



National Audit Office

Report

by the Comptroller
and Auditor General

Cross-government

Outcome-based payment schemes: government's use of payment by results

Our vision is to help the nation spend wisely.

Our public audit perspective helps Parliament hold government to account and improve public services.

The National Audit Office scrutinises public spending for Parliament and is independent of government. The Comptroller and Auditor General (C&AG), Sir Amyas Morse KCB, is an Officer of the House of Commons and leads the NAO, which employs some 810 people. The C&AG certifies the accounts of all government departments and many other public sector bodies. He has statutory authority to examine and report to Parliament on whether departments and the bodies they fund have used their resources efficiently, effectively, and with economy. Our studies evaluate the value for money of public spending, nationally and locally. Our recommendations and reports on good practice help government improve public services, and our work led to audited savings of £1.15 billion in 2014.



National Audit Office

Cross-government

Outcome-based payment schemes: government's use of payment by results

Report by the Comptroller and Auditor General

Ordered by the House of Commons
to be printed on 17 June 2015

This report has been prepared under Section 6 of the
National Audit Act 1983 for presentation to the House of
Commons in accordance with Section 9 of the Act

Sir Amyas Morse KCB
Comptroller and Auditor General
National Audit Office

15 June 2015

This report reviews the government's recent experience of using payment by results (PbR) to deliver public services.

© National Audit Office 2015

The material featured in this document is subject to National Audit Office (NAO) copyright. The material may be copied or reproduced for non-commercial purposes only, namely reproduction for research, private study or for limited internal circulation within an organisation for the purpose of review.

Copying for non-commercial purposes is subject to the material being accompanied by a sufficient acknowledgement, reproduced accurately, and not being used in a misleading context. To reproduce NAO copyright material for any other use, you must contact copyright@nao.gsi.gov.uk. Please tell us who you are, the organisation you represent (if any) and how and why you wish to use our material. Please include your full contact details: name, address, telephone number and email.

Please note that the material featured in this document may not be reproduced for commercial gain without the NAO's express and direct permission and that the NAO reserves its right to pursue copyright infringement proceedings against individuals or companies who reproduce material for commercial gain without our permission.

Links to external websites were valid at the time of publication of this report. The National Audit Office is not responsible for the future validity of the links.

Contents

**Main payment by results (PbR)
schemes considered in this report** 4

Summary 5

Part One

Using outcome-based
payment schemes 11

Part Two

Deciding whether to use PbR 17

Part Three

Designing PbR schemes 22

Part Four

Monitoring and evaluating PbR 31

Appendix One

Our audit approach 39

Appendix Two

Our evidence base 41

Glossary 43

The National Audit Office study team consisted of:
Pauline Ngan and Hannah Robinson under the direction of Victoria Keilthy, and with the assistance of Martin Chong, Benoit Guerin, Ayesha Harris, Nathalie Larsen, Polly Meeks, Henry Midgley, Guy Munro, Marco Neves da Silva, Emma Taylor, Andrew Tuffin and David Wilson.

This report can be found on the National Audit Office website 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

Enquiries: www.nao.org.uk/contact-us

Website: www.nao.org.uk

Twitter: @NAOorguk

Main payment by results (PbR) schemes considered in this report

Welfare to work: Work Programme – Department for Work & Pensions (DWP)

Providers: 17 prime contractors, with about 850 subcontractors (public, private and voluntary organisations)

Aim: to help the long-term unemployed secure jobs sooner, and keep them for longer

Budget: £3.3 billion over 9 years

PbR proportion of contract: DWP expects 80% of payments to be outcome-based over the life of the programme

Family support: Troubled Families – Department for Communities and Local Government (DCLG)¹

Providers: local authorities and partner agencies

Aim: to turn around the lives of 120,000 troubled families in England, extended to a further 400,000 families from April 2015

Budget: £448 million over 3 years (plus £200 million in 2015-16 for the first year of a 5-year programme for up to 400,000 additional families)

PbR proportion of programme: 20% in 2012-13, rising to 60% in 2014-15

Note

1 DWP runs a related programme (European Social Fund Support for Families with Multiple Problems).

Offender rehabilitation: HMP Peterborough and HMP Doncaster pilots¹ – Ministry of Justice (MoJ)

Providers: Social Finance with a range of specialist subcontractors (Peterborough), Serco with subcontractor Catch 22 (Doncaster)

Aim: to reduce reconviction rates

Budget: £8 million over 6 years (Peterborough), £25 million over 3 years (Doncaster)

PbR proportion of contract: Peterborough: 100%; Doncaster: part of wider £250 million prison contract, with 10% of annual value dependent on Serco reaching reduced reconviction target

The offender rehabilitation pilots were superseded in 2014 by the national roll-out of PbR to the probation sector under Transforming Rehabilitation

Note

1 The Peterborough scheme is also known as MoJ's Social Impact Bond (SIB) pilot. MoJ also established 3 other offender rehabilitation pilot schemes around the same time.

International aid: 19 PbR aid projects – Department for International Development (DFID)

Providers: various governments, private sector suppliers and non-governmental organisations

Aims: a range of development issues including water and sanitation, education and health

Budget: approximately £2.2 billion; projects run for between 3 and 12 years

PbR proportion of contract: varies between 9% and 100%

Summary

1 The UK government is increasingly using outcome-based payment schemes – often called ‘payment by results’ (PbR) – to deliver public services. PbR is a mechanism where all or part of the payment depends on the provider achieving outcomes specified by the commissioner. Providers are, to a greater or lesser extent, free to choose the interventions needed to secure the desired outcomes.

Government use of PbR

2 PbR is not a new mechanism, but its use has increased in recent years as part of government’s wider reforms to public services, set out in the Cabinet Office’s 2011 white paper:

“Open commissioning and payment by results are critical to open public services ... Payment by results will build yet more accountability into the system – creating a direct financial incentive to focus on what works, but also encouraging providers to find better ways of delivering services.”¹

3 In the public sector, PbR tends to be used to address complex social issues for which there are no straightforward solutions: for example, getting people on benefits back into work, and reducing reoffending.

Scope of this report

4 This report highlights lessons on the selection and implementation of PbR in the UK public sector. Its intended main audience is commissioners and senior decision-makers within departments, as well as officials at the centre of government with an interest in public service delivery models.

5 We refer to findings from previous published National Audit Office (NAO) reports on the use of PbR for welfare-to-work (DWP’s Work Programme), and family support (DCLG and DWP’s programmes).² We also draw on NAO research on the MoJ’s offender rehabilitation pilots (the precursor to Transforming Rehabilitation),³ and DFID’s use of PbR in international development. Although our evidence is mainly drawn from central government PbR schemes, our findings will also be relevant for local commissioners considering using PbR.

¹ HM Government, *Open Public Services White Paper*, Cm 8145, July 2011, paragraphs 5.4, 5.16.

² European Social Fund support for Families with Multiple Problems (DWP); cross-government Troubled Families Programme (led by DCLG).

³ We do not look at Transforming Rehabilitation in this report, but will return to it in future work.

6 Most operational PbR schemes have still to finish so there is not yet enough evidence to evaluate the effectiveness of either individual schemes or the PbR mechanism itself.

7 This report looks at:

- government's use of PbR for public service delivery (Part One);
- the issues commissioners should consider when deciding whether or not to use PbR (Part Two);
- challenges in designing PbR schemes (Part Three); and
- how government evaluates the effectiveness of PbR schemes, and the PbR mechanism (Part Four).

We have also published a PbR analytical framework. This sets out questions for commissioners that are intended to help them decide when to use PbR, and design and implement PbR schemes.⁴

Lessons for commissioners

Choosing PbR

8 PbR is not suited to all public services. In 2011, government reviewed existing PbR schemes across government and undertook some work to identify areas suited to the use of PbR. It has not updated this to reflect more recent experience in practice. The nature of PbR means it is most likely to succeed if the operating environment has certain features, for example results that can be measured and attributed to providers' interventions. If PbR is applied inappropriately there is a risk that either service quality or value for money may be undermined. When commissioners use PbR for a service to which it is not ideally suited, they need to do so knowingly and take steps to address the consequent risks as far as possible through scheme design (paragraphs 1.9 and 2.4 to 2.6).

9 Commissioners should justify their selection of PbR over alternative delivery mechanisms. Commissioners should consider a range of delivery approaches and ensure they understand the reasons for selecting PbR over alternatives. Where commissioners do identify the reasons for using PbR, as was the case in DFID's scheme to support Rwandan education, it ensures the choice of PbR is open to scrutiny, and helps inform future decisions about PbR's use (paragraphs 2.2 to 2.3).

⁴ This framework is available at: www.nao.org.uk/payment-by-results

10 PbR is a technically challenging form of contracting, and has attendant costs and risks that government has often underestimated. It is difficult to design an effective payment mechanism and forecast the level of performance that would occur without intervention. Furthermore, although PbR transfers some risk to the provider, commissioners need to be aware of the risks they retain, for example that providers fail to meet scheme objectives. In the rush to implement, government has launched some PbR schemes without making best use of pilots to test the planned approach. Commissioners need to consider the time and skills required to design and manage a PbR contract effectively, as underinvestment on these can have negative consequences for the scheme (paragraphs 2.11 to 2.13 and 3.1).

Designing a PbR scheme

11 To get scheme design right commissioners need to understand potential providers' capacity to take on risk. Departments' schemes use a range of providers to deliver services. For example, DWP contracts with large 'prime' providers for the Work Programme, who in turn pass work onto smaller subcontractors (usually smaller, third sector providers). Commissioners should take opportunities at the pre-procurement stage, as DWP has, to engage with potential providers throughout the delivery chain to ensure they understand the ability and appetite of different providers to take on risks associated with PbR contracts. This understanding will help inform the design of schemes, including the appropriate balance between PbR and non-outcome based payments, such as 'fee for service' (paragraphs 3.11 to 3.17).

12 It takes time and effort to design the payment mechanism so it offers appropriate incentives to providers. Commissioners must understand providers' costs in order to create a payment mechanism that offers an incentive to achieve the desired outcomes. If the payment offered is too high, the taxpayer could pay too much for the service; if too low, providers may not bid for the contract. A poorly designed scheme may create perverse incentives for providers, such as welfare-to-work providers prioritising people who are easier to help and 'parking' those who are harder to help. On the Work Programme, DWP uses differential payments for different groups as one way of mitigating this, although experience shows differential payments can be hard to set at the right level (paragraphs 3.19 to 3.25).

13 It is essential that commissioners establish performance expectations at the start of a scheme, taking into account baseline performance and non-intervention rates. Commissioners should aim to define attainable but stretching performance expectations for providers that are above the non-intervention rate (the level of performance that would occur without intervention). This avoids payment for performance that would have occurred anyway. Commissioners need good data on baseline performance, and to carry out robust modelling of likely future performance and sensitivity testing of any assumptions underpinning estimates. Where the commissioner is using PbR for a new service, such as international aid, the lack of historic, comparable data can make this especially challenging (paragraphs 3.5 to 3.7).

Monitoring and evaluation

14 Commissioners need to actively monitor and manage provider performance.

Using PbR does not remove commissioners' responsibility for overseeing providers' performance; poor performance can have a negative impact on service quality, the scheme's overall objectives and the commissioner's reputation. Commissioners need accurate, reliable and timely information on outcomes achieved, and the number of participants attached to the programme, to ensure providers are only paid for results achieved. Good performance information systems can also give early indications of underperformance or undesired provider behaviour (paragraphs 4.7 to 4.10).

15 Commissioners need to plan at the outset how they will evaluate both the effectiveness of the scheme as a whole, and the impact of PbR as a delivery mechanism.

Good evaluation requires commissioners to establish a robust counterfactual before implementation begins. To date, only a few schemes have monitored performance against a counterfactual, for example MoJ's offender rehabilitation pilot in Peterborough and DFID's Girls' Education Challenge. Failure to evaluate schemes means opportunities are missed to gather evidence on the effectiveness of PbR (paragraphs 4.11 to 4.15).

Conclusion

16 While supporters argue that by its nature PbR offers value for money, PbR contracts are hard to get right, which makes them risky and costly for commissioners. If PbR can deliver the benefits its supporters claim – such as innovative solutions to intractable problems – then the increased cost and risk may be justified, but this requires credible evidence. Without such evidence, commissioners may be using PbR in circumstances to which it is ill-suited, with a consequent negative impact on value for money.

17 Government has a growing portfolio of PbR schemes and the Work Programme, one of its largest programmes, has now been running for 4 years. In that time, DWP has made iterative improvements to its approach, offering government as a whole the opportunity to learn lessons for the wider use of PbR. However, neither the Cabinet Office nor HM Treasury currently monitors how PbR is operating across government. Nor is there a systematic collection or evaluation of information about how effectively PbR is working. Without a central repository of knowledge and a strong evidence base to refer to, PbR schemes may be poorly designed and implemented and commissioners are in danger of 'reinventing the wheel' for each new scheme. If PbR is used inappropriately or is executed badly, the credibility of a potentially valuable mechanism may be undermined.

Recommendations

For commissioners

18 **Figure 1** overleaf summarises our framework for assessing PbR schemes, which presents the good practice principles we recommend commissioners follow when using PbR or similar outcome-based payment schemes.

For the centre of government

19 The Cabinet Office and HM Treasury, as the centre of government, should identify a part of government to be the repository of information and expertise about public sector use of PbR. This central PbR function (which could be located in a central department or a delivery department that uses PbR extensively) should support the effective use of PbR across government by:

- identifying and sharing good practice and learning about the design and implementation of PbR schemes;
- developing government's knowledge about when PbR is most likely to be effective;
- providing guidance on the evaluation of PbR schemes and PbR as a delivery mechanism;
- facilitating networking among PbR commissioners; and
- identifying overlaps between individual schemes and coordinating efforts to eliminate duplication.

20 The central function should also gather evidence for the effectiveness of PbR by compiling data on:

- the extent to which PbR delivers the expected benefits; and
- the value for money of PbR schemes compared with alternative delivery mechanisms.

Figure 1

Summary of National Audit Office framework for PbR schemes

Key question	Aim for commissioners	What does good look like?
Overall fit: should you use PbR to deliver this service?	Commissioners only use PbR to deliver public services when it is the model that offers best value.	Commissioners' decisions to use PbR are well-informed and based on: <ul style="list-style-type: none"> ● clearly defined programme objectives; ● a good understanding of the circumstances in which PbR has worked best in the past (and, conversely, when it has been difficult to use successfully); ● clarity about the benefits PbR is intended to bring to the programme at hand; and ● knowledge of the whole-life costs and risks associated with using PbR, and the extent to which the scheme design will be able to mitigate these.
Design: how can you design an effective PbR scheme?	Commissioners develop insight into the operating context before designing their PbR scheme and set clear expectations for performance. Commissioners identify challenging but achievable outcomes on which to base payments. These in turn produce effective incentives for providers.	Commissioners have a good understanding of: <ul style="list-style-type: none"> ● service users/beneficiaries and their needs; ● the delivery chain for the service; and ● delivery, financial and other risks within the system and how they will be allocated through the delivery chain, ensuring risks retained by the commissioner are within their stated risk appetite for the programme. Commissioners set expectations of the likely impact of PbR on: <ul style="list-style-type: none"> ● overall service performance (eg through modelling); and ● individual user outcomes (eg to avoid perverse incentives that might adversely affect long-term outcomes). Commissioners design their PbR mechanism bearing in mind the need: <ul style="list-style-type: none"> ● to translate their objectives into measurable outcomes; ● to structure payments to create appropriate incentives for providers, so the highest payments are matched to the 'best' behaviours or outcomes; and ● for payments to reflect providers' cost base.
Implementation: what do you need to have in place to implement your PbR scheme effectively?	Commissioners are able to monitor the performance of providers and establish clear oversight and intervention mechanisms to minimise the impact of provider failure on public services.	This requires good performance data systems, including ways to independently verify reported performance, to: <ul style="list-style-type: none"> ● minimise delays between achieving results and disbursing payments; ● identify emerging risks to the scheme, either those at individual provider level (eg significant financial problems that might affect services) or wider risks that could threaten achievement of overall objectives; and ● mitigate these risks whenever possible. Commissioners: <ul style="list-style-type: none"> ● ensure all parties are clear about their responsibilities; and ● understand what constitutes underperformance and how to address it promptly.
Evaluation: how can you evaluate the effectiveness of a PbR scheme?	Commissioners evaluate how using PbR has improved service delivery and overall value for money.	Commissioners are mindful of the need to evaluate the effectiveness of PbR schemes from the outset, ensuring the necessary elements are included at the design stage (eg a baseline and an appropriate counterfactual). Evaluation is used both to make iterative service improvements during the life of a scheme and to identify lessons learned at the end. This should cover the scheme as a whole, as well as the effectiveness of the PbR mechanism itself.

Note

1 For more information, the full framework is available at: www.nao.org.uk/payment-by-results

Source: National Audit Office

Part One

Using outcome-based payment schemes

1.1 In this Part we define outcome-based payment schemes and set out their current use in the public sector.

Nature of outcome-based payment schemes

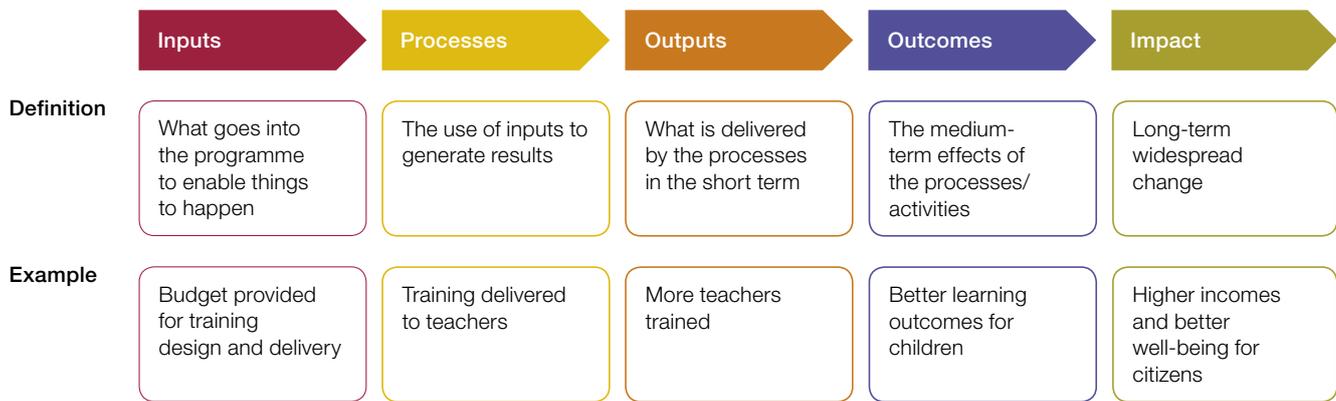
1.2 An outcome-based payment scheme is a way to deliver services where all or part of the payment depends on the provider achieving specific outcomes. Many public service delivery models involve contracts with external providers, and most include terms related to the provider's performance. The distinguishing feature of outcome-based contracting is that the commissioner of the service specifies the desired end result (for example, unemployed people back in work) and the payment to the provider is based wholly or partly on achieving the specified outcome. This outcome-based approach can also apply to non-contractual funding arrangements with local authorities and other public sector bodies.

1.3 Outcome-based payment schemes in the UK public sector are often referred to as 'payment by results' (PbR), although the term is used inconsistently. Some schemes are called PbR but are not outcome-based. For example, some NHS tariffs to pay hospitals for clinical procedures are called PbR, even though the payment is for outputs rather than outcomes. Conversely, under DCLG's New Homes Bonus scheme, local authorities receive bonus payments based on locally delivered increases in housing stock, although the scheme is not commonly referred to as PbR. In this report, we use the term 'PbR' to mean outcome-based payment schemes only.

1.4 In its PbR strategy, DFID sets out the different stages of a 'results chain', illustrating the distinction between inputs, outputs and outcomes (**Figure 2** overleaf).

1.5 The identity of the commissioner and providers in PbR schemes can vary, as can the nature of the payment mechanism (**Figure 3** overleaf). Many schemes include a proportion of upfront payment that is not contingent on the achievement of a specified outcome, and so are not 'pure' PbR. Providers usually need to make an upfront investment, and some form of upfront payment or 'fee for service' can make the scheme more attractive to potential bidders by easing their cash flow and reducing the risk they take on in fulfilling the contract.

Figure 2
DFID's PbR results chain



Source: Department for International Development, *Sharpening incentives to perform: DFID's Strategy for Payment by Results*, June 2014, Figure 1

Figure 3
Nature of public sector PbR schemes

Possible commissioners	Possible providers	Types of PbR payment	Types of non-PbR payment included in PbR schemes
<ul style="list-style-type: none"> Government department Local authority or other local body 	<ul style="list-style-type: none"> Private sector – often contracted directly as 'prime' contractors to deliver PbR contracts. Usually larger organisations with capacity to bear financial risks Third sector – can be directly contracted, but more frequently subcontracted by prime contractors. Usually smaller organisations that can provide specialist support to users Public sector – local government bodies or foreign governments¹ 	<ul style="list-style-type: none"> Payment for outcomes for individual scheme participants Payment for achieving set performance levels or targets across a cohort 	<ul style="list-style-type: none"> Upfront payment to providers for engaging users in a programme ('attachment fee') Payment for delivery of specific elements of a service ('fee for service')

Note

¹ PbR schemes between different public sector bodies are more commonly based on Memoranda of Understanding or other funding agreements rather than contracts, and the provider may invest its own funding in the programme.

Source: National Audit Office

Using PbR in public services

1.6 In the UK, the concept of paying for achieved public service outcomes dates back to at least the 19th century when it was used in education, although it was not always called PbR. The idea of linking payment with results experienced a resurgence in the past decade when government began applying it to health and welfare-to-work programmes.⁵ In 2011, the *Open Public Services White Paper* set out government's intention to use PbR as part of wider reforms to public service provision. Since then government has applied PbR to a wider range of public services, including justice and international development. The white paper stated:

“...it makes sense to build in an element of payment by results to provide a constant and tough financial incentive for providers to deliver good services throughout the term of the contract. This approach will encourage providers to work more closely with citizens and communities to build services that are both more efficient and qualitatively different.”⁶

1.7 Despite central government's support for PbR, neither the Cabinet Office nor HM Treasury currently maintains an inventory of PbR schemes across the public sector. They were unable to tell us how many PbR schemes are in operation or how much money departments have allocated to such schemes, without requesting this information from departments.

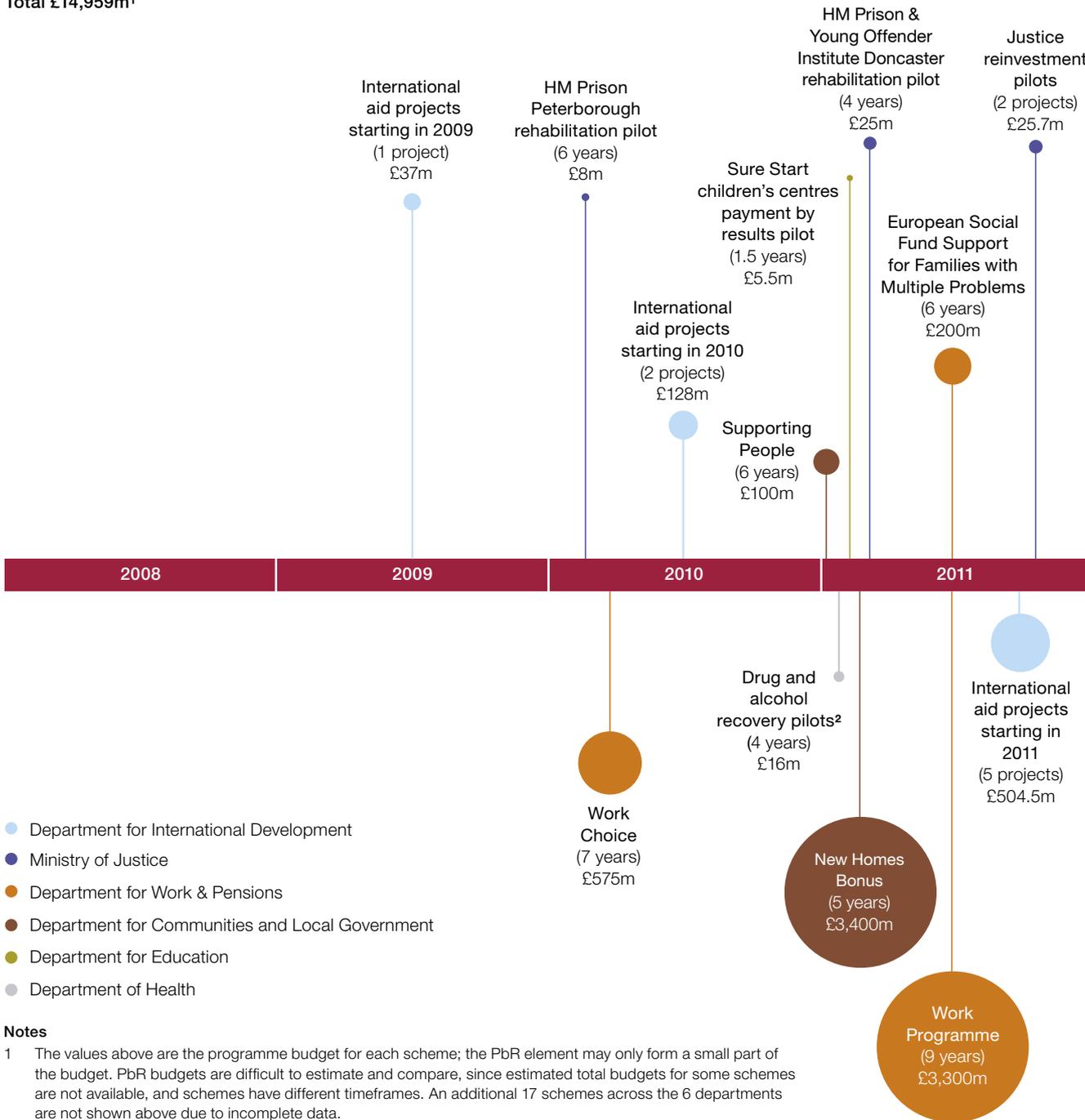
1.8 We looked at 6 central government departments known to be using PbR, and identified 52 schemes containing an element of PbR which were worth a total of at least £15 billion (**Figure 4** on pages 14 and 15). These schemes are operated by the Department for Work & Pensions (DWP), Ministry of Justice (MoJ), Department for Communities and Local Government (DCLG), Department for International Development (DFID), Department for Education (DfE), Department of Health (DoH) and their related bodies. Most of the schemes we identified started between 2011 and 2014. The proportion of PbR payment varies between schemes. Under MoJ's Peterborough offender rehabilitation pilot, 100% of payments are outcome-based; whereas the PbR component of DFID's Rwanda education sector programme is only 9%.

⁵ PbR had limited application in health as part of Commissioning for Quality and Innovation (CQUIN), which aimed to adjust payments based on clinical quality. In welfare-to-work, under the Flexible New Deal 60% of actual payments were outcome-based.

⁶ HM Government, *Open Public Services White Paper*, Cm 8145, July 2011, paragraph 5.12.

Figure 4
UK public service delivery using PbR

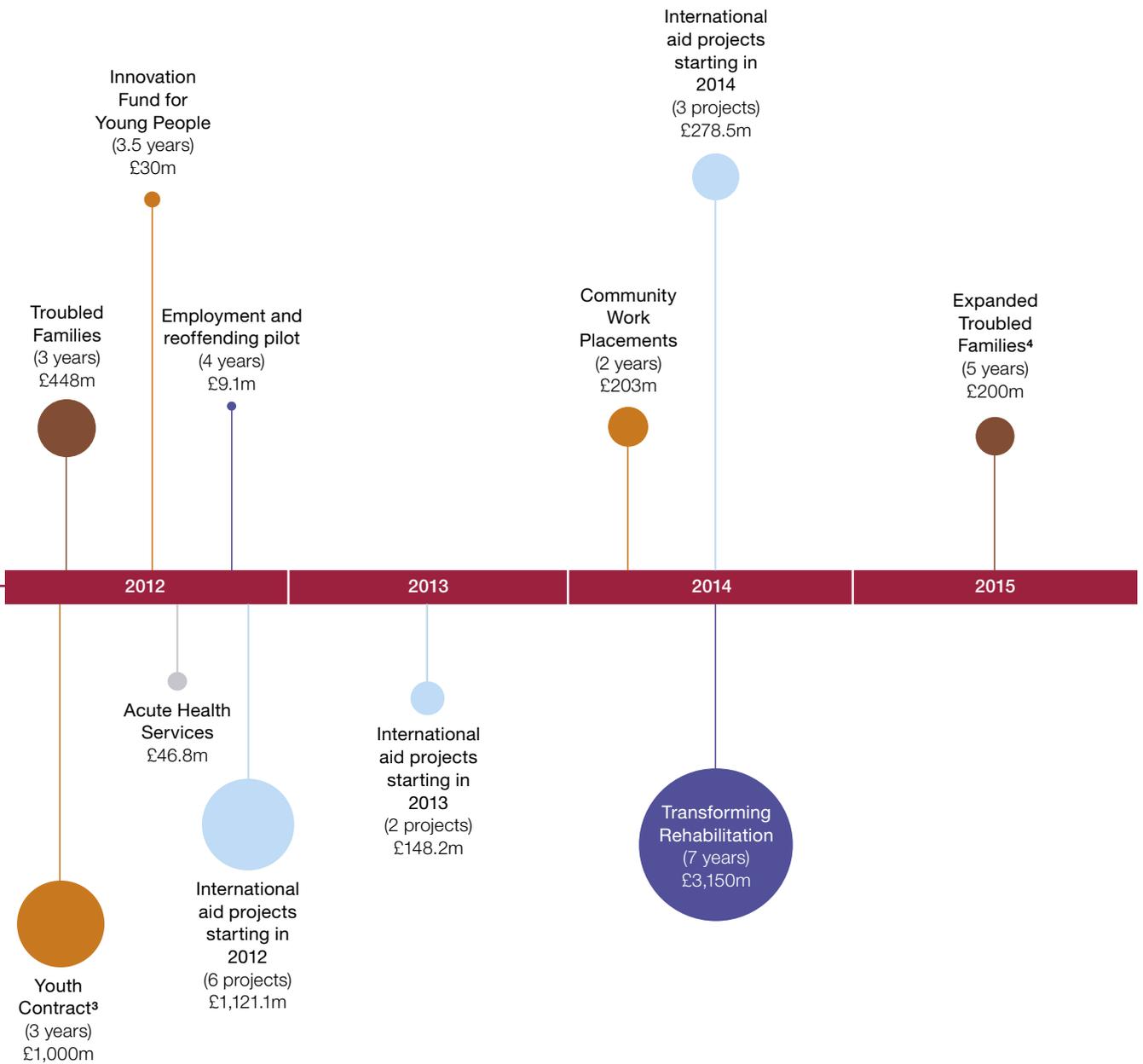
Total £14,959m¹



- Department for International Development
- Ministry of Justice
- Department for Work & Pensions
- Department for Communities and Local Government
- Department for Education
- Department of Health

Notes

- 1 The values above are the programme budget for each scheme; the PbR element may only form a small part of the budget. PbR budgets are difficult to estimate and compare, since estimated total budgets for some schemes are not available, and schemes have different timeframes. An additional 17 schemes across the 6 departments are not shown above due to incomplete data.
- 2 DoH led the drug and alcohol recovery pilots, which also involved the Home Office and MoJ.
- 3 DWP funded £660 million of the Youth Contract, with additional funding from DfE and the Department for Business, Innovation & Skills.
- 4 The expanded Troubled Families Programme's budget of £200 million is for 2015-16 only, with funding for the rest of the 5-year programme yet to be announced.



Central coordination

1.9 Between 2010 and 2014, PbR coordination and lesson sharing across central government departments was mainly carried out by an informal cross-government group of officials. In 2011, the group began work to identify public services where government could make further use of PbR based on its perceived suitability and feasibility in different circumstances. However, the group held its final meeting in March 2014, at which point it felt it had completed its work to support the development and design of PbR schemes. No alternative cross-government PbR network has been established in its place to coordinate ongoing PbR schemes and ensure emerging lessons are shared.

1.10 Neither the Cabinet Office nor HM Treasury monitors how PbR is operating across government departments. Individual spending teams within HM Treasury monitor some PbR schemes, but there is no systematic collection of lessons from using PbR or evaluation of information about whether it is delivering its claimed benefits (discussed in Part Two). Individual commissioners therefore have no central source of comparative evidence about past PbR schemes to inform their choice of delivery mechanism. This increases the likelihood of commissioners using PbR inappropriately.

1.11 Government has a growing portfolio of PbR schemes, and there is some overlap between programmes that could lead to duplication and confusion. Our report on 2 family support programmes led respectively by DCLG and DWP found them to be poorly integrated, despite funding services for a similar group of people. The departments sought to coordinate their efforts through extensive contact, a later agreement and additional resources. However, the existence of two separate programmes addressing one issue caused confusion. Our report found it contributed to the low number of referrals to the DWP's programme, which in turn affected that programme's performance.⁷

1.12 The expansion of PbR across different public services increases the need for central oversight and coordination of data on PbR's use. Such oversight would help to:

- ensure consistency across government's portfolio of PbR schemes, and identify risks of overlap between schemes;
- gather evidence to identify whether, and in what circumstances, PbR offers value for money as a delivery mechanism;
- monitor and evaluate how effectively PbR is working across government; and
- facilitate the sharing of good practice among commissioners about the design, implementation and evaluation of PbR schemes.

⁷ Comptroller and Auditor General, *Programmes to help families facing multiple challenges*, Session 2013-14, HC 878, National Audit Office, December 2013, paragraph 12.

Part Two

Deciding whether to use PbR

2.1 In this Part, we set out the issues commissioners should consider when deciding whether or not to use PbR. We also examine the evidence for the claimed benefits of PbR.

Choosing a delivery mechanism

2.2 When considering how to deliver a public service, commissioners have a range of options to choose from, including in-house provision, fee-for-service outsourcing and, in some cases, PbR. Commissioners need to look at the relative merits of the options available for use on their specific project. They should then present a clear rationale for their chosen approach which takes into account the technical challenges of using PbR and shows its use is value for money.

2.3 The rationale for choosing PbR and the rationale for spending public money pursuing the scheme's objectives should be stated separately; this ensures the choice of delivery mechanism is open to scrutiny, and makes it easier to evaluate whether using PbR had the desired effect on outcomes. It can also help inform future decisions about when and how to use PbR. DFID's business case for its PbR scheme to support Rwandan education stated that PbR was chosen to improve accountability for performance.⁸ Setting down this rationale will help DFID evaluate the project at a later date and test the effectiveness of PbR as a delivery mechanism. Not all business cases explain why PbR was chosen: the Work Programme business case did not make clear why DWP chose PbR over alternative approaches.⁹

Circumstances to which PbR is best suited

2.4 To ensure PbR is used in circumstances where it is most appropriate, commissioners need a good understanding of the available evidence about:

- a** the circumstances to which PbR is best suited;
- b** the benefits of using PbR compared with alternatives; and
- c** the challenges of designing a PbR scheme (discussed in Part Three).

⁸ Department for International Development, *Rwanda Education Sector Programme 2011/12 to 2014/15: Business Case*, paragraph 2.5.

⁹ Comptroller and Auditor General, *The introduction of the Work Programme*, Session 2010–2012, HC 1701, National Audit Office, January 2012, paragraph 1.11.

2.5 As set out in Part One, the structure of PbR schemes can vary. Whilst many involve government contracting with private sector providers to deliver services, PbR can also be used – as for the Troubled Families Programme – to provide incentives for other public sector bodies, such as local authorities. Based on government's experience since 2011, we have used our knowledge of existing schemes that apply these principles to identify the features of services to which PbR may be most suited (**Figure 5**).

Figure 5
Features of services suited to PbR

<p>Clear overall objectives, capable of being translated into a defined set of measurable outcomes</p>	<p>Well-defined, measurable outcomes make transparent the extent of the provider's success, enabling commissioners to monitor the programme and calculate payments due.</p>
<p>Clearly identifiable cohort/population</p>	<p>Before the scheme starts commissioners need to specify which individuals they are targeting, so they can track the impact of the intervention.</p>
<p>Ability to clearly attribute outcomes to provider interventions</p>	<p>Commissioners need to be sure they are rewarding providers for their genuine contribution to desired outcomes. If external factors such as economic conditions are largely responsible for changes in outcomes, PbR may not be appropriate.</p>
<p>Data available to set baseline</p>	<p>To show the impact of the scheme and set effective financial incentives, commissioners need to determine a clear baseline of performance before providers start work.</p>
<p>An appropriate counterfactual can be constructed</p>	<p>To determine the effectiveness of the scheme, commissioners need a clear counterfactual to understand the additional impact of the scheme.</p>
<p>Services are non-essential and underperformance or failure can be tolerated</p>	<p>Commissioners are likely to want closer control than PbR allows of essential services where failure might have dire consequences for public safety or the commissioner's reputation.</p>
<p>Providers exist who are prepared to take the contract at the price and risk</p>	<p>Commissioners will not be able to let the contract if providers do not bid.</p>
<p>Providers are likely to respond to financial incentives</p>	<p>If providers are not motivated by financial incentives, commissioners should question the appropriateness of PbR as a mechanism for delivering the service.</p>
<p>Sufficient evidence exists about what works to enable providers to estimate costs of delivering services</p>	<p>If there is no clear evidence about the activities that are effective in achieving outcomes, providers may be unable to estimate the costs to them of seeking to achieve outcomes, and commissioners will find it harder to price the contract.</p>
<p>Relatively short gap between provider intervention and evidence of outcome</p>	<p>PbR will be less attractive to providers if there is a long gap between the intervention (which requires upfront investment from the provider) and payment for a successful outcome. Providers may consequently prefer a higher fee for service and a lower PbR element if the gap is long.</p>

Source: National Audit Office

2.6 In practice, government has used PbR for services that lack some of the features listed in Figure 5. For example, the intended outcome of DCLG's Troubled Families programme to 'turn around' families could be attributed to other factors such as improving economic conditions or the impact of other government programmes.¹⁰ If commissioners apply PbR to a service to which it is not ideally suited, they need to do so knowingly and take steps to address the possible risks in the scheme design.

2.7 PbR can be an attractive option for commissioners in times of austerity, when there is additional pressure to deliver public services more efficiently and effectively. There is a perception that under PbR the commissioner only pays for success, while the provider is responsible for the upfront investment. There may be advantages of using PbR in appropriate circumstances, including the following benefits claimed by supporters:

- **Cost-effectiveness.** The PbR element of the contract only pays providers for the outcomes they achieve, so public money is better directed towards supporting only those interventions that are successful.
- **Innovation.** By specifying 'what' needs to be achieved rather than 'how', PbR gives greater freedom to providers, which encourages innovation in delivery.
- **Outcomes focus.** PbR focuses providers' efforts on achieving outcomes rather than delivering outputs that may or may not contribute to policy-makers' objectives.
- **Risk transfer.** For the PbR element, providers bear the risk of making upfront financial investments to deliver services without any guaranteed reward if they fail to achieve outcomes.
- **User responsiveness.** PbR schemes are more responsive to users' needs if they are provided by specialist, local-level organisations with a more detailed understanding of service users' needs, or if they give providers the freedom to tailor services to individual needs.
- **Accountability.** PbR contracts can clarify accountabilities because they hold providers to account for the delivery of specified outcomes.

2.8 Commissioners cannot assume that all the claimed benefits of PbR will be realised for every scheme. Therefore, commissioners need to identify the particular benefits they hope to achieve, to ensure design and implementation decisions focus on how best to realise those benefits and to facilitate evaluation. Our 2013 report on the Troubled Families programme found that DCLG chose to use PbR because it would focus the local authority providers' efforts on outcomes and encourage innovation.¹¹ Similarly, DFID's 2014 PbR strategy identified innovation and accountability as benefits of using PbR in international development.¹²

¹⁰ Comptroller and Auditor General, *Programmes to help families facing multiple challenges*, Session 2013-14, HC 878, National Audit Office, December 2013, paragraph 2.12. DCLG told us that the forthcoming Troubled Families evaluation will seek to identify the attributable impact of the programme.

¹¹ Comptroller and Auditor General, *Programmes to help families facing multiple challenges*, Session 2013-14, HC 878, National Audit Office, December 2013, paragraph 2.3.

¹² Department for International Development, *Sharpening incentives to perform: DFID's Strategy for Payment by Results*, June 2014, paragraph 1.11.

Evidence from public service delivery

2.9 A case for using PbR can be made if it can deliver its claimed benefits. However, the evidence that public sector PbR schemes are delivering these benefits is, so far, limited (**Figure 6**).

Figure 6

Evidence of PbR delivering claimed benefits

Claimed benefit of PbR	Examples of PbR delivering benefits	Examples of where benefits have yet to be realised
Cost-effectiveness	Our 2014 report on the Work Programme found it had delivered similar levels of performance to previous welfare-to-work programmes, but at £41 million lower cost (or 2% of the total cost of the scheme). ¹ Performance against DWP's expectations has continued to improve since, in line with wider employment trends.	Recent evidence on the Troubled Families programme highlights the difficulties of demonstrating cost-effectiveness. DCLG estimated an average financial benefit to the public purse across the 7 areas of £11,200 per family. However, DCLG notes that this figure is gross and does not take the cost of the intervention or the system's deadweight (non-intervention rate) into account. ²
Innovation	According to reviews of the Peterborough reoffending pilot providers thought the scheme was more flexible than other funding mechanisms, and encouraged innovation and joint working. ³	The 2014 evaluation of the Work Programme found limited examples of some innovation in service design and delivery at prime or subcontractor level. Providers reported reductions over time in their flexibility to innovate as DWP introduced new performance measures. ⁴
Outcomes focus	The outcomes focus of the Troubled Families programme has led to greater joining-up across local services (eg police and social services) in assisting families facing multiple challenges.	Some projects are measuring outputs rather than outcomes, eg DFID's malaria control programme in Tanzania pays for provision of bed nets rather than declining malaria rates. ⁵
Risk transfer	Under the Work Programme contract, DWP transfers the financial risk of the upfront investment in services to get people into work to its prime contractors (who may in turn pass this risk to subcontractors).	However, contractors do not bear all of the risks of underperformance; departments may have to step in to manage contracts to drive improvements. DWP terminated one of its contracts with Newcastle College Group based on the contractor's underperformance.
User responsiveness	The Work Programme uses a prime and subcontractor model, and subcontractors are often smaller, more specialist providers with experience of working with particular groups, eg ex-offenders or those with mental health issues.	However, the December 2014 evaluation of the Work Programme showed low levels of referrals to specialist subcontractors, and some subcontractors leaving the supply chain. ⁶
Accountability	For the Peterborough reoffending pilot, commissioners held detailed monthly monitoring meetings with providers to monitor progress. ⁷	Some PbR schemes (particularly DFID's) involve complex delivery chains, making it harder to ensure effective accountability for frontline providers.

Notes

- 1 Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraph 19.
- 2 Department for Communities and Local Government, *The Benefits of the Troubled Families Programme to the Taxpayer*, March 2015. DCLG has commissioned an independent evaluation (due to report in late 2015) which it expects to provide additional cost-benefit analysis.
- 3 E Disley and J Rubin, *Phase 2 report from the payment by results Social Impact Bond pilot at HMP Peterborough*, Ministry of Justice, 2014, pp. 5-6.
- 4 Department for Work & Pensions, *Work Programme evaluation: Operation of the commissioning model, finance and programme delivery*, December 2014, p. 121.
- 5 Department for International Development, *Support to Malaria Control in Tanzania Programme: Business Case*, September 2011.
- 6 Department for Work & Pensions, *Work Programme evaluation: Operation of the commissioning model, finance and programme delivery*, December 2014, p. 128.
- 7 E Disley and J Rubin, *Phase 2 report from the payment by results Social Impact Bond pilot at HMP Peterborough*, Ministry of Justice, 2014, p. 37.

2.10 In particular, we found expert opinion differs on the extent to which using PbR promotes innovation. Government has typically used PbR to tackle difficult social problems that lack ready solutions – such as reducing reoffending. Some commissioners hope PbR will give providers the freedom to innovate, which might lead to new, long-term solutions to intractable problems. However, some providers told us that, given the risks associated with it, PbR is best suited to issues to which there are known solutions and where the commissioner's overarching aim is to reduce costs; they indicated that PbR is unlikely to encourage innovation because exploring new approaches is costly and increases the provider's risk. This suggests that where commissioners want innovation, providers are likely to expect additional financial incentive. Figure 6 indicates a mixed picture to date on whether or not PbR does in fact encourage innovation.

Additional considerations

Costs to commissioners

2.11 PbR is particularly resource-intensive because of the technical challenges of setting the right payment mechanism. And, as with other forms of contracting, commissioners will also incur costs of letting the contract, and monitoring and managing providers' performance. Commissioners need to take these costs into account in their consideration of PbR compared with other approaches. They also need to make sure they have the necessary time and skills available to manage the PbR contract throughout its life.

Early identification of risks

2.12 Piloting can help commissioners manage risks by testing their scheme and enabling design weaknesses to be corrected prior to full-scale roll-out. MoJ undertook pilots for their offender rehabilitation scheme to inform later decisions about a possible nationwide scheme. While the use of pilots is good practice, MoJ decided to roll the scheme out nationally before the pilots' results were fully known. More widely, there are many examples of PbR schemes where implementation has been rushed, with very few being piloted or phased in advance of full roll-out.

2.13 Departments told us immediate PbR implementation allows the benefits to be felt quickly by the largest number of beneficiaries. This may be valid in some cases – for example, in DFID's provision of results-based aid to partner governments when piloting in just one region could lead to perceptions of inequity in other regions. However, this argument needs to be balanced against the delivery and reputational risks of rolling out a potentially flawed PbR scheme. Our Troubled Families report concluded that the speed with which DCLG and DWP had implemented their programmes meant neither department had been able to pilot detailed programme elements to identify delivery risks and take mitigating action.¹³

¹³ Comptroller and Auditor General, *Programmes to help families facing multiple challenges*, Session 2013-14, HC 878, National Audit Office, December 2013, paragraph 3.2.

Part Three

Designing PbR schemes

3.1 We noted in Part Two that PbR's benefits may include reducing headline costs, transferring risk and improving service quality (paragraph 2.7). However, PbR is a technically challenging form of specialist contracting, and poor design and implementation can have a negative impact on the quality of public services and value for money. Using government's experiences in 4 key policy areas (welfare to work, offender rehabilitation, family support and international aid), we highlight in this Part some key areas commissioners need to focus on in designing effective PbR schemes.

Setting outcomes and performance expectations

Defining outcomes

3.2 By definition, PbR emphasises the achievement of agreed outcomes that form the basis of payments to providers. Commissioners therefore need to specify clear and measurable outcomes that are aligned with their overarching policy objectives. Otherwise, there is a risk that providers and commissioners will be unable to measure the impact of interventions, or commissioners will pay for interventions that do not contribute to their overarching objectives.

3.3 The Work Programme's overall aims and outcomes are well defined and consistent with each other: they are expressed in terms of getting people into sustained work, especially those in harder-to-help groups.¹⁴ However, other PbR schemes have found it harder to specify outcomes that are readily measurable, and have instead specified measurable outputs. For example, one of DFID's PbR programmes in Tanzania aims to control malaria infection rates, but providers are paid on the basis of the number of bed nets distributed – an output rather than an outcome.¹⁵

¹⁴ Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, Figure 1.

¹⁵ Department for International Development, *Support to Malaria Control in Tanzania Programme: Business Case*, September 2011.

3.4 Ensuring outcomes align with overall aims is particularly hard when schemes have multiple objectives. For example, DCLG's Troubled Families programme aims to 'turn around' families facing multiple problems. While turning around a family might involve addressing a range of underlying issues, DCLG identified a small number of tightly defined national indicators on which to base payments, including securing employment or reducing truancy and antisocial behaviour. In DCLG's expanded programme, from April 2015, it has broadened its measures to make payments based on the achievement of both "significant and sustained progress" assessed against locally defined plans for troubled families.¹⁶

Setting expected performance levels

3.5 When planning a PbR scheme commissioners need to set expected levels of performance, and establish indicators that will demonstrate the effectiveness of the contract. These include setting the baseline of current performance before the intervention starts, and also forecasting the level of performance that would occur without intervention – known as the non-intervention rate or 'deadweight'.

3.6 Commissioners should aim to define attainable but stretching performance expectations for providers that are above the non-intervention rate. If expected performance levels are set too low, commissioners can pay for performance that would have been achieved anyway. Conversely, if performance levels are set unrealistically high, the incentive for providers is weakened and, in extreme cases, may lead to them being unfairly excluded from the market. Commissioners should therefore allow some flexibility in their contracts, so that they can respond to emerging evidence about the accuracy of their original performance expectations.

3.7 Determining baseline and non-intervention rates can be challenging and often requires assumptions and estimates to be made. DFID faces a particular challenge in setting expected performance for its PbR contracts because in many international aid settings the relevant data are unavailable. DFID's PbR project on rural water supply and sanitation in Tanzania estimated that using PbR would lead to a 10% increase in performance. However, the project's business case acknowledged the limited evidence to support the estimate because outcome-based PbR had not been tried before in the sector.¹⁷

¹⁶ Department for Communities and Local Government, *Financial Framework for the Expanded Troubled Families Programme*, March 2015, p. 10.

¹⁷ Department for International Development, *Business Case: Phase 2: Rural Water Supply and Sanitation Programme, Tanzania*, paragraph 94.

3.8 Even when data are available, commissioners can still find it difficult to set appropriate expectations of performance. In designing the Work Programme, DWP estimated the non-intervention rate (in this case the number of participants who would have found employment without the programme's support), as well as the additionality – the extra impact of specific elements of the programme, such as the PbR incentives, or the amount of time providers had to work with participants (**Figure 7**).

3.9 Although the best available at the time, the DWP data were incomplete, especially in relation to harder-to-help participants, and DWP did not adequately test the sensitivity of its assumptions. For example, it did not model performance under different economic scenarios.¹⁸ By December 2012, all Work Programme providers were on performance improvement plans because they were helping fewer participants into work than DWP's minimum expected performance levels.¹⁹ However, performance over the early years of the Work Programme showed that DWP had over-estimated the non-intervention rate, so its initial predictions of minimum performance levels were correspondingly over-optimistic (see note 2, Figure 7).²⁰

Ensuring payments create positive incentives

3.10 Our previous work shows that when designing and implementing PbR schemes commissioners have often struggled to ensure payment mechanisms create positive incentives for providers. In designing PbR schemes commissioners need to:

- understand providers, especially their likely response to 'pure' PbR and non-PbR payments; and
- structure the payment mechanism accordingly to create appropriate incentives and avoid perverse ones.

Understanding providers

3.11 To design an effective payment mechanism that will encourage providers to work towards outcomes, commissioners need to understand the motivations of likely bidders and identify relevant incentives. Some government schemes attract a range of potential bidders both large and small, and from the private and charity sectors. Different providers have different features, for example larger private providers tend to:

- seek to maximise profit;
- have moderate to good access to working capital and a balance sheet that can sustain them in the gap between intervention and payment; and
- have a reasonable appetite for risk.

18 Comptroller and Auditor General, *The introduction of the Work Programme*, Session 2010–2012, HC 1701, National Audit Office, January 2012, paragraph 8.

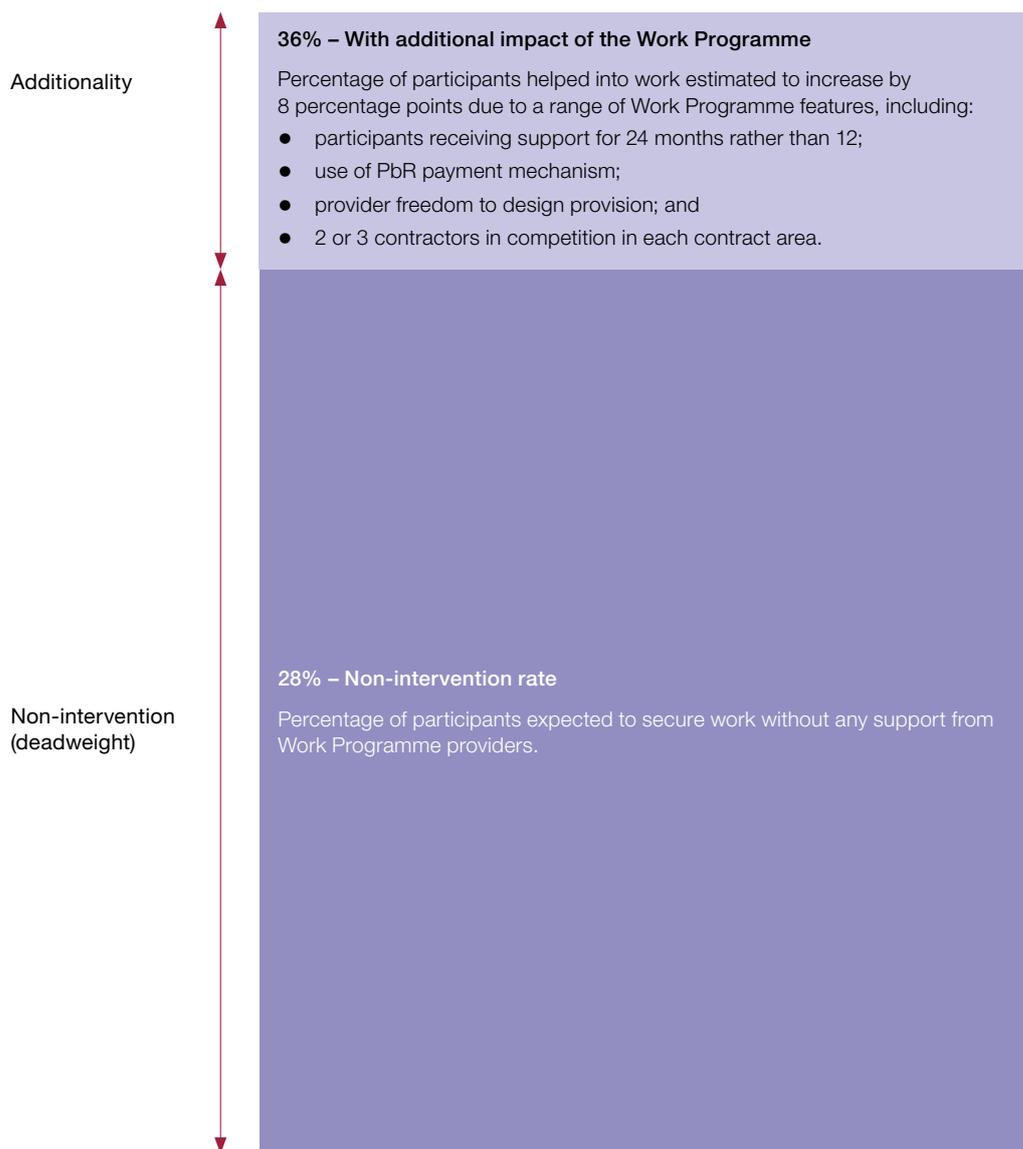
19 Comptroller and Auditor General, *A commentary for the Committee of Public Accounts on the Work Programme outcome statistics*, Session 2012–13, HC 832, National Audit Office, December 2012, paragraph 20.

20 Comptroller and Auditor General, *The Work Programme*, Session 2014–15, HC 266, National Audit Office, July 2014, paragraph 15.

Figure 7
The Work Programme's performance assumptions

DWP developed its performance expectations based on estimates of additionality and non-intervention rate

DWP's performance assumptions



Notes

- 1 Performance expectations in this figure are those DWP set for the entire programme at its start, using weighted averages for the 8 claimant groups existing at the time. They cannot be compared against current performance because DWP has since added new groups and the size of claimant groups has changed.
- 2 DWP later revised its initial performance expectations as additional data became available and it became clear the programme would not achieve original projections.

Source: Comptroller and Auditor General, *The introduction of the Work Programme*, Session 2010–2012, HC 1701, National Audit Office, January 2012, Figure 6

3.12 By contrast, small and medium-size charities:

- must pursue their charitable objectives;
- tend to have a lower appetite for risk and lower working capital;
- find it harder to sustain themselves financially in the period between intervention and payment; and
- may find it harder to absorb losses if outcomes are not attained.

3.13 Commissioners need to consider these differences in order to encourage the best performance from the different types of providers they want to work with. In some cases, commissioners will already have a good knowledge of potential providers (for example, if they have worked with them on other schemes). However, when this is not the case commissioners need to find ways to develop this understanding, for example through pre-procurement engagement with potential providers and, later, competitive dialogue.

3.14 Commissioners should also ensure they consider the whole delivery chain. The voluntary sector has raised concerns about how some prime contractors work with subcontractors, including the extent to which risk is fairly allocated and the amount and types of work referred to subcontractors.²¹ It is not enough therefore for commissioners to understand just those providers with whom they have a direct relationship; they also need to understand how those providers are likely to work with subcontractors further down the delivery chain. DWP has introduced the Merlin standard, a code of conduct that it uses to assess prime providers' treatment of subcontractors.²²

Balance of 'pure' PbR and non-PbR payments

3.15 The nature of PbR contracts means that all or part of the payment is made only after evidence that specified outcomes have been achieved. This deferral in payment is part of the attraction for commissioners, but creates risk for providers who need to finance the upfront investment in the interim. There are reported instances of smaller welfare-to-work providers withdrawing from contracts due to this time lag between investment and payment, although the welfare-to-work market has remained stable overall.²³

3.16 Commissioners need to consider how much of a scheme should be 'pure' PbR payments for outcomes achieved, and how much should be upfront payments in the form of 'attachment fees' or 'fee for service'. In practice, the 'pure' PbR element of schemes varies significantly and can change over the life of a scheme. For example, DWP paid attachment fees in the early stages of the Work Programme, but these were fully phased out by April 2014. In contrast, 100% of payments under MoJ's Peterborough offender rehabilitation pilot were based on outcomes.

21 NCVO, *Stepping Stones: the role of the voluntary sector in future welfare-to-work programmes*, July 2014. Available at: www.ncvo.org.uk/about-us/media-centre/press-releases/703-ncvo-next-work-programme-must-involve-charities-in-design-to-succeed

22 Further information on the Merlin Standard is available at: www.merlinstandard.co.uk

23 HC Work and Pensions Select Committee, *Can the Work Programme work for all user groups?*, First Report of Session 2013-14, HC 162, May 2013. Volume II: Additional Written Evidence, 21 May 2013, Ev. w10.

3.17 It is not in a commissioner's interest for a provider to collapse because of cash flow difficulties part-way through a contract, not least because this incurs unplanned costs for the commissioner. Therefore, in designing a PbR scheme, commissioners need to consider providers' capacity to manage cash flow risk and also need to know the source of potential providers' working capital.

3.18 If commissioners are keen to attract small and medium-sized enterprises and third-sector providers as bidders, they may need to consider a source of social investment to run alongside the PbR contract (for example, a social impact bond provided by a third party, as was used on the Peterborough offender rehabilitation pilot). 'Social investors' are investors who seek a social as well as a financial return. They provide the upfront capital providers need to set up services under a PbR contract, and are paid back when payment is received for the outcomes achieved.²⁴

Creating the right incentives for providers

3.19 In designing the PbR contract, commissioners need to set payments that attract providers to bid, and offer appropriate incentives for them to achieve the desired outcomes. This means setting payments at a level that covers providers' costs and offers an extra amount to cover the providers' financial risk of delivering services when payment is not guaranteed. From the commissioner's perspective, providers would ideally receive payments only for performance above the level of deadweight – in other words, to reward outcomes that have resulted directly from providers' interventions. However, this is difficult in practice because of the challenges of estimating deadweight accurately (see paragraphs 3.7–3.9) and because contracts would be likely to be so unattractive to providers that they would not bid.

3.20 We have seen that providers have been reluctant to bid for some schemes. This suggests that payment levels have inadequately reflected providers' perceptions of the cost and risk to them of taking on the contracts – even when commissioners have actively sought to understand how providers might respond to incentives. For example:

- **Offender rehabilitation**

In May 2012, MoJ tendered contracts for a pilot offender rehabilitation programme at Leeds prison. MoJ closed the competition without a successful bid after all but 1 of the 6 potential providers decided not to compete. The firms reported that the model for the pilot was unworkable. This was despite MoJ taking steps, before tendering, to understand providers' appetite for risk in taking on a PbR contract, including engaging in a process of competitive dialogue with potential bidders.

24 We do not consider social investment in detail in this report, but may return to it in future work.

- **Troubled Families**

Our report on the Troubled Families programme concluded that the incentives were not working as DCLG had anticipated. DCLG lacked information on which to make informed calculations of the cost of interventions, so had to rely on data from older family intervention schemes run from 2000 onwards. In our 2013 survey, 42 of 81 local authority respondents reported investing in services only to the extent needed to secure payments associated with attaching families to the programme (the 'attachment fees').²⁵ While local authorities and other local public bodies made in-kind contributions, local authorities said the payments for achieving longer-term, complex outcomes for the families (the PbR element of the programme) created insufficient incentives for them to make further investment.

Avoiding perverse incentives

3.21 To date, government commissioners have most commonly used PbR to tackle complex social issues for which there are no straightforward solutions. In most of these cases, there are some service users who are easier to help than others, which means providers' costs (and therefore profitability) vary between user groups. This can encourage providers to prioritise those service users who are 'easier to help', because less effort and expense is required to achieve the desired outcomes and hence receive payment.

3.22 Commissioners need to consider the risk of perverse incentives when defining outcomes and designing the payment mechanism. To discourage the neglect of harder-to-help groups, commissioners can distinguish between different groups when setting payments. DFID's Girls' Education Challenge project includes differential payments that reflect the higher costs of reaching marginalised girls. Likewise, DWP's Work Programme payment structure offers providers higher payments for achieving job outcomes for those considered harder to help (specifically, Employment and Support Allowance claimants).

3.23 However, experience so far also highlights the challenge of devising effective differential payment mechanisms. Our 2014 Work Programme report found performance for harder-to-help participants was below DWP's expectations, despite the existence of a differential payment mechanism to encourage providers to work with these individuals (see **Figure 8**).²⁶ Performance for harder-to-help groups on the Work Programme has improved since our report, but there is still limited evidence about the impact of its differential payments.

25 Comptroller and Auditor General, *Programmes to help families facing multiple challenges*, Session 2013-14, HC 878, National Audit Office, December 2013, paragraph 9.

26 Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 16-17.

Figure 8

Work Programme performance for easier- and harder-to-help groups, up to March 2014

	Performance for 'easier-to-help' Jobseeker's Allowance claimants aged 25 and over ¹	Performance for 'harder-to-help' Employment and Support Allowance claimants
	Number of claimants in completed cohorts	Number of claimants in completed cohorts
Referral	316,000	41,000
Attachment and support	310,000 (98%)	40,000 (97%)
Off benefit	40% (proportion of time)	18% (proportion of time)
Job outcome achieved	85,000 (27%)	5,000 (11%)
Sustainment achieved	1,320,000	105,000
Completed programme without an outcome	231,000 (73%)	36,000 (89%)

Note

¹ These claimants are typically referred to the Work Programme after 12 months of unemployment.

Source: Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, Figures 7 and 12

3.24 A further challenge for commissioners is to ensure that a short-term outcome measure does not lead providers to rely on 'quick wins' that have no lasting impact on users' longer-term outcomes. The risk is that providers focus on activities that trigger payments, and that they have no incentive to ensure that results are sustainable. Providers told us that one of the risks of PbR is that it merely prompts providers to cut costs – and that, in a worst case scenario, providers seek to maximise their profit margin on the attachment fee alone by cutting costs as far as possible, and largely ignore the PbR payment element.

3.25 The Work Programme sought to address this by offering providers 'sustainment payments' for every 4 weeks that a participant remained in employment after securing a job. Some of the DFID projects we reviewed also sought to secure sustainability of outcomes in their payment design: for example, under one programme, providers received payments based on the sustained use of improved water, sanitation and hygiene facilities.²⁷

²⁷ Department for International Development, *Business Case and Intervention Summary: Water, Sanitation and Hygiene (WASH) Results Programme*.

Part Four

Monitoring and evaluating PbR

4.1 In this Part, we look at the extent to which government has monitored the performance and evaluated the effectiveness of PbR schemes across the main public services to which it has been recently applied. We also consider the extent to which the effectiveness of the PbR mechanism itself has been evaluated.

Monitoring providers' performance

4.2 Effective data and monitoring systems are needed to ensure:

- payment is made only for outcomes that have been delivered; and
- poor performance or service quality is identified so commissioners are able to address performance issues promptly.

Data to support payments

4.3 PbR contracts need to be underpinned by robust performance information systems. Payments are linked to outcomes achieved and the number of interventions the provider begins (in the case of attachment fees). Commissioners therefore need accurate, reliable and timely information on the numbers of outcomes achieved and participants attached to programmes, to ensure correct payments are made.

4.4 Commissioners also need to consider what proportionate method they will use to validate reported performance, prior to payment. In some cases – for example, the developing countries where DFID operates – there may be no source of independent or third-party data on outcomes. In these circumstances, commissioners will need to use data from providers. This increases the importance of a robust validation mechanism to manage the risk that providers may misrepresent the outcomes achieved to boost payment.

4.5 On the Work Programme, DWP designed the payment regime and let contracts before establishing a reliable way to validate sustainment payments. As a result, it was potentially paying contractors for performance they were not actually achieving. DWP estimates that it paid £11 million in invalid payments up to March 2014.²⁸

4.6 Our analysis indicates that commissioners commonly overlook the need for sound performance monitoring systems to be in place **before** a scheme starts. This is especially the case where implementation is rushed. The offender rehabilitation pilots were hampered because MoJ had not established effective data sharing between its prison and probation systems and the police national computer when piloting began. As a result, providers had to develop temporary, often time-consuming solutions to get the information they needed to operate the scheme.

Monitoring provider performance

4.7 Commissioners need good performance information to receive early warning of underperformance or undesirable provider behaviour that could mean poor value for money. This information also helps commissioners provide assurance to the public and Parliament about the effective use of public funds.

4.8 By linking payments to outcomes achieved, PbR allows commissioners to transfer **some** risk to providers: the provider bears the risk of the upfront investment in the service, only receiving full payment if they achieve the specified outcomes. However, PbR does not enable commissioners to transfer **all** risk. Public criticism about service quality or the methods used by providers to achieve the desired outcomes can have a negative effect on a commissioner's reputation, so early warning of service decline is desirable. This is particularly the case when a provider is likely to fail or exit a contract, which can generate unplanned costs for the commissioner in maintaining continuity of service.

4.9 DWP initially found it difficult to monitor service quality across Work Programme providers. Prime contractors had set their own minimum service standards, and DWP struggled to track the resulting 212 standards. In 2013, DWP introduced 14 standardised quality measures to streamline its monitoring of service quality for each contractor.²⁹

²⁸ Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraph 22. DWP has since recovered the overpayments as part of wider contract change negotiations, which were cost-neutral overall.

²⁹ Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 3.10–3.13.

Managing provider performance

4.10 Because poor provider performance can have a negative impact on outcomes for service users and commissioners' reputation, commissioners need to use the results of monitoring to manage provider performance. DWP has introduced several positive examples of active contract management on the Work Programme:

- **Contract renegotiations:** DWP has revised contractual terms and conditions several times to address limitations in the original contracts. By March 2014, it had introduced over 30 contract variations, which it estimates could save £40 million over the life of the programme.³⁰ Since then, DWP has continued to negotiate changes to contracts to strengthen its oversight of provider performance.
- **Shifting market share:** The design of the Work Programme explicitly encouraged competition in its 18 contract areas, with 2 or 3 prime contractors operating in each area.³¹ If an individual provider performs poorly in a given area, DWP can shift the level of referrals between contractors; it exercised this contractual right in August 2013 and March 2015.
- **Performance improvement plans:** DWP discusses minimum performance and minimum service levels with its prime contractors at least monthly, and can put contractors on performance improvement notices if performance is below expectations. DWP has also issued formal letters to several providers requiring improvement across a range of measures.³²

Evaluating effectiveness

4.11 As well as monitoring and managing performance to improve PbR schemes while they are operating, commissioners need to evaluate the use of PbR. There are two aspects to evaluating PbR, specifically:

- at an **individual scheme level**, PbR's effectiveness in delivering the objectives of the particular scheme to which it has been applied; and
- as a **delivery mechanism**, PbR's effectiveness compared with other mechanisms.

4.12 Good evaluation starts at the planning stage, with a clear statement of what commissioners are seeking to achieve and how value for money will be assessed. However, none of the PbR schemes we looked at were explicit at the outset about how value for money would be evaluated – making it difficult for commissioners and others to assess subsequently whether, in practice, they did offer value for money.

³⁰ Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 4.25-4.26.

³¹ Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 1.9-1.10.

³² Comptroller and Auditor General, *A commentary for the Committee of Public Accounts on the Work Programme outcome statistics*, Session 2012-13, HC 832, National Audit Office, December 2012, paragraph 20.

Evaluating PbR as a delivery mechanism

4.13 When using an approach for which there is little existing evidence of effectiveness – as is the case with PbR – it is important to gather emerging evaluative information and share it promptly with others. Evaluating performance against a robust counterfactual in the form of a control group would enable government to build a strong evidence base for PbR's effectiveness as a delivery mechanism. Evaluation also needs to consider whether using PbR has met the expected or claimed benefits, for example increased innovation, cost-effectiveness or accountability (see Figure 6).

4.14 To date, only a few schemes have monitored performance against a counterfactual, for example MoJ's offender rehabilitation pilot in Peterborough and DFID's Girls' Education Challenge. The relative strength of different comparators that can be used to indicate the effectiveness of PbR as a delivery mechanism are set out in **Figure 9**. Although control groups are resource-intensive to set up, they provide an excellent counterfactual with which to compare performance.³³

4.15 The Committee of Public Accounts has criticised DWP for not using control groups to provide a clear indication that the Work Programme is resulting in improved performance.³⁴ DWP has acknowledged that it would be desirable to compare Work Programme performance against that of control groups. DWP has indicated that it will consider introducing them in future while bearing in mind any ethical issues associated with withholding support – although PbR requires a robust counterfactual to identify the additional impact of providers' interventions, creating a control group (to form that counterfactual) may involve withholding support from people in need.³⁵

³³ National Audit Office, *Evaluation in government*, December 2013, Figure 9.

³⁴ HC Committee of Public Accounts, *The Work Programme*, Twenty-first Report of Session 2014-15, HC 457, November 2014, paragraph 5.

³⁵ HM Treasury, *Treasury Minutes: Government responses on the Eighteenth, the Twenty First to the Twenty Fourth, and the Thirty Third reports from the Committee of Public Accounts: Session 2014-15*, Cm 9013, February 2015, p. 6.

Figure 9

Relative merits of comparators

A control group is the strongest comparator

Comparator	Comment	Example
Control group ¹ with similar characteristics subject to alternative provision (eg in-house provision)	A strong comparator: enables the effectiveness of PbR to be compared with another delivery method. Control group needs to be established before the scheme starts.	DFID's Northern Uganda health PbR project compares results against a control group receiving alternative provision (a contract for provision of inputs such as drugs and personnel).
Control group with similar characteristics subject to no intervention	A strong comparator (although ethical concerns may exist): enables the effectiveness of PbR in tackling a specific issue to be quantified by comparing PbR performance with no intervention. This counterfactual indicates the actual deadweight or non-intervention rate. Control group needs to be established before the scheme starts.	Peterborough offender rehabilitation pilot assesses PbR performance against a national control group of offenders receiving no additional support.
Modelled counterfactual for alternative provision	A moderate comparator: comparisons against a modelled estimate of the likely impact of alternative delivery approaches are dependent on the accuracy of the model, its underlying assumptions and data availability.	DFID plans to evaluate its results-based aid for Rwandan education on the basis of a modelled counterfactual of the costs and benefits of providing the aid through an alternative approach.
Estimate of non-intervention rate, or 'deadweight' (estimate of what would have been achieved without intervention)	A moderate comparator: comparisons against a deadweight estimate are dependent on the accuracy of data and assumptions that underlie the estimate.	The Work Programme's initial estimates of the non-intervention rate were over-optimistic because they were based on assumptions about the state of the economy that did not prove correct.
Baseline of performance at start of scheme	A weak comparator: comparisons against baseline performance do not capture the impact of what would have happened anyway, without the intervention.	Doncaster offender rehabilitation pilot (launched October 2011) compares performance against a baseline of the prison's 2009 reconviction rate.
Commissioners' expectations of performance	A weak comparator: performance expectations are dependent on commissioners' assumptions.	DWP acknowledges that it set initial performance expectations for the Work Programme too high. ²

Notes

1 Ideally any control group used as a comparator should be identified at random, although this may not always be practical.

2 Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraph 15.

Source: National Audit Office

Evaluating individual PbR schemes

4.16 Government-commissioned evaluations of major PbR schemes have so far largely focused on implementation rather than effectiveness or value for money.

- The first 2 Work Programme evaluations were qualitative reviews of how providers were delivering the scheme.³⁶ They presented little discussion of the programme's outcomes apart from noting the ineffectiveness of differential pricing in encouraging providers to work with the harder-to-help groups. The third evaluation examined how providers worked under the scheme and how DWP managed them.³⁷ It included some analysis of outcomes, concluding that by 2013 the percentage of participants entering work appeared to have increased compared with the previous evaluation period.
- The independent evaluation of DCLG's Troubled Families programme is due to report in late 2015. An early non-evaluative report which presented monitoring data on the characteristics of participating families was published in July 2014.³⁸
- Evaluations of the Peterborough and Doncaster offender rehabilitation pilots considered the effects of introducing PbR contracts on service delivery.³⁹ They concluded that the pilots had enabled more individualised support, and reported benefits such as encouraging innovation in how services were delivered.
- None of DFID's projects have yet reached the final evaluation stage but DFID has developed a framework for evaluating PbR schemes, which recommends evaluations include consideration of schemes' impact and value for money.

4.17 Government departments periodically publish performance data on individual PbR schemes, providing further evidence to evaluate their impact. However, so far, the most detailed value-for-money assessment of how effectively a PbR mechanism has worked in delivering public services is our 2014 report on the Work Programme. **Figure 10** on pages 37 and 38 summarises the available evidence on PbR schemes' performance to date.

36 B Newton, N Meager, et al, *Work Programme evaluation: Findings from the first phase of qualitative research on programme delivery*, Department for Work & Pensions, December 2012; P Lane, R Foster, et al, *Work Programme evaluation: Procurement, supply chains and implementation of the commissioning model*, Department for Work & Pensions, March 2013.

37 Department for Work & Pensions, *Work Programme evaluation: Operation of the commissioning model, finance and programme delivery*, December 2014.

38 Ecorys UK, *National Evaluation of the Troubled Families Programme: Interim Report Family Monitoring Data*, Department for Communities and Local Government, July 2014.

39 E Disley, J Rubin, et al, *Lessons learned from the planning and early implementation of the Social Impact Bond at HMP Peterborough*, Ministry of Justice, May 2011; D Murray, J Jones, et al, *Findings and lessons learned from the early implementation of the HMP Doncaster payment by results pilot*, Ministry of Justice, November 2012; E Disley and J Rubin, *Phase 2 report from the payment by results Social Impact Bond pilot at HMP Peterborough*, Ministry of Justice, 2014; E Hichens and S Pearce, *Process evaluation of the HMP Doncaster Payment by Results Pilot: Phase 2 findings*, Ministry of Justice, 2014.

Figure 10

Reported performance on key operational PbR schemes

Programme	Reported performance	Nature of comparator
Work Programme (DWP)	<p>Our 2014 Work Programme report found performance was similar to past schemes, but cost around £41 million less. Among people claiming Jobseeker's Allowance aged 25 and over, 27% moved into employment lasting 6 months or longer. Of the 'harder-to-help' Employment and Support Allowance (ESA) claimants, 11% achieved job outcomes.¹</p> <p>Scheme performance did not meet DWP's higher performance expectations, especially for harder-to-help claimants. DWP's initial expectation was that 22% of ESA claimants would find work. This was later revised to 13% and the actual level of performance was 11%.²</p> <p>The 2014 report examined results for participants who had completed the programme at that time. Outcomes have improved for participants who have since finished.</p>	<p>Performance and cost comparisons with predecessor schemes.</p> <p>Commissioner's expectations of performance.</p> <p>No control group.</p>
Offender rehabilitation pilots (MoJ)	<p>Peterborough and Doncaster pilots showed improvements in reconviction rates for their first groups of offenders.</p> <p>The frequency of reconviction events for the first cohort at Peterborough was 8.4% lower than its comparator (national control group) but did not meet commissioner's expectations.³ Interim data show the frequency of reconviction events for the second cohort is 2.3% lower than its comparator.⁴</p> <p>The reconviction rate for the first cohort at Doncaster was 5.7 percentage points lower than its comparator (2009 baseline) which meant it exceeded its target of a 5.0 percentage point reduction in reconviction rate.⁵ Based on early data, the reconviction rate for the second cohort is 3.9 percentage points lower than the comparator period in 2008-09.⁶</p>	<p>Peterborough: national control group. Also commissioner's expectation of 10% reduction in frequency of reconviction events required to trigger payments.</p> <p>Doncaster: 2009 baseline. Also commissioner's expectation of 5.0 percentage point reduction in reconviction rate.</p>
Troubled Families (DCLG and DWP)	<p>In March 2015, DCLG reported that under its Troubled Families programme, 105,671 of its identified families had been 'turned around' and estimated a potential saving of up to £1.2 billion.⁷ 57 of 152 authorities had turned around 100% of the families in their programmes.⁸ This is DCLG's account of performance based on reporting from local authorities and has not been independently verified. DCLG expect to publish the final programme results in late 2015.</p>	<p>Reported performance to March 2015 is not compared with any estimates of deadweight or comparisons against counterfactuals, although DCLG told us the forthcoming national evaluation will analyse the programme's impact using comparison groups of families.</p>

Figure 10 *continued*

Reported performance on key operational PbR schemes

Programme	Reported performance	Nature of comparator
DFID's international aid projects	Project evaluations are ongoing and not yet at final stage. Annual reviews are published for each project, which assess emerging evidence on effectiveness and value for money.	

Notes

- 1 Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 11, 16, 19.
- 2 Comptroller and Auditor General, *The Work Programme*, Session 2014-15, HC 266, National Audit Office, July 2014, paragraphs 3.3–3.4.
- 3 Ministry of Justice, *Peterborough Social Impact Bond/HMP Doncaster: Payment by Results pilots: Final re-conviction results for cohorts 1*, August 2014, p. 2.
- 4 Ministry of Justice, *Proven Re-offending Statistics Quarterly Bulletin: July 2012 to June 2013, England and Wales*, Annex A: Interim re-conviction figures for Peterborough and Doncaster Payment by Results pilots, April 2015, p. 25.
- 5 Ministry of Justice, *Peterborough Social Impact Bond/HMP Doncaster: Payment by Results pilots: Final re-conviction results for cohorts 1*, August 2014, p. 3.
- 6 Ministry of Justice, *Proven Re-offending Statistics Quarterly Bulletin: July 2012 to June 2013, England and Wales*, Annex A: Interim re-conviction figures for Peterborough and Doncaster Payment by Results pilots, April 2015, p. 27.
- 7 The estimated £1.2 billion saving is based on an assumption that the average savings identified in 7 sample local authorities will be replicated across other areas: Department for Communities and Local Government, 'More than 105,000 troubled families turned around saving taxpayers an estimated £1.2 billion', 10 March 2015, available at: www.gov.uk/government/news/more-than-105000-troubled-families-turned-around-saving-taxpayers-an-estimated-12-billion.
- 8 Department for Communities and Local Government, 'Troubled Families programme: progress information and families turned around', 10 March 2015, available at: www.gov.uk/government/publications/troubled-families-programme-progress-information-at-december-2014-and-families-turned-around-at-february-2015.

Source: National Audit Office

Appendix One

Our audit approach

1 This report reviews the government's recent experience of using payment by results (PbR) to deliver public services. To reach our conclusions, we examined:

- government's use of PbR for public service delivery;
- the issues commissioners should consider when deciding whether or not to use PbR;
- challenges in designing PbR schemes; and
- how government evaluates the effectiveness of PbR schemes, and the PbR mechanism itself.

2 We developed an analytical framework to guide public sector commissioners who are considering using PbR.⁴⁰ The framework sets out the main issues commissioners need to take into account when selecting, designing, implementing and evaluating PbR schemes:

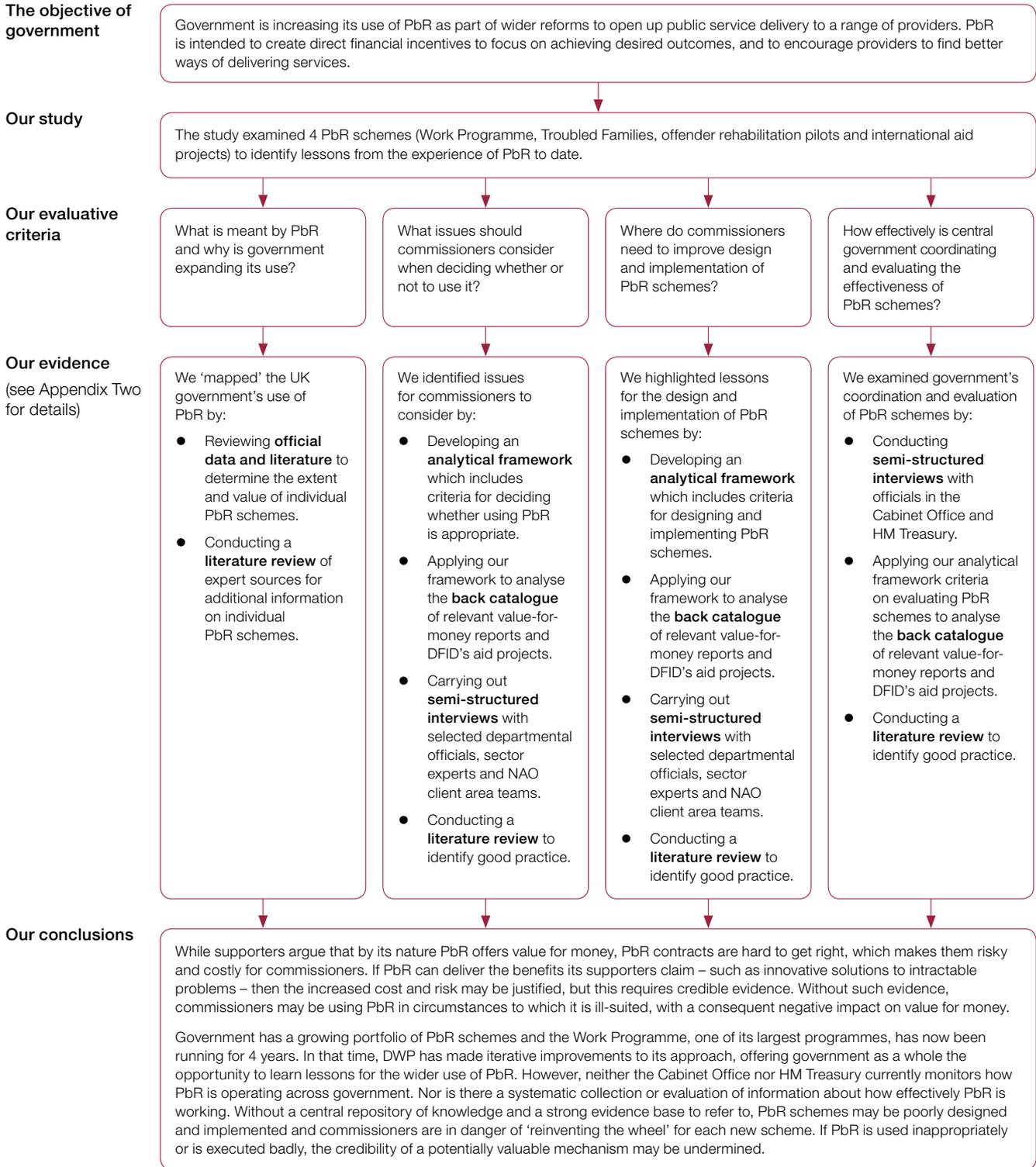
- Overall fit: should you use PbR to deliver this service?
- Design: how can you design an effective PbR scheme?
- Implementation: what do you need to have in place to implement your PbR scheme effectively?
- Evaluation: how can you evaluate the effectiveness of a PbR scheme?

3 We applied the framework criteria to analyse how PbR had been used in 4 areas: the Work Programme, Troubled Families, offender rehabilitation pilots and international aid projects.

4 Our audit approach is summarised in **Figure 11** overleaf. Our evidence base is described in Appendix Two.

⁴⁰ See Figure 1 on page 10.

Figure 11
Our audit approach



Appendix Two

Our evidence base

1 We reached our conclusions on government's use of PbR so far by using evidence collected and reviewed between October 2014 and February 2015.

2 We developed an **analytical framework for commissioners** to identify the issues that commissioners need to consider when deciding whether to use PbR, and when designing, implementing and evaluating schemes.

3 We reviewed the findings of past **National Audit Office (NAO) and Committee of Public Accounts reports and other work** to understand how PbR has been used in the following areas:

- Welfare to work: the Work Programme;
- Family support: Troubled Families; and
- Criminal justice: offender rehabilitation pilots.

We also drew on the expertise of the relevant **NAO client area teams** to gain insights into the experience of PbR schemes so far, and the challenges facing commissioners.

4 We conducted **analysis of the outcome-based international aid projects** operated by the Department for International Development, because this is an area in which government intends to expand the use of PbR.

5 We conducted **interviews** with:

- central departments: Cabinet Office and HM Treasury, to understand the centre of government's perspective on using PbR;
- service delivery departments: Department for International Development and Department for Communities and Local Government, to understand the perspective of departments that have commissioned PbR schemes; and
- representative bodies: Employment Related Services Association, Local Government Association, National Council of Voluntary Organisations and Bond (the international development NGO representative body), to understand the perspective of providers.

6 We reviewed **official documents and published performance data** to identify the extent and value of PbR schemes in 6 service delivery departments.

7 We conducted a **literature review of expert sources**, including international and private sector comparisons, to provide information to support our 'mapping' of PbR schemes and to identify good practice in selection, design, implementation and evaluation of PbR. We engaged ICF Consulting to carry out the literature review.

Glossary

Additionality	Impact of an activity or programme above what would have been achieved in its absence.
Attachment fee	Payment for signing up participants to a programme.
Attribution	Ability to link a specified intervention with the achievement of a specified outcome.
Baseline	The level of performance measured before the intervention begins, against which the intervention's impact can be assessed.
Cohort	Group of people identified to receive intervention.
Commissioner	Organisation which funds or contracts for delivery of a service.
Control group	Group of participants in a programme or trial who receive no intervention or an alternative intervention, which is then used as a comparator to assess the impact of the intervention.
Counterfactual	Method of indicating what would have happened either a) without the intervention, or b) with an alternative intervention, to provide a comparator.
Deadweight	Change in the outcome that would have happened regardless of the intervention (also called the 'non-intervention rate').
Delivery chain	Network of organisations involved in delivering a service to service users (equivalent to a supply chain for products).
Fee for service	Payment based on service levels or outputs delivered, rather than outcomes.
Intervention	Activities undertaken with the intention of producing the desired outcome.
Market share shift	Ability of commissioner to divert referrals (for example jobseeker referrals) to alternative providers if a provider is performing poorly.
Non-intervention rate	See 'Deadweight' above.
Outcome	A result or change experienced by a person, family or community, for example improved parenting.
Output	Unit of service delivered, for example number of people completing a programme.

Payment by results	Practice of paying providers for delivering public services wholly or partly on the basis of results achieved.
Payment mechanism	Basis and terms on which payments are made (including timing of payments).
Perverse incentive	An incentive that has unintended results which go against the desired outcome or aims of the programme.
Prime contractor	Provider who is directly contracted to deliver a service and acts as the single point of contact for the commissioner. Prime contractors may pass on work to subcontractors.
Provider	Organisation which is contracted or funded to deliver the service.
Risk transfer	The shifting of risk from one party to another, for example shifting delivery risk from the commissioner to provider.
Service users	People receiving public services, for example job seekers, ex-offenders.
Social impact bond	A contractual arrangement involving at least three separate legal parties – a commissioner, an investor and a provider – where payment depends wholly or partly on achieving specified outcomes and where some or all of the financial risk of non-delivery of outcomes sits with the investor.
Social investment	The provision of capital for the purpose of generating social as well as financial returns.
Subcontractor	Organisation hired and paid by a prime contractor to provide specific services as part of the overall contract.

This report has been printed on Evolution Digital Satin and contains material sourced from responsibly managed and sustainable forests certified in accordance with the FSC (Forest Stewardship Council).

The wood pulp is totally recyclable and acid-free. Our printers also have full ISO 14001 environmental accreditation, which ensures that they have effective procedures in place to manage waste and practices that may affect the environment.



National Audit Office

Design and Production by NAO Communications
DP Ref: 10694-001

£10.00

ISBN 978-1-904219-85-9



9 781904 219859
