



National Audit Office

Performance measurement practice

The use of sanctions and rewards in the public sector

The following individuals contributed to this report: Sascha Kiess, Diane French, Nick Sloan, Dan Vallance, Danielle Williams.

The National Audit Office scrutinises public spending on behalf of Parliament. The Comptroller and Auditor General, Tim Burr, is an Officer of the House of Commons. He is the head of the National Audit Office which employs some 850 staff. He and the National Audit Office are totally independent of Government. He certifies the accounts of all Government departments and a wide range of other public sector bodies; and he has statutory authority to report to Parliament on the economy, efficiency and effectiveness with which departments and other bodies have used their resources. Our work leads to savings and other efficiency gains worth many millions of pounds; at least £9 for every £1 spent running the Office.

Contents

Executive Summary	2
Part One	
Introduction	5
Part Two	
How sanctions and rewards are currently used	12
Part Three	
The effectiveness of formal sanction and reward mechanisms	18
Annex 1	
Methodology	38
Annex 2	
Different types of sanctions or rewards	40
Annex 3	
Guidance for developing and implementing sanction and reward mechanisms	45
Annex 4	
Performance levers	50
Annex 5	
Bibliography	54

Executive Summary

1 This report is a review of the use of sanctions and rewards based on academic research and an NAO survey of their use in central government. The NAO undertook this to inform the debate on incentivising public sector performance. While many commentators discuss public sector performance, they have not systematically reviewed the use of sanctions for failure or rewards for success. The purpose of this report is to bring together the evidence on the effectiveness of sanctions and rewards, summarise the results of our survey on their use in central government, and to provide a practical guide on how to use them well.

Key findings

2 Government bodies have to manage performance through delivery chains, often having limited control over those in the chain. They usually choose funding or regulatory instruments constructed so that they provide a degree of leverage over performance. But such levers are often insufficient to secure all the desired results. That has led to increased emphasis on sanctions and rewards, including the introduction of high profile schemes such as the 'earned autonomy' of Foundation Trusts and the reward grant attached to achievement in

Local Area Agreements. At a lower level, sanctions and rewards are being considered in relation to a variety of behaviour change initiatives, including sanctions for not recycling, financial incentives for businesses to move freight by water, and criminals being offered reduced prison sentences in exchange for pleading guilty as early as possible.

3 Despite this greater emphasis, there are few empirical assessments of the effectiveness of sanctions and rewards. The evidence that does exist shows that public sector agents – organisations, teams and individuals – *do* respond to incentives, and that those incentives can promote or hinder enhanced effort and performance, depending on how well they have been designed and implemented.

4 As well as reviewing evidence within the academic literature, we conducted a survey of 145 government programmes to understand how sanctions and rewards are currently used in the UK public sector. It identified three broad types of mechanisms: financial (e.g. bonuses or penalties); operational (e.g. granting organisations greater or lesser autonomy from inspection); and reputational (e.g. the publication of league tables).

Use of sanctions and rewards

5 Our survey showed:

- **The overall use of sanction and reward mechanisms is low.** Only around 40 per cent of those who responded to our survey reported using an explicit sanction or reward mechanism somewhere in their programme. To some extent, this reflects a lack of familiarity with the terminology: in a number of cases, follow-up to the survey uncovered that sanction and reward mechanisms were being used where none had been reported. Such responses, however, raise questions over the extent to which programme managers have systematically considered the use of incentives.
- **The use of sanction and reward mechanisms is more common in service delivery programmes.** Nearly 60 per cent of UK government programmes that deliver services reported using explicit sanction or reward mechanisms, compared to just 25 per cent for programmes that aim to build capacity, and 21 per cent that seek to influence outcomes.
- **More sanction and reward mechanisms are aimed at the organisational level rather than the team or individual employee level.** This reflects, among other things, the intrinsic difficulties of measuring the contribution of individual employees to many public sector outcomes.
- **Financial and reputational mechanisms are far more common than operational ones.** Our survey identified 26 financial mechanisms, 21 reputational, but only seven operational.

Effectiveness of sanctions and rewards

6 There have been few published empirical studies of the effects of sanctions and rewards in the public sector, and those that have been done focus on US institutions. There have been a number of further reviews and studies which provide some relevant evidence, however, as well as practitioner or expert judgements. Our survey showed a similar picture.

More than 60 per cent of those using sanctions or rewards (22 out of 35) believed that their use had improved performance. Interviews with the respondents showed, however, that these opinions were based on subjective judgements. We found no quantified evidence of the effect of sanctions and rewards on levels of performance for the programmes in the survey.

7 If the evidence on the effectiveness of sanctions and rewards is limited, it does permit a number of observations:

- **People in the public sector respond to sanctions and rewards, if they regard them as significant.**
- Although public sector workers may well have a variety of motivations to perform well, **financial rewards can improve performance.**
- **If stakeholders have a mix of self-interested and altruistic motives, then the design of sanction and reward mechanisms must recognise that mix,** to avoid reducing overall motivation by focusing on one aspect to the exclusion of the other.
- Because performance in the public sector has a number of dimensions, **it is important that the definitions used for sanctions or rewards capture all significant aspects of performance,** to minimise the scope for perverse effects.
- **If performance is difficult to measure, or where the contribution of a stakeholder is difficult to attribute to performance, then sanction and reward mechanisms which have dramatic financial or personal effects should be avoided.**

Good practice checklist

8 The table below summarises the main practices which emerge from the research and guidance. More comprehensive guidance is available at Annex 3.

Table 1**Guide for incentivising a delivery chain**

Objective

- 1 Clearly define a measurable objective.

Map the strategy

- 2 Identify the activities and the outputs that are needed to achieve the objective (see the example Strategy Map in figure 1). [p. 7 and 47].

Map the delivery chain

- 3 Identify and engage with the key stakeholders, map the performance levers and incentives in use (see the example Delivery Map in figure 2). [p. 8 and 48].
- 4 Use the delivery map to undertake a stakeholder analysis. Identify where the main barriers are, and the evidence on the effectiveness of the current programme, the performance levers, and their associated sanction or reward mechanisms (see Stakeholder Analysis template in Annex 3). [p. 49].

Designing sanction and reward mechanisms

- 5 Understand the full range of sanction or reward mechanisms available – financial, reputational and organisational (see Annex 2). [p. 40].
- 6 Assess the motivations of the key players. Assess the balance between financial and altruistic motivations, and whether they operate at organisational, team or individual level. [p. 26 and 40].
- 7 Identify outcomes over which the players have control or significant influence. [p. 27].

- 8 Link the incentives to performance measures which lead to the desired (long-term) outcomes in a predictable way. A single measure may not capture the relevant aspects of performance, but any set of measures must be kept manageable. [p. 28].

- 9 Ensure the rewards and sanctions are cost-effective. Where they involve financial elements, it may be sensible to model the operation of the system to help define appropriate values or ranges. [p. 30].

- 10 Introduce safeguards to prevent unintended behaviours. Approaches can include maintaining a degree of independence in performance assessment and validating key performance measures. [p. 32].

Measure performance

- 11 Develop and implement data systems that collect and analyse timely and reliable performance data. [p. 33].

Apply the sanction or reward

- 12 Develop the expertise to apply the sanctions or rewards in a timely, consistent and transparent manner, especially when using complex contractual arrangements with delivery agents. [p. 34].

Review effectiveness

- 13 Build regular internal and external reviews of the effectiveness of the sanction or reward mechanism into the programme's overall performance management cycle. [p. 35].

Implement

- 14 Where possible, phase in new sanction or reward mechanisms gradually, or on a pilot basis, to identify and address any dysfunctional behaviour. [p. 37].

Introduction

UK public sector context

1.1 This report contributes to the debate on incentivising public sector performance. It examines the various levers used by government to meet the ambitions set out in Public Service Agreements (PSAs) and other major programmes of work, and asks whether the effectiveness of these levers could be enhanced by the use of explicit or formal sanctions and rewards. While there has been much commentary on public sector performance, there has been limited systematic review of the use of sanctions for failure, or rewards for success, in achieving progress.

1.2 The Government's approach to public service reform since 1997 has changed the relationships between central government, local government, and the many other organisations and citizens involved in delivering government objectives. Cabinet Office (2006) summarised it as comprising four elements:

- Top-down performance management, such as PSA targets, regulation/minimum standards, performance assessment, inspection, and intervention to tackle poor performance.

- Introduction of greater competition and contestability in the provision of services.
- Introduction of greater pressure from citizens (choice and voice).
- Measures to strengthen the capability and capacity of civil and public servants.

1.3 The Government has also promoted the merits of localism – the devolution of performance management to a local level, with less central dictation of priorities:

“Our long term objective has always been to match ambitious national standards with a vigorous local autonomy and flexibility whereby we maximise efficiency, equity and a personalised service for the patient, the pupil and the citizen generally. This opens up a challenging agenda for the second stage of modernisation and reform. It includes far more radical devolution of responsibilities from Whitehall as we give the role of Whitehall a sharper focus. It also demands greater transparency, proper audit and new incentives and an ambitious agenda to encourage local autonomy and managerial flexibility.” (HM Treasury, 2004a).

1.4 The Treasury goes on to argue that this devolution of performance management needs to be accompanied by stronger local accountabilities and incentives for delivering improvements:

*“Credible incentives in the form of **rewards and sanctions** must be made available such that intervention by central government only takes place in cases of clear under-performance that the intermediate tier has not been able to correct.”*

1.5 Examples of explicit sanction and reward mechanisms include: contracts with private sector firms, where financial bonuses and penalties are specified for performance; tying the degree of external inspection to performance, as in the case of the Audit Commission with Local Authorities; and the use of public benchmarking, such as school league tables. In other cases, incentives are not reliant on explicit sanctions and rewards, such as the legislation that requires Local Authorities which underperform against air quality targets to publish action plans.

Delivery chains

1.6 Most government objectives cannot be delivered by government Departments alone, but require a complex network of organisations, including central and local government, government agencies, and bodies from the private and third sectors. In the literature these links have been analysed by reference to ‘Principal-Agent’ theory, which describes the issues faced by one party (the “Principal”) when it needs another party (the “Agent”) to deliver its objectives. If Departments are the Principal, then the Agents are the employees, organisations or citizens that Departments need to work with to deliver their objectives. NAO (2006) lists four basic types of linkages or relationships within these delivery chains. Different types of link need to be managed in different ways. For example:

- **Internal links;** where one part of the chain directly manages another. These can be managed through a standard management review process.

The Principal sets objectives for the Agent, who then carries out the activities and reports progress back to the Principal. If performance is poor, the Principal works with the Agent to improve performance. However, this can be very resource-intensive.

- **Contractual or regulatory links;** where one part of the chain defines through law, and/or funding, how another part operates. These are often managed more at arms-length, where the Principal formally delegates responsibility for certain outcomes to the Agent. The Principal will often use sanctions or rewards as a way of motivating the Agent to deliver on objectives, avoiding the need for resource-intensive performance management. Given the move towards localised performance management in the public sector, there has been an increase of interest in the potential for using sanctions and rewards.
- **Links of common purpose;** where two bodies have parallel missions to work towards the common good, such as between two government Departments, or between a Local Authority and a charity. These rely more on partnership working and trust, as there is no clear Principal or Agent. However, these links can be strengthened through more formal instruments, such as Memoranda of Understanding, in which parties set out their commitments to each other. As with contracts, they may include Service Level Agreements to clarify and strengthen these commitments.
- **Links to the wider community;** where one part of the chain has no formal authority over the groups or individuals with which it wishes to work, and is therefore reliant on persuasion. Here there is a need to rely on “softer” levers, such as providing formal education or information as a means of persuading Agents to behave in a certain way. Advertising campaigns or public information leaflets are commonly used for this purpose.

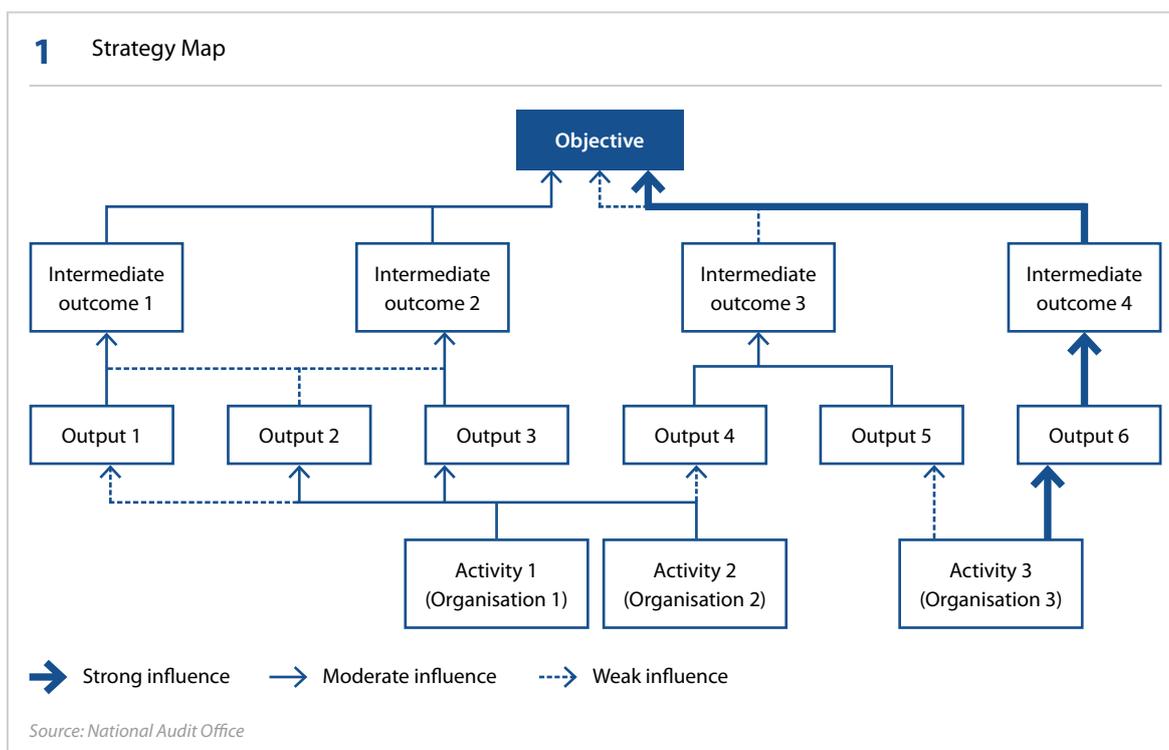
Strategy and stakeholder mapping

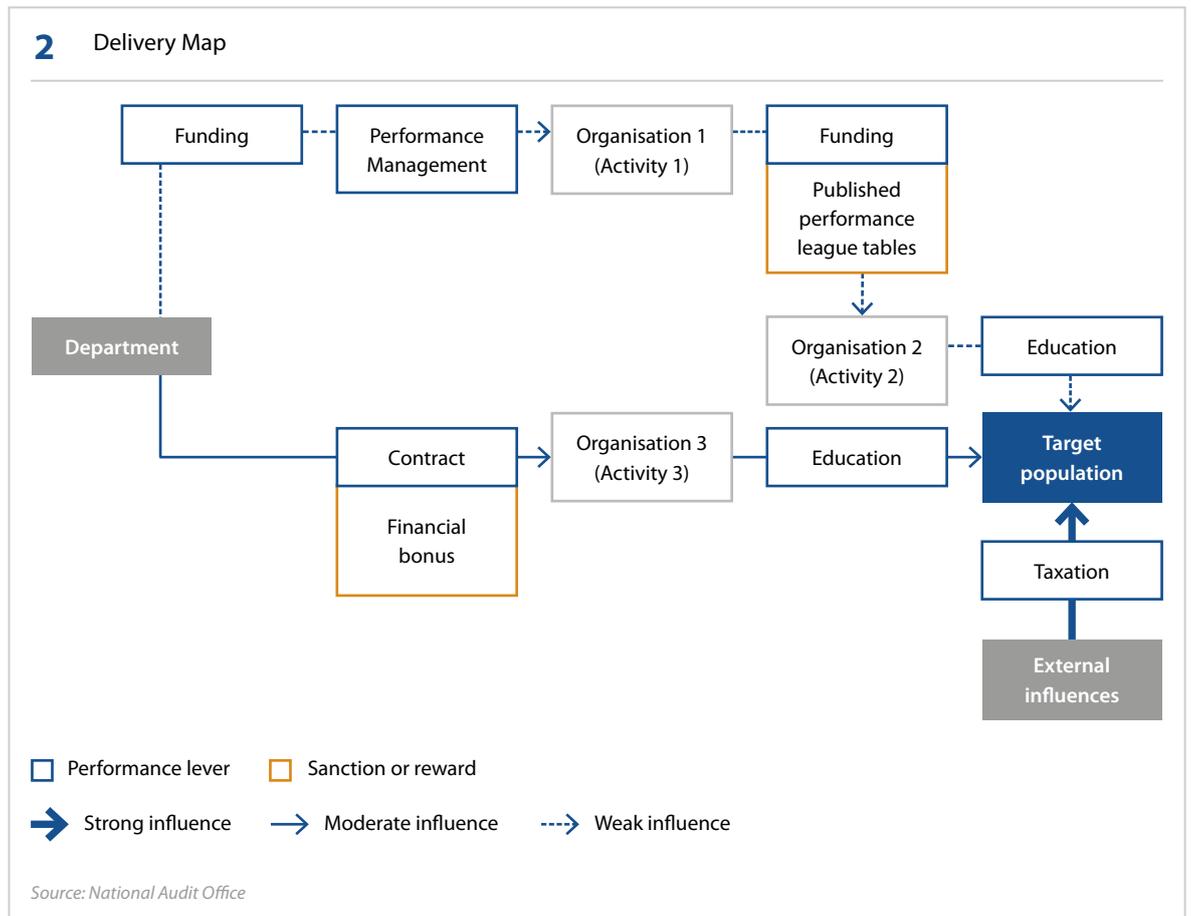
1.7 It follows that designing effective incentives depends on a good understanding of the context, identifying the various performance levers available, understanding the motivations of the delivery agents, assessing the strengths and weaknesses of the levers, and identifying opportunities to strengthen the levers to achieve the desired levels of performance. These steps to planning the delivery of government programmes were formalised within Departments' Delivery Agreements, which were published for each of the 30 Public Service Agreements under the government's 2007 Comprehensive Spending Review. The key elements in this process are summarised below.

1.8 Figure 1 shows a generic strategy map for an objective. It shows how an organisation can take an objective, and identify the combination of activities and outputs that should help to achieve it. It should be based on evidence, or at least on expert judgement, on what drives the outcome. This sort of mapping is designed to help the "owner"

of an objective create a programme structure that is likely to achieve the objective in a cost-effective way. A strategy map can also be used to structure the performance measurement framework, as it identifies the outputs and outcomes that are important to measure. These indicators are also likely to feature in any sanction and reward mechanisms.

1.9 Having identified the activities needed to deliver an objective, suitable organisations and stakeholders can be chosen to perform those activities. It is then helpful to map the relationships between them – including the performance levers that are used and any sanctions or rewards that are attached to them. This type of delivery map can help identify where there is insufficient leverage over performance, and where introducing explicit sanctions or rewards is likely to have the greatest impact. It can also be used to ensure coherence of incentivisation across the delivery chain, and as a basis for considering the likely motivations of individual organisations or stakeholders.





1.10 Table 2 illustrates some of the levers which exist and how they might be used with, or without an explicit sanction or reward. (Further discussion of performance levers is available at Annex 4). The main motivations for using sanction or reward mechanisms include:

- To improve the performance of individual units in comparison to overall performance – for example, of an individual hospital relative to other hospitals.
- To improve the performance of a wider system – for example, to improve the performance of the health system to deal with the threat and occurrence of strokes.
- As a means of allocating resources to more efficient providers – for example, to enable funding to be allocated to a number of different providers based on their performance. (Propper and Wilson, 2003).

Table 2**Performance levers**

Performance lever

Taxation – which, as well as raising government revenue, can be an effective means of incentivising certain behaviours by increasing or decreasing the costs associated with them.

Publishing performance data – which, in a competitive environment, helps service users make an informed choice, and so can be used to encourage competition between individuals and organisations to perform better.

Legislation – which sets the overall legal framework for delivery chains and can enable the use of other levers.

Regulation and standards – which can be used to specify the quantity, quality and/or type of service that providers should offer users.

Inspection and review – which is used to check the quality of delivery or the standards of performance.

Targets – which, combined with the publication of associated performance indicators, are intended to focus attention on key outcomes and to enhance transparency and accountability.

Funding – which can be linked to various measures, such as past performance, organisational capacity, and the achievement of targets.

Without explicit sanctions or rewards

Taxes on alcohol, tobacco and fuel help to raise income for the government. Because the amount of tax individuals pay for these goods is directly linked to the amount they consume, it has an intrinsic effect of motivating consumers to buy less, and so can be used as a performance lever.

Train Operating Companies are required to publish their environmental policies, as well as their performance against key environmental indicators. By allowing greater public scrutiny, it is hoped to encourage the companies to develop sound policies, and perform well against them.

The 1995 Environment Act requires Local Authorities to assess their air quality. There are no formal sanctions if pollution is excessive, but the area will be designated an Air Quality Management Area and an action plan must be drawn up.

Ofgem sets price caps for electricity and gas companies, which incentivises efficiency and contributes to the government's objective of maintaining the competitiveness of the UK's energy markets.

Publication of the results of Department Capability Reviews provides greater transparency of the performance of individual government Departments.

Local Authorities were given efficiency targets following the Gershon Review. Performance against these targets was monitored but there was no formal consequence for non-achievement.

Central government Departments are allocated funding dependent upon their performance in the two previous years, but the link between performance and funding levels is not made explicit.

With explicit sanctions or rewards

Lower rates of Vehicle Excise Duty are offered to Heavy Goods Vehicle operators who achieve target levels of emissions and obtain a Reduced Pollution Certificate.

Converting basic performance information into composite indicators – such as Local Authority assessments or hospital star ratings – makes it clear who is performing better or worse overall, and adds further motivation to perform well.

Local transport authorities are entitled to establish statutory Quality Partnership Schemes with bus operators to require them to provide services of minimum standard to ensure contract renewal.

The Environment Agency can withdraw the operating licences of large-scale industrial polluters – such as car producers and hauliers – if they fail to comply with minimum standards.

Linking Local Authorities' Comprehensive Performance Assessment scores to the degree of intervention and inspection by the Audit Commission.

Publishing the ranking of schools against the target for getting 60 per cent of pupils aged 16 to achieve the equivalent of five GCSEs at grades A* to C. Ranking schools against target achievement rewards higher performers by attracting more pupils, bringing greater revenue.

Reward grants for Local Authorities, such as the Planning Delivery Grant – the size of which depends on performance against explicit targets.

Table 2 continued

Performance levers continued

Performance lever

Internal performance and staff management – organisations throughout the delivery chain use a variety of internal performance mechanisms (such as risk management) and staff management systems (such as annual performance reviews) to improve capacity and performance.

Contracts – which formally delegate responsibilities for achieving certain levels of outputs or outcomes to third parties.

Partnerships and Memoranda of Understanding – which can be established where two or more bodies in a delivery chain have links of common purpose or parallel missions.

Education and information – which can be used to reinforce the effectiveness of other levers, and to change attitudes and behaviours.

Without explicit sanctions or rewards

Many staff appraisal systems use assessments of individuals' performance against non-quantified objectives to manage and motivate the desired behaviour.

Informal incentives can be created by the threat of failing to extend a contract beyond its minimum period, such as Capita's contract with the Department for Children Schools and Families to run the National Strategies programme.

Almost all Public Service Agreements are now shared between Departments, and formalised through jointly agreed delivery plans.

Public information campaigns – such as the drink-driving campaign – are aimed at incentivising a change in behaviour.

With explicit sanctions or rewards

Performance Related Pay for Permanent Secretaries is partially linked to performance against specifically quantified diversity targets.

Transport for London contracts contain financial bonuses and penalty deductions for bus operators, related to their performance against agreed metrics.

These types of levers typically do not include formal sanctions and rewards, unless both parties agree to introduce a contractual element.

Not applicable. Educating and informing organisations and individuals is likely to support one or more of the other performance levers described above.

Study objectives and scope

1.11 Against this background, which implies that a degree of delivery chain analysis will have been performed, the objectives of this study are:

- to identify and classify the sanction and reward mechanisms used in both the public and private sectors;
- to assess the evidence on the effectiveness of different sanction and reward mechanisms in different contexts; and
- to provide some guiding principles to be considered when designing and implementing a system of sanctions or rewards.

1.12 We use the term 'incentives' to include all mechanisms by which organisations and individuals are motivated to perform. We differentiate between two forms of incentives:

- **performance levers** – such as the existence of targets or the routine publication of performance data, which may implicitly motivate performance but are not deliberately or primarily designed for sanctioning or rewarding purposes; and
- **sanctions and rewards** – which are designed to strengthen these performance levers by penalising or rewarding certain levels of performance or discrete actions. We identify three types of sanctions and rewards:



- financial – for example bonuses or penalties based on performance;
- operational – this includes granting organisations greater or lesser autonomy from inspection, or promoting individuals on the basis of their performance; and
- reputational – such as the publication of league tables or offender photographs.

This study focuses on the use of sanctions and rewards, but many of the principles and techniques are also relevant to the performance levers that underpin them.

Our approach

1.13 Our approach has combined:

- desk-based research on the levers used in the public sector, focusing particularly on the use and effectiveness of sanction and reward mechanisms;
- a survey of central government Departments on the sanction and reward mechanisms used in some of the government's key programmes; and
- structured interviews with policy leads in Departments, and key delivery officials, for eight of the government programmes that responded to our survey.

Annex 4 describes in detail our methods and the process of selecting the case studies.

How sanctions and rewards are currently used

1. Government programmes were identified through the Public Service Agreements in operation in winter 2007-08. Whilst the PSA framework changed in April 2008, the major programmes of Departments have continued; often now defined as Departmental Strategic Objectives.

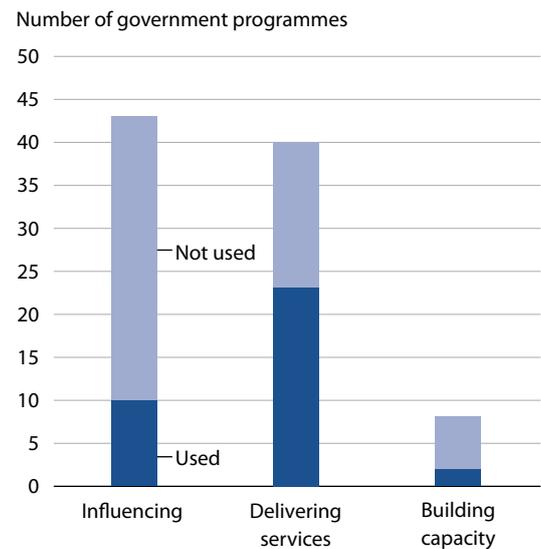
Overview

2.1 This Part describes how some of the government's key programmes are delivered, and presents the results of our survey of government programmes. We use these results to highlight, describe and interpret the various uses of sanctions and rewards.

The use and non-use of sanctions and rewards

2.2 To understand how sanctions and rewards are currently used, we surveyed 145 government programmes¹. We received 91 responses – a response rate of 63 per cent. Of these, 56 (61 per cent of those that responded) reported that they do not use any formal sanction or reward mechanisms. We categorised the surveyed programmes according to whether they were primarily delivered through services, by influence (e.g. to change people's behaviour), or by building capacity. **Figure 3** shows that explicit sanctions or rewards are more likely to exist within delivery chains that are focussed on delivering services.

3 How frequently sanctions and rewards are used for different delivery methods



Source: National Audit Office survey of Departments, and analysis of UK government delivery chains

2.3 This pattern is probably due to the difficulty in attributing actions to outcomes. It is easier to observe, measure, and therefore incentivise the contribution that individuals and organisations make when delivering a service, than when they are influencing outcomes. **Table 3** lists a sample of sanction and reward mechanisms that are used in UK public sector programmes.

Table 3

Examples of sanctions and rewards used in the UK public sector

- Renewal of Air Quality Grants to Local Authorities is dependent upon achievement of the targets specified in the original grant.
- Deductions from out-of-work benefit payments by the Department for Work and Pensions in the event of non-compliance with mandatory activities, such as attending job interviews.
- Bonus funding above a standard amount through the Planning Delivery Grant to Local Authorities, based on performance criteria related to the speed of processing applications.
- Financial incentives offered by the Department for Work and Pensions to recipients of incapacity benefits or lone parent benefits, to help them move into and remain in work.
- Bonus of up to 7.5 per cent paid to teams in Jobcentre Plus on achievement against five performance criteria.
- Green Flag Award for good management of parks and open spaces, given by the Department of Communities and Local Government to Park Management and Maintenance.
- Decreased oversight and greater freedom from central management given to hospitals who attain Foundation Trust status by meeting certain performance standards.

(For a wider list of examples see Annex 2).

2.4 Survey responses showed that more than half of government programmes did not use explicit sanctions or rewards. However, it is possible that this is an overstatement, as further research into those programmes uncovered a number of low profile mechanisms which had been overlooked by those completing the surveys. We selected eight programmes that responded to our survey for further research – four that reported using sanctions or rewards, and four that did not (Annex 1). Of the four programmes that reported not using sanctions or rewards:

- The first example occurred as part of the Department for Communities and Local Government’s objective to reduce homelessness. It reported funding, partnership working, and good relationships with Local Authorities as primary levers, and felt that the strength of these primary levers mitigated the need for formal sanctions and rewards. However, one Local Authority cited a league table – which is circulated by the Department, and ranks all Authorities according to the number and percentage of families they have managed to move from temporary to permanent accommodation – as a big incentive:

“Although the performance measures aren’t perfect, nobody wants to come last”.
- In the second example, the Department for Environment Food and Rural Affairs is renegotiating Service Level Agreements with its three partner agencies – Natural England, the Regional Development Agencies, and the Rural Payments Agency. These Service Level Agreements will enable the Department to reclaim greater levels of funding if the required levels of performance are not met.

2. "Reserve powers" allow Secretaries of State to intervene in Local Authorities.

- The third programme we examined cited the existence of a strong primary lever – in the form of legislation – as a reason for not yet having introduced explicit sanctions or rewards. The Department for Environment Food and Rural Affairs has a legislative lever to improve air quality through the Environment Act. If Local Authorities exceed certain concentration levels in any of the seven pollutants, they need to produce a more detailed assessment of the problem. That may lead to a Local Authority having to produce an Air Quality Action Plan, with annual progress reports. Although the Department has no formal enforcement powers as a regulator, it has the potential to use "reserve powers"² in the case of non-compliance or where it considers that further measures are required. In practice these powers are almost never used, being seen as an extreme step in the context of devolved government. The Department notes, however, that the quality and speed with which the Air Quality Action Plans are produced, and solutions implemented, are variable. The legislative lever alone, although strong in principle, has not in practice led to consistently good performance. The Department is considering the scope to use further sanctions or rewards.
- The fourth programme relates to the Home Office objective to improve the perceptions of the public about drug misuse in their area. Perception measures can be used in two ways: as proxies for a more substantive outcome; or as outcomes in their own right. Initially, the Department viewed public perception of the drug problem as a proxy for the problem itself. It chose not to use any sanctions or rewards around the proxy measure, primarily because at the time perceptions were not a high priority, but also because of the risk of creating incentives to improve perceptions in ways which did not reduce the underlying problem. Subsequently, the Department came to view perception as an outcome in itself. To achieve a sustained improvement in perceptions, research shows that the substantive issues still need to be addressed – whilst also engaging

local communities in developing a shared understanding of the problems, developing local solutions and communicating the outcomes of local action. It has now incorporated the perception measure in the large set of indicators used to provide sanctions and rewards for Local Authority performance. This approach recognises the intrinsic value of improved perceptions, while reducing the likelihood of perverse effects by its inclusion with a broader set of measures.

2.5 One of the issues surrounding the extent to which sanctions and rewards were used concerned the balance between central direction and local choice. In moving to more devolved decision-making, the Treasury identified an increased opportunity to incentivise performance, including the use of sanctions and rewards. This position is supported by the literature and the logic of incentivising, rather than directly managing, performance. In some cases, however, Departments reported that this devolution of priorities and budgets reduced their ability to use sanctions and rewards to incentivise the delivery of national priorities. This highlights the implicit tension in setting national priorities that can be implemented at local discretion. To manage this tension central government needs to develop information systems to analyse whether local needs have been accurately determined and use this information to influence local bodies.

2.6 The cases where sanctions and rewards had been used are discussed throughout the remainder of this paper. The following example shows how sanctions and rewards have been added to primary levers to address a specific performance concern – one of the more common ways they have been brought into use. The National Treatment Agency is a key stakeholder in the government's objective to reduce the harm caused by drugs. For many years, it and the Department of Health used the provision of guidance, the distribution of ring-fenced funding, and performance management as its primary performance levers. However, a review of unit costs showed a substantial variation in the average

amount of money being spent on each patient treated – a variation that could not be explained by local circumstances, and therefore demonstrating a variation in performance. The Agency and the Department of Health concluded that the levers used thus far could be complemented by additional incentives through the allocation process to generate a change in efficiency. For the 2008-09 planning round, the Department of Health changed its funding formula towards a more level contribution per individual treated. This has the effect of rewarding the more efficient areas, and sanctioning the less efficient. This example demonstrates how two organisations have gone through a process of identifying a desired outcome, understood what motivates the delivery chain, measured the performance of the delivery chain, and used an incentive mechanism to address a shortfall – in this case, the variable efficiency of the delivery bodies.

2.7 The discussions we had with stakeholders as part of this exercise yield a number of propositions about use and non-use of sanctions and rewards:

- There may be occasions where the strength of the primary levers over performance, or the nature of the performance measures available, mean that sanctions and rewards are not needed, or not desirable. But these issues have rarely been analysed in depth.
- In the course of running a programme of work, Departments frequently find the levers available to them insufficient to obtain the desired levels of performance.
- Sanctions and rewards do not always need to be high-intensity financial bonuses or penalties, or widely publicised programmes of “naming and shaming”. Even subtle mechanisms, such as the reputational impact of internally published benchmarking data, can have an impact.
- While in principle sanctions and rewards can help resolve tensions between central and local priorities, in practice the ability to use them depends on the precise nature of wider funding and measurement arrangements.

Which performance levers sanctions and rewards support

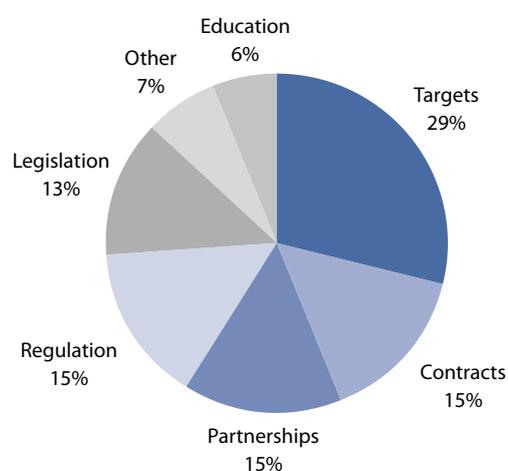
2.8 Our survey shows that targets are the single most common performance lever in delivery chains that use sanctions and rewards (**Figure 4**), with a fairly equal use of contracts, partnerships, regulation and legislation.

Types of sanctions and rewards used

2.9 From our review of the literature and our own survey work we have identified three types of sanctions and rewards:

- **Financial** – these are perhaps the most explicit, in which bonuses are offered, or fines imposed, depending on the Agent’s performance.
- **Reputational** – these try to motivate behaviour through exposing Agents’ performance. This can be done through, among other things, the publication of league tables, the giving out of awards, or “naming and shaming”.
- **Operational** – these can include the amount of audit or inspection burden, or the amount of operational freedoms granted to an Agent.

4 Programmes that used sanctions or rewards also used the following levers



Source: National Audit Office survey of Departments

3. *Tough action against underperforming hospitals and trusts.* Press Release. Department of Health. 4 June 2008.

2.10 Financial and reputational sanctions or rewards are the most common, and very little use is made of operational mechanisms (Figure 5) – although the Department of Health has recently announced the intention to make it easier for the NHS Chief Executive to impose operational sanctions on underperforming Trusts. These would allow the NHS Chief Executive to remove Trust management teams and chairs who fail to meet certain minimum safety, quality and financial management standards, without having to make large payoffs³.

2.11 Within the broad categories of financial, reputational, and operational, we have found many different types of sanction and reward mechanisms. Table 4 lists some of the different types of mechanisms we have encountered – specific examples of these different mechanisms in practice are available in Annex 2.

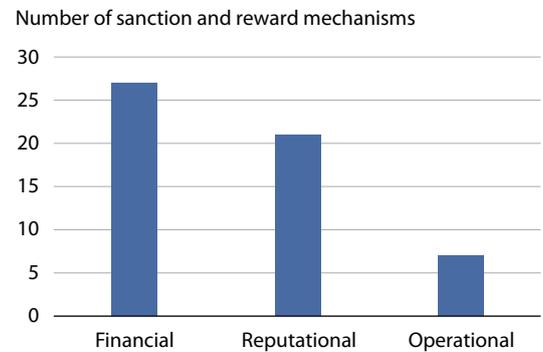
2.12 Reputational mechanisms are more likely to be used when influencing the delivery chain or building capacity, but financial mechanisms are by far the most common form for service delivery, making up 76 per cent of the mechanisms used in that category (Figure 6).

Who uses sanctions and rewards

2.13 We identified four general categories of stakeholders in a standard public sector delivery chain:

- central government Departments;
- intermediary bodies – such as Strategic Health Authorities and regulators;
- delivery bodies – such as hospitals and schools; and
- citizens or participants – often grouped by ‘demographic’, such as school pupils or benefit claimants.

5 The types of sanction and reward mechanisms used



Source: National Audit Office survey of Departments

Table 4

Categories of sanctions and rewards

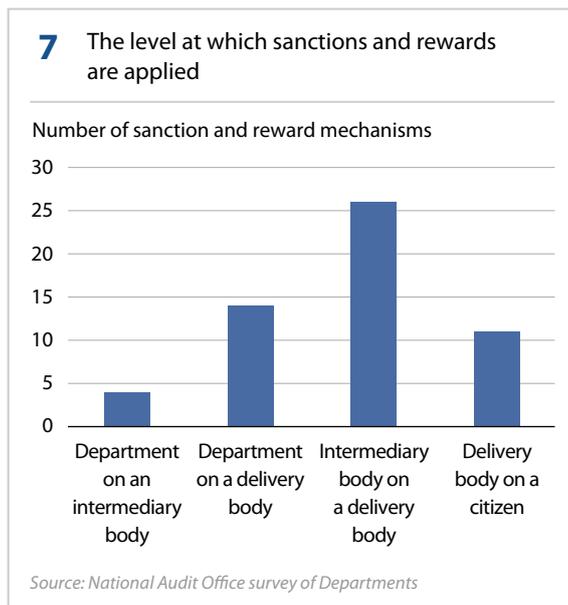
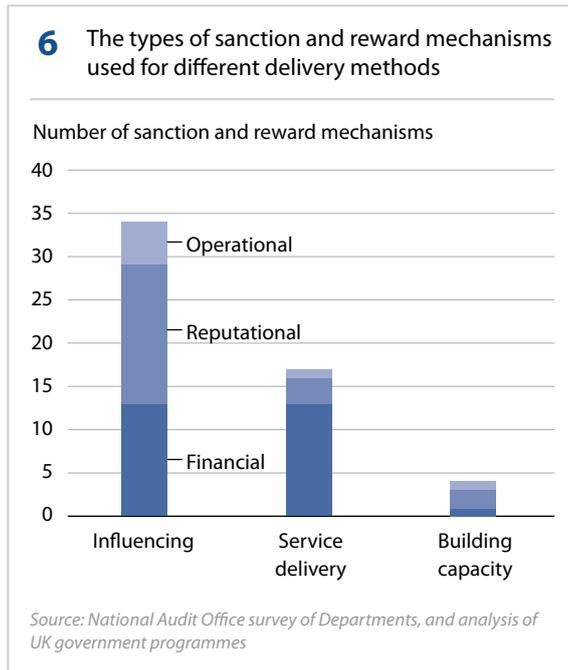
Category	Type of sanction or reward mechanism
Financial	● Contract or grant renewals.
	● Holding back payments.
	● Absolute bonuses or penalties based on reaching certain performance thresholds.
	● Unit-based bonuses or penalties that vary continuously according to the level of performance.
	● Discretionary ad hoc bonuses, in the form of cash or vouchers.
	● Sharing profits or cost savings.
Reputational	● Reduced charges, such as tax concessions.
	● Public scoring and ranking, such as school league tables.
	● Public recognition, through awards or kitemarks.
Operational	● Varying the levels of oversight through audit and inspection.
	● Direct interventions, such as interim-management, dismissal or promotion.

2.14 Our survey (**Figure 7**) shows that most of the sanction or reward mechanisms are used by intermediary bodies to influence delivery bodies. This pattern is understandable, given the relatively large number of intermediary bodies relative to Departments of State.

2.15 Figure 7 does not show to what extent Departments themselves are incentivised to achieve their key objectives. The Treasury describes a number of levers that exist to motivate performance at Department level:

- **Financial.** Performance against Public Service Agreements may have an impact on a Department’s resource allocation in the next spending review, but the relationship between previous performance and future funding is not explicit or formalised.
- **Staff performance.** This makes up one element of the appraisal process – which influences pay – but in a subjective rather than an objective way.
- **Performance management.** Nominated Senior Responsible Officers are made accountable for each Public Service Agreement, and chair a Delivery Board within each Department. They then report to their Cabinet Committees, as well as the Prime Minister’s Delivery Unit.
- **Reputational motivations.** Performance is reported in Department Annual and Autumn Performance Reports, which are scrutinised by Ministers and Department Select Committees, as well as by the public and media.

Our survey shows that Departments, their intermediary bodies, and delivery bodies, use explicit sanctions or rewards along the delivery chains, but we have found little evidence of their existence at a Departmental level.



The effectiveness of formal sanction and reward mechanisms

Overview

3.1 This Part summarises the research on the effectiveness of sanctions and rewards. Based on our literature review, survey, and interviews with government programmes, we then provide guidance on how organisations can manage five key elements of a sanction and reward mechanism: design; measurement; application; review; and implementation.

What is meant by ‘effectiveness’

3.2 There are a number of dimensions in which improved performance may be measured:

- **An increase in volumes:** for example, the increase in the number of hip operations, or in the number of people looked after by Social Services.
- **An increase in efficiency:** the same service may be provided at a lower cost, or more services provided for the same cost.
- **An increase in user satisfaction:** the quality of the public service may increase, leading to higher levels of user satisfaction.
- **An improvement in equity:** the distribution of users of a public service organisation is shifted towards the more disadvantaged – for example, an educational scheme may seek to target more users from minority groups.

3.3 The first three of these dimensions align with the standard “three Es” framework for assessing value for money: economy; efficiency; and effectiveness. A higher volume of inputs for the same cost represents an improvement in economy. Delivering the same, or more, service outputs for lower costs is an improvement in efficiency, while an increase in user satisfaction is one way to measure improvements in the way outputs are translated into outcomes – or effectiveness.

3.4 Equity objectives can sometimes conflict with efficiency objectives or dimensions. For example, increasing equity in educational outcomes may require skewing inputs heavily towards the least able pupils, whereas efficiency considerations would suggest focusing teaching inputs on to pupils in the middle of the ability spectrum. If equity concerns are important to the overall objectives of the public service which the sanction or reward mechanism is intended to support, then they must be considered at the time the mechanism is designed and reviewed.

How effectively sanctions and rewards are used

3.5 Concluding on whether specific sanctions or rewards have had direct impacts on performance is difficult. Our survey results showed that more than 60 per cent of those using sanctions or rewards believed that their use had improved performance, while the remaining 40 per cent were unsure. The various methods for reaching such a conclusion in individual cases are discussed below. There is a body of literature on incentive mechanisms (see Annex 5) that is based on economic theory, and much of this is supported by a theoretical research base, often in the private sector. However, there is less research evidence from the experience of the public sector.

3.6 There are a handful of studies in the literature based on detailed outcome data from randomised trials, almost all based on experiences in the United States. Nonetheless, the clear consensus based on the empirical evidence that exists is that public sector agents – organisations, teams and individuals – *do* respond to sanctions and rewards, and those incentives can promote enhanced effort and performance. **Table 5** overleaf summarises examples of the effective use of sanctions and rewards in the public sector.

3.7 In addition to the research summarised overleaf, two large-scale productivity reviews led by the Treasury in the UK, and the Government Accountability Office in the US, have found evidence to support the use of sanctions and rewards as an important aspect of improving performance.

3.8 The review of ‘Devolving Decision-Making’ undertaken by the Treasury and the Prime Minister’s Delivery Unit in 2003 investigated national and international best practice in health, education, crime, and local government through more than 20 site visits. They identified five critical factors that make performance measurement and management systems particularly effective, one of which was a transparent set of performance rewards and sanctions.

3.9 The Government Accountability Office (GAO) in the US has reviewed accountability for performance in the government’s Grant Management Programme, which constituted nearly 20 per cent of total federal outlays in 2005. Maximizing the extent to which grants achieve their long-term performance goals and objectives is critical to successfully addressing challenges in the US. It concluded:

“As the cases we described illustrate, rewards and penalties are fundamental tools to help drive and motivate desired behaviors, but performance accountability mechanisms are not one size fits all; there is no universal transferable mechanism applicable to all programs”. (GAO, 2006).

The GAO found a number of opportunities to improve the design and implementation of performance accountability mechanisms.

3.10 While our survey showed that more than 60 per cent of those using sanctions or rewards (18 out of 22) believed that their use had improved performance, further interviews highlighted that these opinions were based on subjective judgements, and that there was little quantifiable evidence of the direct link between the sanctions and rewards and the levels of performance. For example, the Department for Communities and Local Government points to the fact that now more than 80 per cent of Local Authorities are achieving their targets on the timeliness in processing planning applications, from only 20 per cent in 2002. However, the Department does not know the relative effects each of a number of sanction or reward mechanisms had on this performance, or, other than anecdotal evidence, whether this has had any detrimental effect on the overall quality of the process.

Table 5**A summary of examples of the effective use of sanctions and rewards in the public sector**

Type	Description of mechanism and research
Financial (Organisations) The Lewin Group et al, 2003	Mechanism: An incentive system in the US – established by the Child Support Performance and Incentive Act of 1998 – provides incentive payments to states that meet one or more of the Act's five outcome-based performance goals and associated targets, and penalises states that fall below threshold levels in certain areas. Research: A statistical regression analysis to explain changes in state incentive payments during the transition to the new system in nine American states.
Financial (Teams) Burgess et al, 2005	Mechanism: A UK scheme – known as the Makinson Incentive Scheme – in which a bonus payment was given to teams responsible for VAT collection within the former HM Customs & Excise. Research: A controlled experiment which compared the performance of two teams that qualified for bonus payments, with a team that did not qualify. It used statistical techniques to remove the effects of uncontrollable factors, such as staffing and local market conditions.
Financial (Teams) Burgess et al, 2004	Mechanism: Team-based financial incentives in the UK's Jobcentre Plus, covering five different targets – including job placements and customer service. Research: Used data from the Agency's performance management system and personnel records, plus matched labour market data over the course of the one year pilot scheme.
Financial (Teams) Burgess & Ratto, 2003	Mechanism: Financial payments to all members of staff – including teachers, cleaning staff and secretaries – in highly performing US schools. Payments were based on a statistically-derived measure of test score gains which removed the impact of external factors, such as pupils' socio-economic backgrounds. Research: Analysis of data for schools in six large cities in Texas between 1991 and 1995, comparing those that used bonus payments to those that did not.
Financial (Teams and individuals) Kahn et al, 2001	Mechanism: Performance-related pay in the Brazilian tax collection authority. This involved the payment of financial incentives based on individual and team performance in detecting and fining tax evaders. Research: Used six years of data in ten tax regions – three years before the reform and three years after. Used statistical modelling to separate out the effect of the bonus scheme from other macroeconomic and policy changes.

Assessment of success

A review of the new incentive system in a sample of nine states found that the median score on each of the five performance measures increased during the time the incentive system was implemented.

Both the teams that qualified for the bonus payments spent more time, and recovered more yield, on all trader groups than the team that did not qualify for the bonus.

The results suggest that the scheme was effective in raising job placements in certain contexts, though the overall average effect was close to zero. The impact of the incentive scheme was greatest in small offices and in districts with fewer offices, but had a negative effect in larger offices.

It also found that the quantity of output increased, but that the scheme had little effect on the quality of service. It suggests that the use of incentive pay is much more cost-effective than a general pay rise or increasing the number of staff.

The results were generally positive, in that pass rates appeared to increase faster in a city that used financial incentives than in those cities that did not.

The findings show that the scheme had a dramatic effect. Fine collections per inspection were 75 per cent higher than in the estimated counterfactual, with impact ranging from 19 per cent to 145 per cent between regions.

Reasons for success

Performance measures:

A survey of states showed that the majority approved of the performance measures that were being used to sanction and reward them, as well as the threshold levels of performance.

Implementation:

States also felt that implementing the scheme over a period of three years helped them to make all the necessary management changes.

There is no discussion about why the scheme was successful.

Team size:

Setting incentives at a small rather than large team level helps to overcome the free-rider problem, as it is easier for peers to monitor each other's performance.

There is no discussion about why the scheme was successful.

Level of reward:

The amounts involved were substantial, frequently providing bonuses over twice the average annual salary.

Table 5 continued**A summary of examples of the effective use of sanctions and rewards in the public sector** *continued*

Type	Description of mechanism and research
Financial (Individuals) Atkinson et al, 2004	Mechanism: Performance-related pay for teachers in England based on pupil scores. Research: A comparison of teachers who were eligible for bonuses against those that were not. It used statistical regression techniques to take account of uncontrollable factors.
Financial (Individuals) Dearden et al, 2004	Mechanism: The Education Maintenance Allowance involved English Local Education Authorities paying a means-tested benefit to 16–18 year olds who remained in full-time education after year 11 – when education ceases to be compulsory. Research: Analysed the participation rates of 16-18 year olds in nearly 7,000 neighbourhoods during a one year pilot of the scheme, comparing participation rates of those pupils who were eligible for the bonus payments to those that were not.
Reputational Hibbard et al, 2003	Mechanism: Public reporting of performance. A report of hospital performance in the US – called QualityCounts – is published and circulated to the public. Research: A controlled experiment, in which one group had the report circulated to the public, one group had the report just made available to themselves, and the final group received no report at all.
Operational Propper et al, 2007	Mechanism: Making performance against waiting list targets a key element of the English NHS performance management system, with failing managers being sacked, and successful Trusts given greater autonomy. This was accompanied by additional performance levers in England, such as escalating targets, and the publication of performance. Research: This was a natural experiment based on the fact that Scotland and England have similar health systems, and so the key difference in performance is assumed to be due to the differences in how they are managed – in this case, the fact that there were stronger sanctions and rewards in England than in Scotland.

Assessment of success

On average, the scheme increased pupil scores by half a grade. However, it is not known whether this improvement was at the expense of other aspects of teaching.

The impact of the subsidy is quite substantial, with initial participation rates (at age 16) being around 4.5 percentage points higher. Full-time participation rates one year later are found to have increased by around 6.4 percentage points.

Quality improvement efforts among the public-reporting hospitals were significantly greater than in hospitals given only private reports. This provides strong evidence that making performance information public stimulates quality improvement in the areas where performance is reported to be low.

The natural experiment allowed an almost direct evaluation of the incentives. However, the study is unable to determine the relative impacts of each sanction or reward.

Reasons for success

There is no discussion about why the scheme was successful.

Understanding the Agents:

The presumption of the policy makers has been that low levels of participation are due to financial constraints, and so financial rewards are required to improve them. This is supported by the finding that the effect of the scheme is largest for children coming from a poorer socio-economic background.

Understanding the Agents:

The report is viewed as affecting the hospital's public image, and therefore is a key motivator to improve areas which need improving.

Level of sanction or reward:

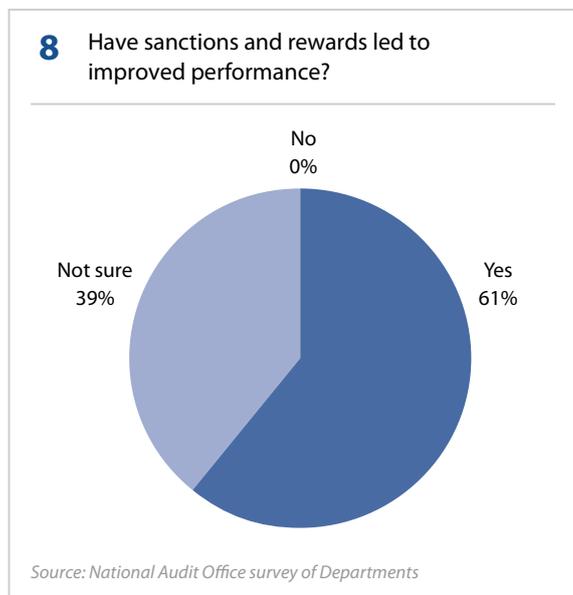
NHS managers were sanctioned more harshly (by being sacked) but also rewarded more (through greater autonomy) in England than in Scotland.

Accompanying performance levers:

Waiting list targets were made tougher, and performance made public in England, but not in Scotland.

4. Although 35 delivery chains reported that they use explicit sanctions or rewards, only 18 answered all the questions concerning design, measurement, application and review.

3.11 Our survey (**Figure 8**) also showed that nearly 40 per cent of delivery chains that used explicit sanctions or rewards do not know whether they have improved performance.



3.12 While there is limited evidence which allows us to isolate the direct impact that sanctions or rewards have on particular dimensions of performance, a review of the literature, combined with our own survey and interviews, identifies a number of lessons on their effective management. We have divided these lessons into the five key elements: designing the mechanism; measuring the 'incentivised' performance; applying the sanctions or rewards; reviewing the results; and implementing or 'rolling-out' the system.

Figure 9 overleaf illustrates this standard process for managing a system of sanctions and rewards.

How to manage sanction and reward mechanisms effectively

3.13 The following sections identify some of the practices within each of the five elements of a sanction or reward mechanism mentioned above, drawing out some general conclusions and guiding principles.

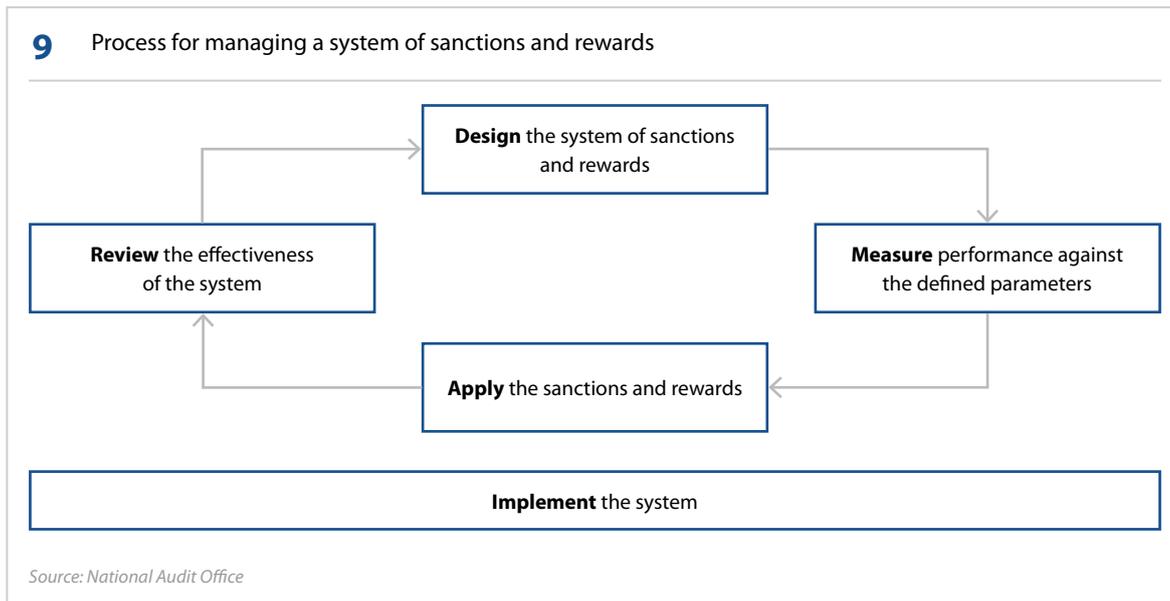
Design

3.14 The design phase of any system is the most critical. Mechanisms need to be carefully tailored to local conditions to be effective. Our survey showed that nearly a quarter of the 18 programme delivery chains who use explicit sanctions or rewards⁴ were unsure about whether their mechanisms were designed using formal research or evidence. One of the key reasons given for this is that the systems were inherited, and the knowledge underlying their design had passed from the corporate memory.

3.15 There is a well-established economics literature which considers the optimal design of incentives. The economic theory is framed in general terms as a problem whereby a "Principal" needs an "Agent" to deliver its objectives, but the Principal cannot fully control (or even observe) the activity of the Agent. In the public sector, such relationships arise throughout delivery chains. This results in policy leaders (the Principals) actively considering what levers they have to motivate the activities required by organisations, employees and citizens (the Agents) to deliver the desired outcomes.

3.16 In the simplest case, when the Agent's output is fully observable, and the value of the output is a perfect indicator of effort, the optimal contract is for the Principal to pay the Agent the full value of their output. But when the Agent's output also depends on some random factor outside their control, the Principal cannot infer the Agent's effort precisely. If the Agent is risk averse, then the optimal incentive scheme consists of a fixed payment – which acts as a risk premium – and a marginal payment proportional to the output – which acts as a reward. The optimal balance between these elements – known as the 'intensity' of the incentive scheme – depends on four factors:

5. See Milgrom and Roberts (1990) and Burgess and Ratto (2003).



- the extent to which additional performance results from extra effort, or the ability of the Agent to influence the outcome;
- the accuracy with which activities can be measured;
- the Agent's tolerance for risk; and
- the Agent's responsiveness to incentives.

3.17 The key insights are that the more risk averse the Agent, the more the output depends on factors outside the Agent's control, and the less accurate the measurement of outputs, the less 'intense' the optimal incentive structure⁵. As a result of these insights, the literature identifies certain design features and considerations which enhance the effectiveness of a system of sanctions and rewards:

- Understand the strategy and delivery chain.
- Involve stakeholders – develop mechanisms in collaboration with the Agents to ensure consensus and appropriateness.
- Understand and consider the motivations of Agents.
- Ensure Agents can influence the outcomes which are being incentivised.

- Select the right performance measure – ensuring the incentivised outcome, or intermediate output, leads to the overall desired outcome.
- Motivate all stakeholders.
- Take into account the team structures of the organisation.
- Ensure sanctions or rewards are of sufficient value or consequence to incentivise the desired behaviour.
- Introduce safeguards to protect against unintended behaviour.

3.18 Each of these features is discussed below in relation to sanctions and rewards, though these lessons are of wider relevance in designing a system of incentives, including the performance levers.

Understand the strategy and delivery chain

3.19 To design an effective system of sanctions or rewards, it is first necessary to understand the context for their use. This requires an understanding of the drivers and the motivations that influence performance against the objective (in the form of a strategy map) as well as an understanding of the various stakeholders and the performance levers that are employed to influence them (in the form of a delivery map). Examples of these maps, along with some guidance, are available in Annex 3.

6. Communities and Local Government (2006).

Involving stakeholders

3.20 Involving the Agents to whom the sanction or reward mechanism will be applied builds engagement and consensus. It can help clarify any potential trade-offs between multiple Principals and objectives, and help the Principal better understand the motivations of the Agents.

3.21 Giving Agents some degree of control of the sanction and reward mechanisms also ensures that targets are achievable and relevant locally. The US Government Accountability Office concluded this was an important factor in the success of a vocational and technical education programme, where individual states were given responsibility for developing their own performance measures (GAO, 2006).

3.22 Similarly, the UK government has recognised the importance of involving Agents in the development of incentive mechanisms. In a report on devolving government, giving local organisations the authority and freedom to create local incentives has been identified as key to driving improvements (HM Treasury, 2004a). The involvement of Local Authorities has been a key feature in the setting of Local Area Agreements, which were introduced as a result of the Local Government White Paper⁶, and is especially relevant given the link to financial rewards in the form of Performance Reward Grants.

The motivation of Agents

3.23 It is difficult to find objective evidence on motivation, but there are studies which show that introducing financial sanctions and rewards can have unintended consequences when at least some Agents have altruistic or other non-financial motivations.

3.24 An experiment at a day care centre in Israel found that the introduction of a small financial penalty for parents arriving late to collect their children actually increased the number of late arrivals (Levitt and Dubner, 2005). Before the fee was introduced, the only penalty was the social or moral

one associated with making the day care providers wait longer. By allowing parents the possibility of paying for late collection, the daycentres reduced the perceived element of social disapproval (Seabright, 2004). This shows that where intrinsic or non-financial motivations dominate – such as social or moral conscience – the introduction of financial incentives may undermine the value of the rewarded task, and have a counterproductive effect. We note, however, that the unintentional effect of the sanction in this example may not necessarily be a bad one. Parents are willing to pay for the additional time their children are being supervised, and the centre is rewarded with additional money to pay for it. If the “sanction” is set at an appropriate level, the net effect could therefore be an overall increase in utility derived from the day care centre.

3.25 Sanctions or rewards can help to focus Agents on tasks or outputs that would have been neglected if intrinsic motivation alone had been relied upon (Prentice et al, 2007). Where intrinsic motivation is high, operational or reputational sanctions and rewards may be more effective. In particular, reputational mechanisms which recognise good achievement, or highlight poor performance, may act to strengthen intrinsic motivation. For example, several stakeholders responsible for school performance argued strongly that teachers are primarily motivated by an intrinsic desire to teach effectively. In this area, Ofsted pointed to the motivational value of their events for the head teachers of schools judged to be outstanding, which allow peer recognition of good performance, and possibly assist career progression.

3.26 Where there is a mixture of potential motivations, incentive schemes should be designed to allow for the possibility of both types of motivation being present. For example, researchers in Sweden found that introducing financial rewards for donating blood reduced the overall number of donors (Gneezy, 2003) – this effect is referred to as “crowding out”, as the financial rewards displace

the social and moral motivation. The researchers found that by allowing individuals to give the payment to charity, the crowding out effect was fully counteracted. This incentive scheme attracted those motivated by self-interest – through the monetary reward – while simultaneously not deterring those who would have donated blood anyway out of intrinsic motivation. In the UK public services context, Le Grand (1995) provides the example of the GP fund-holders scheme, where GPs were allowed to keep any surplus on their funds, but only use this surplus for a purpose beneficial to patients. He argues that this scheme allowed for the motivations of (what he terms) “knights” and “knaves”, in that the surplus can be used to benefit both the GPs and the public – for instance, through improving premises or employing more staff, thus easing work pressure and providing a better or more relaxed service.

Influence over the desired outcome

3.27 A sanction or reward mechanism should incentivise outcomes which can be sufficiently influenced by the Agent’s actions. Unless the Agent can expect their behaviour to affect the outcome, the mechanism provides little motivation to strive for the desired levels of performance. This means that, as far as possible, the sanction or reward mechanism should strip out aspects which cannot be influenced by the Agent.

3.28 A good example of this is the “Contextualised Value Added” measure of school performance in England. By adjusting for prior attainment, this measure helps to account for factors beyond the school’s control, and is generally seen as a better performance measure than the previous “gross output” measure which did not make any kind of adjustment. This approach ensures that even those Agents starting from the lowest base are able to influence measured outcomes. For example, “threshold” schemes, which target absolute levels of performance and do not take past performance into account, may not reward Agents who improve the most as result of starting from a lower base. Allowing

Agents to be measured against their own past performance can help to resolve this issue, so long as it does not unfairly penalise those that performed well in the past.

3.29 If the influence of the Agent on the outcome cannot be satisfactorily isolated, then a lower-intensity sanction or reward mechanism may be more appropriate. High-intensity sanction or reward mechanisms – which link a large proportion of pay, reputation or operational freedoms to performance – can impose too much risk on Agents if final outcomes are beyond their control (Grout and Stevens, 2003).

3.30 If an Agent has limited influence over outcomes, other aspects of performance over which the Agent has more control could be incentivised. For example, the Department for Environment Food and Rural Affairs and the Department for Transport have a shared objective to improve air quality. Local Authorities will have an influence over concentrations of air pollutants in their area, but these may also be impacted by factors which can only be influenced at a national or even international level – such as engine standards, and background emissions from aviation and shipping. Interviews with Local Authorities suggested that incentivising particular actions which are known to have a worthwhile, if minor, impact on air quality – such as awareness campaigns and car pool schemes – may be more effective than trying to design a sanction or reward mechanism which targets absolute levels of air pollutants. However, unless the strategy map (**Figure 1**) linking those outputs to the desired outcome is well understood, the overall effect could be perverse e.g. setting up lots of car pooling schemes that do not get used. This approach should therefore be avoided where links between outputs and outcomes cannot be modelled. If they are used, they may need to be accompanied by activity or quality criteria which act to limit potential distortions e.g. to avoid too many awareness campaigns that start to have diminishing returns.

Select the right performance measure

3.31 The performance measures that are used to sanction or reward performance are important. The wrong choice of metric can lead to Agents focussing their behaviour on some areas to the detriment of others. Where there are several aspects to overall objectives, a successful sanction or reward mechanism should capture all significant aspects of these objectives. It is likely that a range of measures, rather than a single performance measure, will be needed. At the same time, care must be taken to keep the total number of performance measures in a sanction or reward scheme manageable.

3.32 Often the desired outcome is a long term one, for example, that education or training programmes improve lifetime employability. This conflicts with the need for performance measures to be available quickly for incentivising and accountability purposes. This tension can lead to perverse incentives. An example from the US concerns the administrators of a job training programme, who were judged on the employment status of participants soon after they completed the scheme. This led to a situation whereby the scheme administrators were incentivised to discourage participants from seeking further training, regardless of the benefit to the participants' long-term goals, and to move straight into employment (Heckman et al, 2002). Therefore, if short-term outcomes are measured, then these should be demonstrated (empirically or through expert opinion) as having a clear correlation with the desired long-term outcomes. An example of this approach is the Department of Health and the National Treatment Agency incentivising Drug Action Teams to increase the number of users in treatment programmes through linking a proportion of their funding to activity data. In this case, short-term outputs are measured (numbers of people engaged in drug treatment), because outcomes in drug use are long-term, often taking years to emerge. However, there is research evidence that demonstrates a correlation between the activity measured and the outcome (people managing their addiction), meaning that activity data can be used as an acceptable proxy.

3.33 The literature also contains examples of cases where the choice of outcome measure has led to "gaming" or strategic behaviour by Agents, such as anecdotal evidence of "cream-skimming" by schools to select the best students. Threshold measures – measures concerned with achieving an absolute level of performance, such as the number of pupils achieving GCSE grade C or higher – are seen as particularly vulnerable to gaming. This may result in Agents concentrating their effort on those organisations or individuals who are performing just below the threshold, to the detriment of the very good or very poor performers. In some cases this may be addressed by the natural inclination of workers in the public sector to help the most disadvantaged. However, where natural inclination is not enough, there is the option of directly incentivising Agents to specifically prevent widening inequalities in society:

"The Government has recognised this risk and during the 2000 Spending Review introduced targets within Public Service Agreements (PSAs) that directly address inequalities. These targets make explicit reference to the outcomes to be achieved for less well-performing groups as well as the rest of the population." (NAO, 2007).

3.34 If outputs cannot be easily specified, or if there are multiple objectives or tasks where some are more easily measured than others, it is likely that a high-intensity sanction or reward mechanism will not be appropriate (Prentice et al, 2007). In this case, such a scheme is likely to give rise to ineffective incentives, or to encourage the misallocation of effort by Agents. Rather, lower intensity mechanisms should be used, such as operational or reputational schemes, or financial schemes where only a limited proportion of compensation is tied to performance (Burgess et al, 2005).

3.35 Where outcomes are not easily measurable, or where there is a danger of unmeasured outcomes being neglected, a sanction or reward mechanism which combines objective performance measures with subjective reviews may provide a solution. An example of this is the incentive scheme for schools in England, which combines test results with a broad

(but inevitably more subjective) Ofsted assessment of the school's overall performance. Similarly, the Department of Health and the National Treatment Agency incentivise Drug Action Teams to increase the number of users in drug treatment programmes. They do this through tying central funding to activity data – as discussed above – but have supplemented this with other activities. These include Thematic Reviews, which are conducted in partnership with the Healthcare Commission, and look at other aspects of drug treatment to ensure that this measure is not focussed on at the expense of others.

Motivate all stakeholders

3.36 There is a risk that some performance levers and sanction or reward mechanisms only motivate a small proportion of the stakeholders – namely, the best and the worst performing. In 2000 the Government recognised that Public Service Agreements which targeted only aggregate performance figures could be achieved simply by focussing on a few stakeholders, while the rest remained either static or even deteriorated. In the 2000 Spending Review it therefore set targets that directly addressed the particular issue of preventing inequalities among the target population, and raising the standards of all. For example, the Department of Health set a target to:

“reduce adult smoking rates to 21 per cent or less by 2010, with a reduction in prevalence among routine and manual groups to 26 per cent or less”.

However, these kinds of inequalities targets need to be set carefully in order to have the appropriate motivating effect (NAO, 2007).

3.37 Sanction and reward mechanisms run a similar risk. If they only reward the best performers and sanction the worst, then the majority of stakeholders may be content to maintain a mediocre level of performance with no real motivation to improve. In 2005 the Audit Commission sought to address this problem by introducing a “direction of travel” score for all Local Authorities in its annual Comprehensive Performance Assessments. This meant that while there was motivation to achieve a high absolute

score, there was also a motivation for all Local Authorities to be improving their performance year-on-year.

Team-based sanctions and rewards

3.38 Influence over an outcome may be at the level of the team rather than at the level of the individual Agent. Makinson (2000) notes that teamwork better reflects the way in which most public servants work, and measures of output are often only available at the level of the team in the public sector. For example, learning in a school is typically a cumulative process rather than one which can be attributed to an individual teacher, and hence it may be more appropriate to reward a team, or the entire organisation, rather than a particular individual. The Government Accountability Office has concluded that schemes offering significant financial incentives to all staff members in some US schools – including maintenance staff and bus drivers – have contributed to significant performance improvements. This is partly attributed to the camaraderie and collaboration fostered by the scheme among the employees.

3.39 However, the key risk of team-based mechanisms is the issue of free-riding. The larger the team, the more serious the free-riding problem, as it becomes increasingly easy for an individual to disguise their lack of contribution. But free-riding can be mitigated by peer pressure if team members can observe each other's effort. This means that the smaller the team, the more likely team members are to be able to monitor each other and reduce free-riding. Burgess et al (2005) found evidence to support this when they assessed some incentive schemes that were being piloted at HM Customs & Excise and Jobcentre Plus. Where this is not possible or sufficient, free-riding can also be mitigated through incentive schemes which combine both team and individual performance, as they require teamwork, as well as valuable individual contribution.

3.40 The use of team-based measures may also help to overcome the difficulties of retaining intrinsic motivation within a sanction or reward mechanism, discussed in the previous section. A team-based sanction or reward scheme can maintain the collaborative, collective ethos of a public sector organisation without too great a focus on potentially divisive, individual financial gains or achievements. Furthermore, Makinson (2000) notes that it is easier to integrate explicitly financial sanction or reward mechanisms with non-financial incentives in a team framework. For example, having an “office of the month” is less divisive than an “employee of the month”, yet the former can still be rewarded by a shared financial bonus.

Set sanctions and rewards at a cost-effective value

3.41 The aim of a sanction or reward mechanism is to obtain a cost-effective increase in performance. In other words, the cost of the reward for additional units of performance must be proportionate to the value of the additional units of performance – both the value to the Principal, and the increased effort required by the Agent. If there is a disconnect between the costs and benefits in a sanction or reward mechanism, it will not motivate the intended additional effort, and will lead to the inefficient use of resources.

3.42 Whether the mechanism is financial, operational or reputational, the value or consequences of the sanctions or rewards need to be sufficiently high to motivate the desired behaviour. Indeed, the Government Accountability Office concluded that inappropriately valued incentives appear to be one of the main reasons sanction or reward mechanisms fail (GAO, 2006).

3.43 Through our case studies, we found several examples where the size of the sanction or reward was widely seen to be insufficient to have the desired effect:

- Tariffs paid to farmers to provide “environmental goods”, such as retaining hedgerows, under the Rural Development Programme for England have been overtaken by rising

commodity prices. This rise has led to a drop in the uptake of environmental contracts, and an expectation that a number of existing contracts will not be renewed.

- The Vehicle Excise Duty concession available to Heavy Goods Vehicle (HGV) operators in possession of a Reduced Pollution Certificate is much smaller than the capital investment required to achieve the necessary emission standards. This has led to take-up of the scheme being low, with only around 15 per cent of HGVs in possession of a certificate.
- Various UK public sector performance-related pay schemes, reviewed by Prentice et al (2007), were deemed too small. For example, awards of £50 to £150 paid in the Defence Aviation Repair Agency, and a bonus of £400 paid to all staff awarded an “exceeding” performance rating in the former Department for Constitutional Affairs. Makinson (2000) recommended that performance-related pay bonuses should represent at least five per cent of base salary in order to be effective.

3.44 By way of contrast, the relative success of a job training partnership programme in the US is attributed to the size of the rewards, whereby the programme awarded local training centres as much as a 20 to 30 per cent increase in their regular budgets if they successfully trained or placed customers. These awards were particularly valued as training centres could use them more flexibly than regular budget allocations (Heckman et al, 2002).

3.45 As well as rewards being of adequate value, sanctions should correspondingly be stringent enough to incentivise behaviour. Propper et al (2007) discuss the hospital waiting time target scheme in England, which used the dismissal of key managers of hospitals as a sanction for poor performance. In comparison to Scotland – which had a similar focus on reducing hospital waiting times, but without the high-intensity sanction regime – there is evidence that the English scheme significantly reduced waiting times.

Example 1

Department for Environment, Food and Rural Affairs – Rural Development

The Department's Service Level Agreements (SLAs) with its main delivery bodies (the Forestry Commission, Natural England and the nine Regional Development Agencies) included little scope for it to sanction poor performance with respect to rural development. Most of the funding flows from the European Commission, and is subject to their regulations, which if not satisfied raises the possibility of disallowance. The current SLA with the Regional Development Agencies includes a limited sanction, whereby the Agencies are exposed to five per cent of the disallowance risk if they fail to deliver. However, the funding received from the Rural Development Programme for England is small in comparison to the Agencies' other funding streams, and so the Programme is unlikely to be a high priority. The Department is looking to increase the risk percentage shared with the Agencies.

3.46 Affordability constraints can limit the size of rewards, and large financial sanctions for below average performers may not be politically acceptable – such as cutting the funding to a poorly performing hospital. The following examples, from both the literature and our own case studies, show why modelling the costs and benefits of an incentive scheme can help to determine the cost-effective level of a reward or sanction:

- In schools in England and Wales, a new upper pay scale for teachers related to performance was introduced in 1998, but abandoned six years later as it became apparent that the scheme was unaffordable. This was due to a flawed scheme design, which initially brought around a third of teachers into the highest pay scale, but over time brought in far more individuals than expected. This eventually made the scheme financially unrealistic, and resulted in the two top scales being taken away.
- In the introduction of the GP contract, the NAO (2008) found that a lack of understanding of the costs and benefits resulted in a large overspend and a decrease in overall productivity. The new contract for general practice contributed to improved recruitment and retention of GPs, with numbers increasing from around 27,000 to nearly 31,000 since 2003. However, the contract cost the Department £1.76 billion more than it had originally budgeted. The NAO found that in the first two years of the contract, productivity had fallen by an average of 2.5 per cent per year. While the number of consultations with patients had increased, these were not in proportion with the increase in costs. The largest overspend of the contract was due to an underestimation of the amount that GPs would earn from the pay for performance scheme, known as the Quality Outcomes Framework (QOF). In addition, the links to final outcomes were not proven; while there was evidence that the QOF has improved consistency in the quality of care, it is too early to say if overall patients' health has improved as a result (NAO, 2008).
- NAO (2001a) highlights the example of HM Customs & Excise, which developed a model to estimate the current and future levels of smuggling and assess the impact of HM Custom & Excise activities – including the sanctions and rewards – on the level of smuggling. This helps the Department to make resource decisions, allowing it to understand the benefits of achieving particular levels of performance, and so determining the appropriate level of a sanction or reward that should be associated with a particular level of performance.
- The Department for Transport carried out a cost-benefit analysis to provide an indication of the size of any concessionary level of Vehicle Excise Duty to Heavy Goods Vehicle operators that would be necessary in order to incentivise the uptake of low emissions Euro V vehicles. This analysis balanced the expected costs of the scheme against monetised benefits to health and the environment to determine if the scheme was cost-effective.

7. *Ambulance times misreported*. The Guardian. 15 August 2006.

8. *A&E wards accused of fiddling waiting times to hit targets*. Daily Telegraph. 6 October 2006.

Introduce safeguards against unintended behaviour

3.47 Experience has shown that, particularly for high-intensity sanction and reward mechanisms, undesirable manipulation of the system may occur. Examples include waiting list targets in the health service, where NHS Trusts manipulated data to show better results than in fact had been achieved (NAO, 2001b). In the case of response time targets, also in the health sector, authorities exploited lax definitions in the targets to measure response in partial ways⁷, or altered their procedures so that they could legitimately claim that response had started earlier in the process⁸.

3.48 Several mechanisms have been used to counter such behaviour, including the validation or audit of underlying data systems, independent assessment of performance achieved, and tighter specification of what constitutes “performance”. Academics have also suggested introducing a degree of uncertainty over the quantum of, or precise basis for, sanctions and rewards, to prevent over-calculated management responses.

3.49 However, the existence of incentives – whether they are performance levers or explicit sanctions or rewards – may not necessarily lead to dysfunctional behaviour. Kelman and Friedman (2007) suggest three circumstances in which dysfunctional behaviours do not arise – or are at least limited:

- When additional effort in targeted activities may complement rather than substitute for other desired, but untargeted, activities. For example, in order to improve their performance in admitting patients from Accident & Emergency, hospitals reengineered processes that also improved the performance in other areas.
- Self-limitation, whereby the dysfunctional response provokes a countermanding action. For example, unnecessary admissions to inpatient hospital wards from Accident & Emergency create a burden for managers of inpatient wards, who then object to those types of admissions.

- When management take actions to limit dysfunctional behaviours, such as adding new dimensions to a target, or by cultivating the public sector ethos among employees. For example, having a target to reduce elective surgery waiting times at the same time as having a target to reduce Accident & Emergency waiting times may have helped to limit hospitals from substituting their efforts from elective surgery to Accident & Emergency.

These issues need to be considered at the design stage of a programme, and responses selected and then reviewed after initial operation.

Conclusions

- The Principal needs to understand the strategy for achieving the objective – as well as the delivery chain that will help to achieve it – before considering the use of performance levers and the accompanying sanctions and rewards.
- The sanction and reward mechanism should be developed together with the Agents so as to ensure consensus and appropriateness.
- The sanction and reward mechanism should consider the motivations of Agents.
- The Agents must be able to influence the outcome which is being incentivised.
- The incentivised outcome – or intermediate output – must lead to the overall desired outcome from an organisation.
- All stakeholders should be motivated to improve.
- The sanction and reward mechanism should take into account the team structures of the organisation.
- The sanctions or rewards must be of sufficient value or consequence to incentivise the desired behaviour.
- If there is a risk that sanctions or rewards could lead to unintended behaviour, mechanisms must be put in place to minimise this.

Measure

3.50 This stage of the process is only relevant to sanction and reward mechanisms that are linked to measured levels of performance. Sanction and reward mechanisms can be categorised according to whether they are linked simply to compliance, or whether they are linked to measured performance. For example, concessions from Vehicle Excise Duty are offered to vehicles that hold a Reduced Pollution Certificate. This reward mechanism relies on the process of assessing whether a vehicle fulfils the requirements for the certificate – a simple yes or no decision. However, the allocation of funding to Drug Action Teams through Pooled Treatment Budgets needs more complex data systems that capture activity data in drug treatment centres throughout the country.

Development of systems to collect performance data

3.51 Once the performance measures to be incentivised have been chosen, it is necessary to ensure adequate systems are in place to collect data on the chosen measures. Data systems should ensure that data are well-defined, timely and reliable⁹. The NAO has regularly found that less than 50 per cent of the data systems underpinning Public Service Agreements to be fit for purpose (NAO, 2004-07).

3.52 The Government Accountability Office (2006) points to the importance of efforts by the Office of Management and Budget (the body responsible for improving the performance of federal programmes) to improve agencies' abilities to develop high quality, results-based performance measures as a prelude to introducing incentive schemes. Similarly, the UK government has identified accurate and timely publication of performance data as essential to strengthening incentives through sanction or reward mechanisms (HM Treasury, 2004a).

Example 2

Department of Health and National Treatment Agency – treatment for drug misuse

The National Treatment Agency is responsible for increasing the number of problem drug users in effective treatment programmes. As a result of substantial investment into data systems, the Agency has access to robust and extensive data which it can use to allocate funding on a fairer basis. Using a common, trusted data system has been critical in getting buy-in from stakeholders for this new funding mechanism, which rewards those who are using less central funding per individual treated with more money, and sanctions the relatively inefficient areas with less. Although there are moves towards tying elements of the funding to the levels of drug use in an area – rather than the volume of treatment activity – the quality of data at a local level is not good enough yet.

9. *Choosing the right FABRIC, A Framework for Performance Information.* HM Treasury, Cabinet Office, National Audit Office, Audit Commission, Office for National Statistics.

Analysing and reporting the performance data

3.53 Once the collected data have been validated, they need to be analysed, interpreted and then reported in a clear and concise way so that all stakeholders are able to understand performance. Our survey showed that 94 per cent of respondents believed that their analysis of performance data was robust enough to make reliable decisions when applying sanctions or rewards. However, other work undertaken by the NAO on the quality of data systems underpinning major government programmes has found significant weaknesses (NAO, 2004-07).

3.54 For New York's Compstat programme – which allows detailed crime patterns to be analysed and crime reduction plans to be developed – a 25 person team is responsible for central analysis and quality assurance of the data collected from each local precinct. This then paves the way for the weekly accountability meeting (HM Treasury, 2004a).

Conclusions

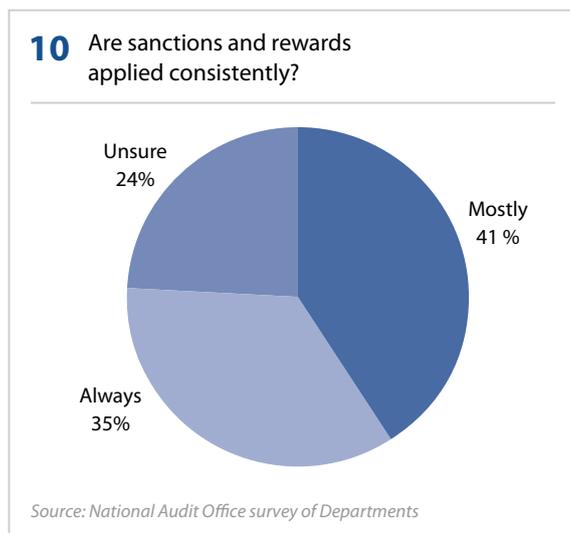
- Data systems should ensure that data are well-defined, timely and reliable.
- Data systems need to be supported by a reliable management structure that analyses and reviews the data.

Apply

3.55 For a system of sanctions or rewards to have the desired incentivising effect, it must be applied in a timely, consistent, and transparent manner.

3.56 The Department for Communities and Local Government told us that as of 2008 it is using a Housing and Planning Delivery Grant to incentivise Local Authorities to provide housing and to put local plans in place. However, the housing element is based on average delivery over the last three years, which means that current actions to increase housing supply will only gradually influence grant allocations, which may weaken the incentivising effect.

3.57 If rewards or sanctions are not awarded as expected, Agents learn that their additional efforts are not worth the cost or risk. The Government Accountability Office (2006) notes that if rewards are paid indiscriminately, or sanctions are not levied as expected, Agents can learn that no additional effort is required to benefit. **Figure 10** shows that over 40 per cent of delivery chains that responded to this question acknowledged that they are not always applied consistently, with nearly a quarter not even knowing whether they were or not.



3.58 Consistency in application is one of the principles highlighted by Macrory (2006) in his report on effective regulatory sanctions. He notes that failure by regulators to follow up low-level enforcement actions – such as warning letters – with the threatened sanction, means that they will not be taken seriously and credibly by firms. Consistency is particularly important in the public sector, where Agents often have more than one Principal trying to influence them. Given the many, and potentially conflicting, objectives held by different Principals, it is possible that Agents may play off one against another (Marsden and French, 2002). Issues with multiple Principals can be mitigated by ensuring that levels of performance which trigger a sanction or reward are clearly defined and understood, and are consistently applied without exception.

3.59 Sanction or reward mechanisms need to be transparent, not only for those to whom they will be applied, but for the motivation of other stakeholders. If an Agent does not know why they were sanctioned or rewarded for something, then the motivational value of the sanction or reward is diminished, as the Agent is not clear what is expected of them. As far as other stakeholders are concerned, if they cannot see how sanctions or rewards are applied to others, they may not feel motivated to behave accordingly. Senior staff at the Dorset and Somerset Strategic Health Authority believe that transparency is one of the key characteristics that made their incentive scheme successful:

“The rewards and sanctions structure is not a private bargain between management teams, in the manner of some NHS contracting or commissioning. Instead, the system is highly visible externally. That visibility goes beyond organisational boundaries and sends a message to frontline staff. Once given to a Trust, staff will reasonably expect to see how the money is spent. The fact that the money flows throughout a financial year makes it simpler for Trusts to plan investments arising from their success. Crucially that planning is in the knowledge that reward is guaranteed if they meet their targets. As a result, the link between achievement and benefit is both immediate and transparent.” (HM Treasury, 2004a).

Conclusions

- Sanctions and rewards need to be applied in a **timely** manner to have the desired motivating effect.
- Sanctions or rewards need to be applied fairly and **consistently**.
- The process of applying sanctions or rewards needs to be **transparent**.

Review

3.60 Once a system of sanctions or rewards is in place, the Principal should evaluate whether it is having the desired effect on performance. This is complicated by the many other factors that impact on performance, making it difficult to isolate the unique contribution of the sanctions or rewards themselves. However, particularly where there is significant risk to the success of the policy – or a large amount of resource is involved – such evaluation is essential. Without it, flaws in the design, measurement, and application of the system may be causing unintentional behaviour, leading to the failure of the policy and wasted resources.

3.61 Prendergast (2001) found that unmeasured outcomes in a Los Angeles police force sanction scheme led to a focus on only some aspects of desired outcomes. The scheme was designed to combat violence against criminals during arrests, yet led to a disproportionate fear of disciplinary action in case of complaints from incorrectly arrested citizens. Given the trade-off between arresting the innocent and not arresting the guilty, the scheme resulted in officers being strongly biased towards the latter. Here, the overly punitive system led to the inefficient outcome of letting too many criminals free, or not arresting them in the first place. This hindered the achievement of the institution's overall objective of fighting crime, as evidenced by the rapid increase in homicide rates after the introduction of the scheme.

3.62 In our survey, 60 per cent of the delivery chains that used explicit sanctions or rewards, and responded to this question, believed that they had improved performance. However, interviews with the

organisations involved highlighted little evidence of the direct link between the sanctions and rewards, and the levels of performance.

3.63 Reviews need to be both fit for purpose, and sufficiently regular. With different types of government programmes, and different systems of sanctions and rewards, the way in which effectiveness is assessed needs to be tailored accordingly. Various approaches are available for assessing the effectiveness of sanction and reward mechanisms. Evaluations can be grouped into the following types:

- **Randomised trials or experiments.** These are the most robust assessments, as they use a clearly defined and monitored control group to understand what performance would be without any intervention. This allows the causal relationships between the sanction or reward mechanism and the measured outcomes to be identified. However, this approach is expensive, time-consuming and can be politically difficult – as some arbitrary recipients benefit from the programme while others are excluded. As a result, there are very few such studies.
- **Natural experiments.** In some cases, natural experiments can provide a good alternative to randomised trials. One example in a UK public sector context is the comparison of schemes regarding hospital waiting time targets in England and Scotland, where the English imposed a high-intensity sanction regime while Scotland did not (Propper et al, 2007). Because of the similarity of the two systems, many of the other determinants of outcomes – besides the sanctions – were not relevant to changes in performance, and so differences could be attributed, within reason, to the sanctions themselves.
- **Econometric or statistical analysis.** This can be used to compare performance before and after the introduction of a sanction or reward mechanism, as well as to compare two similar populations, one of which does, and one which does not use a mechanism. It makes the comparison by statistically removing the factors which may affect the outcome, and which vary over time

or between populations – such as the level of funding or the demographics of the population. The robustness of these kinds of studies depends critically on the quality and quantity of data, but even when the data is good, causation is much harder to determine than correlation.

- **Hypothesising.** Data on outcomes, and on potential explanatory variables other than the sanction or reward mechanism, is often difficult, expensive, or slow to obtain. In these cases, the assessment of effectiveness is limited to examining the path of output measures over time, and hypothesising links with the introduction of (or refinements to) sanction and reward mechanisms. Net output measures, which attempt to measure the value added by the programme, are preferable to gross or raw output measures. One example of a net output measure is the “Contextualised Value Added” measure of pupil test performance used in English school league tables since 2005. This measures the change in exam performance for each pupil between key stage assessments in an attempt to get at the value added by the school. This has been introduced alongside the gross output measure of absolute number of GCSE grades, which could be influenced by many factors outside the school’s control.
- **Qualitative evaluation.** Finally, in some cases the effectiveness of a sanction or reward mechanism is assessed on a purely qualitative basis – for example through questionnaires and/or structured interviews. This approach is inevitably subjective, but can provide insights to complement more quantitative approaches.

3.64 As these methods show, robust assessment of effectiveness is difficult. Reviews of the academic literature find only a small amount of empirical evidence as to whether these mechanisms actually achieve the goals of the programme or Department (Burgess et al, 2002). Prentice et al (2007) found only seven empirical studies on UK public sector financial incentive programmes after 1999.

3.65 Nevertheless, where feasible, scope should be built in to revise and update the sanction or reward mechanism in the light of the best available evidence on its effectiveness. For example, in a vocational and technical education programme (GAO, 2006), officials have attributed the programme’s success to the ability of American states to periodically revise their target and measures during annual negotiations of their state plans.

3.66 In addition to the risks of unintended consequences emerging over time, technology may be introduced that alters performance expectations, or the priorities of an organisation may change (GAO, 2006). Certain performance standards may also become the norm and may no longer need to be incentivised. These are all reasons for why such mechanisms need to be continually reviewed.

Example 3

Department for Communities and Local Government – Planning Development

The Department linked the financial reward of its Planning Development Grant (PDG) to the speed with which Local Authorities processed planning applications. This led to significant improvements in performance over the five years PDG was paid. However, by the end, performance appeared to have reached a relatively steady state. Now that the culture of local planning has changed with regards to timeliness, the Department is now introducing a revised system of financial rewards which are linked to the number of houses constructed and putting local plans in place. There will still be some incentive to maintain performance on planning applications because some grant will be abated from Local Authorities who do not meet targets for timeliness.

Conclusions

- Frequent reviews of the sanction or reward mechanism are required in order to assess its effectiveness, amend it for any design flaws, and take into account changing needs over time.

- The most accurate assessment of the effectiveness of a sanction or reward mechanism would include an estimate of what would have happened in the absence of one.
- Estimates of what would have happened are technically difficult, require a lot of data, and can take a long time. Methodologies include the use of control groups and the econometric analysis of comprehensive data sets.
- The difficulty and expense of conducting an estimate of what would have happened means that the effectiveness of a sanction or reward mechanism is generally assessed in terms of short-term measured outputs based on administrative data.
- At the simplest level, gross outputs are a measure of raw outputs of a programme, and have the advantage of being easy to collect and understand. If available, net outcome measures can be more useful in measuring the contribution of a programme to the outcome. These measures seek to estimate the value that is added by the programme.
- In the absence of other data, the effectiveness of a sanction or reward mechanism can be assessed on a qualitative basis, for example through canvassing the opinions of participants in the scheme.

Implement

3.67 Allowing a sanction or reward mechanism to be phased in over a period of time allows a planned and incremental change which is important to the scheme's success. Indeed, Armstrong and Murlis (1998) argue that a hasty implementation is often responsible for a scheme's failure, rather than any faults with the scheme itself.

3.68 A phased implementation allows organisations to ensure that the scheme is effectively designed. There are often unintended consequences associated with a sanction or reward mechanism, which means that allowing sufficient time to test the mechanism is vital before tying it to sanctions or rewards. A phased implementation also allows time for data collection systems to be developed,

and confidence in the quality and reliability of the collected data to be built up. In particular, there is a strong case for a mechanism to be initially implemented on a trial or pilot basis (Armstrong and Brown, 2001). This allows its impact to be accurately assessed, as it can be compared to a counterfactual.

3.69 Furthermore, phasing in a sanction or reward mechanism over time allows Agents the opportunity to develop the necessary knowledge about the mechanism, and the ability to implement it. This is critical to the success of the scheme. For example, the US Air Force found that the most important factor in implementing its successful performance-based contracting programme was employee training that focussed on how the performance mechanism worked (GAO, 2006).

Conclusions

- Phased implementation allows the design of a sanction or reward mechanism to be tested. It also allows data systems to be developed, and those affected to develop the necessary knowledge about the mechanism and the ability to implement it.
- There is a strong case for initially implementing a sanction or reward mechanism on a pilot basis.

Methodology

In order to understand and to scope the key issues for this study, we reviewed various academic literature, and discussed the topic with academics who have worked in the area of incentives. We employed Deloitte to conduct the literature review, which covered both the theory and the practical applications of sanctions and rewards, in both the public and private sectors, in the UK and overseas.

We then conducted a two-stage analysis process:

- 1 We undertook a survey of the policy leads for significant government programmes, asking questions on the use of sanctions and rewards

in the delivery chains underpinning policy delivery. This provided us with an overview of the different types of sanctions and rewards being used, and the different situations in which they were applied. We piloted the survey with the Department of Health.

- 2 We then selected eight major programmes as case studies. Selection was primarily based on obtaining a mix of delivery chains that did, and did not, use any type of sanction or reward mechanisms. We also tried to get a mix of delivery chain types. The eight we selected were:

Programme	Department	Reported using sanctions or rewards?	Achieved by
Homelessness	Communities and Local Government	No	Service delivery
Planning development	Communities and Local Government	Yes	Service delivery
Educational attainment	Children, Schools and Families	Yes	Service delivery
Rural productivity	Environment, Food and Rural Affairs	No	Influencing
Air quality	Environment, Food and Rural Affairs	No	Influencing
Air quality	Transport	Yes	Influencing
Drug harm	Home Office	No	Influencing
Drug harm	Health	Yes	Influencing



- 3 For the selected programmes we conducted interviews with the Departmental policy leads and key delivery stakeholders to gain a more detailed understanding of the delivery chain. We asked about the use of levers within the delivery chains, views on effectiveness of delivery, and opportunities to enhance effectiveness. Where sanctions or rewards were being used we asked for more detail about how they worked in practice.

Survey

We surveyed all the major programmes of Government as defined by the 2004 Spending Review Public Service Agreement targets in operation at the time of the survey. We achieved a 63 per cent response rate. A review of the respondents and non-respondents did not imply a significant bias to the results; the programmes that responded showed a good mix of Departments and types of delivery chain.

A number of responses were not comprehensive in describing all sanction or reward mechanisms used in the delivery chain. Survey respondents were only asked to outline “formal” sanction or reward mechanisms. In the case studies we found a number of additional sanction or reward mechanisms which had not been included in survey responses.

Different types of sanctions or rewards

Type	Organisational level	Team level	Individual level	End user level
Financial sanction or reward mechanisms				
Contract/grant renewal	<ul style="list-style-type: none"> One year contract renewal for good performance specified in the Department for Children Schools and Families' contract with Capita. The renewal of Air Quality Grants to Local Authorities is dependent on achievement of the targets specified in the original grant. 			
Contract/payment holdback	<ul style="list-style-type: none"> Ten per cent contract holdback by Ontario Realty Corporation for failure by the facilities management company to meet performance standards. A Service Level Agreement with the Regional Development Agencies for funding provided through the Rural Development Programme. The Agencies are subject to five per cent of the disallowance risk for non-compliance with EU regulations. 			<ul style="list-style-type: none"> Deduction from benefit payments to out-of-work benefit recipients by the Department for Work and Pensions, in the event of non-compliance with mandatory activity. Payment withheld by the Department for Environment Food and Rural Affairs for non-compliance by farmers and landowners with the contract terms of the Rural Development Programme.

Type	Organisational level	Team level	Individual level	End user level
Financial sanction or reward mechanisms <i>continued</i>				
Bonus above established baseline	<ul style="list-style-type: none"> Bonus funding above a standard amount through the Planning Delivery Grant to Local Authorities, based on performance criteria related to the speed of processing planning applications. Thirty per cent of the Department for International Development's funding to the International Fund for Agricultural Development made contingent upon performance – system-wide reform is required. Increased funding by the Department for International Development provided to UNICEF or recipient national governments for good performance. 	<ul style="list-style-type: none"> Bonus given to US Job Training Partnership teams on the basis of outcomes, such as job placements and earnings by trainees. Bonus of up to 7.5 per cent paid to teams in Jobcentre Plus on achievement of five performance targets. Glendale School in the US provides financial incentives to all academic and support staff upon the school's achievement of a number of academic, involvement and satisfaction-related performance targets. Team-based performance pay schemes for civil servants in Brazilian Tax Authority based on objective measures of the amount of tax collected, combined with individual financial rewards based on a subjective assessment. Bonus/deduction from employees' salaries based on the over or under-achievement of individual and team targets, by a Hong Kong train operator. 	<ul style="list-style-type: none"> Small performance-related bonuses in different UK government Departments such as the Home Office, Defence Aviation Repair Agency and the former Department for Constitutional Affairs. 	<ul style="list-style-type: none"> Financial incentives offered by the Department for Work and Pensions to recipients of incapacity benefits or lone parent benefits, to help them move into, and remain in work.
Bonus based on stretch target		<ul style="list-style-type: none"> Makinson performance pay incentive scheme in HM Customs & Excise, where teams had to achieve a basic target before being eligible for a bonus based on achievement of a stretched target. 		

Type	Organisational level	Team level	Individual level	End user level
Financial sanction or reward mechanisms <i>continued</i>				
One-off bonus				<ul style="list-style-type: none"> ● Book vouchers and points, which can be converted into prizes, are presented to pupils by some UK schools for educational attainment and school attendance.
Unit-based payment/penalty	<ul style="list-style-type: none"> ● Fifty-five per cent of funding provided by the National Treatment Agency to Drug Action Teams is based on activity data to incentivise the treatment of more people. 		<ul style="list-style-type: none"> ● Piece-rate payment scheme introduced in a US windscreen installation firm, in which workers' pay was linked to the number of windscreens they installed. ● A proportion of GPs' income is related to the quality of their care, based on a sliding scale of achievement through the UK Quality and Outcomes Framework. 	<ul style="list-style-type: none"> ● Renewable Transport Fuel Obligation – a regulatory requirement that a proportion of all road fuel sold in the UK is from renewable sources. Incentivised by fuel duty discount per litre of renewable fuel sold and penalised for each litre not supplied. ● Levy of a NOx landing charge at Heathrow and Gatwick, on airlines, by the Civil Aviation Authority.
Share in savings			<ul style="list-style-type: none"> ● UK GP fund-holders scheme whereby GPs were allowed to keep any surplus on their funds but had to use this for any purpose that was beneficial to patients. 	
Reduction in charge				<ul style="list-style-type: none"> ● Lower rates of excise duty offered to Heavy Goods Vehicle operators who achieve target levels of emissions and obtain the Reduced Pollution Certificate.

Type	Organisational level	Team level	Individual level	End user level
Reputational sanction or reward mechanisms				
Public ranking	<ul style="list-style-type: none"> UK school league tables. Publication of reports by Ofsted, which assess and rank the performance of schools. 	<ul style="list-style-type: none"> Monitoring of Jobcentre Plus district offices' performance against targets, and comparison across districts through a league table. 		
Award	<ul style="list-style-type: none"> Green Flag Award for good management of parks and open spaces, given by the Department for Communities and Local Government to Park Management and Maintenance. Designating "beacon" status to Local Authorities to recognise excellence in performance. 		<ul style="list-style-type: none"> "Outstanding providers" event for teachers hosted by Her Majesty's Chief Inspector and Ofsted. 	<ul style="list-style-type: none"> Pupils rewarded for academic performance through a number of incentive schemes, including letters to parents and Head Teacher commendations for good performance, as well as merit stamps in pupils' diaries.
Publications	<ul style="list-style-type: none"> Publication of the environmental policies submitted to the Office of the Rail Regulator, by the Department for Transport. The scheme is applied to licence holders, such as Network Rail, and Train Operating Companies. 			

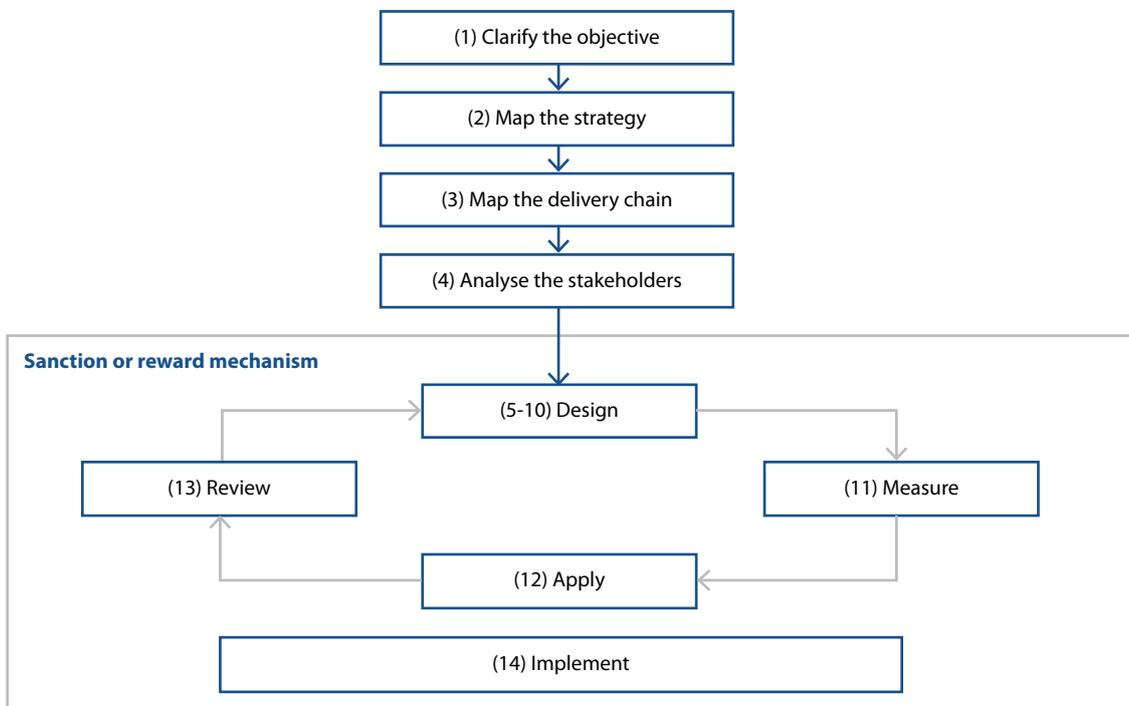
Type	Organisational level	Team level	Individual level	End user level
Operational sanction or reward mechanisms				
Decreased oversight	<ul style="list-style-type: none"> ● Decreased oversight and greater freedom from central management given to UK hospitals who attain Foundation Trust status by meeting certain performance standards. ● Reduced or no inspection from Ofsted on the basis of a school's performance. 			
Dismissal or disciplinary action	<ul style="list-style-type: none"> ● Shutting down schools or putting them into "special measures" on the basis of unsatisfactory performance gauged by Ofsted inspections. 		<ul style="list-style-type: none"> ● Disciplinary action against Los Angeles police officers in case of citizen complaints. ● Dismissal of key managers of hospitals in the UK, on the basis of poor performance, measured in terms of waiting times for elective hospital procedures. 	<ul style="list-style-type: none"> ● Pupils sanctioned for misbehaviour through a demerit system. A pupil with a certain number of demerits is required to undergo "supervised learning" from a teacher.

Guidance for developing and implementing sanction and reward mechanisms

The purpose of this Annex is to provide some practical guidance for organisations that are considering using sanction or reward mechanisms to help achieve their objectives.

Figure 11 outlines a standard process which can be followed:

11 Process for designing and implementing a system of sanctions and rewards to achieve an objective



Source: National Audit Office

Objective

- 1** Clearly define a measurable objective.

Map the strategy

- 2** Identify the activities and the outputs that are needed to achieve the objective (see the example Strategy Map in figure 1). [p. 7 and 47].

Map the delivery chain

- 3** Identify and engage with the key stakeholders, map the performance levers and incentives in use (see the example Delivery Map in figure 2). [p. 8 and 48].
- 4** Use the delivery map to undertake a stakeholder analysis. Identify where the main barriers are, and the evidence on the effectiveness of the current programme, the performance levers, and their associated sanction or reward mechanisms (see Stakeholder Analysis template in Annex 3). [p. 49].

Designing sanction and reward mechanisms

- 5** Understand the full range of sanction or reward mechanisms available – financial, reputational and organisational (see Annex 2). [p. 40].
- 6** Assess the motivations of the key players. Assess the balance between financial and altruistic motivations, and whether they operate at organisational, team or individual level. [p. 26 and 40].
- 7** Identify outcomes over which the players have control or significant influence. [p. 27].
- 8** Link the incentives to performance measures which lead to the desired (long-term) outcomes in a predictable way. A single measure may not capture the relevant aspects of performance, but any set of measures must be kept manageable. [p. 28].
- 9** Ensure the rewards and sanctions are cost-effective. Where they involve financial elements, it may be sensible to model the operation of the system to help define appropriate values or ranges. [p. 30].

- 10** Introduce safeguards to prevent unintended behaviours. Approaches can include maintaining a degree of independence in performance assessment and validating key performance measures [p. 32].

Measure performance

- 11** Develop and implement data systems that collect and analyse timely and reliable performance data. [p. 33].

Apply the sanction or reward

- 12** Develop the expertise to apply the sanctions or rewards in a timely, consistent and transparent manner, especially when using complex contractual arrangements with delivery agents. [p. 34].

Review effectiveness

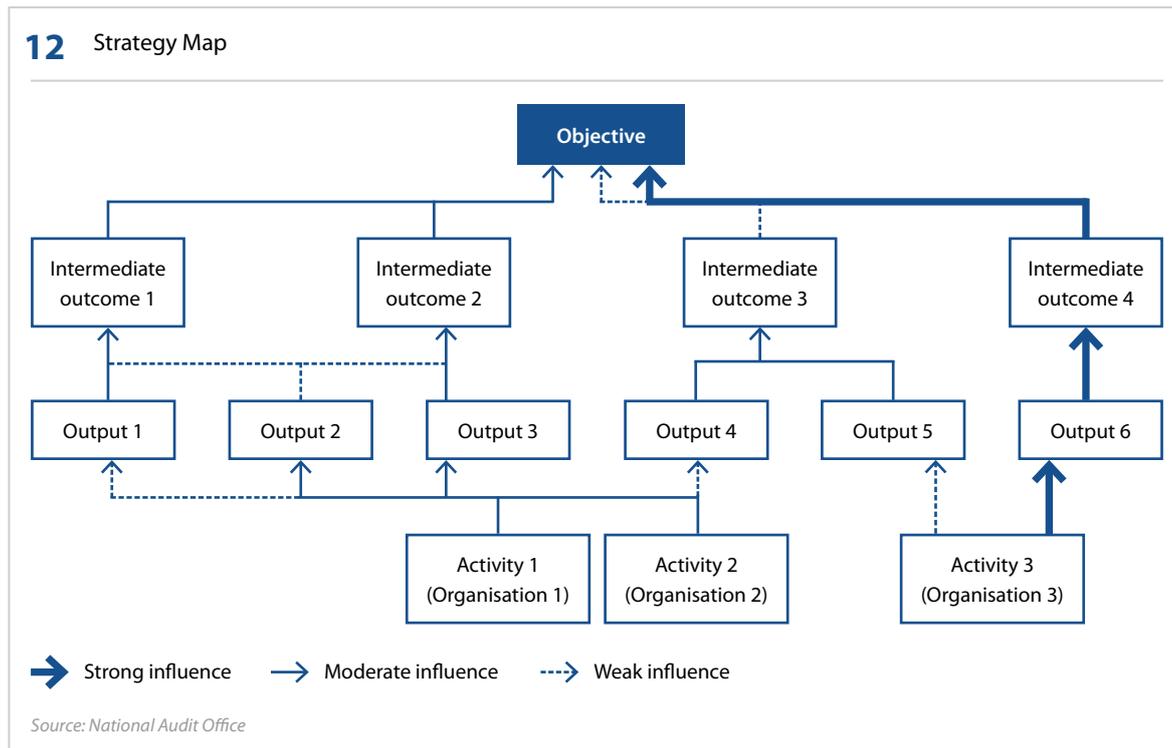
- 13** Build regular internal and external reviews of the effectiveness of the sanction or reward mechanism into the programme's overall performance management cycle. [p. 35].

Implement

- 14** Where possible, phase in new sanction or reward mechanisms gradually, or on a pilot basis, to identify and address any dysfunctional behaviour. [p. 37].

Strategy Map

Figure 12 shows a generic strategy map for an objective. It shows how an organisation can take an objective, and identify the combination of activities and outputs that should help to achieve it. It should be based on evidence, or at least consensus, on what drives the outcome. It can then be used to identify the stakeholders that will deliver those outputs in the form of a delivery map.

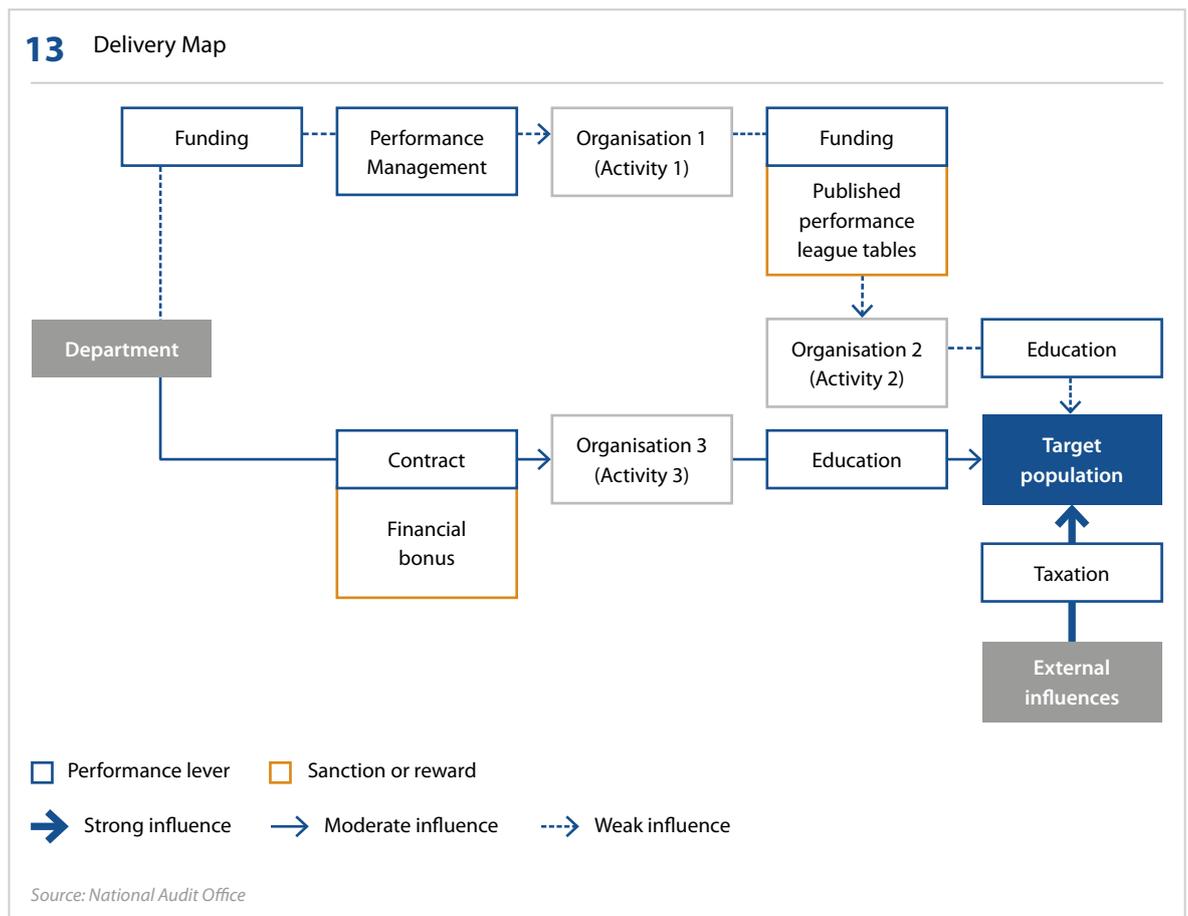


Delivery Map

Figure 13 shows a generic delivery map for an objective which is trying to influence the behaviour of a target population. It shows how an organisation can take an objective and its strategy, and map the relationships between the stakeholders that are involved. It then shows the performance levers that are used to influence those stakeholders, and any sanctions or rewards that are attached to them.

Instructions

- Be clear on the scope of the delivery map. The nature of public services means they are likely to overlap with many other objectives. While it is important to understand all the key stakeholders and influences, it is also important to keep a delivery map focussed on the objective to avoid unnecessary complexity.
- Include external influences, even though they cannot be managed. It is important to know what else influences a delivery chain so that the performance levers and incentives can be designed around them.
- Distinguish between strong influences and weak influences, as this will help to focus the levers and incentives on where they can have most impact.
- Involve as many of the stakeholders as possible to get a full understanding of the influences and incentives.



Stakeholder Analysis

The following framework is a template which can be used to analyse the issues faced by stakeholders within the delivery chain, and develop ideas and recommendations for strengthening the current performance levers with the help of sanctions and rewards. Conducting some of this work with stakeholders is important, not only for generating appropriate mechanisms, but for gaining ownership and consensus.

Instructions

- Transfer the information from the delivery map into this framework.
- Populate the rest of the framework in consultation with stakeholders.

Stakeholders		Performance levers eg. contracts, performance management etc.	Current or proposed sanctions and rewards			Issues	Recommendations
Name	Level		Financial	Operational	Reputational		
Organisation 1	Organisation						
	Team						
	Individual						
Organisation 2	Organisation						
	Team						
	Individual						
Target population	Individual						

Performance levers

Each performance lever we have identified is discussed below, including explanations of some of the implicit sanctioning or rewarding effects associated with them.

- **Legislation.** As well as establishing formal sanctions and rewards, legislation may also generate informal incentives. For example, The Environment Act 1995 requires Local Authorities to review and assess the air quality in their area. There are no formal sanctions with respect to high pollution levels. If levels are judged as being too high, the area will be designated an Air Quality Management Area (AQMA). Furthermore, an action plan will need to be set up by the Local Authorities to reduce pollution. Local Authorities may have the implicit incentive to avoid becoming an AQMA, as this would stigmatize the area, which in turn might lead to negative effects, such as a reduction in tourism. Measures implemented to prevent becoming an AQMA may help the Department for Transport and the Department for Environment, Food and Rural Affairs achieve their common objective to reduce greenhouse gas emissions and improving air quality (HM Treasury, 2007a).
- **Taxation.** While the primary purpose of taxation is to raise government revenue, it is often used as an implicit motivational instrument to support the achievement of a variety of objectives. For example, implementing the right tax scheme may help achieve the Department of Health's objective to reduce adult smoking, and increases in fuel duty may help the Department for Transport and the Department for Environment, Food and Rural Affairs meet their air quality and climate change targets (HM Treasury, 2007a).
- **Targets.** Setting targets may create an implicit motivation for Agents to perform better. This may be because they set a clear expectation, assign accountability, create a fear of losing credibility, and generate external pressure for improved performance. For central government, performance against Public Service Agreement and Value for Money targets is among the factors influencing the funding received by Departments, even though the relationship between performance and funding is not explicitly quantified.
- **Regulations and standards.** Regulation aims to safeguard and/or raise the quality of service provision, and to promote equity by ensuring everyone has access to a minimum standard of provision regardless of their circumstances. For example, the Department of Health's National Service Framework – which specifies minimum standards for health and social services for children, young people and pregnant women. For private sector companies in regulated sectors,

regulation can set benchmarks. For example, in its price review, Ofgem sets a price cap – the path of the maximum price a company may charge for its services during the price review period. By allowing energy companies to keep the profits earned by reducing costs below the levels assumed by the regulator, they have an implicit motivation to become more efficient. In the next price review, Ofgem will take cost reductions from the previous period into account and potentially be able to set a lower maximum price path.

- **Funding.** Receipt of government funding may be only indirectly dependent on meeting certain criteria, which are themselves not always quantifiable. For example, some of the funding for Local Authorities is related to the Local Transport Plan and the council's performance in terms of air quality. Councils therefore have an implicit incentive to improve the quality of their air, but without an explicit link to funding. If a council performs well in terms of air quality, it helps with the achievement of both the Department for Environment, Food and Rural Affairs' objective to improve air quality, as well as the Department for Transport's objective to provide an efficient and reliable interregional transport system (HM Treasury, 2007a).
- **Publishing performance data.** Even without associated sanctions or rewards, published data can be used to drive performance. In some cases there is an implicit motivation among employees and organisations to try to outperform each other. For example, police forces often display weekly or monthly performance data within police stations (HM Treasury, 2004b). The display of this data provides incentives for police officers to increase their performance, either to compare well in relation to colleagues, or for the district to compare well in relation to other districts.
- **Education.** Examples of which include:
 - (i) Active campaigning, and creating an atmosphere of strong social disapproval for offenders of a certain norm, may set informal sanctions for people to disobey this norm. For example, Siegel (2006) argues that the UK campaign to shame drunk drivers created a moral climate with strong informal sanctions. This supported achieving the Department for Transport's objective to reduce the number of people killed or seriously injured in road accidents; bearing in mind that about one sixth of all road casualties involve drivers which are over the legal alcohol limit (HM Treasury, 2004c).
 - (ii) Better nutritional information may incentivise people to improve the quality of their diet. This may also set the incentive for the food industry to increase the quality of its products which in turn would help the Department for Environment, Food and Rural Affairs to achieve its objective to promote a safe food supply chain which meets consumers' requirements (HM Treasury, 2004c).
 - (iii) Ofcom requires fixed-line telecoms operators to submit periodic reports on key performance indicators. These are made available to the public on the TopComm website. This allows consumers to judge the fixed line operators based on these performance indicators and switch to the one which they feel best meets their requirements. This gives companies the incentive to improve their performance against these indicators in order to attract new, or at least not lose existing, customers.
- **Competition and choice.** Since the 1980s there have been concerted attempts to introduce more consumer choice in some public services. For example, parents have more choice over which schools they can send their children to, and GPs have been given more power to select secondary health care providers on their patients' behalf. As long as users or patients can make informed choices, and funding follows those decisions, healthcare providers face incentives to provide high quality, cost-effective services. However, there are problems with attempts to introduce competitive pressure to public services. In the case of schools, the costs to the consumer of switching are significantly higher than in the

case of, for example, utility providers. Children from families that are less well-informed, or less mobile, are at risk of becoming “stuck” in less well-performing schools, furthering the inequity. It also gives schools the incentive to pick the best pupils. Bradley and Taylor (2007) found that reforms in the school system have increased the efficiency of secondary schools (measured in terms of exam performance) but at the expense of increased social segregation.

- **Performance and staff management.** Regular performance reviews can act as strong incentives for delivery chain bodies to maintain and improve service levels. For example, at the central government level, all Departments are subject to Capability Reviews, which require the publication of an action plan by the Department. This incentivises Departments to improve performance, even though the consequences of failing to achieve the stated improvements are not made explicit. On staff remuneration, in most cases in the public sector, individual performance is assessed against a range of non-quantifiable objectives. But even where the formal performance-related component of overall remuneration is small – as it typically is in the public sector – performance can be incentivised by the implicit promise of longer-term career progression.
- **Contracts.** In many cases, contracts will explicitly set out the financial and other implications of performance falling below, or exceeding, target levels. However, contracting bodies can also use a range of implicit sanctions and rewards, such as the prospect of a contract extension subject to satisfactory performance. There are also the reputational consequences to the Agent of having their contract terminated, or not extended.
- **Partnerships and Memoranda of Understanding.** At the central government level, joint objectives can be supported by delivery plans and monitoring systems. In other cases, joint working may be driven by shorter term imperatives, such as the need to respond to a crisis, like flooding. At the local level, partnerships between Local Authorities and key stakeholders have become increasingly common, either on a bilateral basis (such as between a Local Authority and a further education college) or collectively, through umbrella arrangements, such as Local Strategic Partnerships. Another example might be the shared objectives of a prison and a drugs rehabilitation charity. These relationships are not always contractually based; NAO (2006) notes that bus use in Brighton has grown strongly despite the absence of a direct contract between the bus operators and the Local Authority.



Bibliography

- Armstrong, M. & Murlis, H. (1998). *Reward Management*, 4th ed. London: Kogan Page.
- Armstrong, M. (1999). *Employee Reward*, 2nd ed. Institute of Personnel and Development.
- Armstrong, M. & Brown, D. (2001). *New Dimension in Pay Management*. CIPD.
- Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., Wilson, D. (2004). *Evaluating the Impact of Performance-related Pay for Teachers in England*. CMPO Working Paper Series No. 04/113.
- Bradley, S., & Taylor, J. (2007). *Diversity, choice and the quasi market: An empirical analysis of secondary education policy in England*. Department of Economics, Lancaster University.
- Burgess, S., Metcalfe, P. (1999) *The use of incentive schemes in the Public and Private sectors: Evidence from British Establishments*. CMPO Working Paper 99(015).
- Burgess, A., Propper, C. & Wilson, D. (2002). *Does performance monitoring work? A Review of the evidence from the UK Public sector, excluding healthcare*. CMPO Working Paper No. 02/49, CMPO, University of Bristol.
- Burgess, S., & Ratto, M. (2003). *The role of incentives in the Public Sector: Issues and Evidence*. CMPO Working Paper 03/071, CMPO, University of Bristol.
- Burgess, S., Propper, C., Ratto, M., & Tominey, E. (2004). *Evaluation of the Introduction of the Makinson Incentive Scheme in Jobcentre Plus*. CMPO, University of Bristol.
- Burgess, S., Propper, C., Ratto, M., Scholder, S. & Tominey, E. (2005). *Evaluation of the Introduction of the Makinson Incentive Scheme in HM Customs and Excise*. CMPO, University of Bristol.
- Cabinet Office. (2006). *The UK Government's Approach to Public Service Reform*.
- Communities and Local Government. (2006). *Strong and Prosperous Communities – The Local Government White Paper*.
- Dearden, L., Emmerson, C., Frayne, C., Meghir, C. (2004). *Can Education Subsidies stop School Drop-outs? An evaluation of Education Maintenance Allowances in England*. Institute for Fiscal Studies.
- Government Accountability Office (GAO). (2006). *Grants Management: Enhancing Performance Accountability Provisions Could Lead to Better Results*. GAO-06-1046.
- Gneezy, U. (2003). *The W effect of incentives*. The University of Chicago Graduate School of Business.
- Grout, P.A., & Stevens, M. (2003). *Financing and managing public services: an assessment*. CMPO Working Paper No. 03/076, CMPO, University of Bristol.

- HM Treasury. (2004a). *Devolving Decision Making: 1 – Delivering better public services: refining targets and performance management*.
- HM Treasury. (2004b). *Devolving Decision Making: a consultation on regional funding allocations*.
- HM Treasury. (2004c). *SR 2004: public service agreements 2005/2008*.
- HM Treasury. (2007a). *2007 Pre-Budget Report and Comprehensive Spending Review – Meeting the aspirations of the British people*.
- HM Treasury. (2007b). *Review of sub-national economic development and regeneration*. Published in collaboration with the Department for Business Enterprise and Regulatory Reform, and Communities and Local Government.
- Heckman, J., Heinrich, C., & Smith, J. (2002). *The Performance of Performance Standards*. *Journal of Human Resources*, 37(4), 778-811.
- Hibbard, J., Stockard, J., Tusler, M. (2003). *Does Publicizing Hospital Performance Stimulate Quality Improvement Efforts?*. *Health Affairs*, Volume 22, Number 2.
- Kahn, C., Silva, E., & Ziliak, J. (2001). *Performance-Based Wages in Tax Collection: The Brazilian Tax Collection Reform and Its Effects*. *Economic Journal*, 468, 188-205.
- Kelman, S., Friedman, J. N. (2007). *Performance Improvement and Performance Dysfunction; An Empirical Examination of Impacts of the Emergency Room Wait-Time Target in the English NHS*. Harvard University JFK School of Government Faculty Research Working Paper RWP07-034.
- Le Grand, J. (1995). *Knights, Knaves or Pawns? Human Behaviour and Social Policy*. *Journal of Social Policy*, 26(2), 149–169.
- Levitt, S.D., & Dubner, S.J. (2005). *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: HarperCollins Publishers.
- The Lewin Group & ECONorthwest (2003). *Study of the Implementation of the Performance-Based Incentive System*.
- Macrory, R. (2006). *Regulatory Justice: Making Sanctions Effective*. London: Better Regulation Executive.
- Makinson, J. (2000). *Incentives for change: Rewarding performance in national government networks*. Public Services Productivity Panel, HMT.
- Marsden, D. & French, S. (2002). *Performance Pay in the United Kingdom*. 115-147, in Brown, M., Heywood, J. S. (Eds.). (2002). *Paying for Performance: An International Comparison*. New York, London: ME Sharpe.
- Milgrom, P., & Roberts, J. (1990). *The economics of modern manufacturing: technology, strategy and organization*. *American Economic Review*, 80(3), 511-528.
- National Audit Office (NAO). (2001a). *Measuring the Performance of Government Departments*. HC 301, 2000-2001.
- National Audit Office (NAO). (2001b). *Inappropriate adjustments to NHS waiting lists*. HC 452, 2001-2002
- National Audit Office (NAO). (2004-07). *Public Service Agreement: Validation Compendium Reports 1-4*. HC 476, 2004-2005. HC 985, 2005-2006. HC 127 (I & II), 2006-2007. HC 22 (I & II) 2007-2008.

National Audit Office (NAO). (2006). *Delivering Efficiently: Strengthening the links in public service delivery chains*. Published in collaboration with the Audit Commission. HC 940, 2005-2006.

National Audit Office (NAO). (2007). *Targeting Inequalities*.

National Audit Office (NAO). (2008). *NHS Pay Modernisation: New Contracts for General Practice Services in England*. HC 307, 2007-2008

Prendergast, C. (2001). *Selection and oversight in the public sector, with the Los Angeles Police department as an example*. National Bureau of Economic Research Working Paper 37.

Prentice, G., Burgess, S., & Propper, C. (2007). *Performance-pay in the public sector: A Review of the issues and evidence*. Office of Manpower Economics.

Propper, C. & Wilson, D. (2003). *The Use and Usefulness of Performance Measures in the Public Sector*. CMPO Working Paper No. 03/073, CMPO, University of Bristol.

Propper, C., Sutton, M., Whitnall, C., & Windmeijer, F. (2007). *Did 'targets and terror' Reduce Waiting Times in England for Hospital Care?*. CMPO Working Paper 07/179, CMPO, University of Bristol.

Seabright, P. (2004). *Continuous Preferences Can Cause Discontinuous Choices: An Application to the Impact of Incentives on Altruism*. Mimeo, IDEI, Toulouse University.

Siegel, L.B. (2006). *Criminology: Theories, Patterns & Typologies* 9th ed. Wadsworth Publishing.

This report can be found on the
National Audit Office web site
at www.nao.org.uk

**For further information about the
National Audit Office please contact:**

National Audit Office
Press Office
157-197 Buckingham Palace Road
Victoria
London
SW1W 9SP

Tel: 020 7798 7400

Email: enquiries@nao.gsi.gov.uk



Design and Production by NAO Marketing & Communications Team
DG Ref: 008668 | Printed by SLS Print
© National Audit Office 2008