regulatory objectives and alongside other, potentially complementary, interventions.

Ex ante and ex post incentives

Ex ante incentives involve defining behaviour in advance, and typically require ongoing monitoring to assess compliance. Ex post incentives involve an assessment of performance after the event.

The difference between the two approaches can be significant, with ex ante approaches potentially being more relevant where ongoing monitoring has intrinsic utility to the regulator, whereas ex post incentives could be more relevant where the outcome being incentivised is uncertain and where the means of identifying progress toward that incentive is not easy to establish at the time the incentive is introduced.

The choice between the ex ante and ex post incentives is also influenced by other considerations, including the regulatory burden associated with each.

Comprehensive or partial incentive mechanism

An incentive mechanism could be described as 'comprehensive' if it consistently reflects all cost and quality relationships at a point in time and over time. As Joskow³⁰ notes:

as a practical matter this [the use of a comprehensive incentive mechanism] often places very challenging information and implementation burdens on the regulator. Partial mechanisms or a portfolio of only loosely harmonized mechanisms are often used by regulators. Operating and capital cost norms and targets are typically developed separately and the effective power of the incentive scheme applicable to operating and capital costs may vary between them. Separate incentive mechanisms may be applied to measures of quality than to measures of total operating and capital costs. This reality represents perhaps the most significant variation between received incentive regulation theory and incentive regulation in practice.

³⁰ Joskow, P. (2006), 'Incentive regulation in theory and practice: electricity distribution and transmission networks'. Harvard Papers, available at http://www.hks.harvard.edu/hepg/Papers/Joskow_Incentive_2006.pdf (accessed on August 10 2012 at 17.30)

The practical challenges of designing a comprehensive incentive mechanism of the type referred to by Joskow are probably intractable due to asymmetry of information between the regulator and the regulated, but the cautionary note sounded is important to bear in mind.

Individual/multiple benchmarks

A key problem that regulators face when designing an incentive scheme is determining what to measure. As such, multiple benchmarks may be required to monitor the behaviour of the firm, and can be used to monitor a single outcome. However, these particular benchmarks can be treated as individual measures of performance and rewarded individually, or can be combined to a single measure of performance and rewarded in aggregate. These different approaches have different impacts on the incentives to which the firm is exposed. For instance, there is a degree of regulatory risk if it is announced that a number of outcome benchmarks or metrics will be reviewed, but particular focus is in fact placed on a limited subset of these metrics, such as progress towards the completion of a building. In practice, it may be hard for regulators to commit to 'standing by' if some benchmarks are progressing in the right direction, but others are not. The specific issues arising with regard to multiple benchmarks, and, more generally, what to use as the benchmark to trigger an incentive, are considered in section 6 below.

Combination of elements

A combination of elements of the foregoing could be deployed. For example, a partial measure could be associated with multiple benchmarks, while a comprehensive mechanism could be high- or low-powered.

5.2 Specifics of design

This section considers some specific details of the design of incentive mechanisms relevant to outcomes in the water industry. It is helpful to begin with a cautionary remark from Joskow (2006) concerning the feasibility of identifying the 'right' incentive in the context of the wider concerns of regulation:

integrating these incentive mechanisms into a package that gives the correct incentives on all relevant margins remains a considerable challenge for incentive regulation in practice.

This section considers some specific practical considerations that are relevant to the design of incentives, as follows:

- individual or multiple benchmark;
- targeted incentives;
- the level of financial incentives;
- the symmetry of incentives;
- the balance of risk;
- trade-offs and risk;
- practicality.

5.2.1 Individual and multiple benchmarks: what are incentives linked to?

Incentives could be applied at any of the three following levels.

- **At the level of overall performance** (encapsulating all outcomes)—monitoring of success would involve a holistic, high-level assessment of the company's overall performance, with penalties only being applied if the company were deemed not to be performing satisfactorily. This approach could allow for a trade-off between outcomes, such that if a company were to under-deliver against one outcome but over-deliver against another, and its overall performance were deemed to be satisfactory, it might not be penalised.

- In theory, the need for this type of holistic framework is obviated by the application of two-sided incentive mechanisms if the marginal social value and willingness to pay for components of the overall performance is understood, and the regulator can measure efficient incremental costs. If either or both of these conditions fail, trade-offs between components can become complements or substitutes to pure two-sided incentives.
- **At the level of each individual outcome**—assessment would be made at the level of each individual outcome, with penalties applied if the company were deemed not to be performing against any one of the outcomes. The penalty would differ according to which outcome was the subject of under-delivery (or a reward would be granted in the case of over-delivery). This approach could capture a trade-off between the measures of the outcome being assessed made at the level of financial incentive that could be attached to each, such that if a company under-delivered against one measure, but over-delivered against other measures for the same outcome, it might not be penalised.
- **At the level of each individual measure contributing to assessment of delivery of the given outcome**—this would represent a more detailed analysis of performance, looking at each of the company's outcome measures and determining their success or failure. The company would face penalties for failing to meet any one of the measures.

Although setting incentives at the level of each individual measure would lend itself to the most detailed analysis of how the company is performing in relation to achieving each outcome (highlighting where it is doing well and where it needs to improve), there are two main reasons why such an approach might not be appropriate: it may be unlikely to promote simplicity and transparency in the regime, and it could impose a heavy regulatory burden. For example, if all 21 companies were to target ten outcomes, each with three associated performance measures, Ofwat would have to assess delivery against 630 measures. If the outcomes are bespoke, these measures might not be standardised across companies, thus requiring Ofwat to interpret a large amount of data periodically. This level of regulatory involvement would be contrary to the FPL policy of avoiding this type and intensity of intrusion where possible.

Such an approach might not be consistent with the procedural proportionality element of the regulator's guidelines on the characteristics of a good incentive.³¹ Moreover, the narrower the focus of the incentive, the more likely it is that companies will focus on measures—potentially targeting outputs and inputs—rather than outcomes. This could lead to the regulatory regime essentially resembling output regulation, reducing the ability of companies, customers and Ofwat itself from realising the intended benefits of outcome-based regulation. This could particularly be the case for innovation, if incentives are attached to narrow performance measures that do not account for the implementation of innovative processes.

These disadvantages suggest that it might be more appropriate to set incentives at a broader level. In practice, the incentives could work across the three levels outlined above. For example, in the first instance incentives could be linked to an assessment of overall performance but with scope for additional incentives attached to more narrow performance measures in high-priority areas, such as security of supply and customer satisfaction. This would grant the companies some flexibility in performance across outcomes as a whole, while ensuring that they are still penalised for under-delivery in those areas identified by customers as high priority. This would depend on these high-priority areas being subject to an appropriate degree of controllability, and in practice both priority and controllability 'hurdles' would need to be passed.

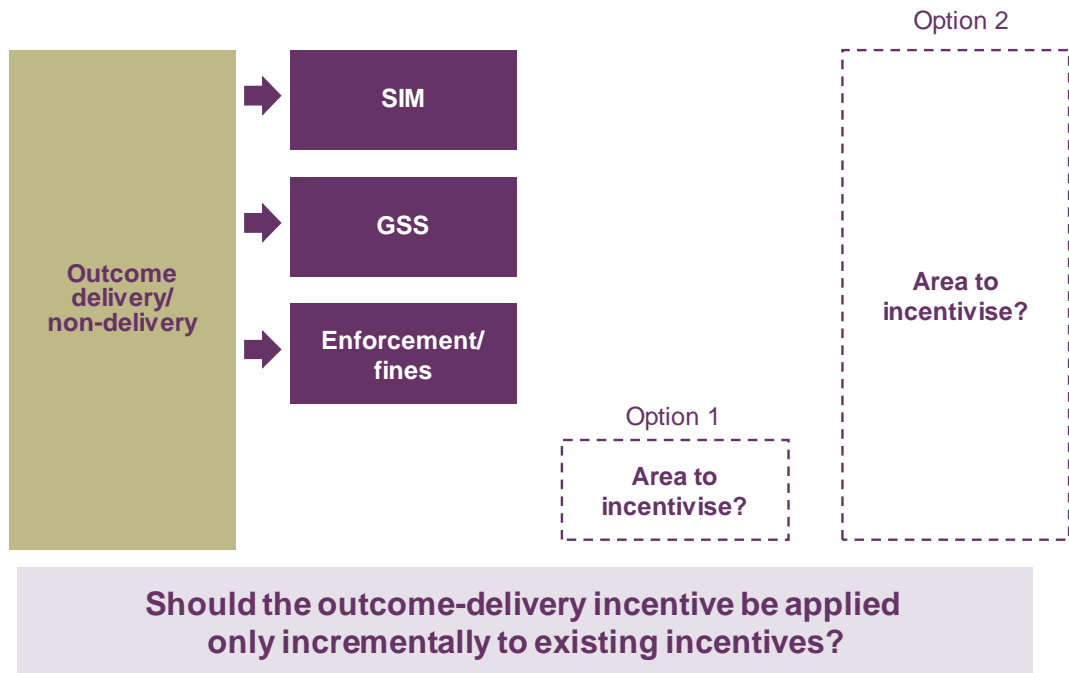
³¹ Ofwat (2010), 'The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper', October, p. 11.

5.2.2 Targeted incentives

A further consideration in the design of outcome delivery incentives is the interaction of the incentives with the wider incentive package. In particular, it might be necessary to establish whether existing regulatory measures and inherent commercial incentives are sufficient (ie, incentive-compatible) to incentivise the delivery of outcomes. Insofar as there are already regulatory incentives in place that reward/punish the company for delivering outcomes, and these are sufficiently strong to incentivise it to do so, additional incentives may not be required.³² For example, companies are already likely to face inherent (underlying) reputational incentives in relation to security of supply and safety, since these are highly visible and widely reported on. As a consequence, additional incentives may not be necessary.

Outcome-delivery incentives could thus be applied in an incremental manner, building on the existing incentives regime and taking account of the underlying (commercial) incentives. Incentives could then be applied only to those areas/outcomes where incentives do not already exist. This is represented by ‘option 1’ in Figure 5.1. Alternatively, incentives could be introduced that would make outcome delivery incentive-compatible, even in isolation from any other incentives. These incentives could be implemented at the level of the outcomes, rather than applied to individual measures of success, as the existing incentive mechanisms are. This is represented by ‘option 2’ in Figure 5.1.

Figure 5.1 Targeted incentives (an example)



Source: Oxera.

In theory, adding the incremental option could be the most appropriate approach as it helps to avoid double-jeopardy (of a company being penalised twice for the same offence) and prevent unnecessary measures from being introduced. It is noted that Ofwat has committed to avoiding precisely these measures.

³² Or, at least, they may not need to be strong. As Ofwat has noted “the strength of the regulator’s incentive required to change company behaviour will depend on the underlying incentive that the company faces. The less our regulatory incentives align with the underlying incentive the stronger our incentive will need to be to produce the behaviour change we seek.” Thus, if the underlying commercial incentives to deliver an outcome are strong, any regulatory incentives need not be.” See Ofwat (2010), ‘The role and design of incentives for regulating monopoly water and sewerage services in England and Wales – a discussion paper’, October, p. 34.

The choice between incremental improvements and wholesale replacement of existing incentives needs to be guided by considerations of practicality, including the effect that major changes will have on the effectiveness of the incentives involved, and on the ability of companies to anticipate, plan and adapt to changes. A decision to replace the existing incentive mechanisms with an overarching incentive regime based on an overall assessment of performance would therefore require close engagement with the industry.

5.2.3 The level of financial incentives

The level of the overall financial value of an incentive has an obvious and direct impact on the extent to which an incentive may be expected to deliver its desired effect. There are a number of areas of regulatory practice and thinking that are relevant to this type of consideration.

One useful strand of thinking from which to proceed is to consider the literature on regulatory sanctions. Sanctions are typically understood to be a form of penalty or enforcement. While it is recognised that the notion of financial incentives being considered here is broader and more refined than that of a threat of a penalty, regulatory debates on the appropriate form of sanction contain insights that are relevant for the present context. For example, the Macrory principles,³³ which influenced the Regulatory Enforcement and Sanctions Act 2008, stated that a sanction should:

- aim to change the behaviour of the offender;
- aim to eliminate any financial gain or benefit from non-compliance;
- be responsive and consider what is appropriate for the particular offender and the regulatory issue;
- be proportionate to the nature of the offence and the harm caused;
- aim to repair the harm caused by regulatory non-compliance where appropriate;
- aim to deter future non-compliance.

Incentives should also be designed to encourage compliance (or the desired outcome) rather than penalising non-compliance. Macrory also recommended that regulators have a published policy for sanctions, and report annually on their activities.

Reviewing this list for points salient to the issue of incentives in general (as opposed to sanctions in particular), one could conclude that incentives need to be proportionate in their incidence, and transparent in their motivation and application. It is also important to note that trade-offs between the principles above are inherent—the financial gain from non-compliance will generally not be equal to willingness to pay, or the overall marginal social harm. Hence, the principles can define guidance as to the upper and lower bounds of sanctions, but trade-offs among the principles will generally be needed to arrive at a particular sanction level.

What does the concept of proportionality mean in the context of establishing a particular financial value for an incentive? A number of factors are relevant. First, in most conceivable circumstances, a (negative) financial incentive should not mean that the company concerned is unable to continue to operate because of lack of funds, except in the exceptionally rare circumstances where an offence is so serious as to merit contemplation of bankruptcy. An example would be a major, ongoing and avoidable failure to supply drinking water. In addition to financing concerns associated with a large one-off penalty, repeated smaller fines may have a significant ongoing cumulative effect such that credit ratios are breached, for example.

The second factor is that a positive financial incentive should not, in some circumstances, deliver benefits to the company that exceed the willingness to pay of customers for the change or outcome associated with the incentive. For example, a financial reward for

³³ Macrory, R. (2006), 'Regulatory justice: sanctioning in a post-Hampton world', May, Department of Business, Innovation and Skills.

delivering a quality improvement should not impose costs on customers in excess of their willingness (or potentially capacity) to pay, since this reward would in effect be a transfer to shareholders for the amount in excess of willingness to pay. However, where resource costs are being moved from one part of the economy to another (and so there are no aggregate welfare implications), such a reward could still be supportable.

That said, in a number of relevant contexts such a consideration would not apply. An example considered below is where customers may have a low willingness to pay for climate mitigation measures, but nevertheless Parliament has decided that the costs of water sector-related mitigation should be borne by water users. These two cases may be considered as representing a type of upper and lower bound on the value of financial penalty/reward, and within this range other considerations will be relevant.

Ultimately, the level of financial incentive will need to operate on a company's incentive-compatibility constraint. This could suggest that the financial consequences of under-spending or not delivering should, as a minimum, be NPV-neutral. However, reduced outcomes, but with customer bills that still reflect privately discounted NPV calculations, will still reduce welfare at Green Book discount rates, in which case a lower minimum could be relevant.³⁴

This could be regarded as the minimum amount that a company would need to pay as a penalty in order to remove any incentive for non-compliance, since the size of the penalty would offset the gain (in an NPV sense). A larger penalty would normally be appropriate, and set at the level of the value loss by customers or society because of the delay in delivering, or failure to deliver, the outcome. A larger penalty may also be justified as a deterrent to other companies.

This type of calculation depends on the availability of an estimate of the economic value of the loss or delay of the associated outcome, which requires the loss in consumer welfare to be calculated.

If a symmetric incentive³⁵ were adopted, some proportion of the value (to customers) of the higher level of the outcome delivered (over and above the level agreed when setting prices) could be paid to the company to reflect the incremental benefits gained by customers.

The following considerations need to be taken into account.³⁶

- The basic approach to setting penalties described in the previous paragraphs does not consider whether the non-delivery of the outcome is due to 'bad luck' or the company's lack of effort. Absent other protections, the less the company can control the outcome, the greater the likelihood that it will be punished. However, the converse situation also applies: companies might benefit from 'good luck'. In practice, attempts to distinguish between legitimate mitigating circumstances and companies of different types are complicated by the presence of asymmetric information, which can lead to the payment of information rents.
- Related to the above, companies have suggested that any incentive mechanism would need to be sophisticated enough to distinguish between a situation in which the company has spent money 'in good faith' to achieve an outcome and one where it has spent money inadequately or inappropriately. This is important if an innovative approach is to be encouraged (ie, in accepting a level of failure).

³⁴ See the discussion in Joint Regulators Group (2012), 'Discounting for CBAs involving private investment, but public benefit', July.

³⁵ Symmetry is discussed in the next sub-section.

³⁶ Some of these points reflect discussions with companies undertaken for this study.

- It has been suggested by companies that in a case where financial penalty were to be adopted, it might be appropriate to use a cap-and-collar approach to reduce the company's financial exposure. The same logic may be relevant to capping the upside (ie, not to reward shareholders for sums significantly in excess of the cost of capital).

The transmission mechanism associated with the financial value of an incentive would therefore be fairly direct, as companies can readily relate the size of the value concerned to other relevant metrics (eg, cash flow), and assess the impact of the fine on strategic objectives (eg, the desired size of dividend distributions).

The relationship between the level or value of a financial incentive and characteristics of different outcomes is arguably simpler than for other specific elements of incentive design, since the trade-offs involved relate simply to the impact of the financial penalty or reward on the affected organisation, and the anticipated marginal benefit to the outcome rendered by the incentive.

In summary, the advantages and disadvantages of larger or smaller financial incentives, whether intended to reward or penalise companies, depend on the broader considerations relevant to incentive design. For example, a large or small penalty will have limited impact if the outcome is not meaningfully under the control of the company. Similarly, an ambitious 'all or nothing' financial incentive may not incentivise desired behaviour (and therefore desired outcomes) if companies do not feel they can realistically undertake the effort required to meet the target.

In practice, given that the level of a financial incentive is readily observed, it should be possible for Ofwat to assess how companies perform in attempting to meet any associated target. This would give an opportunity for a certain amount of observational learning to emerge concerning the appropriateness of different financial values.

The foregoing discussion has established a range of bounds on the appropriate level of penalty or reward that is connected with an incentive mechanism. These range from incremental cost and willingness to pay, to values beyond willingness to pay where there is a public interest consideration not adequately captured by willingness to pay (eg, in relation to climate change mitigation). Financial incentives can also be used to support and complement other measures.

This leaves the question of whether the failure risk of too high a financial incentive is greater than the failure risk of too low incentives. A few considerations are relevant to this question. For example, in setting parameters such as the cost of capital in price determinations, regulators typically allow a degree of 'headroom'. This is done for a number of reasons; for example, a methodology that involves taking the average value of market interest rates in a period of falling yields will result in a gap between that average and the rates applying at the time of the decision is made. These include to reflect a degree of regulatory uncertainty as to what the precise value should be, and to reflect the asymmetric consequences of getting the WACC figure wrong (ie, bankruptcy versus slightly higher returns).

Similar considerations could be used to define the failure risk associated with high- and low-powered incentives. In some cases, the cost of getting this decision wrong could have systemic and long-lasting consequences that go beyond those intended; in other cases the consequences will be less severe.

A penalty, or sequence of penalties, that obliges a company to earn returns below its cost of capital for a prolonged period could have serious consequences. This would suggest that the failure cost of too harsh a penalty would be greater than the failure cost of too low a penalty. Similar conclusions apply in the case of a reward.

As well as depending on the nature of the outcome, financial values can reflect the nature of the companies concerned. For example, a large financial penalty may have the same effect

on a company with modest gearing as a modest penalty on a highly geared company. Other company risk characteristics would also be relevant in establishing the consequences of too high or too low an incentive, such as the risks associated with a large CAPEX programme.

The precise financial value could also be influenced by the extent to which Ofwat perceives there to be a risk of gaming. Even well-designed outcome incentive schemes may face this risk, which could be mitigated by attempting to structure reward and penalties so as to avoid over-rewarding mechanisms vulnerable to gaming.

The financial value of an incentive in the case of a breach could also reflect any remedial steps undertaken by the company since the time when the failure became apparent, and could also reflect the degree of cooperation with Ofwat in developing a remedy.

Financial values could also reflect the practice in public–private partnership (PPP) contracts to introduce a multiplier penalty for repeated bad performance. Thus, if a target to support a particular outcome is not met within a certain period (eg three months), the penalty for failure to comply for the next three months could be 1.5 times the penalty in the first three months. Another feature of PPP contracts is that weightings may be assigned to performance in different areas, and the monetary reward or penalty arising could be the weighted sum of performance, with the weights potentially subject to change over time depending on performance and overall priorities. This could be useful in mitigating the risk associated with too high or too low a penalty in different areas.

5.2.4 Symmetry

The symmetry of an incentive refers to whether companies should be rewarded for good performance as well as penalised for poor performance (sometimes referred to as a two-sided incentive), or whether firms should instead only be penalised for poor performance (sometimes referred to as a one-sided incentive). An Oxera report for Royal Mail in 2005 found that, in the context of service quality performance incentives, at that time, most regulators tended to use two-sided incentive mechanisms.³⁷

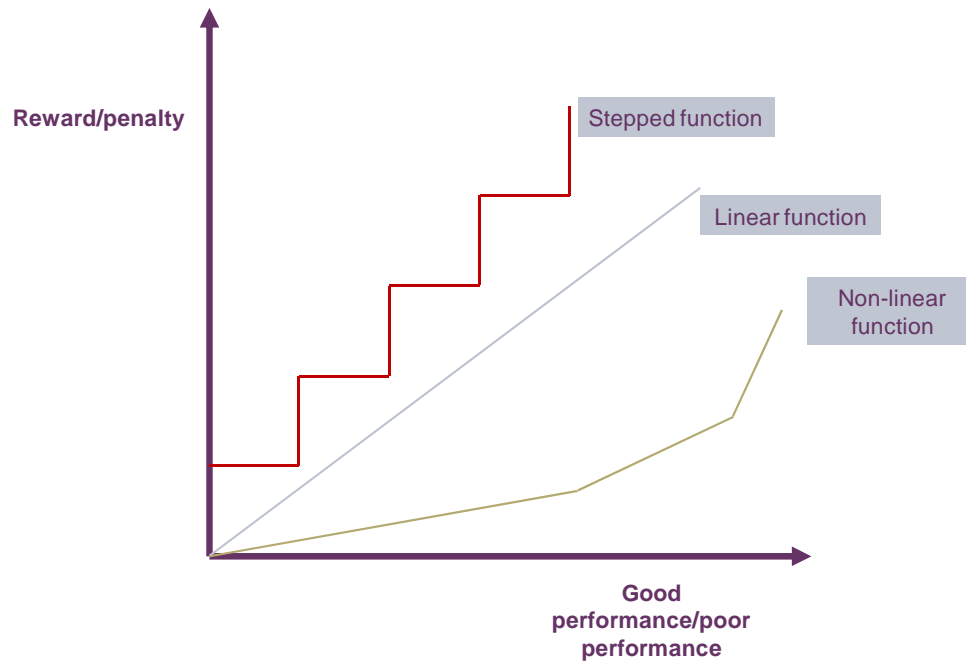
The shape of the reward and punishment function may also differ between incentive mechanisms, introducing another form of asymmetry into the relationship. This type of argument was recognised in Ofwat's 2010 discussion paper:

For example, if we want the companies to meet a given standard, emphasising the penalty associated with not meeting that standard may be appropriate. On the other hand, if we want the companies to achieve outstanding performance, then we may place greater emphasis on rewarding that performance.

Some examples of possible shapes for functions, which complement those contained in the Ofwat 2010 report on incentives, are presented in Figure 5.2 below.

³⁷ Oxera (2005), 'Regulatory precedents for setting service quality incentives: financial levels', September.

Figure 5.2 Possible functional forms for incentive mechanisms



Source: Oxera.

Different shapes for the reward/punishment function have different implications for the operation of the incentive mechanism concerned.

- Penalties that increase (as a function of performance) faster than rewards increase may impart a high incentive to meet the relevant target.
- Graduated payments reduce the complexity of monitoring, and may be suitable in circumstances where it is expensive or not feasible to collect precise performance information. Care needs to be taken to ensure that companies do not face an inefficient incentive to stay at a particular 'level' if there would be merit in raising performance to the next level. On this point, Joskow notes the following:

At first blush, the use of hard caps and floors on the realizations of sliding scale mechanisms that place kinks in the incentive structure are hard to rationalize from a theoretical perspective and appear to have poor incentive properties for realizations near to the kinks in the incentive contract. Caps and floors may be justified as reflecting outcomes that were not contemplated (bounded rationality) in the level and structure of the target performance norms and the distribution of profits around these targets³⁸

- A compromise between linear and stepped or graduated payments could be the use of a non-linear function that delivers higher rewards (or more severe punishments) for better performance (or worse performance). This would mean that the slope of the line would change for different levels of performance.
- The impact of the incentive scheme could be absolute or proportionate. An absolute scheme applies if the regulated firm achieves or fails to achieve certain benchmarks, and does not change according to how much the firm exceeds or fails by. A proportionate scheme varies according to the level of out- or underperformance. In effect, the distinction is simply one of the shape of the incentive function.

³⁸ Joskow, P (2006), op. cit., p. 54.

Another dimension of the symmetry of incentives that is relevant is symmetry through time. For example, incentives could be designed such that their impacts changed with the time elapsed since a particular date, and therefore the symmetry and functional shape of an incentive could change. For example, operators in the electricity distribution sector in the Australian state of Victoria were given progressively more stringent quality of service targets for the control period from 2001 to 2005. Interestingly, it could not be demonstrated that users were willing to pay for the higher quality, so the targets were left at the level of the previous year's control.³⁹ This demonstrates the interaction between the flexibility of an incentive scheme that gets tighter over time, allowing for a degree of 'learning' as to the desired level of willingness to pay.

Some of the points to note that emerge from a consideration of symmetry, and which are relevant to the transmission mechanism of incentives, are as follows.

- The issue of symmetry arises in a number of ways, beginning with an assessment of whether the incentive is two-sided, the functional shape of the incentive, and whether the symmetry is preserved over time. Each of these considerations may be expected to have a different impact on outcomes.
- Two-sided mechanisms may encourage greater innovation by encouraging companies to take risks to earn higher returns, whereas this incentive may be dampened under a one-sided mechanism.
- In the absence of other safeguards, rewarding outperformance may lead to over-delivery—ie, outcomes beyond customers' willingness to pay. This could be addressed by tying the shape of the payment function to willingness to pay.
- The choice of a one- or two-sided incentive should be related to the characteristics of different incentives. For example, where it is felt that compliance is attainable with modest effort and should reasonably be expected of all companies under all normal circumstances, a one-sided (punishment only) mechanism could be appropriate. Another characteristic of outcomes is their impact on risk, and this should be reflected in decisions on symmetry.
- In this context Ofwat⁴⁰ notes a finding of academic literature, which notes that asymmetry may be appropriate where it is critical to avoid unfavourable outcomes, or where average performance is acceptable but outperformance may be desirable.

The appropriateness of different types of symmetry will vary according to the characteristics of outcomes in a number of ways. Outcomes relating to the financial strength and health of the organisation may lend themselves to less symmetric mechanisms if certain indicator variables that are relevant to the overall outcome can be taken as a type of baseline (eg, gearing levels). On the other hand, a more multi-factor outcome, such as customer experience, may merit some form of two-sided approach and symmetry to allow for differences in how companies might improve the customer experience.

The advantages and disadvantages of different forms of symmetry are therefore specific to the type and characteristics of the outcome being incentivised. An asymmetric mechanism that punishes more severely than it rewards may be expected to encourage greater compliance than a pure asymmetric mechanism that only punishes, although this conclusion only applies if wider aspects of the incentive design (eg, controllability) are congruent.

³⁹ Essential Services Commission (2006), 'Victorian Electricity Distribution Business – Comparative Performance: Report 2005'.

⁴⁰ Ofwat (2010), 'Allocating risk and managing uncertainty in setting price controls for monopoly water and sewerage services – a discussion paper', p. 38.

5.2.5 Risk balance

Regulatory incentives derive some of their properties from exposing companies to some amount of risk—indeed, this could perhaps be described as one of the fundamental features of incentive mechanisms. However, there is a need to ensure that the risk associated with different mechanisms is allocated appropriately (risk allocation), and that companies are not exposed to undue amounts of risk (risk mitigation).

The first step in assessing the risk characteristics of an incentive mechanism is to characterise the dimensions along which such risk might vary.⁴¹ An initial list of steps to undertake this characterisation is as follows:

- evaluate risks introduced by the incentive scheme to the different operations of the company;
- identify which parties the risk might fall on;
- identify any additional changes imposed by the incentive mechanism that might mitigate this risk;
- review additional mechanisms introduced to mitigate higher risks or to compensate for higher risks;
- consider what mitigating provisions would be appropriate from a conceptual standpoint;
- quantify the impact of risks introduced by the incentive schemes incorporating the impact of the mitigating provisions.

The transmission of incentives from a risk perspective also depends on how incentive mechanisms affect systematic and idiosyncratic risk. In particular, if additional systematic risk is introduced, this should be reflected in the allowed rate of return through the impact of a higher cost of equity. Absent any regulatory intervention to compensate this risk, the companies to which the risk is allocated may take other steps to reduce their exposure.

The structure of returns from an incentive scheme can also affect this aspect of the incentive transmission mechanism, such as the impact on the mean expected cash flows, and the expected volatility of the impact on cash flows.

As a theoretical principle, risk should be allocated to the party best able to bear that risk. Ofwat's discussion paper on the allocation of risk⁴² lists the different parties to which various types of risk are allocated under the current structure of the water industry in England and Wales. This list highlights the multifarious characteristics of the parties involved in the industry.

If an incentive system shifts too much risk beyond the company's control, the company may choose not to partake in the market. This might ultimately lead to a situation where companies will require greater compensation to participate. When analysing an incentive mechanism, it is therefore necessary to go beyond the simple theoretical principle and instead consider broader considerations. For example, is the allocation of risk consistent with wider regulatory objectives? Does it also correspond to how the benefits associated with the risk are distributed? Does the allocation reflect the flexibility available to different parties to respond to changes in the level and nature of risk? An integrated, consistent approach to regulatory incentive mechanisms is therefore needed to examine these issues in the round.

Two examples of simple mechanisms to mitigate the risk associated with incentive mechanisms are the use of price control re-openers, and of deadbands. Re-opener provisions are intended to allow the regulator to carry out reviews of the company's revenue allowance, either upon request by that company or at its own discretion. Re-openers do not guarantee an automatic revenue adjustment, but rather leave this decision to the regulator.

⁴¹ This reflects Ofwat's FPL papers on risk, including Ofwat (2010), 'Allocating risk and managing uncertainty in setting price controls for monopoly water and sewerage services – a discussion paper'.

⁴² See *ibid.*

Re-openers can therefore be used to manage risks associated with incentive mechanisms by giving the affected company grounds to re-open the price control in the event that the incentive mechanism has an unforeseen, material impact. An important design consideration would be to limit the circumstances under which a re-opener could be requested, while still having an incentive that provides sufficient coverage of relevant risks.

In effect, this involves finding a balance between rules- and discretion-based approaches. Ofwat's existing approach to re-openers is primarily rules-based, and does not allow notified items in areas where the risks are deemed to be covered by the indexation of cost forecasts and the regulatory capital value (RCV), or are considered to be part of general operational risk (since this risk is captured in the cost of capital determination). In addition to relevant changes of circumstance, companies can apply for a re-opener under a substantial effects clause, which in essence is a risk- mitigation mechanism.

Deadbands can also be used to help mitigate the risks associated with incentive mechanisms. The basic operation of a deadband is to permit a margin to exist around a target level of an incentive. This permits minor variations in target performance, such that no adjustments to revenue are made. A deadband also offers customers a degree of flexibility before they are required to pay more money.

In terms of outcome characteristics and choices relating to risk balance, and the advantages and disadvantages of different approaches, introducing a deadband into incentive mechanisms may be helpful for a number of reasons. It can be used to reflect inherent uncertainty about the appropriate level of the target under consideration, and therefore mitigates regulatory risk. It also permits the company a degree of discretion before a revenue adjustment or other change is given effect.

An outcome reflecting climate change prerogatives might merit some form of a deadband to reflect regulatory uncertainty over what an optimal outcome might be, and uncertainty over how companies might improve performance in this area.

Re-openers may not be relevant for outcomes relating to the quality of the water environment (when other incentive mechanisms could be relevant), but would be relevant for an outcome relating to financial health, since a re-opener could be used as an effective means to restore an organisation to financial health, for instance by permitting a higher price cap to be levied.

5.2.6 Trade-offs

The possibility of trade-offs has been raised in a number of places in the foregoing text of this section. The principle alluded to has been the need to apply an understanding of customer preferences and willingness to pay in order to mediate between any competing outcomes, or between measures that need to be traded off against each other in reaching a particular outcome.

In practical terms, more specific guidance will be needed to go beyond an appeal to preferences and willingness to pay. The practical consequences of the necessity of trade-offs can perhaps best be understood by first defining the types of trade-off that might arise, and why they arise, before considering what could be done about them.

A basic reason as to why trade-offs arise is that the boundaries between different outcomes may not be precisely defined, which may mean that some trade-offs are inevitable. Even if it is possible to define outcomes in an exact fashion, some of the inputs, outputs and associated measurement indicators will overlap, and may conflict with one another.

A simple example is where there is an relationship between costs, and reducing one cost puts upward pressure on another cost, in a type of 'waterbed' effect. A familiar regulatory issue, and one being addressed in the FPL project more widely, is the choice between CAPEX and OPEX solutions, which requires some means of ensuring that the trade-off chosen by companies is reasonable and in the interests of users.

A possible means of proceeding when designing incentives is for Ofwat to identify the general trade-offs that might be expected to apply in all cases (such as the choice between CAPEX and OPEX), the types of trade-off that may apply between outcomes, and the specific trade-offs that apply in the context of a particular outcome.

The basis for mediating the trade-offs identified must reflect some form of hierarchy of priorities. Customer interests are obviously central to any ordering of priority, but it is important to note that broader considerations should play a role. For example, if surveyed, customers may accord a relatively low priority to climate change mitigation, but this may still serve as a legitimate outcome for Ofwat to pursue in conjunction with the companies, Defra, and other stakeholders. In turn, this could suggest that trade-offs could be further categorised according to whether they are intended to be directly supportive of the expressed intentions and preferences of customers (eg, quality), or whether they are reflective of wider policy concerns (eg, climate change).

Ofwat would then need to work with stakeholders to establish a hierarchy of priorities that can be used to mediate between trade-offs, such that any trade-off reaches the most desirable (or least undesirable) set of outcomes. This would appear to require implicit values for outcomes.

5.3 Incentive process issues

A number of incentive process issues must be addressed in the practical application of incentives, including the following:

- should some form of common incentive mechanism apply across all outcomes?
- should the division of responsibility for actions between regulator, regulated company, and other stakeholders be the same across outcomes?
- should incentives be the same for each company, or reflect their individual circumstances?
- does responsibility for assessing performance against the outcome sit with the company or Ofwat?

The answers to these and other questions are a consequence of the types of approach taken to the issues highlighted in the foregoing. For example, if a comprehensive rather than a partial approach is to be taken, the company (or Ofwat) will need a means of understanding and making the necessary trade-offs to reach the overall desired outcome. These trade-offs would need to respect customer preferences, which may be difficult to establish in all circumstances, especially where preferences change over time.

Other points to emerge from discussions with companies include the consideration that simplicity of incentives may more readily support engagement with customers than more complex schemes. An incentive set to reflect the value to customers of outcome delivery or under-delivery may be easier to understand for customers. This may require the companies' research into customers' preferences (willing to pay for services) to be linked to the measure of success/outcomes.

Processes themselves need to be flexible in order to address emerging issues, and this flexibility can be supported by transparency.

A fundamental point from earlier parts of this section is that successful incentives need to operate on the incentive-compatibility constraints of individual companies. This suggests that where all companies share a common constraint, it is appropriate to introduce an incentive that is the same, or operates in the same way, across all companies. Where there is a difference in whether and how an incentive-compatibility constraint would bind, and therefore the transmission mechanism involved would differ, company-specific mechanisms may be appropriate. For example, all companies have a financial incentive to make returns for shareholders (although the incentive is weaker for Glas Cymru as a company limited by

guarantee, it is still present to some degree), and therefore an incentive that directly hits the 'bottom line' is of clear relevance. Commonality across incentive mechanisms can operate across different dimensions—a desire to have an outcome based on the same basic level of water quality could use the same type of incentive mechanism across different companies.

A difference in incentives between companies could arise if companies differ in their ability to meet a particular outcome. This also suggests that the division of responsibility among stakeholders need not be the same across different outcomes, and that it may be possible in some cases to rely on companies to report their own progress in meeting outcomes (where a self-reported figure could be readily audited), while in other cases regulatory and/or stakeholder oversight of delivery may be necessary.

However, a wide breadth of differentiation by company may increase the regulatory burden, and may reduce the clarity of incentives. Ultimately, Ofwat would need to decide how much differentiation is 'too much' in the sense of diluting incentives because of complexity.

5.4 Conclusions on incentive design

Table 5.1 summarises some of the key observations of this section, together with some summary advantages and disadvantages. The discussion below the table considers in some more detail the selected elements.

Table 5.1 Summary of discussion

Element	Definition	Implications: advantages and disadvantages
Individual and multiple benchmarks	The question of individual and multiple benchmarks arises because there is a choice among at least the following means of proceeding: applying incentives at the level of overall performance (ie, many outcomes together), at the level of each individual outcome, and at the level of each individual measure	<p>Detailed benchmarking at the level of individual measures increases complexity and the regulatory burden, and may lead to a narrow focus on these individual measures</p> <p>At the other extreme, a challenge with assessing overall performance in a holistic sense is that there is a need to understand trade-offs made between outcomes, whether an acceptable or minimum standard has been reached on all outcomes</p> <p>This suggests that assessing performance at the level of each individual outcome may have some merit, although an overall, holistic view may still be required to understand whether weak performance on one measure is compensated by strong performance on another</p>
Targeted incentives	The introduction of any one specific incentive needs to be congruent with and supportive of the wider set of incentives. This raises the question of whether incentives can be introduced piecemeal, or whether 'all or nothing' comprehensive reforms are more appropriate	<p>Incrementality may have an advantage in being more targeted, and avoiding the double-jeopardy of a company being penalised twice for the same offence</p> <p>However, it may lead to a narrowing of focus, and the failure to adopt a comprehensive and holistic focus. A further important consideration is the need for interventions to be practical: wide-ranging changes that are intended to be holistic may be represent more disruption to the industry than is warranted compared with a more modest incremental intervention</p> <p>Companies have also noted that many incentives are already present in one form or another, so care needs to be taken to understand the existing structure of incentives</p>

Element	Definition	Implications: advantages and disadvantages
The level of financial incentives	Financial incentives can take large or small positive (rewards) or negative (punishment) values	<p>As financial incentives have a direct and fairly clear impact on companies (through their impact on returns), their level is of considerable significance. It is therefore important that companies are not penalised or rewarded excessively for events beyond their control. Additionally, any incentive mechanism needs to be able to distinguish between 'good faith' efforts and inappropriate spending</p> <p>There is a range of design issues associated with how a financial incentive is structured, including whether a cap-and-collar approach would be merited, perhaps to allow for uncertainty about what the precise range should be</p> <p>It is suggested that the value of a financial incentive be related to the value of the outcome (if this is well-defined)</p>
Symmetry	Symmetry captures the notion of whether rewards should be of similar proportion to penalties; whether these should be constant over time; and whether performance of one type should receive greater reward or punishment than that of another type	<p>As Ofwat's paper on incentive indicates, various functional mechanisms are available to capture different notions of symmetry and asymmetry. In practice, it will be important to match the particular functional form with the specific outcome being incentivised. For example, symmetry could change over time in order to incentivise incremental improvements over performance from a particular baseline</p> <p>Companies have expressed concern that one-sided punishment mechanisms will tend to lead to less risky solutions being adopted</p>
Risk balance	Incentive mechanisms derive a significant part of their impact from exposing companies to risk. However, the overall amount of risk needs to be balanced between companies and other stakeholders	<p>A basic principle is that risk should be borne by the company best able to manage it. Companies have expressed concern that linking rewards and penalties to harder-to-measure outcomes may increase their risk overall.</p> <p>As Ofwat's risk paper makes clear, there are a large number of stakeholders in the water value chain, and there is a need for companies to be exposed to reasonable, but not excessive, levels of risks</p> <p>Practical mechanisms that have fairly well-established advantages in managing risk include re-openers and deadbands</p>
Tradeoffs	Trade-offs are an unavoidable part of almost all aspects of regulatory activity, and will need to be addressed in designing incentive mechanisms to support particular outcomes	<p>One means of addressing trade-offs is to use customer preferences to mediate among different activities, benchmarks or outcomes. This will require a close understanding of the trade-offs actually involved</p> <p>It may also be necessary to define a broader hierarchy of priorities to account for the fact that customer preferences in some areas may not provide sufficient guidance</p> <p>Companies have expressed a desire to maintain some form of flexibility to respond to trade-offs</p> <p>Trade-offs, or the principles necessary to decide on them, will need to be decided by Ofwat working in collaboration with other stakeholders</p>

Source: Oxera.

The following considers some of the pros and cons of selected elements of incentive design. In some cases there is no pro or con associated with the element of incentive design per se; for example, there is no pro or con of symmetry in itself, but rather how symmetry and asymmetry are used to relate the characteristics of different outcomes with the form of incentive mechanism being applied.

Level of financial incentives

Financial incentives have a fairly direct advantage, in that their impact on a company can usually be perceived fairly directly by management, which means that the transmission mechanism should also be direct. Financial rewards and penalties will also generally be readily intelligible to end-users, although the case for rewarding good performance may be harder to communicate to these stakeholders than the case for penalising poor performance.

The advantages and disadvantages of setting too high or too low a penalty or reward were discussed above. In essence, once the upper and lower limits of the value of the incentive are defined, the exercise becomes one of identifying the consequences of failure on the low or high side. This will be a function of the outcome under consideration, the type of incentive involved (eg, are the relevant parameters readily controllable?), and the type of company involved (eg, is the company highly geared and vulnerable to a penalty?). The nature of the outcome matters because it may be desirable in some cases to advance a large reward or impose a large penalty if the outcome is of especially high priority.

There are also likely to be advantages to making the value contingent on performance, such as the example cited above from PPP contracts of escalating a penalty for ongoing infringement.

Symmetry

The degree of symmetry in an incentive can have significant impacts on the transmission and ultimate impact of incentives. One-sided, punitive incentives may be warranted for outcomes that reflect statutory or other legal obligations, where non-compliance is not an option. One-sided incentives may suit outcomes where delivery at the highest level may be difficult, but not infeasible, to attain. This could encourage companies to strive to meet the highest possible standard in the hope of a reward. A one-sided mechanism could also be more appropriate in cases where there is a higher degree of confidence over its target level. For example, this is the case when companies can reliably predict and attain the optimal level of the outcome, reducing the overall risk related to that outcome.

Two-sided incentives can be appropriate in a variety of circumstances. They can be used where there is ex ante uncertainty concerning the feasibility of meeting particular targets, and there both punishment and reward can be used as part of the same mechanism so as to avoid excessive generosity or harshness. This consideration may be especially relevant for innovation, the results of which are inherently uncertain.

6 Next steps

6.1 Overview

The preceding sections set out, at a high level, the key considerations around incentive options and design.

The high-level assessment of incentive options indicated that a financial incentive appears to be the option that Ofwat may want to consider introducing, possibly in combination with other options. In particular, the specific characteristics of the outcomes, of which the delivery is being incentivised, may affect the conclusion about which is (are) the most suitable incentive option (s).

Financial incentives may be suitable for outcome delivery, provided that full consideration is given to specific design issues (eg, one- versus two-sided incentives, trade-offs between measure of success or outcomes allowed for in assessing successful delivery, use of deadbands, etc) and to the use of other incentive options in combination or as alternatives.

The success of any mechanism in transmitting incentives and delivering desired outcomes will depend on its feasibility and practicality. In turn, these principles can be supported by assessing whether incentive mechanisms are simple and readily understandable, transparent in their motivation and operation, and operate coherently and consistently with wider regulatory objectives.

Indeed, one of the overall outcomes of the FPL research programme could be a clearer and simpler set of incentives. Complexity makes the operation of the transmission mechanism more burdensome, and may lead to late, expensive or undesired outcomes.

As part of the next steps that Ofwat could consider how to ensure that proposed incentive mechanisms are feasible and practical. In this context, the following steps could be undertaken:

- engagement with companies, UKWIR and their advisors and testing of ideas;
- continuous assessment to ascertain whether the mechanisms are likely to work as expected;
- tests to understand whether individual incentive mechanisms are really necessary, or whether one simple mechanism could achieve the same goal.

In particular, as an initial step, it may be appropriate to focus the engagement with companies on testing practical ideas for incentive implementation. Oxera will support Ofwat in developing a detailed design of the incentive, starting from the options being proposed by Ofwat in the consultation paper on wholesale incentives. The example below illustrates how an incentive could work in practice, focusing on how, depending on the state of the world, price adjustments could be implemented.

6.2 The incentive design in practice

As an illustration, this section presents an example of how financial incentives may work. In addition, it considers an example that also includes a reputational/procedural incentive as part of the incentive package. The example assumes that a two-sided incentive is applied.

The first step would be to assess a level of service desired within the regulatory period (eg, the desired outcome/milestones, or more than one outcome, if this is deemed appropriate).

The second step would be to assess the level of allowed revenues and, more specifically, the allowed returns that relate to that level of service.

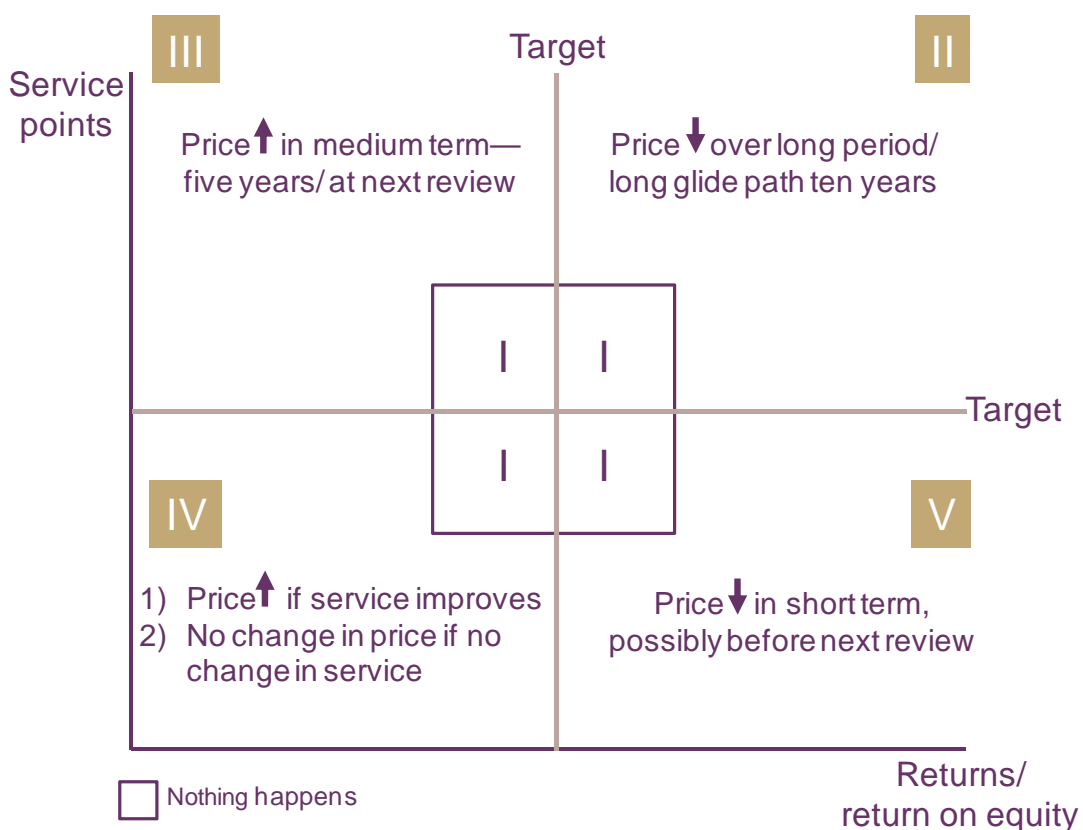
The service level that customers want and are willing to pay for, and the associated allowed revenues and return, effectively form the regulatory settlement for the delivery of that service (outcome/milestone or outcomes). This can be illustrated as a point with a given level of service provided to customers and a given level of return earned by the company.

When establishing the regulatory settlement, companies and customers may also need to agree the points when incentives are to be triggered, and the level of rewards or penalties that a company can expect when achieving or failing to achieve the level of service agreed in the regulatory settlement.

During the course of the regulatory period, the 'state of the world' may turn out to be different from that assumed in the regulatory settlement. The sections below illustrate the mechanics of how the financial incentive could result in price adjustments and the timing of these adjustments.

Differences can arise both in terms of the level of service delivered and the returns earned. There are fundamentally five possible scenarios relative to what was agreed in the regulatory settlement: (I) returns and service levels are broadly as agreed; (II) returns are higher and service is higher; (III) service is higher and returns are lower; (IV) service is lower and returns are lower; and (V) returns are higher and service is lower. These scenarios are illustrated in Figure 6.1.

Figure 6.1 An illustration of possible implementation of two-sided financial only incentive mechanism



Source: Oxera.

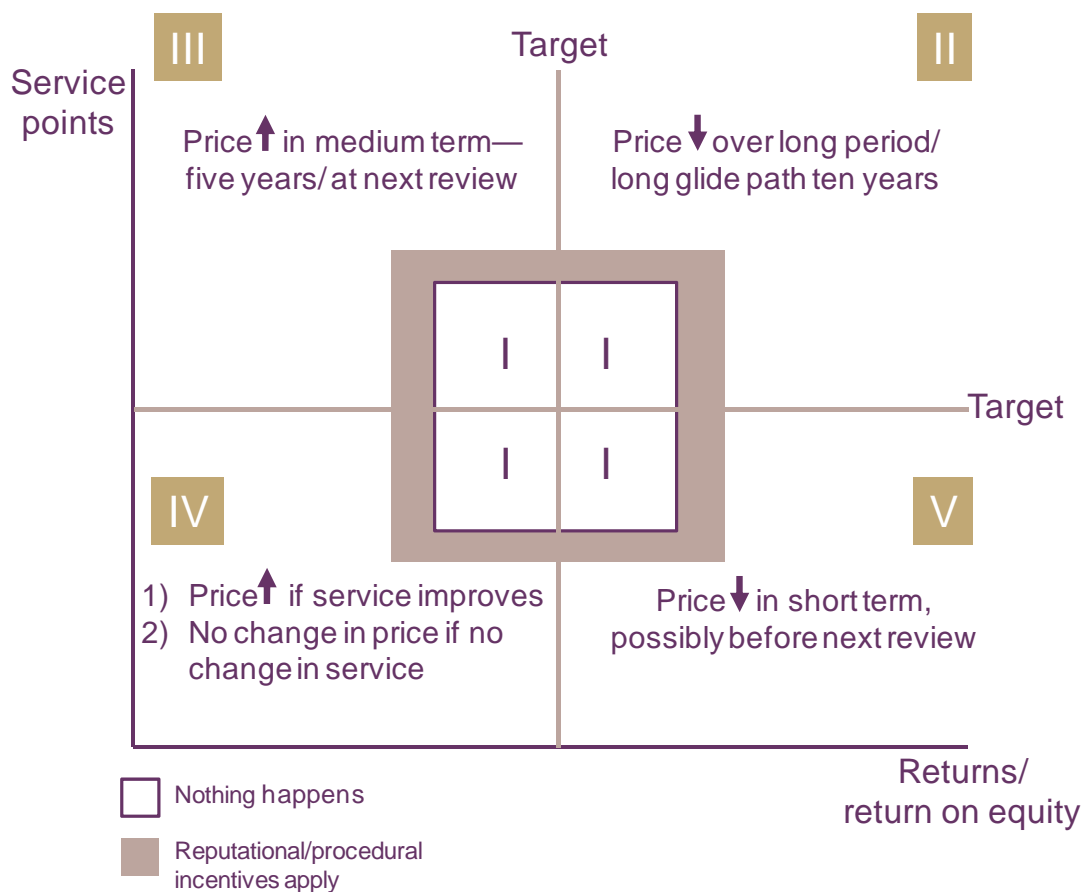
The implementation of the incentive could work as follows:

- I. If the actual performance of the company on services and returns locates it close to the combination of service levels and associated returns agreed at the regulatory settlement (in the figure, this is any point within the central box), no action is taken. In practice, this is the adoption of a deadband.
- II. If the company earns higher returns but also delivers a higher level of service (for which customers are willing to pay), the claw-back of those higher returns may be profiled over a period of time (eg, ten years) to allow the company to retain a reward. In other words, this is assumed to allow the company to retain a return in order to provide it with a positive financial incentive (reward) for service over-delivery. This reward would be over and above the reward allowed for the cost efficiency via the menu.
- III. If the company earns lower returns but provides a higher level of service, the price could be adjusted to reflect that higher level of service (again, assuming that customers are willing to pay for it) over the medium term. The time lag is aimed at ascertaining whether the delivery of a higher level of service is arising from 'good luck' (ie, outside the company's control), or is actually due to the company's performance. If the over-delivery persists, it could be assumed that the higher service levels are due to the company's performance, and it may therefore be appropriate to reward the company. The level of the reward would depend on what was agreed for over-delivery at the price settlement and on any implications of the menu.⁴³
- IV. If the company earns lower returns and provides lower services, there could be two options:
 - the company manages to improve its service level within a reasonable time, and prices are allowed to increase thereafter. In effect, underperformance would be considered to be due to 'bad luck' (ie, outside the company's control) in this instance, and, by improving its service, the company would show its capacity to manage the event outside its control. In this context, the company is provided with the return deemed appropriate at the regulatory settlement;
 - the company continues to under-deliver, and prices and returns are left unchanged (the company is penalised for underperformance). In this instance, underperformance is assumed to be the responsibility of the company and its management. As such, the company would be penalised and not allowed to increase returns via price increases.
- V. If the company earns higher returns but delivers a lower level of service, this could be interpreted as 'gaming', in which case the price would need to be reduced to bring the level of return in line with the level of service being delivered (and this could take place within the regulatory period).

If reputational/procedural incentives are deemed to be appropriate, their use can be incorporated into the above example. As in the case of financial only incentives, the five possible scenarios relative to what was agreed in the regulatory settlement are: (I) returns and service are broadly as agreed; (II) returns are higher and service is higher; (III) service is higher and returns are lower; (IV) service is lower and returns are lower; and (V) returns are higher and service is lower. However the implications of (I) differ from the previous case, as illustrated in Figure 6.2 below.

⁴³ For example, the menu incentive may allow for some over-expenditure incurred in the previous period and related to the over-delivery of the outcome to be recovered in the new period. In this case, customers would be paying, in part, for the delivery of a higher standard of the outcome. The delivery incentive reward may be set equal to the customers' valuation of this higher standard—ie, their maximum willingness to pay for it (maxWTP). If the reward is given without considering the mechanics of the menu, customers could end up paying for the delivery incentive (maxWTP) plus some of the cost incurred to achieve this higher level of the outcome. Customers would pay in excess of their maxWTP. Therefore, it appears appropriate to consider the interaction between the menu and the outcome delivery incentive.

Figure 6.2 An illustration of possible implementation of a two-sided financial and procedural/reputational incentive mechanism



Source: Oxera.

In this case, the incentive mechanism would allow for a 'buffer' zone beyond the deadband before the company is given a financial penalty. The structure of this incentive mechanism may be particularly suitable when the degree of controllability by the company is deemed to be only partial—for example, because the delivery is likely to be influenced by external events or other agents' actions (eg, environmental outcomes).

In this context, it may be more appropriate to use options other than financial incentives; although, if there is a marked deviation from the service levels target, the company could still be subjected to a financial penalty. The assumption in this case is that, despite the limited controllability over the outcome, except in extreme circumstances, the company may be still in a position to mitigate/manage the risk posed by external events/other agents to the delivery of the target level of service.

As previously noted, a credible prospect of a financial incentive may lead the company and its management to do all it can to manage the risk and keep delivery and its pace within a reasonable (according to its customers) level.

Similarly, only in the case of substantial outperformance of the agreed service levels would the company be given a financial reward, while, for a smaller outperformance, no action would be taken, or procedural/reputational incentives would be given. This is assumed to reinforce the customers' protection embedded in the previous model (as financial rewards for outperformance are given only with a time lag) and to avoid paying the company for its 'good luck'.

A1 Ofgem experience

Ofgem has adopted what it terms an outputs-based approach as part of its RIIO (Revenue = Incentives + Innovation + Outputs) framework for the next transmission and gas distribution price controls commencing on April 1st 2013. Despite the different use of terminology, this approach has similarities with Ofwat's proposed outcomes approach. As part of Ofgem's outputs-based approach, incentives have been agreed which apply around (non-) delivery. These incentives include enforced standards, reputational measures, and financial rewards and penalties. This appendix outlines some of the key messages from Ofgem's RIIO-T1 outputs strategy document,⁴⁴ its final proposals for fast-tracked companies⁴⁵, and the initial proposals for the non fast-tracked companies⁴⁶, which are deemed to be of relevance to Ofwat's outcomes-focused approach.

Outputs

- In formulating its strategy for RIIO-T1 and RIIO-GD1, Ofgem identified a set of output categories which was tested within stakeholder working groups. These output categories will be included in the operators' licences at the start of the price control, and cover:
- safety and reliability;
- connections;
- network availability;
- customer satisfaction;
- environmental outputs;
- wider works outputs.

Process

- The structure of output delivery incentives in terms of their type (eg, financial or reputational), form (eg, symmetrical or asymmetrical) and size (ie, the percentage change in allowed revenues) was largely determined by Ofgem. Companies were, however, given 'some scope to propose additional or alternative outputs and incentive arrangements providing they justify that these address the specific needs of their stakeholders, ie as evidenced from their stakeholder engagement process.'⁴⁷
- In practice, when they put forth measures in their RIIO-T1 business plans, the companies largely followed the representative output measures and incentives outlined in Ofgem's outputs strategy document. This has meant that the targeted outputs and delivery incentives are largely the same for each of the companies, as opposed to being set on a bespoke, company-by-company basis. Ofgem then advised on those areas where companies needed to undertake further work in this context, prior to the final proposals.⁴⁸ Following the final proposals, there is still ongoing work with some of the output measurements (eg, the design of a customer satisfaction survey).
- There will be a mid-period review of outputs which will allow for changes to agreed measures if such changes are justified by changes in government policy or legislation. New outputs could be introduced where these are required to meet the needs of customers and other network users. Output measures will only be revised outside of the

⁴⁴ Ofgem (2011), 'Strategy for the next transmission price control – RIIO-T1, Outputs and Incentives', March 31st.

⁴⁵ Ofgem (2012), 'RIIO-T1: Final Proposals for SP Transmission Ltd and Scottish Hydro Electric Transmission Ltd, Final Decision – Supporting document', April 23rd.

⁴⁶ Ofgem (2012), 'RIIO-T1: Initial Proposals for National Grid Electricity Transmission and National Grid Gas', July 27th.

⁴⁷ Ofgem (2011), 'Strategy for the next transmission price control – RIIO-T1, Outputs and Incentives', March 31st, p. 7.

⁴⁸ Ofgem (2012), 'RIIO-T1: Final Proposals for SP Transmission Ltd and Scottish Hydro Electric Transmission Ltd, Final Decision – Supporting document', April 23rd, p.3.

mid-period review where Ofgem ‘identifi[es] an error, or the measurement/reporting of an output does not meet the intended purpose.’⁴⁹

- It is worth noting that Ofgem has acknowledged there are additional reporting requirements involved in their approach to enable them to monitor companies’ performance against the output measures.⁵⁰ This could thus increase the regulatory burden.

Delivery incentives

- Insofar as possible, Ofgem has tried to set out an ex ante framework for the type and size of the output delivery incentives, and where it has been unable to ‘set out a mechanistic reward/penalty, [it has] set out rules for how [it] will determine the size of the reward/penalty in the light of a company’s performance.’⁵¹
- The regulator has applied the following delivery incentives in its final RIIO-T1 proposals for the fast-tracked companies (SPTL and SHETL) and in the initial proposals for NGET:

Table A.1 Incentive types adopted by Ofgem

Output	Incentive	Type
Safety	Compliance with safety obligations set out by Health and Safety Executive (HSE)	Enforced standard
Reliability	Based on Energy Not Supplied metric but with NGET set a higher target level than SHETL AND SPTL. Licence condition on minimum performance standard and financial penalty with a 3% collar SHETL required to pay customers compensation for being off supply for longer than six hours	Enforced standard with financial penalty
Availability	Required to prepare and maintain a Network Availability Policy	Reputational
Customer satisfaction	Measured by customer satisfaction survey. Symmetrical incentive valued at +/-1% of allowed revenue	Financial
Connections	One-sided penalty, must meet existing legal requirements with a penalty of 0.5% of allowed revenue for failure to meet timing requirements (ie, late delivery)	Enforced standard with financial penalty
Environmental	Combination of reputational and symmetrical financial penalties. Eg, differences from target subject to rewards/penalties based on non-traded carbon price	Reputational and financial

Source: Oxera based on Ofgem (2012), ‘RIIO-T1: Final Proposals for SP Transmission Ltd and Scottish Hydro Electric Transmission Ltd, Final Decision – Overview document’, April 23rd, pp.15-16; Ofgem (2012), ‘RIIO-T1: Initial Proposals for National Grid Electricity Transmission and National Grid Gas – Outputs, incentives and innovation supporting document’, July 27th, pp. 5-24.

- For the most part, the target outputs and measures are the same for all three of the electricity transmission companies. For example, all three companies must meet a reliability target with a measure based on the Energy Not Supplied (ENS) metric. However, the incentives are not applied equally across the companies. For example, SHETL will be required to pay customers compensation for being off supply for longer than 6 hours, whereas SPTL and NGET will not; and NGET faces a slightly higher (ie, less challenging) target level of Energy Not Supplied than do SHETL and SPTL. NGET’s outputs fall under the same output headings but involve different measures and incentive types, as it is a gas transmission company.

⁴⁹ Ofgem (2011), ‘Strategy for the next transmission price control – RIIO-T1, Outputs and Incentives’, March 31st.

⁵⁰ Ibid., p.4.

⁵¹ Ofgem (2010), ‘Consultation on strategy for the next transmission and gas distribution price controls – RIIO-T1 Outputs and incentives’, supplementary annex, December 17th, p. 5.

- In determining the incentives to deliver outputs, Ofgem noted that financial incentives were not applicable to all output measures, and in particular should not be applied to those outputs which are not within the control of the regulated company.

We have not proposed financial incentive mechanisms for all output measures... [for] output measures where the network companies have a low level of control over performance, such as the proportion of renewable energy transported, we propose to require the TOs to report on their performance in order to provide a reputational incentive, but do not propose a financial incentive.⁵²

- For example, Ofgem has only set reputational incentives (and not a financial incentive) for the network availability and safety output categories. With regards to network availability, Ofgem considered the potential to implement financial incentives but decided against such an approach, referencing the fact that ‘it is important that a company has information on and control over what it is being incentivised to deliver’ and the difficulty in setting a baseline against which performance could be measured.⁵³

Box A1.1 Ofgem guidelines on the use of financial incentives

As set out in the RIIO handbook, Ofgem only deems it to be appropriate to apply financial incentives when:

- there is clarity on the primary outputs to be delivered;
- there is confidence in the data used to measure performance;
- it considers delivery of the primary output to be important;
- there are not already incentives in place on the network company through other schemes or obligations.

Furthermore, Ofgem’s guidelines require that, when applying financial incentives, ‘the strength of any incentive should take account of the network company’s degree of control over the output.’

Source: Ofgem (2010), ‘Handbook for implementing the RIIO model’, October 4th., p.76; Ofgem (2010), ‘Consultation on strategy for the next transmission and gas distribution price controls – RIIO-T1 Outputs and incentives’, December 17th, p. 53.

- Ofgem also noted that the appropriate incentive design will differ depending on the type of output.

The structure of the incentive mechanism, for example whether it is symmetric/asymmetric, and the basis for setting the reward/penalty will depend on the output measure.⁵⁴

- Consequently, the customer satisfaction and environmental output incentives are symmetrical with rewards (ie, higher allowed revenues) as well as penalties, while the incentive for connections is one-sided, with a penalty for failing to meet the existing legal requirements, but no additional reward for compliance. In general, there are only additional rewards for over delivery where companies are able to demonstrate that this is valued by customers.

We do not expect to provide any additional revenues associated with the over delivery of outputs where this is not valued by consumers. However, in other cases, where the

⁵² Ofgem (2010), ‘Consultation on strategy for the next transmission and gas distribution price controls – RIIO-T1 Outputs and incentives’, supplementary annex, December 17th, p. 5.

⁵³ Ofgem (2012), ‘RIIO-T1: Final Proposals for SP Transmission Ltd and Scottish Hydro Electric Transmission Ltd, Final Decision – Supporting document’, April 23rd, p.5.

⁵⁴ Ofgem (2010), ‘Consultation on strategy for the next transmission and gas distribution price controls – RIIO-T1 Outputs and incentives’, supplementary annex, December 17th, p. 5.

company can demonstrate the incremental output that was delivered was desired by network users, we will recognise the efficient costs associated with this output in setting allowed revenues.⁵⁵

- As outlined in the RIIO strategy document, the rewards and penalties adopted by Ofgem are intended to be proportional to the value of the output:

We have sought to ensure that the rewards/penalties associated with incentive mechanisms reflect the value that consumers place on the output, and that incentivised output measures fulfil the requisite criteria (ie controllable, material etc.) to ensure companies and consumers do not face windfall gains or losses.⁵⁶

- However, in practice there have been concerns around proportionality relating to timely delivery incentives for wider works outputs.⁵⁷ Ofgem has adopted a penalty collar up to 10% of allowed revenue for such financial penalties. SPTL proposed a nominal collar of £10m, while SHETL proposed a collar of 5% of base revenue in the year that the construction of the project commenced.

⁵⁵ Ibid.

⁵⁶ Ofgem (2011), 'Strategy for the next transmission price control – RIIO-T1, Outputs and Incentives', March 31st, p. 3.

⁵⁷ Ofgem (2012), 'RIIO-T1: Final Proposals for SP Transmission Ltd and Scottish Hydro Electric Transmission Ltd, Final Decision – Supporting document', April 23rd, p. 7.

Park Central
40/41 Park End Street
Oxford OX1 1JD
United Kingdom

Tel: +44 (0) 1865 253 000
Fax: +44 (0) 1865 251 172

Stephanie Square Centre
Avenue Louise 65, Box 11
1050 Brussels
Belgium

Tel: +32 (0) 2 535 7878
Fax: +32 (0) 2 535 7770

200 Aldersgate
14th Floor
London EC1A 4HD
United Kingdom

Tel: +44 (0) 20 7776 6600
Fax: +44 (0) 20 7776 6601