PAYMENT BY RESULTS

2016

Payment by Results:
Lessons from the Literature

RUSSELL WEBSTER

FOREWORD

This literature review was undertaken as the first stage of a project funded by the Oak Foundation to develop an interactive tool to assist commissioners and providers decide whether payment by results (PbR) might be an effective approach to commissioning a particular public service.

The Oak Foundation is an international charitable trust, whose mission is to address issues of global social and environmental concern, particularly those that have a major impact on the lives of the disadvantaged.

I would like to take this opportunity to thank all of those involved in the project so far for their advice and guidance; particularly the members of the project's Advisory Board (AB), who have contributed their time for free: Kate Aldous (AB); Tim Bennett; Shane Britton (AB); Prof Chris Fox (AB); Tom Gash; Les Hems; Daria Kuznetsova (AB); Alex Martin; Marcus Roberts (AB); Ben Robinson (AB); Hannah Robinson (AB); and Eileen O'Sullivan.

Special thanks to Tom Mason for his role as an exemplary peer reviewer. Needless to say, I remain solely responsible for any errors or failings in the text.

You can find further information on this project here and a wide range of, free-to-download, resources on payment by results on my website/blog.

Russell Webster

January 2016

CONTENTS

Click here to learn more about the interactive PbR tool

1: ABOUT THIS REVIEW

Payment by Results is becoming an increasingly common commissioning model with over 50 schemes containing an element of PbR worth a total of at least £15 billion of public money in the UK (National Audit Office, 2015).

As a number of commentators (e.g. Audit Commission, 2012; Hunter & Breidenbach-Roe, 2013) remark, there is not sufficient evidence to draw firm conclusions about whether PbR is an effective commissioning approach for public services. However, there are enough good quality reviews and evaluations to identify key learning points which could prove invaluable for commissioners and providers as they design and operate PbR contracts. Battye (2015) shares a similar view; arguing that empirical evidence on the effectiveness of PbR is lacking and evaluations should focus on:

Payment by Results is becoming an increasingly common commissioning model with over 50 schemes containing an element of PbR worth a total of at least £15 billion of public money in the UK (National Audit Office, 2015).

As a number of commentators (e.g. Audit Commission, 2012; Hunter & Breidenbach-Roe, 2013) remark, there is not sufficient evidence to draw firm conclusions about whether PbR is an effective commissioning approach for public services. However, there are enough good quality reviews and evaluations to identify key learning points which could prove invaluable for commissioners and providers as they design and operate PbR contracts. Battye (2015) shares a similar view; arguing that empirical evidence on the effectiveness of PbR is lacking and evaluations should focus on:

> ***"explanation (looking within programmes; surfacing and testing their theories) and refinement (using findings to improve programmes; using theories to look across programmes) to advance the debate on PbR from 'for or against?' to 'when, how, for whom and under what circumstances?"***

One of the key themes to emerge from this literature review (e.g. Lagarde et al, 2013; ARTD, 2014) is the limited amount of discussion and negotiation which takes place between commissioners and providers at all stages of the process: from considering whether PbR might be an effective approach, though specifying outcomes and incentives, to designing the contract itself.

The principal objective of this literature review (and the project as a whole) is to encourage and facilitate this dialogue by helping stakeholders to:

- ***Clarify the purpose of an individual PbR scheme;***
- ***Identify critical success factors;***

- ***Identify common issues which cause problems and difficulties; and***
- ***Be aware of ways to address or at least mitigate these problems.***

The findings of this literature review will inform a series of workshops with researchers, commissioners, providers and service users to design a prototype interactive tool to help all parties think through the main issues of using a PbR contracting model. The tool will then be piloted and launched in the late summer/ early autumn of 2016.

Structure of this review

The next chapter (2) explores the concept of payment by results and develops the working definition of PbR for this review. Chapter 3 provides an overview of the literature analysed for this review. Chapter 4 looks at the background to payment by results and explains why it has become such a contentious approach. The chapter acknowledges that there is a lack of a definitive evidence base on whether PbR works, although it still suggests there is value in presenting an overview of both more and less successful PbR schemes to date. Chapter 5 explores the different purposes and rationales for the development of PbR-funded public services. Chapter 6 mines the literature to identify a range of factors which are associated with successful PbR schemes as well as those found to be linked to common failings. Finally, Chapter 7 presents the conclusions of this review.

Jargon

Payment by results has generated a good deal of jargon particularly around the issues of setting outcome measures and the gaming of these measures. Where terms such as "***black box***", "***creaming and parking***", ***counterfactuals*** etc. are used in the text for the first time, readers are referred (by hyperlink) to the glossary of PbR jargon at the end of this review.

2: DEFINING PbR

Payment by results is an increasingly popular method of paying providers of public services predominantly because of its promise of improved effectiveness and more efficient use of public money.

Inevitably, this popularity has resulted in a wide range of approaches with quite different characteristics all being described as PbR schemes. Crowe et al. (2014) conclude that there is much less PbR in practice than is claimed. Some researchers and policy makers use a range of parallel terms including:

pay-for-performance	**performance contracting**	**outcome commissioning**	**financial incentives**
pay for success	**outcome-based payment schemes**	**prospective reimbursement**	

For the purpose of this review, I have sought to concentrate on the main characteristics of PbR schemes used across the broad social care sector in the UK including initiatives such as the Work and Troubled Families Programmes, the new probation contracts[1] as well as schemes to help homeless people[2] and those dependent on drugs and/or alcohol[3].

The two most important characteristics of these schemes are:

- ***A focus on commissioning outcomes rather than outputs or activity; and***
- ***The alignment of contract rewards (and/or penalties) to incentivise the achievement of these outcomes.***

Therefore, the working definition of PbR for this review is:

"a commissioning approach to the delivery of public services where contract payments are wholly or partly dependent on the achievement of specified outcomes."

1 The MoJ has contracted with the 21 new (as of February 2015) Community Rehabilitation Companies to reduce reoffending rates within a PbR framework.
2 DCLG(2014)&DCLG(2015)
3 DrugScope & RSA (2013)

Click here to learn more about the interactive PbR tool

3: THE LITERATURE REVIEWED

Criteria

The selection criteria for the studies reviewed were developed to reflect the purpose of the literature review – to facilitate dialogue between commissioners and providers by helping stakeholders to:

- ***Clarify the purpose of an individual PbR scheme;***
- ***Identify critical success factors;***
- ***Identify common issues which cause problems and difficulties; and***
- ***Be aware of ways to address or at least mitigate these problems.***

Three main criteria were used to identify relevant literature:

- ***Studies focused on payment by results schemes which met the working definition quoted above***
- ***Studies in English and predominantly in the UK (in order that the eventual interactive tool reflects the way in which public services are commissioned in this country)***
- ***Studies published from 2008 onwards (in order to reflect the significant changes in the approach to the commissioning of public services since this date).***

Exceptions were made for studies which were highly recommended within the literature or by commentators or researchers of renown within their relevant field (either commissioning of public services or a key social care sector such as worklessness, substance misuse etc.).

Owing to the perceived controversial nature of PbR and its conflation with issues of privatisation and cuts in public services, there are a large number of articles which raise concerns about PbR without either a systematic review of the literature or the first hand evaluation of a particular PbR initiative. On the whole this type of "PbR-phobic" literature has been avoided, particularly because there are a substantial number of more rigorous studies which advance a raft of criticisms of the payment by results approach.

Identifying studies

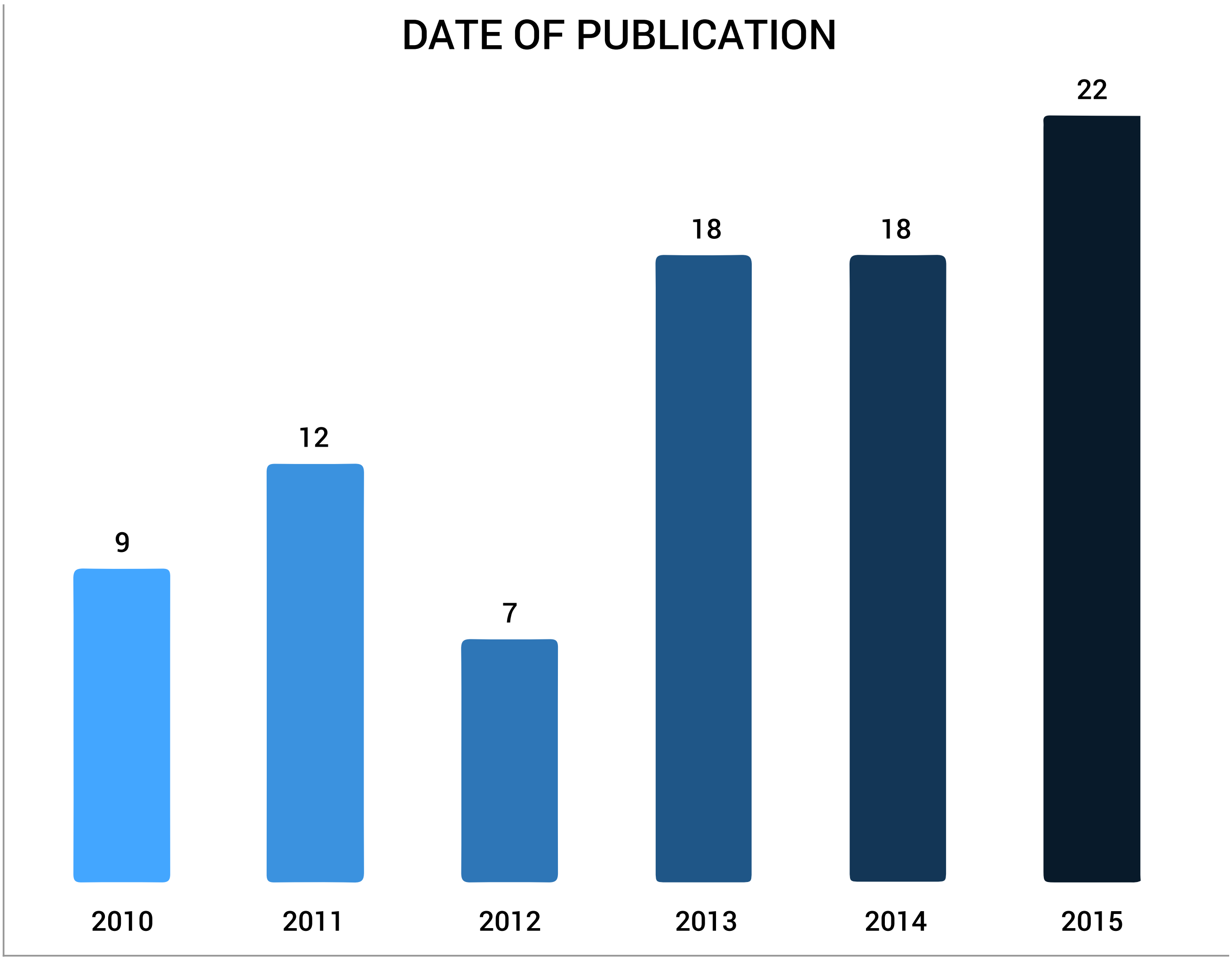
Studies were identified using three main methods:

- ***A search of Google Scholar, Google Web and Academic Search Complete using the search terms: "payment by results", "pay for success", "payment based on performance" and "payment for performance".***
- ***A request for PbR studies and grey literature on the author's website which was publicised via the LinkedIn PbR group and nine other mailing lists across the broad social care sector.***
- ***A series of requests to leading academics and researchers in the PbR/outcomes literature field in both the UK and Australia.***

The studies reviewed

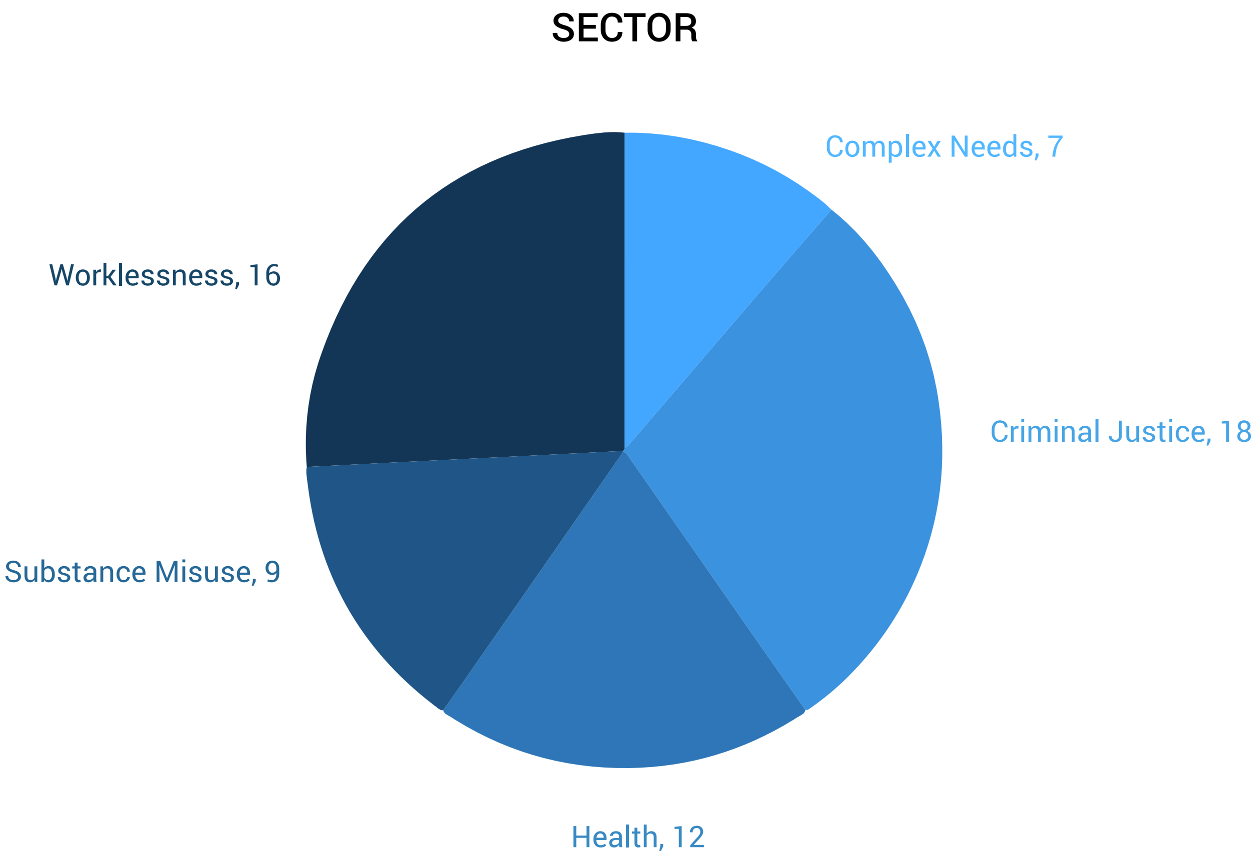
A total of 93 studies were analysed for this review. A large majority (75) focused primarily on payment by results in the UK, seven focused on PbR in Australia, six on the USA and five on PbR internationally.

Seven studies were published prior to 2010; 28 between 2010-2012 and the other 58 between 2013-15. The chart shows the increase in available literature over the last three years reflecting the growth in PbR-commissioned services.



Almost one quarter of these studies (22) dealt with payment by results generally with the rest focusing on PbR in a particular social care sector.

Three studies focused on housing/homelessness; two each on international development and children and one on defence. The remaining 62 studies focused on five main sectors as shown below:



Of the total 93 studies, six were systematic reviews of PbR studies and a further six were reviews of other kinds. Twenty eight were evaluations of PbR initiatives and one a feasibility study. The remaining 52 documents were either discussion papers (36) or policy statements (16).

Click here to learn more about the interactive PbR tool

4: A BACKGROUND TO PbR

A controversial model

Payment by results is an increasingly common commissioning model; found across a number of disparate sectors – from defence (Australian Department of Defence, 2010) to international development (Clist & Dercon, 2014) – and growing in popularity all over the world with initiatives particularly common in Australia, the United States and the UK.

The PbR "movement" can be traced back at least as far as the 1990s and even earlier; the first big wave was probably the trend for payment for performance in US healthcare (Petersen, 2006). PbR has also been integral to the design of government initiatives to get unemployed people back to work in Australia and the UK since the 1990s[4].

However, it has become much more widespread, in the UK in particular, since 2010 when it was championed by the government in general[5] and the Treasury[6] in particular. This rapid growth of PbR schemes has taken place at the same time as substantial cuts in public services. The consequence of this is that for many commentators, PbR has become synonymous with government budget cuts (Community Links, 2015) or the privatisation of public services (Policy Exchange 2013). Some, such as Whitfield (2015), go as far as to say that payment by results and the related trend in social investment is driven primarily by financial institutions whose motivation is to privatise the welfare state and open up new markets for global corporations.

Others such as Gash et al. (2013) note that the growth in PbR has taken place at the same time as a shift to the prime provider model which has also been the object of much criticism.

A lack of evidence

Many commentators have criticised the growth in the PbR commissioning model despite the lack of an evidence base to guide whether and in what circumstances it can be effective.

Lagarde (2013) summarises this view and helpfully highlights the difficulties of evaluating PbR as an approach when the purpose and mechanics of different initiatives vary so markedly:

> ***"The jury is still out on whether the benefits of P4P [payment for performance] schemes in public services exceed their costs. Indeed, it may not be wise to attempt to reach an overall verdict about P4P since the performance of each scheme is highly likely to be the product of the interaction between the detail of its design, the specifics of the setting, and the particulars of the provider and the service being delivered."***

4 In the UK, the Government created Training and Enterprise Councils (TECs) in 1990-91 which soon development "an increasing emphasis on output-related funding – linking payment to measurable outcomes such as job placements" Lourie (1997). In Australia, the contracting out of Australia's public employment service on a PbR basis was announced in 1996 and commenced in 1998.

5 David Cameron (2011) Speech on Big Society delivered on 14 February. London: Cabinet Office

6 CabinetOfficeandHMTreasury(2010)Callforevidenceonpublicservicereform.London:HMTreasury

In fact, not only do schemes vary in terms of their design and the context in which they are implemented – making comparison(s) between schemes difficult; but many individual PbR schemes have been implemented in a way that does not aid their evaluation.

The National Audit Office (2015), in a detailed exploration of the government's use of payment by results, concluded that PbR currently lacks a credible evidence base and baldly stated that without this evidence base:

> ***"commissioners may be using this mechanism in circumstances to which it is ill-suited, to the detriment of value for money."***

There are four additional reasons why, despite the plethora of UK PbR schemes, rigorous evaluation has proved so problematic.

Firstly, some substantial PbR pilots were cancelled or modified before being completed – two examples are the original probation PbR pilots[7] and the Peterborough Prison "through- the-gate" reoffending scheme[8].

Secondly, original evaluation plans which intended to compare PbR pilot areas with other areas became undone because of the very substantial changes in the funding and organisation of public services. This issue affected the Department of Communities and Local Government sponsored evaluation of the Supporting People PbR pilot (DCLG, 2014). The Justice Reinvestment pilot final report (Wong et al. 2015) was frustratingly unable to provide information on the outcomes of the second year of the two year PbR pilot.

Thirdly, as the NAO (2015) report points out; many of these new initiatives are not yet mature enough for a full evaluation:

> ***"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."***

Finally, many schemes in health care (such as the Quality and Outcomes Framework) were implemented nationally, making it difficult to attribute changes in performance over time to the implementation of PbR.

Does PbR work?

Of course, this lack of evidence means that it is not possible to give a meaningful answer to the question of whether PbR works. However, it may still be helpful to summarise briefly some of the main findings to

7 These were curtailed by Justice Secretary Chris Grayling in 2013 when he decided to move immediately to a national roll-out of PbR-based contracts for the new Community Rehabilitation Companies (CRCs) – effectively the wing of the new probation service working with low and medium risk offenders.

8 Again, this pilot was ended early when the Transforming Rehabilitation initiative effectively transferred the responsibility of reducing reoffending to the new CRC.

Click here to learn more about the interactive PbR tool

date. Contradictory as they are, they still provide a helpful context to the later exploration of critical success factors and common failings.

Overall

McNeil and colleagues (2015) in an Institute of Public Policy Research paper summarise:

> ***"So far, there is a lack of clear evidence that PbR can perform better than other payment systems, either in terms of money saved or improved outcomes for end users."***

Sector by Sector

The table on the following page summarises some of the key findings of evaluations of PbR schemes on a sectoral basis; somewhat unoriginally, schemes with mainly positive outcomes are highlighted in green, those with mainly negative outcomes are highlighted in red and those with mixed outcomes highlighted in amber:

Sector	PbR Scheme	Reference	Comments
Children's Centres			
	PbR trial	Frontier Economics 2014	PbR feasible but would require considerable modification
Criminal Justice			
	Peterborough Prison	MoJ (2015), M0J (2014)	Performance improving but below target
	Doncaster Prison	Pearce	Initial moderate outcomes, subsequently poor
Health			
	242 US medical groups	Mullen	P4P found improved performance for cervical cancer screening but not other outcomes
	Review of 17 US studies	Petersen	12 of 15 studies found positive impact on quality, 4 found unintended negative consequences (gaming)
	260 US hospitals	Werner	More than half hospitals met high performance targets initially, although those in control group did catch up after 4-5 years.
Homelessness			
	London Rough Sleepers	DCLG 2015	Performing well on main target of keeping service users in stable accommodation
	Supporting People	DCLG 2014	Costs down & outcomes up, but not possible to attribute to PbR
Substance Misuse			
	Delaware	McLellan	Increased uptake of treatment services and engagement in that treatment by stimulating innovation.
	UK Drug & Alcohol recovery pilots	Mason	Pilot areas outcome rates worse than non PbR areas because service users were remaining in treatment longer, which may turn out to be positive.
Troubled Families			
	UK Scheme	NAO 2013	Inconclusive, final evaluation of first phase currently awaited.
Worklessness			
	Work Programme	NAO 2014	Similar performance to previous welfare to work schemes but 2% cheaper. Did not succeed with harder-to-help groups of claimants.
	Australian & US Workfare	Finn 2010	Positive outcomes and reduced costs frequently claimed but may be offset by reduced quality and at expense of most disadvantaged.

PbR not the only approach

The lack of a robust evidence base leads many of those reviewing payment by results to remind their readers that, despite PbR's growing popularity, other commissioning approaches will still often be more appropriate. Both the Royal Society for Arts (Sturgess, 2011) and the National Audit Office (2015) state this specifically; Community Links (2015) recommended a ***blended*** funding model with PbR only ever part of the payment mechanism and Revolving Doors Agency (2015) despite highlighting the potential benefits of PbR argue that it is "***often not the most appropriate commissioning model***" for services for people with complex needs.

In similar vein, Fox & Albertson (2012) are not convinced that PbR is an appropriate commissioning approach in the reducing reoffending sphere and recommend the further exploration of justice reinvestment and personalisation as alternative ways of incentivising improved outcomes and stimulating innovation.

5: THE INTENTIONS OF PbR

Payment by results can be seen as part (or possibly the end point) of a movement away from more passive methods of funding public services such as block contracts. Policy makers and commissioners have increasingly seen payment systems as an important lever for influencing provider behaviour in line with a range of different intentions.

One way of measuring if a given PbR scheme works is to assess whether it met the primary intention for which it was commissioned. The next section of this review considers the different intentions, purposes or rationales of PbR schemes.

The (now defunct) Audit Commission (2012) in one of the first authoritative reviews of payment by results placed great emphasis on understanding the purpose of schemes. The first chapter of its report is entitled "A clear purpose" and starts:

> *"A clear purpose is important as it will shape design and implementation. Although there may be subsidiary objectives, schemes usually have one or more of these main aims:*
>
> *• improving outcomes or service quality;*
> *• reducing costs or improving value for money;*
> *• or stimulating innovation or transformational change."*

The Audit Commission also notes that it may be too ambitious to try to achieve all these aims in one PbR scheme because:

> *"The challenges associated with designing the scheme to achieve the right balance between risk, cost, change and outcomes vary with different aims."*

Similarly, Hunter & Breidenbach-Roe (2013) in their review of a number of PbR contracts found that the intention and purposes of PbR were not always clear:

Click here to learn more about the interactive PbR tool

> *"Often there are a number of conflicting drivers underpinning the move to commission on a PbR basis and this has resulted in poor implementation and the inappropriate application of PbR. ... [We found a] Number of cases where PbR has been used purely as an alternative way of paying for the same service, rather than as a method for improving outcomes by developing new forms of service provision."*

This theme is developed by a number of other commentators including Fox and Albertson (2012) in their examination of the applicability of PbR to the criminal justice sector and Battye (2015) in his paper aimed at refocusing the evaluation of PbR.

The table below summarises the different rationales for payment by results schemes:

Rationale	Commentator
Improve outcomes	Audit Commission, Battye, Nuffield Trust
Improve service quality	Audit Commission, Battye, Nuffield Trust, Sitra
Reduce costs/improve value for money	Audit Commission, Battye, Fox & Albertson, Sitra
Defer payment	Fox & Albertson
Stimulate innovation	Audit Commission, Battye, Fox & Albertson, Sitra
Transfer of risk	Battye, NAO, Nuffield Trust
Encouragement of new market entrants	Battye, Fox & Albertson, Sitra
Reduced commissioning costs	Battye
More/better focus on outcomes	Battye
Reduce inequalities	Battye

As stated previously, there is as yet no clear evidence base to conclude whether PbR routinely improves outcomes. However, the literature does provide some insight into the other seven rationales commonly ascribed to PbR schemes and the findings are summarised in the next section.

Improve service quality

Interestingly, there is limited evidence on whether PbR schemes improve service quality with relatively few evaluations placing much emphasis on the views of the end-users of PbR- commissioned services. However, DCLG (2015), in its second interim evaluation of the scheme to tackle rough sleeping in London, found that the PbR approach had stimulated providers to develop an effective delivery model. This "navigator" model prioritised individualised support provided by keyworkers who were incentivised by PbR to ***"go the extra mile"***. Both providers in this pilot scheme (funded by Social Impact Bonds) had met their targets on helping street homeless people to access and remain in stable accommodation.

On the other hand, there were numerous criticisms (e.g. Rees et al, 2013) of the way in which the contracts for the Work Programme encouraged providers to "park" claimants who were harder to help into work and provide them with a low quality, low intensity service.

Click here to learn more about the interactive PbR tool

Pearce (2015) found a similar situation in the Doncaster Prison resettlement PbR pilot where the specified binary outcome (paying the providers on the simple fact of whether prisoners re-offended rather than the frequency or seriousness of that reoffending) resulted in support being withdrawn from released prisoners as soon as they committed a new offence.

Reduce costs/improve value for money

As discussed above, one of the core reasons for the increasing use of PbR has been its use as a mechanism to drive down the cost of public services. The official evaluation of the Work Programme (NAO, 2014) found that 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). Even when overall costs are not reduced, making payment contingent on outcomes typically has the effect of deferring payment until at least the subsequent year's budget.

However, many commentators warn against using PbR solely or predominantly for this purpose. The NAO (2015) cautions in its conclusion that:

> *"commissioners may be using PbR in circumstances to which it is ill-suited, with a consequent negative impact on value for money."*

Johnson (2015) summarises the learning from Australian workfare schemes and recommends strongly that such schemes should not be competed on price for three main reasons:

- ***It encourages "gaming" with would-be contractors over-promising and subsequently under-delivering.***
- ***It drives down quality, and, consequently, performance.***
- ***It strongly encourages providers to respond to under-priced contracts by "creaming" (focusing on the easiest to help end users who may have found work without government intervention) at the expense of parking (providing a very minimal service to) those who are harder, and therefore more expensive, to help.***

A common theme in the literature (e.g. Sturgess, 2011; Haldenby et al, 2014) is ***"the winner's curse"***, a phrase used to describe the phenomenon where a bidder (typically for a very large government contract) prices their offer at such a low level to win the contract that they subsequently find it impossible to deliver an effective service within budget. Sturgess (2011) suggests that this is a particular risk when a service is commissioned on a PbR basis for the first time, recommending that in these circumstances:

> *"fixed-price competition based entirely on quality may be more appropriate".*

Afscme et al. (2015) note that in the USA, where Social Impact Bonds or other types of private or philanthropic investment vehicles are increasingly used to fund "Pay for Success" schemes, the complexity of schemes often entails transaction costs beyond the loan and interest of funding, including the cost of legal services, evaluation, programme administration and loan management. They note that these costs may incur considerable extra costs compared to a direct commissioning approach and that many State governments are effectively paying a considerable premium in order to access private finance up-front.

Stimulate innovation

One of the main potential advantages of payment by results according to many of its proponents is that it stimulates innovative approaches to entrenched social problems. Perhaps the classic model of PbR is one where the commissioners stipulate the outcomes they desire and then wait to see whether providers deliver them, safe in the knowledge that if providers fail to meet the contract requirements, they will not be paid. This model typically espouses a "black box" approach where providers are free to design the service in any way they see fit. This freedom is the factor that facilitates the ability to innovate combined with the fact that providers will not need to invest time and resources in regular performance monitoring and meetings with commissioners, being judged instead solely on the outcomes they achieve.

To date, the evidence on whether PbR stimulates innovation is very mixed. There are a number of positive examples:

- ***Frontier Economics (2009) found that the new St Giles Trust through-the-gate prison resettlement model brought a tenfold return on investment.***
- ***McLellan et al. (2008) found that the PbR approach to substance misuse provision in Delaware, USA fostered the adoption of new, evidence-based clinical interventions and an expansion of opening hours.***
- ***The DCLG (2015) evaluation of the London rough sleepers pilots praised the innovative navigator model (discussed above).***

However, these examples are in the minority. The DWP-commissioned study of the outcome-based contracting of the Work Programme (Hudson, 2010) did find innovation in provider-led pathways; but these were much more about reducing operational costs and achieving performance efficiencies rather than developing new, more effective ways of helping long-term unemployed people into work.

Bond International (2014) in their review of the meaning of payment by results for the UK voluntary sector found that PbR had stifled innovation in the international development arena because the benefits of flexibility were more than offset by the threat of financial loss – in other words, providers were too anxious about not being paid to move away from tried and trusted models.

KPMG (2014) found this to be particularly true for Australian schemes funded via Social Impact Bonds where they found a contradiction between the goals of service innovation and building an investment bond on a sound evidence base.

Similarly, NCVO (Sheil & Breidenbach-Roe, 2014) found that although champions of PbR claim that the approach is an opportunity for the voluntary sector, cherished for its culture of innovation, in practice government PbR contracts have often favoured the largest providers with a ***"blunt instrument"*** approach.

Boyle (2011) argues that PbR encourages such close attention to measurement systems that the focus of commissioners and providers is often much more on numbers than on new ways of working.

Sturgess (2011) in his review for the RSA argues that payment by results tends to encourage three categories of innovation:

- ***Identification of those beneficiaries for whom particular service models will work best;***
- ***Creation of effective management processes (for example, through joining up fragmented supply chains) enabling services to be tailored to different classes of beneficiary; and***
- ***Encouragement of much greater co-production on the part of beneficiaries."***

The NAO (2015) summarises this situation and acknowledges that current PbR models do not seem automatically to stimulate innovation and recommends that where commissioners are specifically seeking new ways of working, they should consider incentivising them.

Transfer of Risk

The Audit Commission (2012) describes how all PbR schemes allow commissioners to transfer practical and financial risk to providers; by linking payment to defined results, commissioners ensure they do not pay for poor performance. The provider (or, in the case of Social Impact Bonds, a third party investor) is often expected to cover early start-up and running costs until payments for success reimburse this outlay. This transfers the financial risk of setting up a new scheme from the commissioner to the provider. The use of social investment (both the Peterborough prisoner resettlement and the London rough sleepers PbR pilots were funded by Social Impact Bonds) allows commissioners to invest in new approaches without needing to find new money from budgets already under pressure.

However, the Audit Commission emphasises that it is not possible to transfer all risk, be that risk reputational, practical or financial. Commissioners retain their responsibility for local citizens receiving a good quality and effective service. They must also be diligent in ensuring that the terms of a PbR contracts are not too generous so that the public continues to receive good value for money.

Certainly politicians have made much of the advantages of payment by results transferring risk to providers, enabling them to argue that new government PbR initiatives will ensure proper value for public money. Chris Grayling, the government minister responsible for both the Work Programme (as Minister of State for Employment) and the Transforming Rehabilitation new probation contracts (as Justice Secretary) argued consistently that one of the core advantages of PbR is that new providers would not be paid in the case of poor performance.

Perhaps the best example of how risk can be transferred is the Work Programme which compares the performance of 18 prime contractors operating 40 contracts in 18 contract package areas (CPAs) – there are two to three contracts in every CPA. This ***"yardstick competition"*** (Sturgess, 2011) allows providers to be more fairly judged and commissioners to reward successful providers by allocating them a greater share of the contract. In the case of the Work Programme, the DWP regularly reviews performance and shifts market share as a result; for instance, in August 2013 substantial components of the contracts in 10 of the 18 CPAs were transferred from poorly performing providers to companies achieving better outcomes.

Clist & Dercon (2014) developed 12 principles for payment by results in international development in which they argue that risk transfer is not a rationale for PbR per se; stating that the primary purpose of transferring risk should be to sharpen performance incentives for the provider, and not for the commissioner to offload risk.

The NAO (2015) cautions that although PbR transfers some risk to providers, commissioners need to be aware of the risks they retain, in particular overall responsibility for a public service, and recommends that, wherever possible, commissioners pilot new PbR schemes.

Encouraging new market entrants

A core objective of some payment by results scheme is the desire to create a new market, or open up an existing market to new providers.

Transforming Rehabilitation (TR) was a Ministry of Justice (MoJ) initiative which restructured the existing public sector probation service into two segments between 2013-2015. While the National Probation Service was created as a public body to work with high-risk offenders, 21 local Community Rehabilitation Companies were formed to work with low and medium risk offenders with PbR-based contracts for these CRCs put out to competition. Although voluntary sector providers were encouraged to compete, 20 of the 21 contracts were awarded to private sector prime providers, with voluntary sector partners apparently in quite subsidiary roles[9].

Conversely, as Sitra (Knight-Markiegi & Quinn, 2013) point out, the Children's Centres PbR pilot scheme did aim specifically to open up this market to more voluntary sector providers.

However, a number of commentators have noted that, in practice, a number of PbR contracts have not succeeded in attracting the number of new providers anticipated. This is frequently related to the amount of risk that providers are required to entertain in order to provide services under a PbR contract. The NAO (2015) reports that the MoJ tendered contracts for pilot offender rehabilitation programme at Leeds prison but closed the competition without a successful bid after five of the six potential providers decided not to compete on the basis that the pilot model was unworkable.

Both Roberts (2011) and Bond International (2014) note that the pressure on cash flow which frequently attends PbR contracts (because providers typically do not get paid until they have achieved the stipulated outcome targets) makes it difficult for smaller private and third sector organisations to compete effectively.

Bond International and the National Association of Youth Justice (2011) also raise concerns that this sort of competition may not only reduce the diversity of potential providers, but can potentially also inhibit open relationships and learning between erstwhile partners who have become competitors.

The NAO (2015) recommends that commissioners take the time to understand the provider markets, the landscape of provision and providers' appetite for the risk inherent in payment by results before deciding to use a PbR approach.

Reduced commissioning costs

As stated above, what is sometimes seen as the pure form of payment by results involves a very reduced

9 The 21st was the probation staff "mutual" led partnership which was successful in the Durham and Teesside Contract Package Area.

role for commissioners who stipulate outcomes and allow providers to seek to meet them in any way they see fit. Since payment is dependent on achieving these outcomes, in theory commissioners do not need to invest the same level of time and resources in performance and contract management.

However, this approach has been very rare in practice. Very many commentators (e.g. Battye & Sunderland, 2011; DCLG, 2015 and NDEC, 2014) caution about the amount of initial work required of commissioners in designing outcomes and the mechanisms for measuring, verifying and auditing them.

Roberts (2011) notes that there is a tendency for bureaucratic systems to develop in order to manage PbR contracts, a finding echoed by Hunter & Breidenbach-Roe (2013) in their review of a number of PbR contracts who reported over-intrusive monitoring by commissioners and recommended that duplicating existing contract stipulations often runs counter to the purpose of a given PbR scheme. In similar vein, the NAO (2015) comments that:

> ***"PbR is particularly resource-intensive because of the technical challenges of setting the right payment mechanism."***

The scale of the problem is helpfully summarised by Lagarde (2013):

> ***"The jury is still out on whether the benefits of P4P schemes in public services exceed their costs."***

More/better focus on outcomes

There is unanimity within the evidence base that PbR contracts result in much greater focus on outcomes, for the simple reason that commissioners need to define outcomes which ensure that provision meets their objectives and providers must achieve outcome targets in order to get paid. One example is the evaluation of PbR in Children's Centres (Frontier Economics, 2014) where the authors recorded:

> ***"a shift in thinking towards focusing on monitoring outcomes rather than outputs."***

Sturgess (2011) notes that commissioners' outcomes payment systems will have a powerful impact on the service provided, potentially locking out some approaches to service provision. Similarly, Hudson (2010) in an analysis of the Work Programme for the DWP, found that the contract was shaping the nature of the support provided.

Many commentators highlight the importance of aligning outcome measures very closely with desired objectives. Lagarde et al (2013) found that workers will tend to focus on incentivised tasks almost like a form of "tunnel vision".

Rees and colleagues (2013) in the Third Sector Research Centre's analysis of the voluntary sector's experience of the Work Programme found that the behaviour of prime providers was completely determined by the specified payment outcomes. In this case, the lack of a sufficient financial incentive to work with individuals who were very unlikely to find work without the input of extensive and expensive resources resulted in what was seen by providers as the rational response of "parking" these service users; in effect denying them a service.

Reducing inequalities

The reason that Battye (2015) includes "reducing inequalities" as one of the possible outcomes of payment by results schemes is that PbR contracts can be designed to incentivise providers to focus on particular groups – an approach which can also address the problems of "creaming and parking".

Alshamsan and colleagues (2010) undertook a systematic review of the impact of pay for performance on inequalities in UK health care. Their analysis of 22 studies found some weak evidence that the use of financial incentives reduced inequalities in chronic disease management between socio-economic groups but the main conclusion was that inequalities between age, sex and ethnic groups persisted after the use of incentives.

Karve et al. (2007) in a review of the performance of 3,449 US hospitals found that pay-for- performance, although designed to improve care, may have disproportionally penalised hospitals caring for large minority populations – thereby compounding existing disadvantage. This finding seems to be related to the form of P4P used by Medicare which provided financial rewards to hospitals whose care performance ranked in the highest quintile (relative to their peers), but reduced funding to hospitals that ranked in the lowest quintile.

A number of commentators caution that the gaming behaviour which is often an unintended consequence of introducing PbR (e.g. Gravelle et al., 2010) can result in the most vulnerable/hard-to-help service users being neglected as they are less profitable for providers. The DCLG (2015) second interim evaluation of the London rough sleepers PbR scheme highlighted that the challenge for the rest of the pilot is for providers to maximise their return from achieving their targets while continuing to support vulnerable rough sleepers who had not yet moved into stable accommodation.

Having explored what the literature says about the different rationales for PbR schemes, the next chapter looks at what can be learnt from the evidence base about critical success – and failure – factors.

6: CRITICAL SUCCESS & FAILURE FACTORS

This chapter explores the factors identified in the literature as being associated with the successful – and unsuccessful – operation of payment by results schemes. These factors are often related to the different purposes or rationales of PbR initiatives discussed in the previous chapter.

In many cases, of course, factors relating to success and failure turn out to be the opposite sides of the same coin. For this reason, the chapter is not structured around the impact of different factors but, rather, on a chronological basis: starting from the consideration of whether a market may be appropriate for a particular kind of public social care provision, through the process of defining outcomes and setting incentives, concluding with monitoring and evaluation. The chapter is structured into these four main stages:

Understanding the market

Defining outcomes

Setting prices and incentives

Monitoring and evaluation

In terms of planning and delivering a PbR scheme, these stages are, of course, inter-related with the definition of outcomes and setting of incentives particularly closely linked.

Before starting this discussion, it is perhaps useful to repeat that many commentators suggest that commissioners carefully consider the relative merits of different commissioning approaches for a specific service and do not decide on a PbR approach until they have looked at these four stages in detail. Indeed, the NAO (2015) identifies ten features suited to a PbR commissioning approach, although most commentators are not so prescriptive, or limiting. The ten NAO criteria for a successful PbR scheme are:

Clear overall objectives, capable of being translated into a defined set of measurable outcomes
Clearly identifiable cohort/target population
Ability to clearly attribute outcomes to provider interventions
Data available to set baseline
An appropriate counterfactual can be constructed
Services are non-essential and underperformance of failure can be tolerated
Providers exist or are prepared to take the contract at the price and risk
Providers are likely to respond to financial incentives
Sufficient evidence exists about what works to enable providers to estimate costs of delivering services
A relatively short gap between provider intervention and evidence of outcome

Understanding the market

It has already been discussed that one of the prime drivers of PbR schemes is the desire to change a market, typically to open a public sector market to private (in particular) and voluntary sector providers. This section examines lessons from the literature for both commissioners and providers when a PbR approach is being used for a service for the first time.

When PbR commissioning is led at a national level (e.g. Work Programme, Transforming Rehabilitation), there has been a recent tendency to make contract package areas large in order to have sufficiently large cohorts of service users to validate the outcome measurements which generate payments. Indeed, Mulheirn (2013) argued that the MoJ's proposed cohorts for TR (which divided England and Wales into just 21 Contract Package Areas, amalgamating many of the existing 35 probation trusts) were still ***"relatively small."***

Clearly large contracts are likely to be more attractive to large providers and commissioners need to make decisions about contract size in this knowledge. The NAO (2015) urges commissioners to be clear about their expectations for prime providers' supply chains to ensure that smaller organisations (some of whom may have local expertise which cannot easily be replicated, at least in the short term) can be involved in service delivery. The NAO also suggests that commissioners may need to establish a safety net for the most vulnerable providers.

Rees (2013a) was one of several commentators who criticised the way in which many voluntary sector providers were apparently used as "bid candy" in the Work Programme tendering process only to have very little involvement in eventual service delivery. The same commentators also condemned the way in which Prime Contractors sub-contracted to much smaller organisations under the same PbR terms as the main contract. This was felt to be unfair since small scale organisations found it difficult to cope with the cashflow difficulties of only receiving payment many months after having delivered the contracted service.

Hudson and colleagues (2013), in the official DWP evaluation of the influence of outcome- based contracting in the Work Programme noted that there was considerable imbalance in supply chains; one of the results of which was that end providers had very little contact with local commissioners.

These criticisms about Work Programme supply chains were heeded by the MoJ in the TR competition whose terms (MoJ 2013) required Prime bidders to abide by the Merlin Standards (emqc 2013) and encouraged them to hold all or most of the financial risk and not pass this down to small sub-contractors. The NAO (2015) encourages commissioners to understand the market before deciding on the form of the contract; in particular commissioners are urged to consider whether potential providers have sufficient financial resources to bid for a contract which requires considerable initial investment and where payments are delayed until the achievement of outcome measures has been verified. Commissioners are also encouraged to assess what level of risk potential providers may be willing to consider taking on.

One of the most obvious and productive ways for commissioners to gauge the market is to discuss possible contract forms with potential suppliers. Lagarde (2013) and ARTD (2014) both recommend that commissioners negotiate the details of schemes with providers with the latter arguing that this approach promotes a culture of shared accountability which is a major contributor to effective provision.

Community Links (2015) and Mental Health Australia (2015) both recommend that one way in which commissioners can mitigate risks in new PbR markets is by operating a blended model where not all contract payments are dependent on achieving outcomes.

Defining outcomes

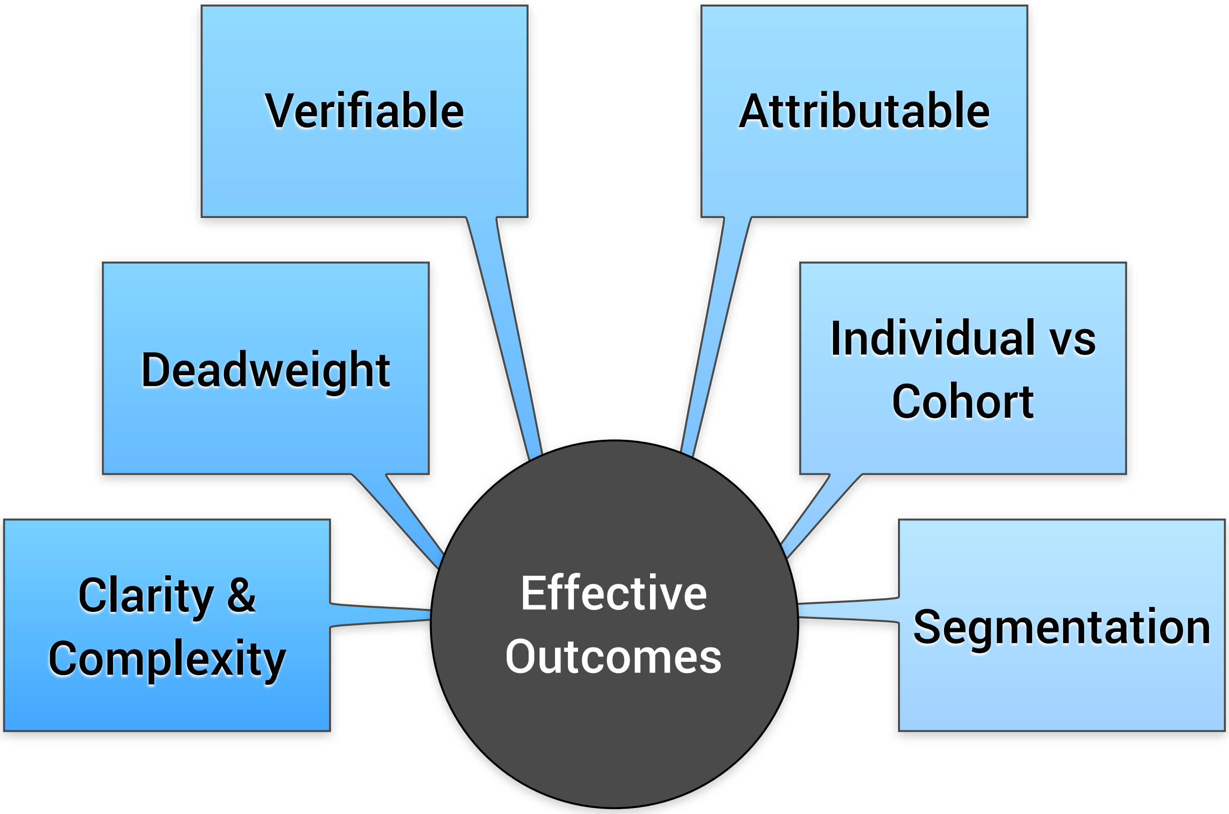
There is a very strong consensus (e.g. Clist & Dercon, 2014; Werner, 2011) within the literature that the outcomes set in a PbR contract (and the related incentives, see next section) have a very strong influence on the way in which the service is designed and delivered.

The NAO (2015) summarises the research base on PbR outcomes as being inconclusive:

> ***"There is no clear evidence of what works best in structuring outcomes".***

Nonetheless, a number of clear themes emerge from the literature of factors which influence the appropriateness of the outcomes set. Before discussing these themes in turn below, it is important to note that many studies (e.g. DCLG, 2015; Social Finance, 2012 & GHK, 2011) caution about the amount of time and money which commissioners have to invest in defining and verifying outcome measures, and which providers must spend on recording them.

They advise that there is a considerable risk of the savings generated by PbR schemes being outweighed by the additional costs of the bureaucracy involved in outcome measurement.



Clarity and Complexity

Unsurprisingly, many commentators (e.g. Reform, 2011) prize clarity in outcome measures with Wong and colleagues (2015) stressing the importance of ensuring that measures are meaningful to providers; and, ideally, capable of being understood and monitored using existing data recording systems. McLellan et al. (2008) describe the impact of devising meaningful outcome measures in their account of a PbR approach to improving drug and alcohol treatment outcomes in Delaware, USA which:

> ***"had intuitive appeal to all parties associated with the process. Patients, program [sic] administrators, clinicians and policy makers immediately understand the essence of the intervention. Providers proudly talked about how the performance contracts could work for them clinically and financially. The strategy as implemented simply makes economic, managerial and clinical sense."***

McLellan also emphasised the importance of an uncomplicated reporting system.

There is much less agreement in the literature about how simple or complex outcome measures should be.

Several government departments have placed emphasis on keeping outcomes simple and understandable (Crowe et al., 2014); a view also espoused by this author (Webster, 2013). Boyle (2011) agrees; arguing that the more complex measurement becomes, the less helpful it is, going as far as to construct his own version of Boyle's law:

> *"When you use numbers as the basis for payment, they become irrelevant to the broader objectives of the service"*

Community Links (2015), in a paper based on its own experience of delivering employment services under several PbR contracts, contend that PbR works well with smaller contracts with straightforward outcomes, but caution that the approach risks forcing providers to focus on a narrow range of outcomes rather than working holistically.
Conversely, other commentators point out that many PbR schemes are designed to tackle entrenched social problems with end-users often requiring co-ordinated and extensive interventions from a range of providers, making it hard to define outcome measures which are both simple and accurate. Crowe and colleagues (2014) in a major Institute for Government study which included a focus group approach to consulting with commissioners and providers found that:

> *"[although] keeping outcomes simple and understandable [is] a stated goal of government, [this] could undermine the process of co-creating them with the community. Most focus group participants criticised the use of high-level or macro outcomes as too crude and inflexible to adequately represent outcomes for a very wide range of individuals – including those with different multiple needs and barriers to employment."*

Roberts (2011) points up the difficulty in getting outcome measurements right for schemes aimed at helping people recover from dependency on alcohol and/or drugs; arguing that there is a need for a range of outcomes which take into account ***"distance travelled."*** Social Finance (2012) in their own paper looking at the same client group suggest specific outcome measures, albeit within the context of the seemingly hard-to-measure ***"life progress"***.

Keeping to the substance misuse sector, McLellan (2008), in his evaluation of a successful payment by results scheme across drug and alcohol treatment provision in Delaware, USA, found that proxy indicators can sometimes be as effective as outcomes. By combining three measures – length of time in treatment, active participation in treatment (defined as number of sessions per month varying according to the stage of the programme) and programme completion – McLellan found that the performance contracting approach resulted in clinical innovation and improvement and better treatment outcomes (although long term recovery was not measured). In particular, the specification of an active participation in treatment measure meant that providers did not merely retain end users in service for a longer period of time, but effectively engaged them in treatment.

A good illustration of the debate around whether outcome measures should be simple or complex is found in the literature on the MoJ's approach to measuring reductions in reoffending for the new probation contracts under the Transforming Rehabilitation initiative.

Initially, the MoJ proposed that reoffending should be measured using a binary approach - paying the providers on the simple fact of whether individuals re-offended or not. There was considerable expert criticism of this approach (e.g. Probation Chiefs Association, 2013) stating that the process of desisting from crime often takes considerable time with offenders reducing the frequency and seriousness of their offending. The eventual payment mechanism included both binary and frequency (but not seriousness) measures in an attempt to devise a more reliable approach to measuring reoffending rates (details in MoJ 2013b).

A number of commentators (including Fox & Albertson, 2012) note that even though simple measures may be desirable, it is often necessary to go through quite a sophisticated and complex process analysing current and likely outcomes before arriving at a clear definition of an appropriate outcome measure. Measures which are found fit for purpose are often co-produced by commissioners, providers and, sometimes, service users; or, if not, co- produced often the product of lengthy discussions and negotiations (Lagarde et al., 2013).

The far-reaching consequences of outcome measures is highlighted by Pearce and colleagues' (2015) evaluation of another reducing reoffending scheme based at Doncaster Prison which used only a binary measure. Since the provider could not earn income from working with any released prisoner who had already re-offended, this resulted in support being immediately withdrawn from these service users, in effect denying them a service to help them desist from crime.

Deadweight

Deadweight is a straightforward technical term which refers to outcomes which would have happened anyway without any public funded intervention. By way of illustration, a proportion of offenders will go straight without the help of the probation service, and a number of long-term unemployed people will find work without any official information, advice and guidance.

Therefore, the NAO (2015) and others recommend that commissioners ensure that they factor deadweight into the outcome measures they set to avoid paying for achievements which would have occurred anyway. One way in which the NAO recommends avoiding paying for deadweight is to construct a counterfactual in the form of a control group so that the impact of a PbR scheme can be fully measured. However, counterfactuals for PbR schemes are relatively rare, although the Peterborough Prison reoffending (MoJ, 2015) and drug recovery pilots (Mason et al. 2015) both compared the performance of their respective pilots against national performance.

However, these sorts of control groups are very expensive to set-up and, obviously, impossible for national initiatives (since service users in all areas are receiving the new intervention). In these latter cases, performance can either be measured by yardstick comparison between different providers (as in the Work Programme) or against a historical baseline (new probation contracts).

Verification

When commissioners devise a contract where payment is mainly contingent on providers meeting outcome measures, they need to be confident in the data relating to whether these measures are achieved. Sometimes, this is not a problem. For instance, the new Transforming Rehabilitation contracts specify reoffending targets for all offenders. Community Rehabilitation Companies, the new providers, have no influence over the offenders they work with (who are "selected" by the courts) nor the reoffending rate which is measured in a long established manner by a reliable third-party (the Police National Computer).

However, for some services, verifying outcomes is much more challenging. In the drug and alcohol recovery pilots, the contract design required new service users to be classified into one of five categories with different payment tariffs according to the severity of their substance misuse problem. It would obviously be unreasonable to make the same payment for a binge drinker who required a small number of brief interventions as for a dependent heroin and crack cocaine user who might require a very intensive service for several months or even years.

In most of the pilots, commissioners did not trust providers to categorise new service users and introduced an independent third-party to make this assessment when the new service user made contact with the treatment service. A number of substance misuse experts (including DrugScope/RSA, 2013 and this author, Webster, 2012) criticised this approach for two main reasons:

- ***It introduced extra cost and bureaucracy into the treatment system.***
- ***It placed an additional barrier in the way of those seeking help for their dependency on drugs and/or alcohol – making it no longer possible for them to access treatment directly (itself, a key indicator of a successful outcome).***

Commentators recommend that the process and costs of verification are borne in mind when outcome measures are decided. Hunter & Breidenbach-Roe (2013) caution against using too many outcome measures, highlighting the cost of collecting and verifying targets. Similarly, Clist & Dercon (2014) in their 12 principles for PbR in international development stated that:

> ***"The most tangible cost of PbR relative to other forms of aid is the cost of verification."***

Attribution

Payment by results purists argue that it is unimportant whether outcomes can be attributed to the work commissioned by providers or not. They contend that as long as commissioners specify the outcomes they desire, how these are achieved are immaterial – if reoffending or unemployment drops, why worry about the reasons?

The converse of this view is, of course, that providers bear all the risk of PbR schemes and if targets are missed because of factors outside of their control, they will not be paid for the services they have delivered.

In the real world, views tend to be more pragmatic. The operation of the Work Programme has been very challenging because of substantial changes in the economy. Lane and colleagues (2013) in an official DWP evaluation found that six months after the Work Programme went live (October 2011), the estimated number of referrals leapt from 2.5 to 3.3. million unemployed people. This not only meant that it was much harder to get long term unemployed people into work at a time when job losses were rising all over the country, but it also caused serious capacity problems for providers who were only just developing their services.

A number of commentators, including Hedderman (2013), have argued that reconviction rates are not an accurate indicator of real-life reoffending since the official data can also be affected by police clear-up rates, or the time taken to prosecute serious cases in particular. She points out that national and local initiatives focusing on a particular type of crime can cause spikes in reconviction data.

There is, however, a consensus that the most problematic issue relating to attribution is where a number of agencies are working with the same service user – as is often the case in PbR schemes which are frequently targeted at people with complex needs (offenders, homeless people, drug and alcohol dependents etc.). The Revolving Doors Agency (2015) summarises the current situation:

> *"Different providers still too often receive different pots of money to work towards different outcomes with the same clients. This can make for very confusing and expensive provision and arguments over which intervention achieved a certain outcome."*

Frontier Economics (2014) make a similar point in their evaluation of the use of PbR in children's centres:

> *"Attribution of changes in measures to individual or groups of centres is inherently problematic because many services are delivered in conjunction with other agencies; other agencies deliver similar services or services with similar objectives; children and families often use more than one centre; and there may be considerable time lags between the use of centres and outcomes."*

As yet, there does not appear to be a solution to this issue apart from the Prime Provider model where the lead delivery agent is responsible for designing a supply to chain to meet the full range of end users' needs and for negotiating a share of the payments for success.

Individual vs cohort

Outcome measures can be set at an individual or cohort level. Both sorts of measures exist in the UK.

The Work Programme, for instance, makes payments on an individual basis. The DWP initially paid an attachment fee for every new unemployed person worked with and then paid for a series of results; initially for finding a job and then for sustaining employment for every six month period up to a maximum of two years. The attachment fee was phased out after the end of the third year of the contract, leaving all payments to be made on a results basis. A similar individual measurement approach was adopted in the drug and alcohol recovery PbR pilots and the Troubled Families scheme.

However, the Doncaster and Peterborough prison reoffending pilots and the new Transforming Rehabilitation probation contracts are all based on a cohort approach with providers expected to reduce the overall reoffending rate of all those offenders they work in comparison to the national rate (for the former prison pilots) or the local historical rate (for the probation contracts).

Fox and Albertson (2012) are among the commentators who emphasise that if commissioners adopt the cohort approach, they must ensure that groups are large enough for changes to be statistically significant. This can be problematic in practice since, even if the overall cohort is large, outcome measures are sometimes calculated quarterly on much smaller numbers to ensure that providers have sufficient cashflow (this is the model for the Transforming Rehabilitation contract, MoJ, 2013b).

Segmentation

Although many PbR contracts are large, covering thousands of individual end-users; there has been an increasing trend to segment the user population in the recognition that different sub-groups require a different approach and, therefore often different outcome measures. When commissioners adopt a cohort approach, Fox and Albertson (2012) again caution that a balance must be struck between having a sufficiently tailored approach while maintaining sufficiently large cohorts to ensure that outcomes measures are reliable and contracts viable.

Johnson (2015) in his summary of lessons from Australian PbR worklessness schemes highlights two key issues relating to segmentation:

> ***"Unemployed people are not just one single homogeneous mass. They have different needs, generally requiring more intensive assistance as they move further from work."***
>
> ***"It is probably easiest and most effective to have different contracts (and providers) for different groups of jobseekers. This enables greater focus and facilitates contract/performance management."***

Carter and Whitworth (2015) argue that the Work Programme approach to segmentation is not sufficiently granular, doesn't distinguish client groups with different needs and therefore encourages "creaming and parking."

Hunter & Breidenbach-Roe (2013) note the importance of having particular regard to the hardest-to-help groups in any contract; arguing that without clear incentives providers may fail to work with service users who require more complex, sustained or costly interventions.

Setting price and incentives

Of course, it is not just the outcome measure itself which has the dominating influence on the way in which a PbR-contracted scheme is delivered, but the price attached to that outcome. Werner (2011) in his evaluation of PbR in the US healthcare sector summarises:

> ***"Most experts agree that pay-for-performance incentives must be designed very carefully, to give hospitals sufficient motivation to improve quality without producing unintended consequences, such as reducing access to care for disadvantaged populations."***

Gaming

Gash and colleagues (2013) note that gaming is prevalent in many public markets and that it is often accentuated by fiercely competitive environments, an apposite description of the Work Programme and Transforming Rehabilitation procurement processes. Many commentators have discussed at length how gaming is intrinsic to payment by results contracts. The definition of gaming and the different forms it takes vary considerably.

The most blatant and obvious forms of gaming, perhaps more properly termed fraud, are described by Sturgess (2011) who includes the infamous example of a number of hospitals which, in order to improve their performance on Accident and Emergency (A&E) waiting times, bought more expensive trolleys so that they could be designated as beds and patients could be redefined as having been admitted as soon as they were placed on them.

Boyle (2011) repeats this example and describes the manipulations adopted by other hospitals to circumvent the same targets. Some hospitals kept patients in ambulances outside the hospital so that the four-hour target period did not start until patients were brought onto the hospital premises; others had special units where A&E patients could be taken to remove them from the target, sometimes for days at a time.

A common gaming technique in PbR contracts is known as ***"creaming"*** (AKA ***"cherry picking"***, ***"low-hanging fruit"*** or ***"cream-skimming"***) where providers focus their resources on the end-users who are easiest to help. The flip side of ***"creaming"*** is ***"parking"*** where the hardest to help end-users receive no, or a very low level of service.

There is a consensus amongst official evaluations and third party research on the Work Programme that gaming is wide-spread within its contracts. Indeed, Rees and colleagues (2013b) interviewed a number of providers with long experience of welfare to work provision who argued that gaming was embedded in the Work Programme and could be seen as a rational response to PbR since a proportion of customers would always be very unlikely to get a job.

The DWP did try to avoid this problem by segmenting the long term unemployed population and setting a higher incentive payment for providers who succeeded, for example, in helping people with disabilities into work. However, these incentives were not seen as sufficient and did not reflect the need to invest a much greater level of resources to achieve the outcome measures for this group (House of Commons Work and Pensions Committee, 2015).

Responses to gaming

Finn (2010) recommends that the risks of gaming in worklessness programmes can be reduced through contract design and oversight. He suggests that the inclusion of measures related to job retention, wages and benefits, and earnings gains, for example, all help diminish any incentive to place participants into poor quality jobs.

Another mechanism recommended to counteract "creaming and parking" is the target accelerator model devised by Mansour & Johnson (2006). The basic principle of this model is that it starts with a finite, or capped, number of people being required to achieve outcomes. They recommend that this should be a high, and therefore significant, percentage of a given target group. The premise is that the further a provider reaches into that group, the harder they become to help and the more it will cost to achieve the outcomes.

The example advanced by Mansour & Johnson suggests that commissioners might not pay any incentive for the first 10% of employment outcomes, on the assumption that these jobseekers would have found work without intervention, but would pay £600 per outcome for the next 15% of jobseekers, £1,000 for the following 15% and so on, increasing the incentive as providers ***"penetrated"*** further into the cohort to people with harder-to-meet needs.

Another method of encouraging targeting on hard-to-help individuals (a key objective of PbR schemes since achieving outcomes with this section of the target group results in the biggest savings to public expenditure) is the ***"prize"*** approach described by Sturgess (2011) as an approach used in Employment Zones, a precursor to the Work Programme:

> ***"Prizes can act as 'last mile incentives' for providers to continue to achieve outcomes beyond the level where the cost of helping each jobseeker begins to outweigh the amount of the outcome payments. In Employment Zones, providers could earn a bonus for placing a certain proportion of their jobseekers in employment. The amount of the bonus and the proportion of jobseekers providers were required to place in order to earn the prize were not disclosed. This ensured that providers had an incentive to find employment for as many jobseekers as possible, rather than simply finding work for those whom it was profitable to place based on the amount of the outcome payments."***

The dangers of competitive pricing

A number of commentators highlight the dangers of using PbR as a mechanism solely to cut costs rather than to stimulate innovation and improve efficiency and effectiveness. Mansour & Johnson (2006) describe what they see as the significant dangers in competitive or fixed pricing (the procurement model used for the Work Programme, Transforming Rehabilitation and many other UK government PbR initiatives):

> ***"In a competitive pricing exercise, providers are obviously incentivised to compete on the basis of price. After a point (which is very quickly reached), it is not the size of their profit margins which are being squeezed but, in order to achieve cost savings, it is the expenditure on premises, staff salaries and client resources. The quality of the service procured is seriously compromised. ...***
>
> ***Providers will offer lower and lower unit prices in order to secure a place in the market, but in the process will generate unsustainable pricing levels which, in turn, will produce an unstable and low quality market. They may actually offer a level of performance which they will not be able to deliver, knowing that subsequent contract management is likely to be 'soft' on them. They may base their offer on insufficient evidence or they may lack commercial awareness, discovering only after contract-award that their offer is ultimately untenable. There are considerable risks in all these for the procurer."***

The DCLG (2014) evaluation of the Supporting People PbR pilot also noted the importance of commissioners setting the right outcome measures and sticking to them (rather than being "soft"):

> *"There was too much flexibility in how PbR terms were applied, for example commissioners were very lenient in the application of targets and were prepared to meet providers' request for new indicators to measure progress, rather than stick to the agreed PbR outcomes."*

A similar approach was noted in the first year of the drug and alcohol recovery PbR pilots[10].

Incentives and reinvestment

McLellan et al. (2008) attributed the successful operation of a PbR initiative designed to improve outcomes in the Delaware, USA substance misuse treatment system to a model which introduced incentive payments on top of existing contracts.

A number of commentators (Wong, 2015 & Webster, 2013) highlight the importance of a reinvestment approach where providers who achieve targets and realise savings to public expenditure are given a proportion of those savings to reinvest in their service to further improve outcomes.

Reinvestment also tackles the criticism (Bond International, 2014) that PbR may undermine intrinsic motivation to act by replacing it with extrinsic/monetary incentives – particularly in the international aid arena.

However, reinvestment can imply that the same provider retains their contract to reap the fruits of their good performance which may not be possible depending on the legal parameters of specific contracts and the duties of competitiveness placed on many commissioners of public services.

Contract structure

Related to the pricing of outcomes is the overall structure of the contract both in terms of its length and the delay until, and interval between, incentive payments.

Contract length

Sturgess (2011) sets out the main parameters for consideration of contract length; arguing that contracts must be long enough for an effective intervention to be delivered and to ensure that sufficient providers are encouraged to take the risk associated with PbR and enter the competition. However, he also cautions against very long contracts which may weaken competitive pressure and, consequenntly have the tendency to thin the market since fewer tenders mean less frequent competition and fewer opportunities for new providers to enter the market. Of course, short-term contracts and more frequent competition bring their own disadvantages; causing disruption to service delivery and imposing significant costs on bidders which must ultimately be recouped. Sturgess recommends that where longer contracts are preferable, commissioners can impose minimum performance standards which allows them to terminate contracts in the case of unsatisfactory performance, a subject explored in more detail in the next section on monitoring

10 Intelligence gained from attendance at Drugscope & RSA (2013) providers summit, conducted under Chatham House rules.

and evaluation.

Delay between intervention, impact and payment

The issue of cashflow is discussed by very many researchers (e.g. Sheil & Breidenbach-Roe, 2014; Roberts & Singleton, 2011) who highlight the fact that many small providers will be deterred from bidding for contracts where outcome payments are delayed for many months or years.

Many schemes (e.g. the first three years of the Work Programme and the early years of the Transforming Rehabilitation contracts) seek to mitigate these effects by operating a partial PbR model with a fee paid for initial work undertaken.

Many schemes working with people with long term complex problems (such as offending, substance misuse, mental health, homelessness) may need to work with individual service users for periods of well over one year to achieve sustainable outcomes. However, delaying outcome payments for this length of time (inevitably extended further by the measurement and verification processes) is impractical. Even the largest organisations cannot operate without income for long periods of time without incurring unreasonably high levels of additional costs. Sturgess, again, (2011) argues that PbR does not work well in settings where there is a long delay between intervention and outcome for three main reasons:

- ***Extraneous variables become more problematic – it's more difficult to determine whether the intervention itself achieved the outcome.***
- ***Outcome payments cannot be withheld for very long periods if commissioners wish to sustain a diverse and commercially viable market.***
- ***Over extended periods, monitoring costs will outweigh the benefits of ensuring that outcomes have truly been achieved.***

There is no consensus on how to resolve this issue. Many (such as Revolving Doors Agency, 2015 & Roberts, 2011) argue for the use of milestones as a way of recording ***"distance travelled."*** While several others (including Boyle, 2011 & Webster, 2013) cite the most popular version of Goodhart's law, that ***"when a measure becomes a target, it ceases to become a good measure"***; in other words, it is likely that providers focus on meeting the milestones (e.g. keeping someone in treatment) rather than the outcome (helping them become and remain drug free).

When PbR is used as a way of improving outcomes in an existing system, utilising established monitoring tools, this problem can be overcome. McLellan and colleagues (2008) found that paying incentives monthly gave immediate feedback, rewarding progress and promoting buy-in, as well as avoiding cash flow problems.

Monitoring and evaluation

Theoretically, a pure payment by results approach should require minimal monitoring and evaluation. Provided that commissioners have set the right outcomes and are confident that the prescribed way of measuring them is accurate and robust, then the success of a scheme can be gauged simply by whether the provider achieves the outcome targets.

Real world situations mean that commissioners rarely feel able to just ***"wait and see"*** whether providers (and therefore the contract itself) are a success. Nevertheless, there is an expectation (e.g. Make Justice Work, 2012 & Battye & Sunderland, 2011) that commissioners have a lighter touch to performance monitoring since many schemes encourage providers to innovate and use any approach they like to achieve outcomes (the so-called ***"black box"***). A key factor which allows providers to adopt different approaches is the release of resources from onerous performance monitoring systems into service delivery. A number of commentators (Hunter & Breidenbach-Roe, 2013; Drugscope & RSA, 2013) have noted, however, that this light touch performance monitoring on the part of commissioners has rarely been evident, in large part because many of the PbR schemes have been high-profile and/or politically sensitive.

The risks of ***"parking"*** – whereby harder to help (and therefore typically more vulnerable) service users receive a minimal service – mean that it is often important for commissioners to require a minimum delivery standard. In the Work Programme (Lane et al., 2013), bidders were required to propose their own standards.

It is the role of commissioners to decide how to assure that these minimum delivery standards are met; although most commentators recommend a third party approach.

From a provider perspective, Community Links (2015) note that ***"close monitoring and regular re-profiling are essential elements in effective management of PbR contracts"*** and disclose that they strengthened management information systems for this purpose.

7: CONCLUSIONS

This chapter summarises the conclusions of this literature review and examines the implications for commissioners, providers and other stakeholders (including service users) involved in PbR-contracted services.

Effectiveness

There is a consensus that the evidence base is not able to give a clear indication as to whether payment by results works. There is also growing agreement that the different purposes and structures of schemes make it probably unrealistic to reach a binary assessment about the effectiveness of the approach.

The association of many PbR schemes with very robust cost-cutting and/or the privatisation of previously public markets has caused considerable controversy and confusion which has enabled researchers working from the same data to reach opposing conclusions about the same initiative. By way of illustration, those who saw the Work Programme as being primarily about helping long term unemployed people back to work with the least investment of public finances will assess it as very successful. Conversely, those who thought it was designed to get people with entrenched difficulties such as disability or addiction into work, will concluded that it has failed.

The main lesson which commissioners can draw from the research around the different rationales for PbR schemes is that payment by results can be applied to many different ends, but rarely for more than one purpose within the same initiative.

Critical success (and failure) factors

Nonetheless, the rapid proliferation of PbR schemes has helped to generate useful knowledge for those involved in their commissioning, delivery or evaluation. There are clear lessons at each of the four stages of: understanding the market; defining outcomes, setting price and incentives; and monitoring and evaluation.

Understanding the market

The key lessons from the literature are:

- ***When considering introducing a payment by results approach to a particular sector/service, commissioners should ensure that they align the level of risk they wish to transfer with the "risk appetite" of current and potential providers.***
- ***Commissioners should be aware that large contracts will attract primarily large providers and they will need to proactively prescribe any requirement that small/medium/local providers are to be involved in service delivery.***
- ***It is extremely risky to procure a PbR service for the first time on a competitive pricing basis. There are many benefits in co-designing new PbR contracts with potential providers and service users.***

Defining outcomes

There is no clear evidence of what works best in structuring outcome measurements, but commissioners will benefit from attending to eight key issues:

- ***Outcome measures should be both clear and meaningful to providers.***
- ***Where possible, outcome measures should be simple although straightforward measures are not always appropriate in schemes where service users have complex needs and outcomes may take several years to achieve.***
- ***Measures should always take into account deadweight to ensure that any PbR contract results in improved performance.***
- ***Wherever possible, outcome achievements should be verified by existing monitoring systems which are neither overly onerous nor expensive to operate.***
- ***Commissioners need to decide how important the attribution of outcomes is and to ensure that there is an equitable system for paying for results when several providers are involved.***
- ***As much as possible, outcomes should reflect the behaviour of providers and not external factors (otherwise providers can be inappropriately rewarded or penalised).***
- ***Both individual and cohort outcome measures can be valid; where cohorts are used, they must be sufficiently large for results to be meaningful.***
- ***For many services, segmentation of the target group will be vital to avoid the "parking" of less profitable service users.***

Prices and incentives

The price and incentives structure will determine the form and quality of any PbR- contracted service. Gaming is to be expected and can be managed by segmentation with higher incentives where required and differential pricing using such mechanisms as target accelerators.

The dangers of competitive pricing are well-established and commissioners should be aware that they retain overall responsibility and cannot transfer all risk to providers.

Reinvestment can be a powerful incentive associated with continuous improvement in quality outcomes.

Commissioners should pay careful attention to the length of contract and, in particular, the delay between the delivery of service and payments for results.

Monitoring and evaluation

The monitoring and evaluation of PbR schemes should, typically, be ***"light touch"*** with the emphasis on verifying whether providers have achieved specified outcome measures. However, some form of quality assurance, ideally involving the views of the end-users, may be important as insurance against vigorous gaming practices.

Implications

Perhaps the most important implication from the research base is that there is currently a very limited understanding of how the theoretical advantages of payment by results play out in its practical application. It seems unlikely that this situation will change any time soon; Van Herck (2010) undertook a systematic review of 128 studies into payment for performance in the US healthcare sector (which had been operating for over 15 years) and still found that it was too early to draw firm conclusions about gaming and the unintended consequences of PbR contracts. It is, however, abundantly clear that outcome measures and, in particular, their associated pricing and incentives structure have a very powerful impact on the way in which a particular service is delivered.

There is a consensus that unexpected, often perverse, consequences are commonplace as providers focus wholeheartedly on these measures in a manner which is either described as ***"rational and efficient"*** or ***"gaming the system"***. Irrespective of their perspective on this issue, many commentators recommend:

- ***The piloting of new initiatives (NAO 2015) and avoidance of hasty implementation of new PbR schemes (Wong, 2015).***
- ***The need to be flexible and modify contracts in the light of provider behaviour (Finn, 2010).***
- ***The value of mixed models with not all income linked to the achievement of outcome measures (Community Links, 2015).***
- ***The importance of transparency (Gash et al. 2013).***

Perhaps the most important final lesson from the literature is that developing a new PbR service is likely to be a process of trial and error (Mental Health Australia, 2015), that it is unreasonable to expect to get the contract right first time (Sturgess, 2011) and good practice to build in the potential for constant adaptation (Finn, 2010).

Glossary of Payment by Results Jargon

Bid Candy refers to the experience of many small voluntary and community sector providers in the Work Programme who were included in the tenders of many large private corporations ("the prime providers") to make these bids appear more attractive and socially responsible. A number of these smaller organisations found that they did not receive sufficient referrals and that, in some cases, the management charges levied by the "primes" were disproportionately high.

Binary PbR shorthand which refers to one of the ways in which reoffending is measured. A binary approach is very attractive, predominantly because of its simplicity – easy to measure, and easy to understand. If a released prisoner does not commit a crime, (or, rather, does not get caught and convicted), the project gets paid. If s/he is convicted, the project doesn't get paid. A binary measurement is also particularly attractive to government because anyone who does not reoffend incurs no costs whatsoever to the criminal justice system. However, criminal justice experts emphasise that most offenders rarely just stop committing crime, but tend to offend less seriously and less often over a period of time.

Black Box refers to the fact that because commissioners are buying outcomes rather than activities, providers are free to operate the service in any way they see fit since they will only be paid if they achieve the results set by commissioners.

Recently, there has been much debate about whether the contents of the "black box" should be shared publicly. Some providers think they should be able to keep their methods secret in order to preserve their Intellectual Property, others think that operational approaches should be shared for the benefit of society.

Cashable Savings is a term relating to whether the hoped-for improved performance of a PbR funded scheme results in tangible and immediate savings to the public purse. For instance, when the Work Programme finds someone a job, that person comes off benefits and start paying taxes – a double whammy for the Treasury. However, if a PbR reoffending scheme results in 100 criminals stopping committing crime, there are no immediate cashable savings since although prison is an expensive resource, there would need to be a large and sustained reduction in the crime rate before a prison could be closed and money saved.

Cohorts The defining characteristics of most PbR schemes is their large-scale. This is, typically, for two main reasons. Firstly, because governments are often looking for economies of scale to reduce public expenditure. The second reason is that PbR schemes are all about improving outcomes and you need a sufficiently large scheme – and a good size "cohort" of end users for any outcome monitoring to be regarded as statistically significant.

Counterfactual PbR commissioners do not want to pay for outcomes which would have happened anyway (see deadweight), therefore they might develop a counterfactual which models what would happen without a specific intervention, emphasising the importance of setting outcome targets in a real-world context.

The Ministry of justice would not want to pay out on a PbR scheme which reduces reoffending by 2%, if in the rest of the country reoffending dropped by 4%. One of the major early criticisms of the Work Programme was that providers were only placing an average of 3.5% long-term claimants into work despite the counterfactual suggested by the evidence base that around 5% of long-term unemployed typically find work without receiving any intervention.

Creaming means prioritising end-users who are most likely to achieve a provider's outcome targets and therefore bring in payments, leaving those individuals with more complex needs/problems without the service they need.

Deadweight PbR is about driving innovation and improved performance. The aim is to pay for achievements, not activity. Therefore, contracts are carefully drawn so that providers are not paid for "deadweight". Deadweight is simply a term that refers to outcomes which would have happened anyway. The one-year reconviction rate for short-term prisoners is currently 59%. The Ministry of justice will not want to pay providers for helping to prevent reoffending among the 41% who already go straight. Therefore, PbR targets will always be set to improve current performance.

Financial Cliff Edge describes a difficulty with the structure of some PbR contracts which set an absolute target. For instance, a provider may be expected to reduce reoffending rates by 5% and, if it does so, receive a high value payment for this result. However, if it reduces reoffending by only 4.9% it will receive no money at all. A Policy Exchange (2013) report on PbR described this as a "financial cliff edge" and argued, quite reasonably, that this approach is not commercially viable.

Frequency The "binary" approach to measuring reoffending in a PbR scheme is a simple yes/no indicator of whether an offender commits further crime. The alternative is to use a "frequency" approach which measures a reduction in the total number of further crimes committed. Frequency is regarded as a more realistic (although more complex) method of measurement because we have a large evidence base that tells us most criminals who stop offending do so over a period of time. PbR schemes which use a frequency measure incentivise providers to work with more prolific offenders.

Market Warming PbR involves a sea change in the relationship between commissioners and providers and many commentators regard it as commissioners' responsibility to steward the market and ensure that there is a good range of potential providers from the public, private and third sectors. The term "market warming" refers to this process, in particular commissioners communicating their strategic goals and aspirations of the commissioning process at an early stage.

Parking means providing a minimal or no service to those end-users who are unlikely to achieve a provider's targets and therefore bring in a payment by result. For example, a released prisoner who quickly reoffends may still have a range of unmet needs but, because s/he is unable to contribute towards a reducing reoffending PbR target is provided with only a low level service.

Perverse Incentives Every PbR conference features the phrase "the devil is in the detail". It is generally agreed that unless PbR contracts are very carefully written, it is entirely possible that they will encourage the wrong sort of activity by providers. For example, even though Work Programme providers receive a higher payment for getting jobs for hard-to-place individuals, their need to maintain a steady cash flow has perversely incentivised them to prioritise working with the easiest-to-place jobseekers (see "Creaming" above).

Risk Transfer Perhaps the key driver of PbR for governments is the fact that payment by results contracts transfer financial risk away from the public purse to the provider organisation. In a pure PbR model, the provider only gets paid if they achieve outcomes which are usually defined in a way that saves money to the government. If they don't achieve these outcomes, they don't get paid and the government saves money. Most current schemes are not 100%, PbR although the Work Programme became so in 2014/15.

Social Impact Bonds are a form of outcomes-based contract in which public sector commissioners commit to pay for significant improvement in social outcomes (such as a reduction in offending rates, or the number of unemployed people found jobs) for a defined population. Private (or philanthropic) investors finance the bonds which provide the working capital for providers. In this way, the public sector passes the financial risk over to the investors who only receive a return on their investment if the outcomes are achieved.

Bibliography

AFSCME, Center for Effective Government, EOI, ITP, In the Public Interest et al. (2015) A guide to evaluating pay for success programs and Social Impact Bonds. www.inthepublicinterest.org

Alshamsan, R. et al. (2010) Impact of pay for performance on inequalities in health care: systematic review. Journal Health Service Research Policy. 2010 Jul;15(3):178-84.

ARTD Consultants (2014) DHHS funded community sector outcomes purchasing framework. Community Sector Relations Unit Department of Health Human Services, Hobart, Tasmania.

Audit Commission (2012) Local payment by results. London: Audit Commission

Australian Department of Defence (2010) Next generation performance-based support contracts: achieving the outcomes that defence requires, discussion paper. Canberra: ADOD

Battye, F. (2015) Payment by Results in the UK: Progress to date and future directions for evaluation. Evaluation 2015, Vol. 21 (2) 189-203

Battye, F. & Sunderland, J. (2011) Thinking about...Payment by Results. GHK

Bond for International Development (2014) Payment by results: what it means for UK NGOs. London

Boyle, D. (2011) The pitfalls and perils of payment by results. Local Economy 2011 26:627. Sage.

Cabinet Office and HM Treasury (2010) Call for evidence on public service reform. London: HM Treasury

Cameron, D. (2011) Speech on Big Society delivered on 14 February. London: Cabinet Office

Carter, A. and Whitworth, E. (2015) Creaming and parking in quasi-marketised welfare to work schemes: Designed out of or designed into the UK Work Programme? Journal of Social Policy, 44 pp 277-296.

CESI, NIESR, IES & SPRU (2014) Work Programme evaluation: Operation of the commissioning model, finance and programme delivery, Research Report No 893. London: Department for Work & Pensions.

Clinks (2010) Payment by results: what does it mean for voluntary organisations working with offenders?

Clist, P. & Dercon, S. (2014) 12 principles for payment by results in international development. London: Department for International Development.

Community Links (2015) Payment by results, Issues that Matter No. 5. London: Community Links

Crossley, S. (2015) The troubled families programme: the perfect social policy? Centre for Crime and Justice Studies: Briefing 13 November, 2015.

Crowe D., Gash, T. & Kippin H. (2014) Beyond big contracts: commissioning public services for better outcomes. London: Collaborate at London South Bank University

Dar, A. (2015) Work Programme: background and statistics. House of Commons Library Briefing paper Number 6340, 26 June 2015.

D'Souza, C. & Gill, S. (2015) Lambeth assessment and referral service: Benchmarking exercise. London: Unpublished.

Department for Communities Local Government (2015) Qualitative evaluation of the London homelessness social impact bond: second interim report. London: DCLG

Department for Communities and Local Government (2014) Supporting people payment by results pilots: final evaluation. London: DCLG

Department of Communities and Local Government (2012) The Troubled Families programme: Financial framework for the Troubled Families programme's payment-by- results scheme for local authorities.

DrugScope & RSA (2013) Drug and Alcohol Recovery Payment by Results (PbR) pilots – National Service Providers Summit

emqc (2013) Merlin Standard. Derby: emqc – merlinstandard.co.uk

Finn, D. (2010) Outcome-based commissioning: Lessons from contracting out employment and skills programmes in Australia and the USA. London: UK Commission for Employment and Skills

Flodgren, G. et al. (2011) An overview of reviews evaluating the effectiveness of financial incentives and changing healthcare professional behaviours and patient outcomes. Cochrane database of systematic reviews 2011, Issue 7. Art. No.: CD 009255

Fox, C. & Albertson, K. (2012) Is payment by results the most efficient way to address the challenges faced by the criminal justice sector? Probation Journal 59(4) 355-373

Fox,C. & Albertson, K. (2011) Payment by results and social impact bonds in the criminal justice sector: New challenges to the concept of evidence-based policy? Criminology and Criminal Justice. http://crj.sagepub.com/content/early/2011/08/06/1748895811415580

Frontier Economics & the Colebrooke Centre (2014) Payment by Results in Children's Centres Evaluation. London: Department for Education

Frontier Economics (2009) St Giles Trust's Through the Prison Gate: an analysis of economic impact, London: Pro Bono Economics.

Gash T., Panchamia N., Sims S. & Hotson, L. (2013) Making public service markets work: professionalising government's approach to commissioning and market stewardship. London: Institute for Government

Gravelle, H., Sutton, M., & Ma, A. (2010) Doctor Behaviour under a Pay for Performance Contract: Treating, Cheating and Case Finding? The Economic Journal Vol. 120, Issue 542, pp. F129-F156, February 2010.

Grimwood, G., Edmonds, T.' McGuiness, F., Powell, T., Roberts,N & Wilson, W. (2013) Delivering public serices: the growing use of payment by resuls. House of Commons Library, Home Affairs Section SN/HA/6621

Haldenby, A., Harries, R. & Olliff-Cooper, J. (2014) Markets for good: the next generation of public service reform. London: Reform Research Trust.

Hedderman, C. (2013) Payment by Results: Hopes, fears and evidence. British Journal of Community Justice, Vol.11 (2-3)

Hoggett, J., Ahmad, Y., Frost, E., Kimberlee, R., McCartan, K. and Solle, J. and Bristol City Council (2014) The troubled families programme: A process, impact and social return on investment analysis. Project Report. University of the West of England, UWE Repository.

House of Commons Committee of Public Accounts (2014) Programmes to help families facing multiple challenges. Fifty-first report of session 2013-14

House of Commons Justice Committee (2014) Crime reduction policies: a co-ordinated approach ? Interim report on the Government's Transforming Rehabilitation programme. London: www.parliament.uk

House of Commons Work and Pensions Committee (2015) Welfare-to-work: Second report of Session 2015-16. London: The Stationery Office Limited.
Hudson, Phillips, Ray, Vegeris & Davidson (2010) The influence of outcome-based contracting on provider-led pathways to work. London: Department of Work and Pensions

Hunter, D. & Breidenbach-Roe, R. (2013) Payment by Results contracts: a legal analysis of terms and processes London: BWB/NCVO

Johnson, R. (2015) Dos and Don'ts from Down Under.
http://buyingqp.com/2015/10/07/dos-and-donts-from-down-under/

Karve, A. et al. (2007) Potential unintended financial consequences of
pay-for-performance on the quality of care for minority patients. American Heart Journal March 2008

Keohane, N. & Downing, A. (2013) Breaking bad habits: Reforming rehabilitation services. London: Social Market Foundation.

Knight-Markiegi, A. & Quinn, A. (2013) Comparing payment by results across public services and in housing related support. London: Sitra.

KPMG (2014) Evaluation of the joint development phase of the NSW Social Benefit Bonds trial. Sydney: Government Advisory Service

Lagarde, M.; Wright, M.; Nossiter, J. & Mays, N. (2013) Challenges of payment-for- performance in health care and other public services – design, implementation and evaluation. London: Policy Research Unit in Policy Innovation Research

Lane, P.; Foster, R.; Garinder, L; Lanceley, L. & Purvis, A. (2013) Work Programme evaluation: procurement, supply chains and implementation of the commissioning model. London: Department for Work and Pensions.

La Valle, I.; Gibb, J; Brzyska, B.; Durbin, B; Sharp, C.; Aston, H & Rutt, S. (2011) Feasibility study for trials of payment by results for children's centres. London: NCB

Lourie, J. (1997) Training and Enterprise Councils. London: House of Commons Library Research Paper 97/48

Mansour, J. & Johnson, R. (2006) Buying quality performance: procuring effective employment services. London: Ingeus/Work Directions

Make Justice Work (2012) Just Results: payment by results and community sentencing

Mason, P., Fullwood, Y., Singh, K. & Battye, F. (2015) Payment by Results: Learning from the Literature. Birmingham: ICF International.

Mason, T. et al. (2015) The impact of paying treatment providers for outcomes: difference- in-differences analysis of the payment by results the drugs recovery pilot. Addiction, 110, 1120 – 1128

McLellan, A.; Kemp, J.; Brooks, A. & Carise, D. (2008) Improving public addiction treatment through performance contracting: The Delaware experiment. Health Policy 87 pp. 296-308

McNeil, Clare & Hunter, Jack. (2015) Breaking boundaries: towards a "troubled lives" program for people facing multiple and complex needs. London: Institute for Public Policy Research

Mental Health Australia (2015) Discussion paper: options for commissioning and funding of mental health services. Deakin: Mental Health Australia

MoJ (2015b) Annex A Interim reconviction figures for Peterborough payment by results pilot. London: MoJ Statistics Bulletin 29 October 2015

MoJ (2015) Proven re-offending statistics quarterly bulletin July 2012- June 2013. London: MoJ Statistics

Bulletin. Annex A provides interim reconviction figures for Peterborough and Doncaster PbR pilots.

MoJ (2014) Peterborough Social Impact Bond, HMP Doncaster: Payment by results pilots. Final re-conviction results for cohorts 1. London: MoJ Statistics Bulletin.

MoJ (2013a) Principles of competition – Transforming Rehabilitation Programme. London: MoJ.

MoJ (2013b) Transforming Rehabilitation Programme – payment mechanism: market feedback and development considerations. London: MoJ.

Moss, Ian (2010) The state of commissioning: preparing Whitehall for outcomes-based commissioning. London: Institute for Government

Mulheirn, I. (2013) Paying for results? Rethinking probation reform. London: Social Market Foundation.
Mullen, K.; Frank, R. and Rosental, M. (2009) Can you get what you pay for? Pay for performance on the quality of healthcare providers. Cambridge, MA: National Bureau of Economic Research, working paper 14886

National Audit Office (2015) Outcome-based payment schemes: government's use of payment by results. London: HC 86 Session 2015-16

National Audit Office (2014) Report by the Comptroller and Auditor General The Work Programme HC 266 Session 2014-15 2 JULY 2014

National Audit Office (2013) Programmes to help families facing multiple challenges: report by the Comptroller and Auditor General HC 878 Session 2013 – 14 3 December 2013

National Association of Youth Justice (2011) Payment by results and youth justice system

National Drug Evidence Centre, University of Manchester (2014) Evaluation of the drugs and alcohol recovery payment by results pilot program: interim summary report. Manchester: NDEC

Newton, B. et al. (2012) Work Programme evaluation: Findings from the first phase of qualitative research on programme delivery. London: DWP.

Nuffield Trust (2014) The NHS payment system: evolving policy and emerging evidence. Research Report.

Pearce, S. et al. (2015) HMP Doncaster payment by results pilot: final process evaluation report. London: Ministry of Justice Analytical Series

Petersen, L. et al. (2006) does pay for performance improve the quality of healthcare? Annals of Internal Medicine Volume 145 Number 4 pp 265 – 272

Policy Exchange (2013) Expanding Payment-by-Results: Strategic choices and recommendations. London:

Policy Exchange.

Probation Chiefs Association (2013) Transforming Rehabilitation briefing. London: PCA

Rees J., Taylor R., & Damm C (2013a) Does sector matter? Understanding the experiences of providers in the Work Programme. Birmingham: Third Sector Research Centre.

Rees, J.; Whitworth, A. & Carter, C. (2013b) Support for all in the UK work programme? Differential payments, same old problem... Birmingham: Third Sector Research Centre

Reform (2011) Administering justice by results: Reform roundtable with Crispin Blunt MP, Parliamentary Under-Secretary of State for Prisons and Probation

Revolving Doors Agency (2015) Adding Value? Reflections on payment by results for people with multiple and complex needs

Roberts, M. & Singleton, N. (2011) Payment by Results co-design group consultation draft outcomes for piloting PbR for recovery - Joint Response from DrugScope and the UK Drug Policy Commission.

Roberts, M. (2011) By their fruits: applying payment by results to drugs recovery. UK Drug Policy Commission

Sheil, F. & Breidenbach-Roe, R. (2014) Payment by results and the voluntary sector. London: NCVO
Social Finance (2012) Social Impact Bonds: Enabling long term recovery from addiction. London: Social Finance.

Sturgess, G. (2015) Contestability in public services: an alternative to outsourcing. Australia and New Zealand School of Government

Sturgess, G. et al. (2011) Payment by Outcome: A commissioner's toolkit. London: 2020 Public Services Trust at the RSA

Van Herck, P et al. (2010) Systematic review: Effects, design choices, and context of pay- for-performance in healthcare. BMC Health Services Research 10:247

Webster, R. (2012) Second Commandment of PbR: First do no harm. www.russellwebster.com

Webster, R. (2013) The 10 commandments of payment by results. Amazon.

Werner, R.; Kolstad, J; Stuart., E. & Polsky, D. (2011) The effect of Pay-For-Performance in hospitals: Lessons for quality improvement. Health Affairs, 30, no.4 (2011):690-698

Whitfield, D. (2015) Alternative to private finance of the welfare state: The global analysis of social impact bonds, pay-for-success and development impact bond projects. Adelaide: Australian Workplace Innovation and Social Research Centre and European Services Strategy Unit

Wong, K. Ellingworth, D. & Meadows, L. (2015) Local Justice Reinvestment Pilot: final process evaluation report. Ministry of Justice Analytical Series.